ELAN - Linguistic Annotator

version 7.0

ELAN - Linguistic Annotator: version 7.0

Table of Contents

List of Figures

List of Tables

ELAN - Linguistic Annotator

version 7.0

The Language Archive, MPI for Psycholinguistics, Nijmegen, The Netherlands

This manual was last updated on 2025-06-30.

The latest version can be downloaded from: https://archive.mpi.nl/tla/elan.

Original Author: Birgit Hellwig

Updates for version 2 and higher: Dieter Van Uytvanck

Updates for version 3.2.0 and higher: Micha Hulsbosch

Updates for version 4.0.0: Aarthy Somasundaram and Micha Hulsbosch

Updates for version 4.1.0: Maddalena Tacchetti and Aarthy Somasundaram

Updates for version 4.1.1 and higher: Aarthy Somasundaram

Updates for version 4.4.0 and higher: Jeroen Geerts

Updates for version 5.0 and higher: Han Sloetjes

Updates for version 6.9 and higher: Divya Kanekal

Introduction

ELAN (EUDICO Linguistic Annotator) is an annotation tool that allows you to create, edit, visualize and search annotations for video and audio data. It is developed at the Max Planck Institute for Psycholinguistics, Nijmegen, The Netherlands, with the aim to provide a sound technological basis for the annotation and exploitation of multi-media recordings. ELAN is specifically designed for the analysis of languages, sign languages, and gestures, but it can also be used by anyone who works with media corpora, i.e., with video and/or audio data, for purposes of annotation, analysis and documentation.

ELAN supports:

- the display of speech and/or video signals, together with their annotations;

- the time linking of annotations to media streams;

- the linking of annotations to other annotations;

- an unlimited number of annotation tiers as defined by the users;

- different character sets;

- the export as tab-delimited text files and multiple other formats;

- the import of multiple file formats;

- search options.

This manual helps you to understand and use the features of ELAN.

Part I is the user's guide. It is organized around the following topics:

- ELAN documents

- the screen display and the navigation through a document

- annotations

- working modes

- search options

For each topic, basic information is given. Following that, the use of features is explained in a step-by-step way. It is recommended that you read relevant chapters before starting to work with your own data.

Part II is the reference guide, it provides brief information on the following topics:

- mouse options

- menu items

- shortcut keys

- key concepts of ELAN

An overview of the differences between the succeeding versions of ELAN can be found online via: https://archive.mpi.nl/tla/elan/release-notes.

Notation Conventions

The following notation conventions are used:

- Menu items, icons and screen displays are written in a bold typeface (like this line).

- (Shortcut) keys are written in **SMALL CAPS**.

📒 **Note**

Information on troubleshooting is marked with an icon like this one.

Recent changes to this manual

2025-06-20: updates for 7.0, new audio spectrogram detail window and hunspell spell checking added to the transcription mode

2024-11-21: updates for 6.9, improvements on the ELAN Lexicon Service for Global Signbank and improvements on interlinearization mode

2024-05-21: updates for 6.8, changes concerning the automatic backup system and a new option in the delimited text import function

2023-10-26: updates for 6.7, a new preference and subtitle export option

2023-08-10: updates for 6.6, function to modify the boundaries of all annotations, new options in the import from FLEx, preliminary Dark Mode for the cross-platform Look and Feel

2023-04-13: updates for 6.5, new participant oriented functions: add dependent tiers to the tier hierarchy of each participant, delete all tiers of selected participants, import of annotations from an Annotation Server, export of statistical information on the annotated portion of recordings, recognizer that interacts with Whisper

2022-06-21: updates for 6.4, WebLicht with toolchain file, opening linked external resources in the browser or other default application, export of annotations to an Annotation Server, audio extraction from video with FFmpeg (Linux and macOS), forcing of video aspect ratio for vertical (portrait) videos too

2022-01-24: updates for 6.3, new annotation by entering start and end time, first implementation of a Spectrogram viewer, export of before and after column in multiple file search

2021-06-22: updates for 6.2, import of WebAnnotation JSON files, simple statistics with word list export, more extensive and accessible logging from native libraries, CV entry colors in annotation density plot

2021-03-02: updates for 6.1, waveform view without wave file, sorting of single file search results, new "no annotation" constraint in structured multiple file search

2020-11-25: updates for 6.0, added export to time-aligned interlinear text and export to WebAnnotation JSON, inter-annotator reliability based on Fleiss' kappa, support for loading remote EAF files, new preference options

2020-03-16: updates for 5.9, introduction of an annotation density view and export function and of a document properties viewer, updates concerning the possibility to annotate remote media over http(s), changes in configuration of analyzers in interlinearization mode, updates on the analyzers themselves and on lexicon editing, introduction of a new option in media clipping

2019-10-08: updates for 5.8, updated information on available media players, added and changed a few import sections, introduction of EAF validation and file locking options

2019-04-09: updates for 5.5, removed the separate Export clip using m2-edit, minor improvements in the import of separated text files and the export to FLEx, integration of a fully native AV Foundation media player

2018-12-05: updates for 5.4, added import of a Toolbox dictionary, preparation for inclusion of JavaFX media player (and of AV Foundation based player on masOS) and the related move from Java 6 to Java 8

2018-08-20: updates for 5.3, added updating of multiple files with elements from a template, added an option to Merge Transcriptions to not overwrite existing tiers, some changes in dealing with (external) controlled vocabularies

2018-03-28: updates for 5.2, Interlinearization Mode added configuration of the rendering, added a Play Selection panel, added a configuration panel for the Whitespace analyzer, added support for a custom sort order for lexical entries

2017-12-20: updates for 5.1, added Copy annotations from Tier to Tier, added Java Sound player and clipping options, added Modify Annotation Time by typing option, added new data extraction options from timeseries tracks

2017-10-25: updates for 5.0, additions to Interlinearization Mode, updates to Segmentation Mode, new options in tab-delimited text export

2017-04-10: updates for 5.0.0-beta, added Signbank lexicon connection, Media displayer for annotations and controlled vocabulary entries, Spell Checker (preliminary)

2017-01-04: general update for 5.0.0-alpha, added Interlinearization Mode, font size setting added to the preference panel

2016-05-12: general update for 4.9.4, added section on working with tier sets, some additions to the preferences panels

2015-11-25: general update for version 4.9.2, added Korean language, Tab-delimited text / Traditional text / Interrater text export changes/improvements, copying of video coordinates changed, basic drag 'n drop file opening/creation function added, ability to create selections in Segmentation mode added.

2015-05-20: general update for version 4.9.0, added inter-annotator functionality, added WebLicht processing, general maintenance.

2014-12-03: general update for version 4.8.0, added play media section in preferences, context-menu in timeline viewer reduced, option to deactivate tier in Tier name viewer added, Tokenize tier is more flexible.

2014-10-25: general update for version 4.8.0 beta, introduction of ELAN comments section

2014-05-07: general update for version 4.7.0, added multi-lingual CV, frame-skipping options in preferences, added independent audio-level control, added audio signal-viewer easy switch option, improved Multi-file editor.

2013-11-04: general update for version 4.6.2

2013-05-13: update for version 4.6.0, added multiple annotation selection, new alignment-view in multiple file search, FLEx import & export added

2013-01-07: minor fixes to text

2012-12-10: general update for version 4.5.0

2012-07-12: general update for version 4.4.0

2011-08-15: general update for version 4.1.2

2011-04-23: general update for version 4.1.0, introducing the transcription mode

2010-12-23: general update for version 4.0.0

2010-03-12: general update for version 3.8.1

2009-08-20: general update for version 3.8.0, among others the possibility to change shortcuts

2009-04-02: general update for version 3.7.2.

2009-02-03: general update for version 3.7.0, a viewer for integrated display of metadata and a find-and-replace function for multiple files

2008-08-19: general update for version 3.6.0, an extensible Audio Recognizer framework for semi-automatic segmentation and annotation

2008-05-19: general update for version 3.5.0, preliminary support for ISO Data Categories and simplified creation and application of a translation for ELAN's user interface

2008-03-06: general update for version 3.4.0, new customization options and support for timeseries data in csv/tab-delimited text files

2007-12-10: general update for version 3.3.0, among others the options for exporting ELAN data are expanded

2007-10-04: general update for version 3.2.0, among others the structured search through multiple annotation files (This version of the manual is the first to be made from Docbook source to enable an easy generation of PDF and HTML.)

2007-03-27: some cross-references have been corrected

2007-02-22: removal of obsolete section "menu items" in the reference guide

2007-02-20: references to sections are corrected, new screenshots, keyboard shortcuts updated, a lot of small corrections

2007-02-08: general update for version 3.0, among others the new search facilities were added

Chapter 1. ELAN documents

Basic Information

Media Files and Annotation Files

Every ELAN project consists of at least two files: one (or more) media file(s), and one annotation file.

1. One (or more) media file(s):

   - zero or more video file(s) (`*.mpg`, `*.mp4` etc.)

   - and/or one (or more) audio file(s) (`*.wav` etc.)
   The video file allows you to view the video and listen to the sound. If you want to view the waveform as well, you might need to create an additional `*.wav` file through a conversion program that converts the audio data from a video file into a `*.wav` format. As of version 6.1 for Windows and macOS and 6.4 for Linux, attempts are made to show the waveform directly from the audio track of the video (but a `*.wav` file can still be more accurate and perform better). In the case of multiple video files, the audio of the first selected video file is played (audible). The kind and number of supported video formats depend upon the media framework you are using. If your media framework supports a format, the same probably goes for ELAN.

   The following media frameworks are known to work:

   📋 **Note**

   > Which frameworks are actually available depends on the ELAN variant and the operating system.

   - Windows (in order of preference, DirectShow/Microsoft Media Foundation being the best solution):

     - JDS (Java Direct Show)

       - Java - Microsoft Media Foundation (.mp4, .m4a,. m4v (win 7 and higher), .wmv, .wma, .asf)

     - JavaFX Media Framework (.mp4, .wav)

     - VLC Player Library (Java wrapper for VLC Player 3.x)

     - Java Sound (.wav)

   - MacOS (in order of preference):

     - Java - AV Foundation Framework

     - JavaFX Media Framework (.mp4, .wav)

     - VLC Player Library (Java wrapper for VLC Player 3.x)

     - Java Sound (.wav)

   - Linux (in order of preference):

     - VLC Player Library (Java wrapper for VLC Player 3.x)

     - JavaFX Media Framework (.mp4, .wav)

     - Java Sound (.wav)

📝 Note

- There are several issues with the VLC based player. Depending on the platform options like zooming in or frame stepping might or might not work. Spherical or 360-degree .mp4 videos currently don't work on macOS. It may sometimes be necessary to resize the video area to enforce an update of the video image.

- For `*.mov` files (i.e., Cinepak-Quicktime-Movies) it is important that these are self-contained files, i.e., the video information needs to be contained within the `*.mov` file itself. If this is not the case, ELAN will not be able to display the file.

- Unlike other media files, the playback rate of Windows Media Audio (WMA) files cannot be altered.

2. One annotation file:

- an annotation file created by ELAN (`*.eaf` , "EUDICO Annotation Format")

- or an imported annotation file. See the the section called "Import from" section for the supported formats.

All information (e.g., the tier setup, the time alignment, the annotations) is saved to the annotation file only – never to the media file(s).

📝 Note

Take care when editing a media file. Afterwards you probably will want to resynchronize its alignment with the corresponding annotations, as described in (the section called "Synchronizing video files").

Although it's not compulsory it is a good practice to use a common name for media files and the annotation file. So, it is recommended to use e.g. `a.eaf` next to `a.mp4` and `a.wav`.

Imported files also do not need to have the same name as their media files, and they can be located in different directories. All imported files can ultimately be saved as ELAN files (`*.eaf`).

All annotation files (`*.eaf`) can be exported as text, FLEx, Toolbox etc.

Special ELAN data folder

ELAN creates a special folder to store settings, preferences and configuration files and to cache content downloaded from internet. The name and location of this folder depends on the operating system:

- on Windows: <user_home>\.elan_data

where the <user_home> folder resolves to (depending on the Windows version) something like C:\Users\user_login_name

- on Linux: <user_home>/.elan_data

where <user_home> is just the user's home folder

- on Mac OS X: <user_home>/Library/Preferences/ELAN

where <user_home> resolves to /Users/user_login_name. To access the Library folder in the Finder you can hold down the Option (Alt) key when clicking the Go menu. Library will then be visible in the list.

Apart from that ELAN expects to find specific files and folders in its installation folder.

How to start and manage documents

To start ELAN, do the following:

1. Double-click on the ELAN icon (on your desktop, in the Start menu or in the Dock).

   The initially empty window is displayed in which you can open different kinds of documents.

2. Click on the File menu (see Figure 1.1, "File menu").

   The main options in the File menu for creating or opening a file are:

   • New (the section called "Creating a new document")

   • Open (the section called "Opening an existing document")

   • Import several formats are suported (e.g. Toolbox, FLEx, Praat), all described in the the section called "Import from" section

   • Import Multiple Files As for batch conversion of files to ELAN format, described in the the section called " Multiple file import options" section

Figure 1.1. File menu

3. Click on:

- Open... in case you want to open an ELAN file (`*.eaf`) (the section called "Opening an existing document")

- New... in case you want to open a media file in ELAN (e.g. `*.mp4`, `*.mpg`, `*.wav`). This is not for opening an existing annotation file (`*.eaf`) (the section called "Creating a new document").

- Import and then on one of the formats listed in the submenu (the section called "Import from").

Other dialog windows will appear and prompt you to enter the names and locations of the different files. Then the ELAN window appears and displays the selected files.

4. As an alternative to the above, it is possible to open an `.eaf` file by drag-and-drop onto an ELAN window or to create a new document by drag-and-drop of a media file onto an ELAN window.

Once you have started ELAN and opened a document, use the File menu to open, create or import a second document. When done with a document use Close (the section called "Closing a file") to close it or Exit (the section called "Exiting ELAN") to close all files and exit ELAN.

Selecting a user interface language

Elan supports various user interface languages. You can set an interface language at any given time. To do so, choose Options > Language and select one of the available languages.

📝 Note

The selected Language does not influence the content of the produced or edited `*.eaf` files in any way.

At present Catalan, Chinese, Dutch, English, French ,German, Japanese, Portuguese, Russian, Spanish, Swedish and Korean language modules are available. However, new languages can be easily added. If you want to provide a translation for a different language, please contact the ELAN development team.

Alternatively, you can immediately incorporate a new translation as follows. In the directory `locale` under the directory where ELAN is installed, you will find the files `ElanLanguage.properties` and `SearchLanguage.properties`. These files can be used as a basis for your translation. Copy the files to the directory .elan_data (Linux and Windows) or Library/Preferences/ELAN (on Mac OS) in your home directory and simply edit the entries in the files. To view the result of the translation, click Options > Language and select Custom.

Creating a new document

In case you have a media file but no annotation file (`*.eaf`, `*.txt`, `*.trs`), click on File > New. Next, a New Transcription dialog window (see Figure 1.2, "The New Transcription dialog window: media") will be displayed, e.g.:
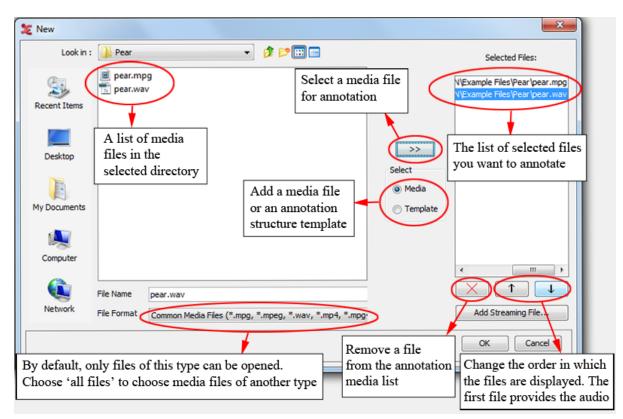
Figure 1.2. The New Transcription dialog window: media

Do the following:

1. Click on the Look in pull down box (on the top left of the window) and browse to the directory that contains the media files.

2. If you want to use media files of another type than the most common formats, select All Files or one of the other format filters in the Files Format dropdown menu. Whether or not a media type is supported depends on your software configuration.

3. Double-click on a displayed media file (e.g. pear.wav) (`*.mp4`, `*.mpg`, `*.wav`, etc.) to select it. It now appears in the rightmost box. Alternatively, you can click on the media file name and click on the >> button afterwards.

4. If you want to use a predefined set of tiers (a template), select the Template radio button and choose the template (i.e. `*.etf`) to be used:

Figure 1.3. The New Transcription dialog window: templates

5. Beside media files on disk you can also add a remote file over e.g. the HTTP(S), RTSP (Real Time Streaming Protocol) or any other protocol a media player framework might support. Click on Add Remote File... and enter or paste the full URL of the remote media. Click on OK.

6. Click OK to open the new annotation document; otherwise click Cancel to exit the dialog window without creating a new file.

An ELAN window containing the new document appears.

> **📒 Note**
>
> The actual appearance of the window(s) shown for starting a new transcription can differ considerably depending on the operating system.

> **📒 Note**
>
> By default the audible sound will be produced by the first media file in the list. Even if both a video file and a `.wav` file have been selected, the audible sound is from the video, if that file is the first in the list. This can be changed in the Volume section of the Controls tab (right top corner of the window).

Alternatively, you can start a new project by simply opening ELAN. Instead of choosing File > New, just browse to the files you wish to work with from within your explorer (e.g. the Finder in macOS, or Windows File Explorer), select them all and then drag n' drop them onto the ELAN main window. A new transcription will be opened containing the selected media-files.

Set the author and license information of a document

An ELAN document can have an author. To set the author, click Edit > Set Author.... Enter the name of the author and click OK. It is also possible to add information concerning the license policy or policies that apply to the document, click Edit > Document Properties...

Figure 1.4. Document properties window

A License can consist of a URL, a license text or both. Multiple Licenses can be specified.

• The New and Remove buttons allow to add a new license or remove an existing license

• The Import button opens a file dialog to select a file from which the text will be read and copied

• The Default button shows a list of more or less standard licenses (Creative Commons, GNU) to choose from

Synchronizing video files

Sometimes one has two or more video files of the same recorded scene, e.g. when 2 different cameras were used. In that case it may happen that both recordings don't start exactly at the same moment. In order to fix this, one should synchronize the videos. Synchronizing is the process of working out the gap (offset) between the start times. When played together, the start point of the earlier source is shifted forwards to this offset. This can be done in two different ways:

- If you know the amount of offset for a video, you can enter it by activating the **Linked Files** dialog window (via Edit > Linked files…
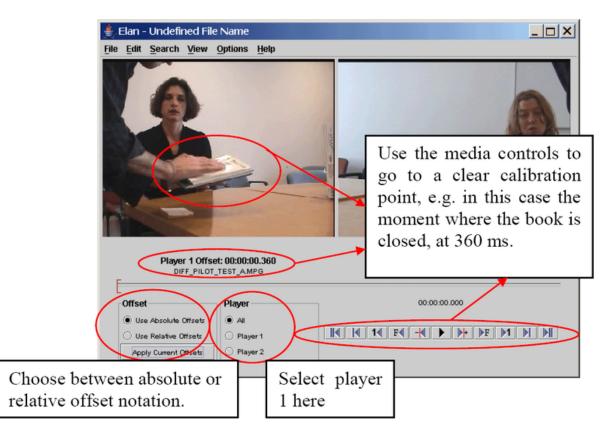
Figure 1.5. Synchronizing video files: Enter offset in 'Linked Files'

Double-click the offset time for the video you want to alter and enter it in milliseconds. Click Apply to save.

- If you do not know the offset time, please follow these steps to synchronize your videos:

1. Open a new document with 2 (or more) video files by selecting both files in the New Transcription dialog window (as seen above).

2. Select the pull down menu Options > Media Synchronization Mode.

3. Make a choice about how the time codes should be displayed:

   a. **Absolute Offsets**: for every video its own timing is being shown.

   b. **Relative Offsets**: the video of player 1 is appointed to be the "master", i.e. the time position of the other videos will be determined as starting point of this file, which starts at 00:00:00.000.

4. Select the radio button Player 1. You can now choose a moment in the video which is easy to calibrate (some clear anchor point, in both of the videos). For instructions how to navigate through the video file, see the section called "How to navigate through a document".

Figure 1.6. Synchronizing video files: Offset of player 1

5. Now select Player 2 and go to the same calibration point.

📋 Note

See the section called "Changing the linked media files" for changing the order of the videos, i.e. the order of appearing in Player 1, Player 2, etc.
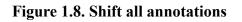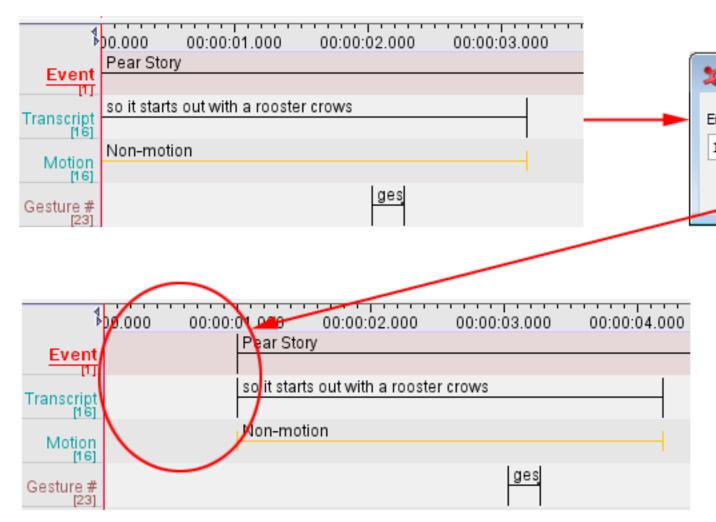
Figure 1.7. Synchronizing video files: Offset of player 2

6. Finally, choose Apply Current Offset. By selecting the play button both videos will be played together now, so you can check if the synchronization between them is correct. If not, please repeat step 3-5 until the result is satisfactory.

7. Leave the synchronization mode by selecting Options > Annotation Mode. Now you are ready to start entering annotations.

8. By double clicking on a video, it will be placed in the leftmost video window (which is also the biggest one in case there are 3 videos).

📄 **Note**

If you changed the media file synchronization of a file that already is annotated, you might want to move the annotation units all together to the right (later, positive value) or to the left (earlier, negative value) on the time axis. This can be done using the Annotation > Shift all annotations … menu (see also the section called "Activating and deactivating the Bulldozer mode or Shift mode"):

Figure 1.8. Shift all annotations

This process won't delete any annotation. If the annotations are shifted to the left, the maximum shift will be restricted by the leftmost annotation unit.

Opening an existing document

In case you have an ELAN file (`*.eaf`) of a media file (`*.mp4`, `*.mpg`, `*.wav` etc.), click Open in the File menu.

The Open dialog window will be displayed:

Figure 1.9. Open dialog window

Do the following:

1. Browse to the directory that contains the ELAN file (`*.eaf`).

2. Double-click on the annotation file to open it.

   An ELAN window containing the document will appear.

Alternatively, instead of clicking File Open, you can drag and drop an `*.eaf` file directly from your file-explorer (E.g. Finder in OSX or Windows explorer) onto the ELAN main screen. The document will open and an ELAN window with the document will appear.

You can only open files of the ELAN annotation format (`*.eaf`). If you try to open a file of a different format, the following error message will appear:

Figure 1.10. Error message: no eaf file

📄 Note

If ELAN cannot find the associated media files (`*.mpg`, `*.mpeg`, `*.mp4`, `*.mov`, `*.wav` etc.), it will check if these files exist in the directory of the `*.eaf` file. If they are still not found there, it will ask you where the media files are located.

Figure 1.11. Locate media file

Opening a remote ELAN file (`*.eaf`)

To open an existing ELAN file which is available online, at an accessible remote location, select File > Open Remote File... and enter or paste the address or URL in the box that appears.

Figure 1.12. Open remote file

Click OK to open the file. Due to network latency opening a remote file can take considerably longer than opening a local file.

Re-open recently accessed files

A convenient way to reopen a file that you have previously been working on is the File > Open Recent File dropdown menu. The maximun number of files shown in the list, can be set in the User Interface panel of the Edit Preferences window (See the section called "Editing preferences").
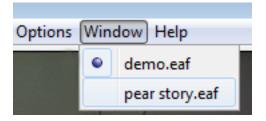
Restoring a backup file

Backup files created by the Automatic Backup system (see: the section called "Creating automatic backups") can be opened and restored by selecting File > Restore Backup File.... This will open a file dialog which allows to select and open a file with the extension `*.eaf.001` (etc.). If the file successfully opens, it can be saved via the Save and Save As actions, both will show a Save As file dialog.

Switching between multiple files

If you open up several annotation files, you can get an overview of the currently opened documents under the Window dropdown menu:

Figure 1.13. Windows dropdown menu

Click on one of the files to select it. Or use the keyboard shortcuts **SHIFT+DOWN** or **SHIFT+UP** to activate the next or previous window in the list.

Saving a document

All documents can be saved as ELAN files (`*.eaf`, ELAN Annotation Format). This includes documents that were created by ELAN itself (see the section called "Creating a new document") as well as documents that were imported into ELAN from Shoebox/Toolbox, CHAT, Transcriber etc.(see the section called "Import from"). To save a document as an ELAN file:

1. Either use the menu options:

   a. Click on File menu.

   b. Click on Save or Save as.

   c. The file chooser will suggest a file name, e.g. based on a linked media file.

2. Or use the shortcut key **CTRL+S**.

📝 **Note**
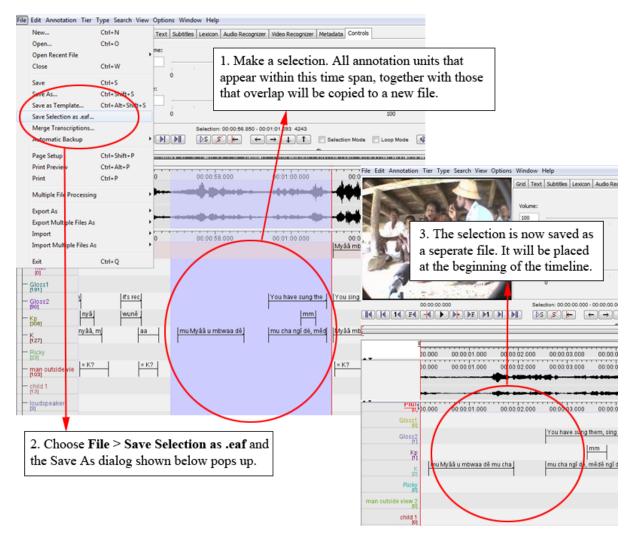
Apart from the `*.eaf` file, a `*.pfsx` file will be written as well. This file contains user- and document- specific settings like the font size used to display text. The `*.pfsx` file can, however, be safely removed as it does not contain any annotation data.

You also can save in .eaf version 2.7 This is the old version of .eaf (prior to ELAN 4.7) If you have used the controlled vocabularies for instance in ELAN 4.7, and save to eaf version 2.7, you may lose some information (colors may not be remembered for instance).

Saving a selection as .eaf file

Apart from saving a whole document you can also store the contents of a certain time span to an .eaf file, using the following steps:

Figure 1.14. Save selection as .eaf file

📄 Note

If annotation units overlap with the selection, they will be shrunk until they fit within the selected interval.

Figure 1.15. Save As Dialog

1. Enter a file name for the new `*.eaf`.

2. Check Clip media for the selection using the script to also clip the media for the selection made and link the new clipped media in the new `*.eaf` file. (For more details on *clipping* the media see the section called "Media clip using script" .)

3. Click on save to create a new `*.eaf` file.

Validating an EAF file

Sometimes a file won't open or behaves strangly when edited. Although there might be messages in the log file, it is often unclear what the problem is. It is possible to validate an `*.eaf` file and receive a report of errors found. This process performs an XML validation and additionally checks the file for the most common other errors encountered in `*.eaf` files. To start the process:

1. Choose File > Validate EAF File...

2. In the Open dialog select the file to check and click Open

3. A small progress indicator will be shown

4. Finally a window displaying the process report appears:

Figure 1.16. Validation report window

```
Process Report                                                    ×
┌─ Process Report ─────────────────────────────────────────────────┐
│ Checking file: C:\Temp\schema\Val_Test.eaf                       │
│                                                                   │
│ ++ Start of XML validation by SAXParser                          │
│                                                                   │
│ Resolved schema to local resource: /mpi/eudico/resources/EAFv3.0.xsd │
│                                                                   │
│ ERROR:   cvc-datatype-valid.1.2.1: '' is not a valid value for 'NCName'. │
│ Line:    9                                                        │
│ Column:  58                                                       │
│ (...)                                                             │
│ Received 0 warnings and 7 errors                                 │
│ -- End of XML validation by SAXParser                            │
│                                                                   │
│ Checking contents of the EAF file                                │
│ ++ Start of Tier Types (a.k.a. Linguistic Types)                 │
│ -- End of Tier Types                                             │
│ ++ Start of Tiers                                                │
│ Checking tier: CHI                                               │
│ ERROR: annotation with id="a1" refers to non-existing time slot with id="ts1" │
│ ERROR: annotation with id="a1" has no TIME_SLOT_REF2 (end time) │
│ Checking tier: FA1                                               │
│ Checking tier: vcm@CHI                                           │
│ Checking tier: xds@FA1                                           │
│ -- End of Tiers                                                  │
│ ++ Start of Controlled Vocabularies                             │
│ -- End of Controlled Vocabularies                               │
│                                                                   │
└───────────────────────────────────────────────────────────────────┘
  Save    Close                          Font Size: ...   —    +
```

5. The report can be saved as a text file

Not all possible errors are detected, just the most common ones:

- consistency of a tier's type and its place in a tier hierarchy

- consistency of tier type and type of annotations (alignable or reference)

- consistency of several annotation properties (time slot references, validity of time values, overlapping annotations etc.)

- some controlled vocabulary properties

Errors in the file are not repaired by this process; if the file still opens in ELAN, it might be possible to apply the necessary changes, otherwise the file might need to be corrected in a text- or XML-editor.

Merging transcriptions

Under some circumstances it might be useful to combine the contents of two separate transcription files into a single one. To achieve this, follow these steps:

1. Choose File > Merge Transcriptions

2. A dialog window appears:

Figure 1.17. Merge transcriptions dialog window

3. If one of the files to be merged is currently opened, select Use current transcription. Otherwise choose Browse… and select the first `*.eaf` file.

4. Choose the second file by clicking the Browse... button.

5. Check Append Annotations to choose one of the options below. Otherwise the annotations will be added to the very left of the first `*.eaf` file (i.e. as a result the second file's annotations are followed by the first file's annotations).

   • Select after the media in first source file to append the annotations of the second file after the media duration of the first `*.eaf` file.

   • Select after the last annotation in the first source file to append the annotations of the second file after the end time of the last annotation of the first `*.eaf` file (please note, the last annotation does not always end at the time the video file ends but can occur before that time).

   • Select after the given time position to append the annotations of the second file after a provided point of time in the first `*.eaf` file (hence after a *given time position*).

6. Check Add linked media and secondary files if you would like to add the media files from the second source to the list of linked files from the first source. This is helpful if you are merging two different projects which contain different media files.

7. Enter a file name for the result of the merge operation.

8. Click on Next. A **Merge Transcriptions** dialog window will appear:

Figure 1.18. Merge transcriptions dialog window

The tiers of the first source are shown as a reference; these don't have to be selected because the first source is always copied completely. The sort buttons allow to list the tiers alphabetically, ascending or descending. The second list shows the tiers of the second source. They can be selected individually or all at once through the Select All button. This list of tiers can be sorted as well, independently of the tiers of the first source.

9. Select the tiers of the second source file that you want to merge with the first file.

10. If there are common tiers (tiers with the same name) in both files and you want to allow annotations of the second file to overwrite those of the first, make sure Allow existing annotations to be overwritten is checked. If this option is not checked, only those annotations of the second source will be added to the tier in the first source, that do not overlap any existing annotations.

11. If there are tiers with the same name in both files and you don't want to merge those tiers, you can select the Merge transcripts without overwriting existing tiers option. The tiers from the second source will then be copied with a suffix added to the name. E.g. if there is a tier named **Event** in both files, the tier from the second source will be copied as **Event-1** (continuing numbering until a unique name is found).

12. Click on Finish to start merging.

13. When the merge procedure has been finished you can choose whether to open the result immediately in a new ELAN window:

Figure 1.19. Open new transcription

Saving a template

A template offers the possibility to reuse the same document setup for more than one media file. This includes:

- tier types (see the section called "How to define a tier type")

- controlled vocabulary (see the section called "Associating a tier type with Controlled Vocabularies")

- tiers (see the section called "How to define a tier and its attributes")

Saving a template is done as follows:

1. Click on File menu.

2. Click on Save as Template…

3. Choose a file name ending in .etf

4. Click on Save

See the section called "Creating a new document" for instructions on using a template.

When saving a template a preferences file is created alongside of it. This preferences file will be used when a new document is created on the basis of the template.

Changing the linked media files

As from ELAN version 2.4, the possibility exists to explicitly change the list of media files that are linked from an `*.eaf` file. This option can be handy if, for instance, you have moved media files to another location after the last time you edited an ELAN file or renamed one or moreof the linked media files.

Activating the **Linked Files** dialog window (via Edit > Linked files…) will get you the following screen:

Figure 1.20. Linked files dialog window

The following options are available on the *Linked Media Files* tab:

- Add…: add a link to a new media file from the current `*.eaf` file.

- Remove: remove the selected media file.

- Update…: specify a new location of the selected file. This is especially useful if the checkbox Status is not marked. The latter indicates the media file could not be found while the ELAN file was opened (e.g. because the media files was moved).

- Set Master Media: make the selected media file the *Master Media*.

- Set Extracted from…: indicate that a sound file has been extracted from a video file.

- : moves a file up/down in the linked file list. The file on top automatically becomes the *Master Media* file. The audio file on the highest location is displayed in the Waveform Viewer.

📑 Note

By default the audible sound will be produced by the first media file in the list. Even if both video and audio (e.g. `.wav`) files have been linked, the audible sound is from the first file in the list. This can be changed in the Volume section of the Controls tab (right top corner of the window).

The *Linked Secondary Files* tab shows files that are linked as secondary files. In particular, files that contain data that need to be displayed by the Timeseries Viewer (see the section called "The Timeseries Viewer") are found here, but other files may be linked as well. The following options are available:

- Add…: add a link to a new file to the current `*.eaf` file.

- Remove: remove the selected file.

- Update…: specify a new location of the selected file. This is especially useful if the checkbox Status is not marked. The latter indicates that the file could not be found while the ELAN file was opened (e.g. because the media files was moved).

- Set Associated With...: associate the file with another linked file.

Creating automatic backups

ELAN allows you to create automatic backup copies. To create backups, do the following:

1. Click on File menu.

2. Go to Automatic backup.

3. Click on one of the submenus:

   - `On`

     to switch automatic backup on, application wide (i.e. for all open documents), using the current settings

   - `Off`

     to switch automatic backup off, application wide

   - `Backup Settings...`

     to open the Automatic Backup preferences panel, where related settings can be set. For a description of the settings see: Automatic Backup Preferences [42]

Figure 1.21. Automatic backup

Backup files are saved with the extension `*.eaf.001` (`*.eaf.002` etc.). Such backup files can be opened and restored via the Restore Backup File... menu option (or by renaming its extension to .eaf by removing the `.00n` suffix on the file system. It is possible and advised to use a pool of backup files.

📝 **Note**

Automatic backups can only be made *after a file has been saved*! If you did not save your file before, a warning window will be shown when the backup should be made for the first time, urging you to save the file first.

📝 **Note**

In case of a crash or forced quit of the application, it is best to first create a copy of the backup file(s) of the files that were open at the time of the crash, before reopening those files. When

the original file is damaged (but can still be opened), the damage can propagate to the automatic backup file(s) once the backup cycle starts again.

Printing

1. Printing from within ELAN can be achieved by selecting the File > Print menu.

2. Then a standard Print dialog window is shown, choose OK to start printing.

Figure 1.22. Printing

Configuring the page settings (page setup)

Through File > Page Setup, you can alter the paper size and other settings such as orientation, etc. of the pages to be printed.
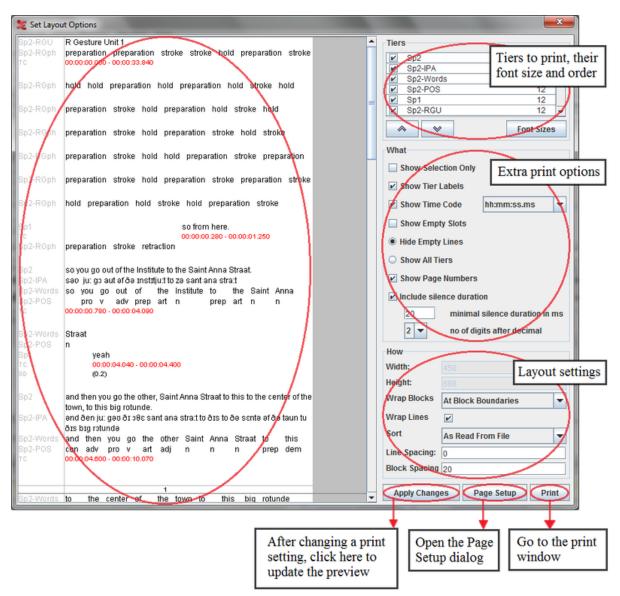
Figure 1.23. Page setup

Previewing the printed pages

The fine tuning of the print result can be done by opening the Print Preview window, which is accessible via the File > Print Preview menu.

Figure 1.24. Print preview

Tiers settings:

- Put a check mark in front of all the tiers that should be printed.

-
  Arrange the order of the tiers with the buttons.
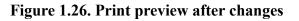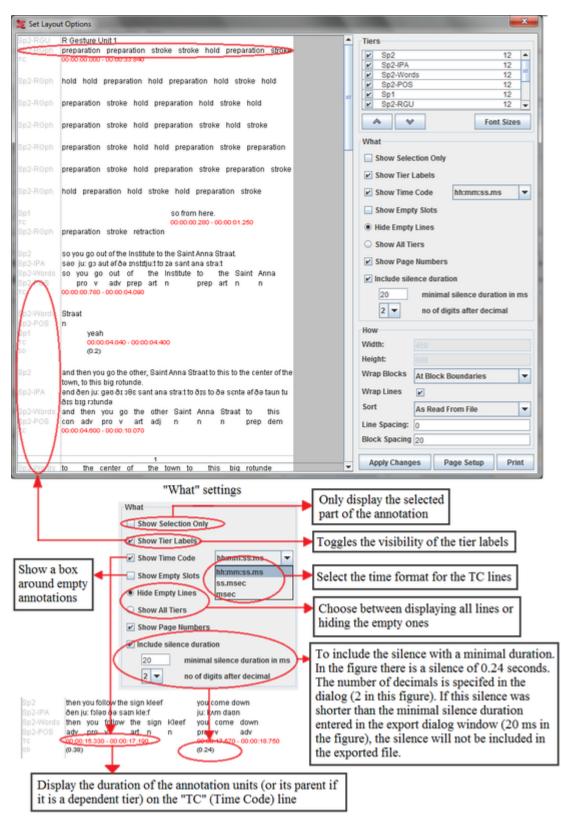
- Advanced Selection Options: click this button to get an advanced selection dialog window (see the section called "How to select tiers").

- The font size of the tiers can be adapted by clicking on the Font Sizes button. A new window will appear:

Figure 1.25. Font sizes
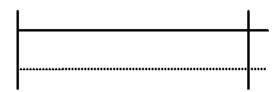
After choosing the desired font size, click on the Apply Changes button in the Print Preview window. After that, the changes will appear:

Figure 1.26. Print preview after changes
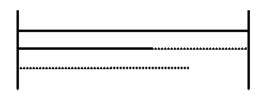
- Width: specify the width of the printed area (in pixels). This value can only be changed by selecting a paper format in the Page setup dialog (see the section called "Configuring the page settings (page setup)").
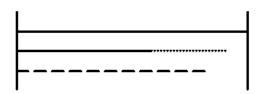
- Height: enter the height of the printed area (in pixels). If you leave this empty, the default height will depend upon the selected paper size.

- Wrap Blocks:

  - *No wrapping*: use 1 line for each tier, only usable for files that contain a small amount of annotations.
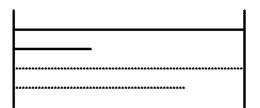
  - *Within block*: wrap blocks, and continue with a new block on the same line if there is space left.

  - *At block boundaries*: wrap blocks, and continue with a new block on the same line if there is space left and if the new block fits on that line.

  - *Each block*: wrap blocks, and start on a new line if a block ends.

- *Sort*: specify in which order the blocks will appear. This is similar to the tier sorting function (see the section called "Sorting tiers").

- *Line spacing*: amount of white space between the lines (default: 0 pixels).

- *Block spacing*: amount of white space between the blocks (default: 20 pixels).

Opening a wave file in Praat

ELAN offers the possibility to open wave files (or a part of them) in Praat (see praat.org [http://www.fon.hum.uva.nl/praat/]). To achieve this, follow the steps below:

1. If desired, make a selection first (See the section called "How to make a selection")

2. Right-click in the waveform viewer

3. Choose Open file in Praat or Open selection in Praat

4. If you haven't specified the location of the Praat program yet, you will have to locate them now in the file dialog

5. The Praat program containing the wave file appears
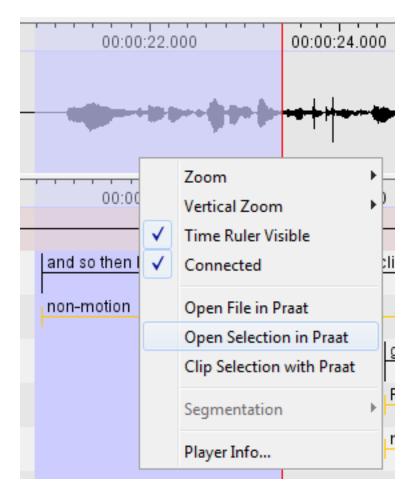
Figure 1.27. Opening a wave pattern in Praat

⊟ Note

> Make sure you are using a recent version of Praat (preferably 6.x or higher), otherwise this feature might not work.

Exporting a selection to a wave file with Praat

Similarly to opening a selection with Praat you can also carve out a selection and save it as a separate wave file. Make a selection (see the section called "How to make a selection") and right-click in the waveform viewer. Choose Clip Selection With Praat. The selected part will now be stored in the same folder as the original `*.wav` file, with a suffix like _23718_25110.wav, the numbers represent the begin and end, expressed in milliseconds. The selection will also be opened in a new Praat screen.

Closing a file

To close a file that is being viewed/edited without exiting ELAN use the File > Close menu option. Alternatively you can also close it by clicking on the arrow in the right upper corner or by pressing **CTRL+W**.

Exiting ELAN

To quit ELAN wand to close all opened windows, press **CTRL+Q** or do the following:

1. Click on File menu.

2. Click on Exit.

If you exit ELAN without having saved the changes (see the section called "Re-open recently accessed files"), the Saving transcription dialog window appears, e.g.:

Figure 1.28. Save on exit

Check mark the files for which you want the changes to be saved. Click OK to save the changes or click Cancel to return to ELAN.

Editing preferences

Preferences can be changed in a dedicated dialog window. Click Edit > Preferences > Edit Preferences. Select a category in the list to the left of the dialog window. The categories and their preferences are:

Figure 1.29. Edit preferences

- Editing

  - `Deselecting the inline text edit box commits changes:`

    If this option is not checked (default) changes made to an inline edit box are discarded if you leave the edit box without explicitly committing the changes. This happens for instance if you click outside the current edit box.

    If this option is checked changes are committed if you leave the current inline edit box.

  - `Enter key commits changes in the inline edit box:`

    If this option is not checked (default) pressing **ENTER** will insert a line break (a.k.a. newline) in an inline edit box. To commit the changes you should hit **CTRL+ENTER**.

    If this option is checked **ENTER** will not insert a line break. It commits the changes as if you pressed **CTRL+ENTER**.

  - `Clear Selection after creating or editing an annotation`

    This option makes ELAN clear the selection after creating or editing an annotation

  - `Clear Selection on single mouse click`

    With this option checked (default), a selection that has been made in the timeline will be cleared when you click outside of that selection.

  - `Create new annotations on the dependent tiers when a new annotation is created`

    If this option is selected, annotations on dependent tiers are created automatically when a annotation on a independent(parent) tier is created. By default , this option is not selected.

  - `Creates annotations aligned with video frames`

    When annotations are created, they can be aligned with the video frames by selecting this option.

  - `Snap Annotations`

    If this option is checked, you can specify the maximum value to snap annotations in (ms).

  - `Always center the active annotation`

    This option if checked, always put the active annotation in the center of the viewer in annotation mode.

  - `When copying an annotation, copy:`

    When you want to copy an annotation as text to an external document , you can either choose to copy the annotation + begintime + endtime, or the annotation only. The default is 'copy the annotation + begintime + endtime'. This will copy the annotation, Tier name, begin and end time.

  - `When copying an annotation use the time format of 'Copy Current Time'`

    This option refers to the first option in the Media preferences category.

  - `Default language for multilingual content:`

    Set the preferred default language used in multilingual content, such as controlled vocabularies. When using multiple CVs, the language set here will be the CVs default language. Multilingual

content is based on ISOcat data categories (DCR). CMDI metadata is displayed in the set language, if this is available from the ISOcat category.

- CV

Figure 1.30. Edit CV preferences

Inline edit box with a CV entry:

- Specify the width of the edit box (in pixels)

  In the text field a minimal width for the edit box can be specified. Useful when the default width cuts off too much of the entry values.

- Show CV entry description in the inline edit box

  When checked, the CV entries will be shown as a two-column table with the description of each entry in the second column.

- Specify the percentage of the width of the first column (the CV entry value column)

  In combination with the above option, this determines how much of the width of the box will be reserved for the CV entry values.

Suggestion panel of the CV inline edit box:

- Look for CV entries that contain instead of start with the search string

  When checked, this option will suggest entries from the CV that contain the given search string. When unchecked, only entries that start with the search string are suggested.

- Also look in the descriptions of the CV entries

  With this option checked, descriptions of the CV entries that contain the search string are also displayed in the suggest panel.

- Ignore case

If this option is checked, case will be ignored when matching the search string entered in the CV suggest panel.

`External Controlled Vocabulary:`

- `After changes in an ECV, don't update the annotation value but update the reference to a CV entry`

  When ELAN opens an .eaf file with links to an external CV, it checks by default if the value of each annotation that has a reference to an entry in an external CV, still corresponds to the value of that entry. If the value of the entry was changed, the annotation value is updated accordingly. If the entry is no longer there, the reference is removed.

  By selecting this option, the annotation value will always remain unchanged. Instead, the CV will be searched for an entry with a value equal to that of the annotation and the reference will be updated to that entry or removed if there is no such entry.

`Suggest child annotation CV entries:`

- `Suggest child annotation CV entries based on the parent annotation value`

  By default, this option is unselected. You need to enable this if you want to automatically filter the CV entries for the annotations on the child tier (having symbolic association with the parent tier) based on the parent annotation value. The child tier needs to have the tier type associated with a controlled vocabulary and symbolic association stereotype. Filtering here means that the words in the parent annotation value are matched agaist the CV entries for the child. The automatically suggested entries depend on the below options. The suggestions are based on the options selected in the above Suggest Panel: Look for CV entries that contain instead of start with the search string, Also look in the descriptions of CV entries and Ignore case.

  - `Suggest CV entries based on any of the words in the parent annotation value`

    If this option is selected, the suggestions are based on all individual words of the parent annotation value.

  - `Suggest CV entries based on the exact sequence of words of the parent annotation value`

    If this option is selected,as the name suggests, suggestions are based on the exact sequence of words of the parent annotation value.

    More info about this can be found here: see the section called "Suggesting CV entries based on the parent annotation value"

- Media

· **Figure 1.31. Edit Media preferences**

- `When copying the current media time, use the format`

  A time format can be selected, determining how the media time will be copied to the pasteboard, e.g. `hh:mm:ss.ms` for the hours:minutes:seconds.milliseconds format.

- `Media navigation:`

  - `Frame forward and frame backward jump to begin of next or previous frame:`

    If this option is not checked (default) clicking the frame forward button (see the section called "The Media Player options") will put the crosshair forward by the amount of ms in one frame. So if the crosshair is in the middle of a frame, clicking frame forward will put the crosshair in the middle of the next frame. The same goes for frame backward.

    If this option is checked the crosshair is put at the beginning of the next (or previous) frame no matter where it is in the current frame.

  - `Classic (pre ELAN 4.7) frame forward and frame backward behaviour:`

If you encounter problems with frame by frame navigation in on-going projects, checking this option will revert ELAN back to the classic behaviour. (unchecked by default)

- Video display:

  - All video's the same size, in a single row:

  When there are three or four video's linked, only one of them is displayed big; the others are small. Check this option If you want all video's to be displayed in the same size and in a single row.

  - Place the video / media in the center

  If selected, then the video is placed in the center and the viewers are on the left and right side of the video. By default the video is the left side of the application.

  - Always keep detached video window on top

  This is selected by default, the detached video window is in front of other windows, including those of other applications. When this is deselected, the video is allowed to disappear e.g. when another application is activated.

- Media location:

  - Set default directory for media files

  Click Browse... to set a default directory. ELAN searches this directory for a media file if it fails to find it using the absolute or relative path the .eaf file refers to or the same directory the .eaf file is in.

- The document's Changed flag is set when the media location has changed

  With this checked, you will notice when a media file has changed its location.

- Media clipping with script:

  - Prompt for a clip file name

  When this is checked (default), ELAN will prompt for a file name to save your clipped media file. You can also choose a location to save the file to. More information on using the script for clipping can be found in the section called "Media clip using script"

  - Use annotation value for clip filename where possible

  When this option is checked, in single clip mode the value of the active annotation (if there is one) is used for the output filename, in multiple media file mode the annotation values that are in the exported tab-delimited text file are used for constructing the output filename. Some obvious problematic characters (e.g. /, \, :) in the annotation value will be replaced by underscores, but this can still fail because of an illegal output filename.

  - Only clip the first media file

  This option is unchecked by default. If errors occur clipping multiple video files and a wave file at once, check this and only the first (or master) media file will be clipped.

  - Clip multiple media files in parallel

This is checked by default. The script will run in multiple instances next to each other when clipping the video files and wave files. When this is causing problems, e.g. incorrect output files, uncheck this option and the script will run each instance after another.

- `Controls:`

  - `Show volume control for individual media players`

  Check this if you want to see and use the volume controls for each media player in the project. (see the section called "The Media Player options").

- `Automatically play media:`

  - `When an annotation is activated`

  When checked, the media will play whenever you activate an annotation, by clicking on it.

  - `When marking current time as begin of new annotation`

  Shift-enter can be used as a keyboard shortcut to mark the start (or end) of an annotation. When checked, the media will start playing when Shift-enter is pressed.

- Metadata

  - Check the metadata keys you want to display in the Metadata tab in the main window by default. (See also the section called "The Metadata Viewer".)

- Platform/OS

  Mac OS X:

  - `Use screen menu bar`

  if checked ELAN will use the screen menu bar in Mac OS. Note that this option enforces the Mac Look and Feel.

  - `Use Mac Look and Feel`

  if selected ELAN will use the Mac OS Look and Feel. If the option `Use screen menu bar` is checked, ELAN will use the Mac Look and Feel, regardless of the selected option here.

  - `Cross Platform Java Look and Feel`

  if selected ELAN will use a cross-platform Look and Feel, which is less similar to *native* macOS applications. When this option is selected two more option become available:

    - `Dark Theme`: this option changes the user interface, after a relaunch, to dark mode (experimental).

    - `Bold Font`: by default this Look and Feel uses a bold font for menus and labels etc. This option allows to switch to a regular font variant.

  - `Use Mac File Dialog`

  if checked ELAN will use a dialog which is similar to the native Mac OS file dialog. Otherwise a platform independent (i.e. Java) file dialog.

  - `Media Framework`

  select either

- `Java - AV Foundation Framework` this integrates a fully native AudioVideo Foundation based player.

  - `Log messages at debug level` this informs the native player to produce detailed messages for the purpose of debugging.

  - `Use the stop mechanism of the framework at the end of play selection` the most accurate stop behavior at the end of playing a selection is achieved when it is performed on the lowest level, by the framework itself. But this function is known to lead to crashes on some versions of the OS, in which case it is best to deselect this option here.

- `JavaFX Media Framework` for (.mp4 and .wav).

- `VLC Player Library*` only partial support on macOS, the video rendering is performed in Java.

- `Java Sound (.wav)` an audio player based on the Java Sound API.

- `Extraction of audio samples from video for audio visualization (waveform, spectrogram)`

  If no wavefile is linked, a listed framework might be able to produce the audio samples from a video.

  - `None, do not extract audio from video` select this to disable audio extraction from video for audio visualization. This is now the default.

  - `Use AV Foundation` uses the *native* macOS AV Foundation framework for audio extraction (there are known issues).

  - `Use FFmpeg` select this option to use FFmpeg libraries for audio extraction. These libraries are part of the ELAN distribution.

Windows:

- `Use Windows Look and Feel`

  if checked ELAN will use the Windows look and feel. If unchecked, an independent look and feel will be used (default).

- `Cross Platform Java Look and Feel`

  if selected ELAN will use a cross-platform Look and Feel, which is less similar to *native* Windows applications. When this option is selected two more option become available:

  - `Dark Theme`: this option changes the user interface, after a relaunch, to dark mode (experimental).

  - `Bold Font`: by default this Look and Feel uses a bold font for menus and labels etc. This option allows to switch to a regular font variant.

- `Media Framework`

  select either:

  - `Java - DirectShow Framework`

    `Java - Microsoft Media Foundation (mp4,.m4a,.m4v(win 7 or higher only), .wmv, .wma, .asf)`

- **Correct the video frame when pausing the player** This is checked by default. If unchecked, the crosshair might not jump to correct the frame, but might also be less accurate.

- **Synchronous interaction with the player** This option applies to the Microsoft Media Foundation based player only (the default player for .mp4 and .wmv files). In this Foundation operations like "start", "stop", "jump to" etc. are performed asynchronously by default. If strange jumps or other quirks are experienced while using the player, the synchronous mode can be tried (it is still experimental). Changing this option requires the file to be closed and opened again to take effect.

- **Log messages at debug level** this informs the native player to produce detailed messages for the purpose of debugging.

- **JavaFX Media Framework** for (.mp4 and .wav).

- **VLC Player Library***

- **Java Sound (.wav)** an audio player based on the Java Sound API.

- **Extraction of audio samples from video for audio visualization (waveform, spectrogram)**

  If no wavefile is linked, a listed framework might be able to produce the audio samples from a video.

  - **None, do not extract audio from video** select this to disable audio extraction from video for audio visualization. This is now the default.

  - **Use Microsoft Media Foundation** uses the *native* Windows Media Foundation for audio extraction.

Linux:

- **Select a Look and Feel**

  - **Cross Platform Java Look and Feel**: if selected ELAN will use a cross-platform Look and Feel, which is less similar to *native* Windows applications. When this option is selected two more option become available:

    - **Dark Theme**: this option changes the user interface, after a relaunch, to dark mode (experimental).

    - **Bold Font**: by default this Look and Feel uses a bold font for menus and labels etc. This option allows to switch to a regular font variant.

  - **System Look and Feel**: if selected a Look and Feel is used which closely resembles that of the host environment.

  - **Nimbus Look and Feel**: another cross-platform Look and Feel, described as 'polished'.

- **Media Framework**

  select either:

  - **VLC Player Library***

  - **JavaFX Media Framework** for (.mp4 and .wav).

  - **Java Sound (.wav)** an audio player based on the Java Sound API.

- `Extraction of audio samples from video for audio visualization (waveform, spectrogram)`

  If no wavefile is linked, a listed framework might be able to produce the audio samples from a video.

  - `None, do not extract audio from video` select this to disable audio extraction from video for audio visualization. This is now the default.

  - `Use FFmpeg` select this option to use FFmpeg libraries for audio extraction. These libraries are part of the ELAN distribution.

  *) The VLC based player requires that the VLC Player, version 3.x or higher, is installed on the system in one of the default locations (this depends on the platform). Spherical or 360-degree video (in .mp4) is supported only on Linux and Windows.

- Preferences

  - `Automatic check for updates`

    Selecting this option will enable the automatic checking for new updates, which will run once a month and will inform you of available updates.

  - `Save transcriptions as old EAF 2.7 format`

    Checking this option will make that your projects are saved in the older 2.7 version of eaf and ensures compatibility with older versions of ELAN. Multi-lingual CVs will not be saved with your project!

  - `Lock EAF files when opening them (creates .lock files, possibly hidden)`

    Locking an eaf file will prevent someone else from opening and editing the same file at the same time (in case the file is in a shared location) or prevent yourself from opening the same file twice. There are two important limitations to keep in mind:

    - this mechanism doesn't prevent anyone from opening and editing the file in another application, e.g. a text editor or an older ELAN version (< 5.8)

    - the locking only applies to the situation where the file is opened in the main ELAN window; multiple file operations in ELAN (e.g. the editing variants of the section called "Editing multiple files and analysis of multiple files") still ignore the lock and can read and write files even when they are already open

    In case of a crash of ELAN, .lock files might have to be deleted manually before the eaf file can be opened again. The .lock file is in the same directory as the eaf file and might not be visible by default in a file explorer.

  - `Preferences location`

    It is possible to specify an alternative directory where preferences (.pfsx) files should be stored. Click Browse... and select the preferred directory for ELAN preferences files. By default a preferences file is stored in the same directory as the eaf file.

  - `Location for other (non-Audio/Video) linked files`

    It is possible to specify a default directory where ELAN can look for missing linked files other than primary media files (e.g. metadata, timeseries files). Click Browse... and select the preferred directory, click the X to remove this preference.

  - `Tier Set`

- Work with tier sets

  When checked, this enables the application of tier sets.

- Set default file path for tier sets

  This allows to specify an alternative location for the tier set file. Click Browse... and navigate to a different location than the default (the ELAN data folder). The X restores the default file location.

- User Interface

  - Number of recent items

    select the number of recently edited items ELAN should remember.

  - Change the UI font size

    allows you to set the font size of menus, labels, buttons etc. The slider allows to specify a value between 100% and 200%. This can e.g. be useful on high resolution displays if the font seems too small. This option requires a relaunch of ELAN.

  - Default font for tiers, annotations and CV entries

    it is possible to specify a preferred font, and its size, for tiers, annotations and CV entries, to be used if no other font is specified (e.g. for specific tiers). Initially ELAN either uses the Arial Unicode MS or the (Java) system's standard font. This option is especially useful if the initial font does not provide glyphs for the language or writing system you are working with. This option requires a relaunch of ELAN.

  - Tooltips

    if checked ELAN will show tool tips with information about the data or about the functionality of ELAN, depending on the position of the mouse cursor.

  - Menu Options

    if the Show annotation count box is checked, the number of annotations per tier will be shown in the menus and tier lists in the viewers.

  - Painting strategy for custom timeline-based viewers and components

    this option lets you choose to set unbuffered or buffered painting of components. Imac retina displays can only handle unbuffered (direct) painting, for instance.

    - Paint to a buffer first, then to the screen.

    - Paint directly to the screen (default).

- Viewers

Figure 1.32. Edit Viewers preferences

- Viewers

  - Horizontal scroll speed

    this sets the speed of the horizontal scrolling, which is done with Shift+scrollwheel or by swiping with two fingers on a laptop. (default 50)

  - Color for symbolic annotations

    You can set the color for symbolic annotations, e.g. non time alignable annotations. The default color is set to orange, you can browse for colors and set them as your favourites.

- Timeline:

  - Active Annotation Bold

    if checked the blue frame of the active annotation has a bold line.

  - Reduced Tier Height

    if checked the height of the tiers displayed in the timeline viewer is reduced. The results is that more tiers are visible.

  - Zoom Slider Visible

    if checked a small zoom slider will be visible in the right bottom corner of the viewer.

- Tier order in dropdown menus

  - Show tiers in alphabetical order:

  with this options selected, most viewers with a tier selection dropdown menu will present the tier names in alphabetical order. Many windows and some viewers have their own tier ordering controls and ignore this setting.

- Subtitles

  - Number of subtitle viewers:

  select the number of Subtitle Viewers you wish to display in the Subtitle tab.

- Select Viewers

  If the video is placed in the center, then it is possible to select which viewers should be shown in the left and right pane of the video. Select either Left to the video or Right to the video for each viewer. The order of the viewer in the table also determines their sort order in the tab pane. To sort the viewers use the buttons Move Up and Move Down to rearrange their sort order.

- Automatic Backup

Figure 1.33. Edit Automatic Backup preferences

The Automatic Backup settings are now global, application wide settings, i.e. applying to all open documents.

- Automatic Backup

  On / Off: with these checkboxes automatic backup can be switched on or off, application wide.

- Number of backup files

  This can be set to a value from 1 to 5. This sets the size of the pool of backup files the program iterates over. The extension of the files will be `*.eaf.001` (... `*.eaf.005`). Together with the backup interval setting this determines how long it will take before a backup file is overwritten by a new version.

- `Backup Interval>`

  Sets how often each open file will be backed up, the choice ranges from every minute to every 30 minutes.

- `Backup Files Location`

  - By default backup files are stored in the same folder where the source `.eaf` file resides. This option allows to specify a folder where all backup files will be stored. Use the Browse button to select a destination folder, use the X reset button to restore the default setting.

Import and Export options

Export as

ELAN offers various export options. To export, click on File > Export As and one of the options.

- Toolbox file (the section called "Toolbox file(UTF-8)")

- FLEx file (the section called "FLEx files")

- CHAT file (the section called "CHAT files")

- Tab-delimited text file (the section called "Tab-delimited text file")

- Tiger XML (the section called "Tiger XML")

- Interlinear text file (the section called "Interlinear text file")

- HTML file (the section called "HTML file")

- Traditional transcript file (the section called "Traditional transcript files")

- Time-aligned Interlinear text file (the section called "Time-aligned Interlinear Text")

- Praat TextGrid file (the section called "Praat TextGrid file")

- WebAnnotation JSON file (the section called "WebAnnotation JSON")

- Alphabetical list of words (the section called "Alphabetical list of words")

- SMIL clip (the section called "SMIL clip")

- QuickTime text (the section called "QuickTime Text")

- Subtitle text (the section called "Subtitle Text")

- Tiers for recognizers (the section called "Tiers for recognizers")

- Media clip using script (the section called "Media clip using script")

- Image from ELAN Window (the section called "Image from ELAN Window")

- Filmstrip Image (the section called "Filmstrip Image")

- Annotation Density Plot (the section called "Annotation Density Plot Image")

- Shoebox file (the section called "Shoebox file")

Apart from these export options for single files, ELAN also supports multiple file exporting options. More details regarding these options can be found here: the section called "Multiple file export options"

How to select tiers

Figure 1.34. Tier Selection panel in most of the dialogs

Different ways to select tiers :

- By Tier Names

  Select the tiers by checking the boxes before each tier name.

- By Type

  This tab shows a list of the tier types available in the current transcription. Select the types by checking the boxes before each type name. Selecting the types will select all the tiers of the each selected types. To modify the selected tiers switch back to By Tier Names.

- By Participant

  This tab has a list of all the participants in the transcription. Select the participants by checking the boxes before each type name. Selecting the participants will select all the tiers of the each selected participants. To modify the selected participant switch back to By Tier Names.

- By Annotators

  This tab has a list of all the annotators in the transcription. Select the participants by checking the boxes before each annotator name. Selecting the annotators will select all the tiers of the each selected annotators. To modify the selected tiers switch back to By Tier Names.

- By Languages

  This tab has a list of all the languages in the transcription. Select the language(s) by checking the boxes before each language name. Selecting the languages will select all the tiers of the each selected language. To modify the selected tiers switch back to By Tier Names.

📄 Note

To select multiple tiers, press **Shift** and click on the successive tiers or click and drag the mouse along the tiers to select them

Other options :

•   To sort the selected order of tiers use the ⌃ and ⌄ buttons to move the tiers up and down in the table.

•   Show only root tiers : Check this option to show only the root tiers in the transcription.

•   Select All : click this button to select all the boxes in the current tab.

•   Select None : click this button to de-select all the boxes in the current tab.

•   OK : click on OK to select the tiers

•   Close : click to close the dialog or cancel the changes

Toolbox file(UTF-8)

Similar to exporting a document to Shoebox (see the section called "Shoebox file") ELAN data can be exported to a Toolbox document with an UTF-8 encoding. This export provides more options for output customization.

To export a file into Toolbox, do the following:

1. Click on File menu.

2. Click on Export as > Toolbox File (UTF-8)...

    The Toolbox Export dialog box appears:

Figure 1.35. Toolbox Export dialog window

Only the left part of ELAN tier names containing an @ are identified as tier markers for Toolbox. These markers form a block in the exported file. The right part of the ELAN tier names are identified as participant names. These are exported with the marker ELANParticipant see the figure below:
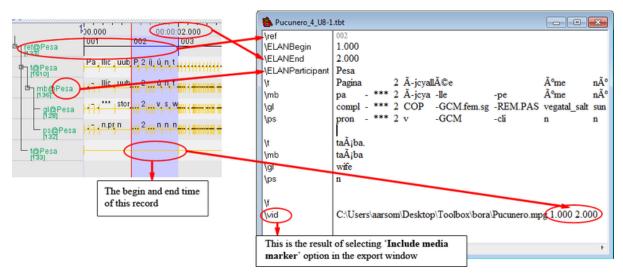
Figure 1.36. ELAN file and exported Toolbox file

If you use a Shoebox *.typ file to specify the Toolbox database type ELAN extracts the database type name from the first line of the type file (e.g. the database type name Text in \ +DatabaseType Text) and puts is in the first line of the exported file (e.g. _sh v3.0 400 Text).

When there is only one root tier (tier without a parent tier) in the transcription (e.g. ref) this will be used as the record marker by default. When there are multiple root tiers "\block" will be added as record marker. In both cases it is possible to specify a custom record marker instead.

Some options not touched up in Figure 1.35, "Toolbox Export dialog window":

- By first selecting a tier(the section called "How to select tiers") and then selecting Insert blank line after this marker you insert a blank line after the selected marker every time the marker is printed in the exported file. The tier name is colored blue in the dialog box.

- By selecting Wrap block you can let ELAN wrap a whole block if one of the lines in a block is longer than a specified number of characters (default is 80 characters). A block in this context refers to the markers that are part of the interlinearization.

- When Wrap blocks is selected it is also possible to select Wrap lines. This applies to long marker lines that are not part of the interlinearization. There are 2 variants: when Wrap to next line is selected the line is split into 2 or more lines that immediately follow each other, regardless of their position in the record. When Wrap to end of block is selected everything beyond the first wrap is placed at the end of the record. Note that wrapped interlinearization blocks are grouped as much as possible.

- When Include empty markers is selected all markers will be printed in each record, whether there is content or not. When this option is not selected a marker will not be printed in a record when it has no content.

- By selecting Add master media time offset to annotation times you can add to the annotation times the time offset from the master media that originated from the synchronization of media files (see the section called "Synchronizing video files").
  Make a choice and click on OK to continue.

3. Specify the name and directory of the exported file.

4. Click Save to export the file; otherwise click Cancel to exit the dialog box without exporting the file.

   The file is exported as a `*.txt | *.sht | *.tbt` file.

If there already exists a file of the same name, ELAN will ask you whether or not it should overwrite the existing file.

5. Open the exported file in Toolbox.

It contains the following information:

a. All tiers and annotations.

Each ELAN parent annotation (including all its referring annotations) corresponds to one Toolbox record. E.g., in the illustration below, the ELAN parent annotation "CLLDCh3R02S01.001" corresponds to the Toolbox record "CLLDCh3R02S01.001".

b. The time code information for each parent annotation.

Each ELAN parent annotation (i.e., each Toolbox record) contains the additional field markers \ELANBegin and \ELANEnd (i.e., the begin and end time of the parent annotation).

This time code information allows you to import the Toolbox file back into ELAN, without having to manually re-align the file (see the section called "Shoebox file").

FLEx files

ELAN allows you to export your project to the SIL Fieldwork Language Explorer software, also referred to as FLEx. The data exchange is realized through .flextext files, a file type that defines several container elements and attributes (see below), onto which ELAN's tiers (via their tier type) and annotations have to be mapped. For the configuration of these mappings the complex, multiple step export window described below, is provided. Configuration will be less complicated in case the .eaf was created by importing a FLEx .flextext file. On import, some FLEx attributes are "encoded" in the names of tiers, on export these attributes are reconstructed by "decoding" the tier names. To better understand the options in the user interface, a simplified representation of the structure of a .flextext file follows here.

```
<interlinear-text>
  <item lang="" type="">...</item>
  <paragraph>
    <phrase>
      <item lang="" type="">...</item>
      <word>
        <item lang="" type="">...</item>
        <morph type="">
          <item lang="" type="">...</item>
        </morph>
      </word>
    </phrase>
  </paragraph>
</interlinear-text>
```

All elements can occur multiple times, e.g. there can always be multiple item child elements for any parent element.

🗒 Note

If your .eaf file contains multiple participants, make sure you have given each participant a name value. You can set a participant value under Tier > Change Tier Attributes....

Choosing File > Export as > FLEx file … will give you the following screen:

Figure 1.37. Export FLEx file step 1

In this screen you can specify:

- which tier type corresponds to which FLEx element

- which tiers should be included in the export

- with the Export interlinear-text tier option, if there is a tier corresponding to the `interlinear-text` element and, if so, which tier it is. This determines whether a tier and its dependent tiers provide the contents for `item` child elements of `interlinear-text`.

- with the Export paragraph tier option, if there is a tier corresponding to the `paragraph` element. If so, its segmentation is used for grouping `phrase` child elements, if not, each `phrase` will be embedded in its own `paragraph` element.

Figure 1.38. Export FLEx file step 2

The second screen allows to:

- map tier types to the `item` child element of the correct, corresponding container element

- specify which tiers should be exported as that `item`

- specify with the Select a tier type for 'morph-type' tiers option, which tier type provides the value for the `type` attribute of the `.flextext` morph element. This should be a valid FLEx morph type. If this option is deselected each `morph` element will be exported with attribute `type="root"`.

Figure 1.39. Export FLEx file step 3

The third screen allows to customize the FLEx `lang` (language) and `type` attributes output:

- the upper part of the screen contains a table and two radio buttons. The buttons enable to switch between tiers and tier types mode (the latter is preferred). The contents of the table is updated after a change in choice. The value of each cell in the type and language column can be selected from a pull-down menu.

- the lower part of the screen allows to edit the list of values selectable in those pull-down menus. The type and language radio buttons determine which list is being updated by either adding new values or removing existing values. The list for `type` is based on a FLEx controlled vocabulary, which could be out-of-date at the time of use, therefore new values can be added manually. The list of languages currently is based on "decoding" the tier names and on the content languages of the tiers. The list can be empty, it should be filled manually in that case.

> 📄 **Note**
>
> FLEx requires that for languages that have both a two letter ISO 639-1 code and a three letter ISO 639-3 code, the two letter code should be used. This screen tries to automatically replace three letter codes by their two letter equivalents where needed, but it is good to check the codes.

- For more information on the structure of FLEx, see Figure 1.69, "FLEx to ELAN structure".

The final screen allows you to save the file as a flextext file, so it can be used in FLEx.

📄 Note

On the third-party resources page of ELAN (https://archive.mpi.nl/tla/elan/thirdparty), you can find a teaching-set which covers the aspects of importing from FLEx to ELAN and back to FLEx.

CHAT files

1. Choosing File > Export as > CHAT file … will give you the following screen:

Figure 1.40. Export Chat file

2. Fill in the necessary fields.

📄 Note

Chat labels must be preceded by * (for root tiers) or % (for dependent tiers). While root tiers have to contain exactly 3 characters, dependent tier names can have up to 7 characters.

3. Click on Export…

4. Fill in a chat file name and choose Save.

Tab-delimited text file

All documents can be exported into a tabular format for purposes of further analysis and/or printing. This includes documents that were created by ELAN itself (see the section called "Creating a new document" and the section called "Opening an existing document") as well as documents that were imported into ELAN from any of the supported formats. Do the following:

1. Click on File menu.

2. Click on Export as > Tab-delimited Text ….

The Export as tab-delimited text dialog window is displayed, e.g.:

Figure 1.41. Export as tab-delimited text dialog window

1. Select the tiers to be exported. (the section called "How to select tiers")

2. Select to export a selected time interval only.

3. Add time offset from the master media to the annotation times.

4. Include header lines with media file location info, include the tier and/or participant names from the output file

5. Annotations sharing the same begin and end time are exported in the same row.

6. Select to include the description of the controlled vocabulary.

7. Select time information and format.

8. Add extra time format expressed in hours, minutes, seconds and frame.

3. By default, ELAN exports all annotations, but it is possible to restrict the export process to selected annotations. The following three options are available:

   a. Export only those annotations that correspond to a selected time interval. Do the following:

      i. In the ELAN window, select the desired time interval (see the section called "Making a selection on an independent tier").

      ii. In the Export as tab-delimited text dialog window, click in the box to the left of Restrict to selected time interval. A check mark appears indicating that this option has been selected.

   b. Export only those annotations that are contained on particular tiers. Do the following:

      In the Export as tab-delimited text dialog window, select those tiers that you want to export. A check mark appears next to any selected tier.

   c. Export only those annotations that (a) correspond to a particular time interval and (b) are contained on particular tiers. To do this, combine the two steps under (a) and (b) above.
   By default, the output contains one annotation per row, with the tier name in one of the columns, time information in several following columns and then the annotation value.

4. By selecting Add master media time offset to annotation times you can add to the annotation times the time offset from the master media that originated from the synchronization of media files (see the section called "Synchronizing video files").

5. The option Include header lines containing media file information allows you to add the media-file path information for each media file to the header of the exported file.

6. The option Separate column for each tier gives each tier its own column in the export file. Annotations that have the same begin time and the same end time are exported in the same row i.e. the same tab-delimited line. The following options allow to also have annotations in the same row if they are not fully aligned but do overlap. As a consequence each annotation can be in the output more than once, making annotation counts unreliable.

   • If you check `Repeat values of annotations spanning other annotations` the spanning annotation is put in each row containing an annotation it spans. The spanning annotation is not in a row by itself.

   • The option `Only repeat within annotation hierarchies` limits the previous option. An annotation is only repeated if it is on one of the ancestor tiers in the annotation hierarchy.

   • The option `Sliced annotation output showing temporal co-occurrences` is an alternative way to repeat annotation values based on overlaps. In this export all unique

begin and end times of all annotations in the export are placed in one list, creating new intervals (between each two successive time values). Each interval is exported if there is at least one annotation overlapping that interval and in the column of each tier the value of the overlapping annotation, if any, is exported.

- The option `Include the annotation id` appends the annotation identifier between brackets to the annotation value (e.g. [a13]). This makes it possible to distinguish annotations in the output, which is hard to do in the case of repeated values.

7. Select the time markers you want to export (begin time, end time and/or duration of every annotation unit).

8. Choose the time format (hh:mm:ss.ms, ss.msec, milliseconds and/or SMPTE time code)

📄 **Note**

If you choose the SMPTE (hh:mm:ss.ff) format, the selected video standard (PAL or NTSC) just indicates the way seconds and milliseconds are converted to frame numbers. This is independent of the actual video standard of the associated video(s).

9. Click OK to start the export process; otherwise click Cancel to exit the dialog box without exporting the annotations.

10. Finally you will see a save dialog window. In the Encoding drop down box a text encoding can be selected (either ISO-latin, UTF-8 or UTF-16). In the file format box there are two options, `*.txt` saves a tab-delimited text file, `*.csv` saves the annotations in a comma separated values file, placing all text values between double quotes. Make an appropriate choice and click on Save.

📄 **Note**

Some Mac applications, like TextEdit, have difficulties to load UTF-8 encoded files. This is most noticeable for "special" characters, e.g. IPA. Using UTF-16 is recommended in that case.

A message appears to inform you that the file has been exported.

The contents and the layout of the exported file depends on the selected options. It can be opened with any program that can handle tab-delimited or comma separated texts, e.g., Microsoft Excel.

Figure 1.42. Tab-delimited text

> 📋 **Note**
>
> Some versions of Excel seem to have problems importing tab-separated files (white rectangles are shown instead of the column borders). As a workaround you can open the text file first in a text editor (e.g. Notepad) and copy and paste the content into Excel.

Tiger XML

If your ELAN annotations contain syntactic elements, it is possible to export these to Synpathy (see https://archive.mpi.nl/forums/t/synpathy-software-information/2649). This function is available via File > Export as > Tiger-xml…

First select out of the candidate tiers the one you want to be exported. Afterwards, map the tiers onto the correct description ("word" or "pos"). Finally enter the name of the file (`*.tig`).

> 📋 **Note**
>
> Synpathy is a tool for annotating, analyzing, and graphically editing the syntactical structure of sentences (e.g. Linguistically annotated text corpora), developed at the Max Planck Institute for Psycholinguistics. The application is based on the SyntaxViewer from the TIGER search project developed by the IMS (Institute für Maschinelle Sprachverarbeitung, University of Stuttgart).

Interlinear text file

This function (File > Export as > Interlinearized Text...) is very similar to ELAN's printing system. Therefore more information can be found in the section called "Previewing the printed pages". The

main difference is that the width of the exported text depends in this case on the number of characters that fits on one line.

Figure 1.43. Maximum line width

After selecting an appropriate layout click on Save as and choose a location and file name. These files can afterwards easily be edited with any text editor (preferably using a fixed-with font). Optionally tick the Insert tabs between annotations box if you prefer to have the white space between annotations to be filled with tabs instead of spaces (especially useful when importing a text file into Word). If Insert tabs between annotations is selected, you could also have single tab instead of multiple white spaces. To do that tick Tabs Instead of Spaces box if you prefer to have tabs instead of multiple white spaces.

HTML file

Similarly to the export to interlinear text (see the section called "Interlinear text file") you can also export annotations to a HTML file, through the File > Export as > HTML... menu.

Figure 1.44. Export as HTML

The only extra option for the HTML export is

• Play media : Check this option to play the media file in the exported html file.

> 📄 **Note**
>
> To play the media HTML 5 is required. It is necessary to place the exported html in the same location as the media file in order to play the file from the html export.

Traditional transcript files

In some situations a straight-forward list of the annotation units, one after another, can be handy. For that cause an export option to a "traditional transcript text" has been added to ELAN. In its simplest form it just will create a text file containing the successive annotations of several tiers, in chronological order. This feature can be found under File > Export as > Traditional Transcript Text....

Figure 1.45. Export Transcript Text

"Restrict to the selected time interval' allows you to export only the data that is currently selected. (see the section called "Making a selection on an independent tier").

'Wrap lines' sets a maximum number of characters before the line gets wrapped.

'Merge annotations on the same tier...' makes it possible to merge annotations on the same tier if the gap in between these annotations is less than a certain amount of milliseconds.

You can number the annotations, each wrapped line, and include or exclude tier labels or participant labels in the export.

One of the options enables you to include silences with a minimal duration. The figure shows there is a silence of 0.2 seconds between '*yeah*' on the tier K-Spch and '*and then you go the other ...*' on the tier W-Spch. The first annotation ends at 00:00:04.400 seconds and the next annotation begins at 00:00:04.600 seconds, resulting in a silence of 0.2 seconds. If this silence was shorter than the minimal silence duration entered in the export dialog window (20 ms in the figure), the silence will not be included in the exported file. The silence duration indication can have 1, 2 or 3 numbers of digits after the decimal.

Empty lines after each annotation (block) can also be included or excluded in the generated output file. Lastly, you can set a fixed width (in number of characters) for the tier labels.

The option to use Jefferson-style alignment based on "[" characters in overlapping annotations, can change the position of parts of annotations by vertically aligning corresponding "[" characters. (Alignment of matching "]" characters is not supported yet.)

Time-aligned Interlinear Text

This export function (File > Export as > Time-aligned Interlinear Text...) produces interlinear output but, unlike standard Interlinear Gloss, the formatting is based on time alignment . This is achieved by using a monospaced (fixed width) font in combination with a customizable *character-to-milliseconds* calculation factor. As a consequence, depending on this factor, the export might cut off part of the annotation value.

The export offers a few text styling options (underline, bold, italic) and the output format is (simple) HTML.

Figure 1.46. Export settings
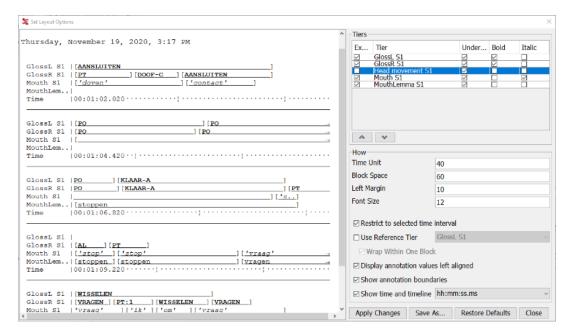
The ouput can be customized in various ways:

• In the top right area of the window is the usual Tiers selection panel. But with additional columns that allow to specify a style per tier. The font style options are underline, bold and italic.

• The remainder of the right area of the window, the "How" panel, contains options to further customize the output:

  • Time Unit the value entered here determines the number of milliseconds one character represents.

- Block Space this is the width of the text block in number of characters. This does not include the margin.

- Left Margin the number of characters for the tier labels.

- Font Size the font size to use for the output.

- Restrict to selected time interval this allows to export only the selected fragment instead of the entire transcript.

- Use Reference Tier when a reference tier is selected, the annotations of this tier are exported, together with overlapping annotations on other selected tiers.

- Wrap Within One Block when a reference tier is used, this option determines whether or not line wrapping is performed within a block. Without wrapping the block width may exceed the specified block space.

- Display annotation values left aligned by default annotations are exported right aligned, with this option the output is left aligned.

- Show annotation boundaries with this option the begin and end boundary of annotations are marked with "[" and "]" characters.

- Show time and timeline with this option a kind of timeline, in text, is added to the output.

- The left half of the screen shows a preview of the output based on the current settings.

After changes in settings the Apply Changes button updates the preview. The Save As... button starts the actual export, currently **html** is the only supported format.

Figure 1.47. Different export settings

Praat TextGrid file

When you wish to work with your annotations in Praat, ELAN enables you to export your annotation to a Praat TextGrid. To do this, click File > Export as > Praat TextGrid.... In the dialog window that appears you can select the tiers you wish to export(the section called "How to select tiers") and specify whether you want to restrict the output to the selected interval.

After clicking OK, you can enter a file name and select an encoding. In addition to TextGrid files in the default encoding for the operating system, ELAN supports Praat TextGrid files with UTF-8 and UTF-16 encoding. Finally click on Save.

WebAnnotation JSON

The preliminary export function File > Export as > WebAnnotation JSON... stores annotations according to the W3C Web Annotation Data Model [https://www.w3.org/TR/2017/REC-annotation-model-20170223/] specifications. This model and format are intended to enable sharing and reuse of annotations across applications and platforms.

Figure 1.48. Export settings and JSON preview

The export window offers a few options to customize the output. Apart from the possibility to select the tiers to export and to only export the selected interval, there are a few format specific options which determine which information is included and how it is structured. After changing settings, the Update button applies the settings and updates the preview on the left side of the window. The Export button initiate the actual export to a `.json` text file.

Alphabetical list of words

Sometimes it can be very useful to have a alphabetical list of (unique) words from one or more tiers. ELAN offers a way to generate such lists. Go to File > Export as > List of Words ... and select the tiers(see the section called "How to select tiers") from which you want to extract the words. The annotations of the selected tiers will be tokenized (split into words) using either a default set of delimiters or a user definable set. Check Count occurrences if you want the list to include the number of occurrences for each token. The Include overall totals in the export file option results in some basic overall statistics at the end of the file. The Include frequency percentages in the export option adds another column to the output, containing the percentage of each unique word (or annotation) of the total word count. After selecting tiers (or better, deselecting unwanted tiers) you can click OK and choose a file name. Clicking Save will save the word list.

SMIL clip

ELAN supports export to SMIL-compliant clips. With a suitable player this enables you to view media files and the associated annotations as a subtitled movie.

> ### 📑 **Note**
>
> For a description of this standard and players see http://www.w3.org/AudioVideo/

Export SMIL for Real Player

1. Select the File > Export As > SMIL > Real Player... menu. This will bring up this dialog box:

Figure 1.49. Export SMIL Real Player

2. Select the tiers you want to export (see the section called "How to select tiers").

3. Check Restrict to selected time interval if you only want to export the current selection. Otherwise the whole media file and associated annotations will be exported.

   • Check Recalculate the begin time of the selected annotations to start from zero if you only want the current selection start time to start from zero.

4. Check Add master media time offset to annotation times to add the annotation times the time offset from the master media that originated from the synchronization of media files (see the section called "Synchronizing video files").

5. Check Minimal duration per subtitle (in ms.) to specify the minimal display duration of a subtitle. For instance, if a annotation is only 0.3 seconds long, but you want to display a subtitle at least 0.5 seconds, enter 500 (ms).

6. Click on Edit Font and Display settings... button. This will bring up this dialog box:

Figure 1.50. Change subtitle text settings

- Click on the respective Browse.. button and select the color from the dialog displayed to set the background color and text color of the subtitle text.

- To set the font of the Text, click on the respective Browse... button and select a font from the font list.

- Font size and the alignment of the subtitle text can be selected from their respective list.

- Click Default button to set the default setting.

- Click on the Apply button to apply the new setting

7. Choose OK to export the clip.

8. Click on the suggested file name to change the location where the SMIL clip will be saved.

Export SMIL for Quick Time

Exporting SMIL for Quick time is very much the same as exporting SMIL for real player (see the section called "Export SMIL for Real Player"). To export SMIL for Quick time, go to File > Export As > QuickTime.... This will bring up a dialog box very similar to export SMIL for Real player . The only extra option which is not available for real player is Merge tiers into one QuickTime text file.If selected, all tiers are merged into one file and if not selected a separate text file will be generated for each tier. It is also possible to set a transparent background for the subtitles. This is done by selecting Transparent background in the dialog (see Figure 1.50, "Change subtitle text settings") which pops up by clicking the Edit Font and Display Settings... button. Finally click on OK to export.

QuickTime Text

Another format you can export to from ELAN is QuickTime subtitle Text. To do this, go to File > Export As > QuickTime Text.... Select the tiers(see the section called "How to select tiers") you want to be included in the subtitles. Optionally specify the following options:

- Restrict to selected time interval: restrict the subtitles to the current selection.

- Recalculate the begin time of the selected annotations to start from zero: recalculates the time of current selection to start from zero

- Add master media time offset to annotation times: add to the annotation times the time offset from the master media that originated from the synchronization of media files (see the section called "Synchronizing video files").

- Minimal duration per subtitle (in ms.): specify the minimal display duration of a subtitle. For instance, if a annotation is only 0.3 seconds long, but you want to display a subtitle at least 0.5 seconds, enter 500 (ms).

- Merge tiers into one QuickTime text file: If not selected a separate text file will be generated for each tier.

- Edit Font and Display Settings... : (see Figure 1.50, "Change subtitle text settings")

- Reuse last custom display settings: when ticked the last used custom font and display settings are automatically applied to the exported text

Finally click on OK. By default the subtitles are stored in a QTtext `.txt` file. If you enter a file name with the extension `.xml` the subtitles are stored in a TeXML - tx3g formatted XML file (the merge tiers option is ignored in that case).

Subtitle Text

Besides the QuickTime subtitle Text (see the section called "QuickTime Text") ELAN can export annotations to there are few other subtitle formats: SubRip (`.srt`), Spruce (`.stl`), Timed Text Markup Language(ttml) (`.xml`) and LRC (`.lrc`) . Click on File > Export As > Subtitle Text... and select the tiers(see the section called "How to select tiers") you want to include in the subtitle file. Specify whether the subtitles should be restricted to annotations in the selected time interval, whether the time of the selected interval should be recalculated to start from zero and if the master media time offset should be added to the annotations times. The third option lets you specify the minimal display duration of a subtitle. For instance, if a annotation is only 0.3 seconds long, but you want to display a subtitle at least 0.5 seconds, enter 500 (ms). Finally there is an option to specify that for each selected tier a separate subtitle file should be created. The default is to export all selected tiers to a single subtitle file.

Figure 1.51. Export as Subtitles text

After you have selected tiers and specified the options, click on OK. Enter a file name in the next window and click on Save.

Tiers for recognizers

Tiers for the recognizers are exported in the AVATech tier format. A brief description of the AVATech tier format can be found in this document: Avatech-interface-spec-2014-03-06.pdf [https://www.mpi.nl/tools/elan/docs/Avatech-interface-spec-2014-03-06.pdf]. Files can be exported as .txt, .csv and xml.

1. Select File > Export As > Tiers for Recognizer... menu. This will bring up this dialog box:

Figure 1.52. Tiers for AVATech recognizers

2. Check Show only root tiers to show only the top level tiers.

3. Select the tiers you want to export. Keep **CTRL** pressed and click to select multiple tiers, press **Shift** and click to select multiple successive tiers.

4. Check Restrict to selected time interval if you want to export the current selection. Otherwise the whole media file and associated annotations will be exported.

5. Check new format to output the tiers to a new, more extensive xml format that supports a separate output scheme of overlapping tiers.

6. Click OK to export the tiers and give a file name, where the tiers can be exported. Also choose the format you want, e.g. txt, csv or xml.

Media clip using script

ELAN supports any command line tool that can extract clips from a video (or audio) file. For that purpose it uses a script file named "clip-media.txt" which can be found in the folder where ELAN is installed. In most cases some configuration needs to be performed in the script file, e.g. which command line tool to use, before clipping can succeed. Therefore ELAN first checks the (see the section called "Special ELAN data folder") for the presence of the "clip-media.txt" file, before trying this file in its installation folder. By copying the customized "clip-media.txt" file to the data folder, the changes are accessible to all versions of ELAN.

Mac OS users will have a default execution line in "clip-media.txt" looking like this:

```
osascript ./scripts/qtp_clip_10_10_export.scpt $in_file $out_file
$begin(sec.ms) $end(sec.ms)
```

Which means that an AppleScript script in the "scripts" folder will be executed when clipping media. There is a PDF file on the ELAN web site to help users with editing the script file.

Windows users can e.g. put a copy of `ffmpeg.exe` (or `ffmbc.exe` for clipping mp4 files) in the folder where ELAN is installed (or modify the execution line such that the full path to `ffmpeg` is included). You can find `ffmpeg` and `ffmbc` online.

If you want to use the syntax for `ffmpeg`, remove the # in front of the line starting with `'ffmpeg.exe -i .........'` If you want to use the syntax for `ffmbc`, remove the # in front of `'ffmbc.exe -vcodec copy.......'` Make sure the syntax you do not want to use has a # in front of it, this comments the line out.

The syntax for `ffmpeg` can be: `ffmpeg.exe -i $in_file -vcodec copy -acodec copy -ss $begin(sec.ms) -t $duration(sec.ms) $out_file`

Where the elements are:

- `ffmpeg.exe` : the path of the application

- `$in_file` : specifies the input file

- `$out_file` : output file

- `vcodec copy -acodec copy` : copy both the video- and audiocodec

- `$begin(sec.ms)` : specifies the begin time frame of the clip

- `$duration(sec.ms)` : the duration of the clip.

Look in the script file for more explanation and examples. If it is not possible to edit the script file due to file permissions, copy `"clip-media.txt"` to the the section called "Special ELAN data folder" (and modify it to use an absolute path to the clipping application).

A few examples for command line tools are:

`C:\ffmpeg.exe -i $in_file -vcodec copy -acodec copy -ss $begin(sec.ms) -t $duration(sec.ms) $out_file`

`C:\ffmbc.exe -vcodec copy -acodec copy -ss $begin(hour:min:sec.ms) -t $duration(hour:min:sec.ms) -i $in_file $out_file`

To clip a media file first make a time selection and choose File > Export As > Media Clip using Script.... A dialog will appear in which you can set the file name and the location to save the clipped file to. You can specify more options for clipping in the Preferences dialog, see the section called "Editing preferences".

📑 Note

If you have more media files to be clipped, typing a file name with a extension in the 'Save as' dialog will use the same extension for all the files that will be clipped. If you want to use the same extension from the original media file for the clipped files, then don't type an extension with the file name in the 'Save as' dialog which prompts you to set the file name and location for the clipped media files.

Image from ELAN Window

To export an image from the ELAN window (i.e. to make a screenshot):

1. choose File > Export As > Image from ELAN Window...

2. Enter a file name and an extension (`*.jpg`, `*.jpeg`, `*.png` or `*.bmp`)

3. click on Save.

> ### 📄 Note
>
> If you are using Windows, it sometimes happens that ELAN's video window is black on the picture created using this function. This can be solved by temporary disabling the hardware video acceleration:
>
> a. Right-click on the desktop
>
> b. choose properties
>
> c. select the Settings tab
>
> d. Click on the advanced… button
>
> e. Select the Troubleshooting tab
>
> f. move the Hardware Acceleration slider tot None
> Don't forget to re-enable the hardware acceleration afterwards, because this has a strong effect on the system's graphical performance.

Filmstrip Image

Figure 1.53. An exported filmstrip image

To export a Filmstrip Image first select the time segment you want the filmstrip of. Then click File > Export As > Filmstrip Image.... In the dialog window (see Figure 1.54, "Exporting to a filmstrip image") you can define the width of each video frame, which frames to include and whether ELAN must add a time code in each frame. Moreover, ELAN can add the waveform, with or without a ruler, and specify the height. You can also specify whether the stereo channel should be displayed separately or merged or blended. Click on OK to generate the image. Finally select a destination folder, enter a file name and click on Save.

An example or an exported filmstrip image can be seen in Figure 1.53, "An exported filmstrip image".

Figure 1.54. Exporting to a filmstrip image

Annotation Density Plot Image

This option allows to save an image of a graphical representation of the density of annotations on selected tiers. This is the same functionality, with the same customization options, as in View > Annotation Density Plot...(the section called "Annotation Density Plot").

Shoebox file

All Shoebox files that were imported into ELAN (see the section called "Shoebox file") can be exported back into Shoebox. In this case, the time code information is kept.

To export a file into Shoebox, do the following:

1. Click on File menu.

2. Click on Export as > Shoebox file ….

ELAN documents

The Shoebox Export dialog box appears. Make a choice and click on OK to continue.

Figure 1.55. Shoebox Export dialog window

- By selecting Wrap block you can let ELAN wrap a whole block if one of the line in a block is longer than a specified number of character (default is 80 characters).

- By selecting Add master media time offset to annotation times you can add to the annotation times the time offset from the master media that originated from the synchronization of media files (see the section called "Synchronizing video files").

3. Specify the name and directory of the exported file, e.g.:

Figure 1.56. Name and directory of exported file

4. Click Save to export the file; otherwise click Cancel to exit the dialog box without exporting the file.

The file is exported as a `*.txt` | `*.sht` | `*.tbt` file.

If there already exists a file of the same name, ELAN will ask you whether or not it should overwrite the existing file, e.g.:

Figure 1.57. File Exists

5. Open the exported file in Shoebox.

It contains the following information:

a. All tiers and annotations.
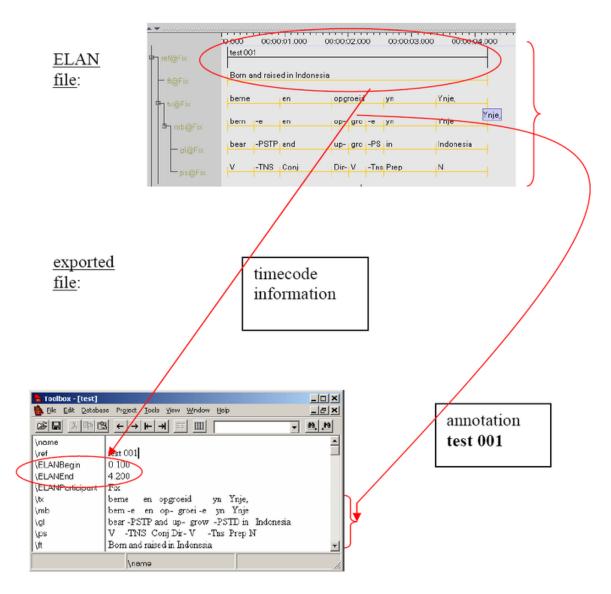
Each ELAN parent annotation (including all its referring annotations) corresponds to one Shoebox record. E.g., in the illustration below, the ELAN parent annotation "Ligya-001" corresponds to the Shoebox record "Ligya-001".

b. The time code information for each parent annotation.

Each ELAN parent annotation (i.e., each Shoebox record) contains the additional field markers \ELANBegin and \ELANEnd (i.e., the begin and end time of the parent annotation).

This time code information allows you to import the Shoebox file back into ELAN, without having to manually re-align the file (see the section called "Shoebox file").

Figure 1.58. ELAN file and exported file

Import from

ELAN supports importing file from :

- Toolbox file (the section called "Toolbox file")

- Fieldworks Language Explorer (FLEx) file (the section called "Fieldworks Language Explorer (FLEx) file")

- CHAT file (the section called "CHAT file")

- Transcriber file (the section called "Transcriber files")

- CSV / Tab-delimited Text file (the section called "CSV / Tab-delimited Text files")

- Subtitle Text or Audacity Label file (the section called "Subtitle / Audacity Label file")

- Praat TextGrid file (the section called "Praat TextGrid file")

- WebAnnotation JSON file (the section called "WebAnnotation JSON file")

- Tiers from recognizer (the section called "Tiers from recognizer")

- Shoebox file (the section called "Shoebox file")

There are also options in ELAN available to import multiple files at once. More details regarding these options can be found here: the section called " Multiple file import options"

Toolbox file

ELAN supports the import of documents from Toolbox, allowing you to link transcribed and/or interlinearized documents to the time axis of media files. In order to import from Toolbox, you need at least the following two files:

- the Toolbox file (`*.txt, *.sht, *.tbt`);

- the media file(s) (`*.mpg`, `*.mov`, `*.wav` etc.);

Optionally you can use the corresponding Toolbox database type file (`*.typ`). If this is not available, one has to provide a list with field markers (= tier names).

📝 Note

If you do not know the Toolbox database type file, do the following:

1. Open the Toolbox `*.txt | *.sht | *.tbt` file in Toolbox. Make sure it is the active window (click on it to activate it).

2. Click on Database menu.

3. Click on Properties …. The Database Type Properties dialog box appears. The name of the database type is displayed in the header, e.g.:

Figure 1.59. Database type properties dialog window

4. Locate the directory of the database type file (e.g., "texts.typ" in the above illustration). It is probably located in the directory "My Shoebox Settings".

Importing Toolbox files with a TYP file

To import a Toolbox file into ELAN, do the following:

1. Click on File > Import > Toolbox File. The Import Toolbox dialog box appears.

2. Specify the name and directory of the two files, e.g.:
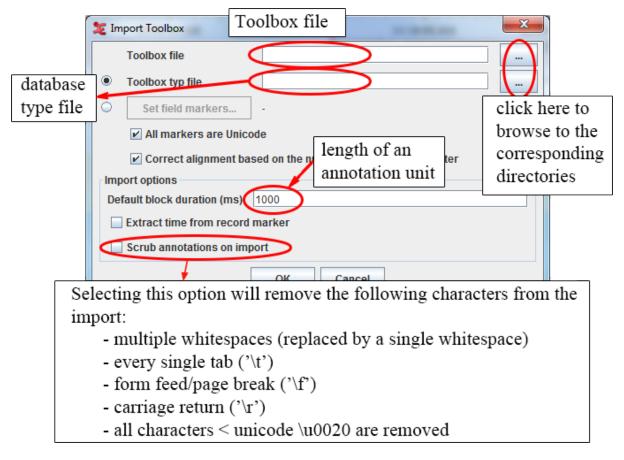
Figure 1.60. Import Toolbox file

3. Like `*.eaf` documents, the Toolbox file and the media file(s) do *not* necessarily need to have the same name, and they do *not* need to be in the same directory (see the section called "Basic Information").

   If the Toolbox file contains both aligned (i.e. containing time information) and non-aligned records, the aligned ones will maintain the timing, whereas the location of the non-aligned records will be interpolated automatically.

4. Click OK to import the file; otherwise click Cancel to exit the dialog box without importing the file.

   An ELAN window containing the imported Toolbox file appears.

Importing Toolbox files without a TYP file

Instead of using a Toolbox `*.txt` | `*.sht` | `*.tbt` file, there is also an option in ELAN to define the field markers yourself when importing a Toolbox file.

1. select the Set field markers and click on the button in the import dialog. The following window appears:

Figure 1.61. Set Shoebox/Toolbox field markers

2. Now fill in a field marker as used in the Shoebox/Toolbox `*.txt|*.sht|*.tbt` file

3. Optionally select a parent marker (see the section called "Basic Information: Annotations, tiers and tier types")

4. Optionally select a stereotype (symbolic subdivision or association, see the section called "Basic Information: Annotations, tiers and tier types")

5. Choose a character set (Latin-1, SIL IPA or UTF-8) for the tier (only available with Shoebox import! Toolbox charset is UTF-8)

6. Click on Add.

7. Repeat step 2-6 for all field markers.

8. If the selected marker designates a participant, check the Participant Marker checkbox. If you don't want the selected marker to be imported, tick Exclude from import.

9. finally choose Close and click on OK in the import Shoebox file dialog

📑 Note

Some markers are already 'built-in' in ELAN and must not need to be set: ELANBegin, ELANParticipant, ELANEnd.

Loading and storing Markers

Once you have manually created a set of field makers, you might want to reuse them later on. ELAN provides support for this:

• To save a set of field markers, select the Store Markers… button. This will display a save dialog. Enter a file name, and press save.

• The same way you can open a stored field marker set by clicking on Load Markers…

Figure 1.62. Store markers

Connecting the transcription to a media file

Once the import has succeeded, you can add a reference to a media file via the Edit > Linked Files…
menu, as described in the section called "Changing the linked media files". If the imported Toolbox
file was exported from ELAN before, you won't need to establish the link to the media file(s) again,
as in that case the location information is stored in the file.

About the import process

ELAN imports Toolbox files according to the following conventions:

1. The Toolbox field markers are imported as ELAN tiers. The tier label is identical to that of the field
   marker, except for the added extension @'Speaker-ID'.

   This addition is necessary because ELAN and Toolbox differ in how they code information about
   multiple speakers:

   • In ELAN, each speaker is coded on a separate tier.

   • In Toolbox, all speakers are coded using the same field, and their identity is specified in a separate
     field.
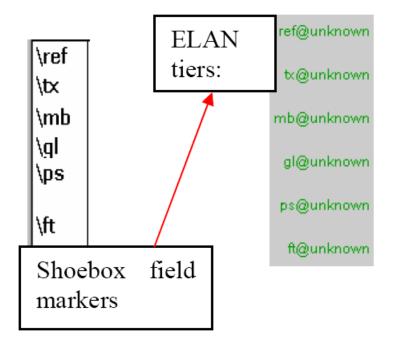
Figure 1.63. Toolbox field markers and ELAN tiers

When importing texts by multiple speakers, ELAN splits each Toolbox field into several ELAN tiers (one for each speaker) and adds the speaker-ID to the tier label.

If speaker information is not specified in the Toolbox file, the extension @unknown is added.

The following screenshot illustrates how ELAN treats texts by multiple speakers:

Figure 1.64. Multiple speakers in ELAN

Note that ELAN can only read speaker information if:

• A marker is defined as a Participant marker in the Set field marker dialog (see Importing Toolbox files without a TYP file above), or if:

• It is coded in a Toolbox field labelled \EUDICOp or \ELANParticipant (see illustration above). If this field is not present, or if speaker information is coded in a different field, ELAN will assume

that there is only one speaker. I.e., if you have multiple speakers and if you want ELAN to assign them to separate tiers, do the following:

1. For every Toolbox record, add the field marker \EUDICOp.

2. For every Toolbox record, enter the relevant speaker-ID into this field.

📑 Note

When the file is exported back to Toolbox (see the section called "Toolbox file(UTF-8)"), the extension @'Speaker-ID' is automatically dropped from the field marker, and the Toolbox records are sorted according to their record marker (e.g., in the above illustration, "test 001" is sorted before "test 002" etc.)

3. Based on the information contained in the Toolbox database type file, the tiers are brought into a hierarchical relationship and are assigned to tier types (see the section called "Basic Information: Annotations, tiers and tier types" for details of tier hierarchies and tier types). For every tier name a corresponding tier type with the same name is created. All of these tier types are connected with a stereotype in such a way that it fits with the original Toolbox structure.

   • The Toolbox record marker is assigned to the stereotype None, i.e., it is an independent, time-alignable parent tier.

   • The transcription and parsing fields of Toolbox are assigned to the stereotype Symbolic Subdivision, i.e., they are referring tiers that can be subdivided into smaller units.

   • All other fields are assigned to the stereotype Symbolic Association, i.e., they are referring tiers that cannot be subdivided into smaller units.

If you define the markers yourself, then there also is the possibility to choose the Time Subdivision stereotype. For example:

Figure 1.65. Time Subdivision

4. If you import a Shoebox record, all SIL IPA characters are converted into Unicode characters during import. If you export the file back into Shoebox (see the section called "Shoebox file"), the Unicode characters will be converted back into SIL IPA characters. This does not apply to Toolbox records.

5. Initially, unless it had the time code information, the imported Toolbox file does not contain information about timing. Instead, ELAN automatically assigns each Toolbox record to a three second time interval, as in the following illustration:

Figure 1.66. Fixed time intervals

The time alignment has to be done manually for each Toolbox record. Do the following:

1. Activate the Bulldozer mode: Click on Options > Propagate Time Changes > Bulldozer Mode (see the section called "Activating and deactivating the Bulldozer mode or Shift mode" for the three available modes).

> 📑 **Note**
>
> If you do not activate the Bulldozer mode, you will inadvertently overwrite and thereby delete existing annotations. Make sure that Bulldozer Mode is enabled in the Options > Propagate Time Changes menu.

2. Click on the first annotation on the parent tier (i.e., the first Shoebox record). It appears in a dark blue frame.

3. Modify the boundaries of that annotation, so that they are aligned with the correct time interval (see the section called "Changing the boundaries of an existing selection and annotation" for ways of modifying boundaries).

4. Press **CTRL+ENTER** to apply the new time interval.

The parent annotation (together with all its referring annotations) is assigned to the new time interval. All other parent annotations are moved to the right.

5. Repeat steps 2 to 4 for each parent annotation.

The following screenshot illustrates steps 1 to 4:
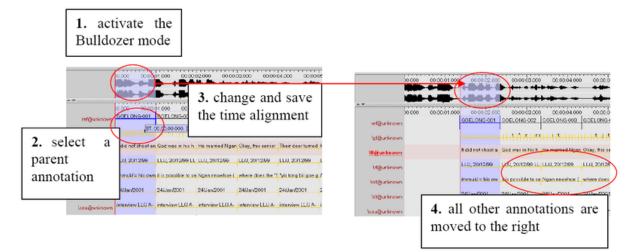
Figure 1.67. Time alignment

After you have done the time-alignment, you can export the file back to Toolbox – in this case, the time code information will be kept (see the section called "Toolbox file(UTF-8)"). If you then re-import the file back into ELAN, ELAN automatically assigns the Shoebox records to their correct time intervals.

An imported Toolbox file can be saved as an ELAN file (see the section called "Re-open recently accessed files"), exported back into Shoebox (see the section called "Toolbox file(UTF-8)"), or exported as a tab-delimited text (see the section called "Tab-delimited text file").

Fieldworks Language Explorer (FLEx) file

ELAN can import documents from the SIL Fieldworks Language Explorer (FLEx). This involves a few steps:

1. Click File > Import > FLEx File.... Select the `.flextext` file and relevant media files by clicking the ...-buttons.

2. In the import window select the `.flextext` file exported from FLEx. Optionally also add media files here (if not already in your `.flextext` file). There are several options to customize the import:

   - `Include "interlinear-text" element`: the top-level `interlinear-text` element can contain a title and other information. If selected, these will be converted to one or more tiers.

   - `Include "paragraph" element`: a text may contain multiple `paragraph` elements, each containing one or more `phrases`. If selected, this option allows to ignore the `paragraph` layer when importing.

   - `Import participant information from "Note" field`: if the FLEx file contains a `note` item type containing the name or code of the participant/spaeaker, this option makes that it will be stored in a tier's `partipant` attribute.

   - `Smallest time-alignable element`: when the `word` element is selected here, the time-alignment for that level will be lost when exported again from ELAN to FLEx. In `.flextext` time alignment is stored on the `phrase` level.

   - `Specify the top-level "phrase" item type`: by default the `<item type="txt">...</item>` child element is converted to the parent tier of each level. Here it is possible to specify an alternative for the `phrase` level, e.g. `segnum`.

   - `Use the "speaker" attribute as tier prefix`: by default tier hierarchies for different speakers in the file are prefixed in ELAN with `A`, `B`, `C` etc. This option allows to
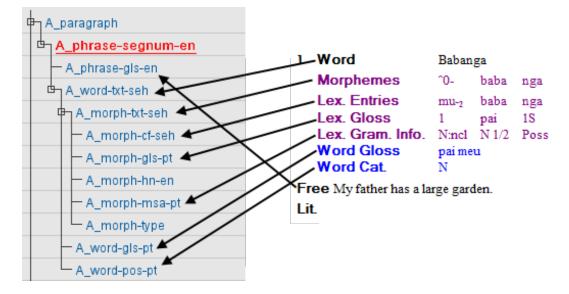
specify that the contents of the `speaker` attribute of a `phrase` element should be used instead. Note that this can interfere with the conversion, the encoding and decoding, from tier name to `.flextext` element.

3. It is possible to have tier types created simply for all major elements (`phrase`, `word`, `morph` etc.) or, more fine-grained, for each combination of major element plus `item type` up to a combination of major element, the type and the language.

4. Finally, set a duration per `phrase` element in milliseconds. This has to be set if the FLEx export files do not contain timestamps. When importing a FLEx file that was edited in ELAN before and exported as a `.flextext` file, time duration information has already been stored in the file.

Figure 1.68. Import FLEx file

Figure 1.69. FLEx to ELAN structure

The tier structure created after import in ELAN is roughly like in the example above. The mapping of the FLEx structure onto ELAN tiers follows the schema: **<Speaker>_<element>-<item-type>-<language>** Where the Speaker prefix is a generic label (A, B, C, ...).

FLEx tiers and their representation in `.flextext`:

Word	<word>	<item type="txt">
Morphemes	<morph>	<item type="txt">
Lex. Entries	<morph>	<item type="cf">
	<morph>	<item type="hn">
Lex. Gloss	<morph>	<item type="gls">
Lex. Gram.	<morph>	<item type="msa">
Word Gloss	<word>	<item type="gls">
Word Cat.	<word>	<item type="pos">

📝 Note

> On the third-party resources page of ELAN (https://archive.mpi.nl/tla/elan/thirdparty [https://archive.mpi.nl/tla/elan/thirdparty#documentation]), you can find a workflow description covering importing from FLEx to ELAN and back to FLEx.

CHAT file

It is possible to import CHAT files (used in e.g. the Childes project) in ELAN:

1. Select File > Import > CHAT File …

2. Select the Chat file

3. Click on Open

Some remarks about this import feature:

• supported are old CHAT files and CHAT-UTF8, not XML CHAT

• existing media alignment in %snd tiers is maintained in ELAN:

- when no media alignment is present at all, each CHAT utterance gets a default interval of 1 second assigned

- when partial media alignment is present, the time interval is equally distributed over preceding unaligned utterances

- overlapping utterances of the same participant are corrected as good as possible

- CHAT dependent tier names are mapped to ELAN Tier Types

- ELAN tier names are either CHAT participant labels or CHAT tier names, followed by '@participantName'

Remaining issues:

- '<' and '>' characters in CHAT cause parsing errors when the imported file is saved as EAF file

Transcriber files

The feature to import Transcriber annotation files into ELAN works as follows:

1. Choose File > Import > Transcriber File …

2. Select the transcriber file (`*.trs`) and click on Open

3. If the associated sound file cannot be found, a dialog will be shown asking you to locate it. When this request is cancelled, one can choose to open the annotation file without the sound, or to stop the whole import process.

The transcriber tiers will be mapped on the ELAN equivalents:

- *Section* becomes a independent tier and *turn* becomes a referring tier of *section* (see also the section called "Basic Information: Annotations, tiers and tier types").

- Events are embedded into the annotation text.

CSV / Tab-delimited Text files

A CSV (Comma Separated Values) or Tab-delimited Text (or Tab Separated Values) file is a text file in which one can identify rows and columns. Rows are represented by the lines in the file and the columns are created by separating the values on each line by a specific character, like a comma or a tab. CSV or Tab-delimited Text files can be compared to spreadsheets like the ones in Microsoft Excel in that they also have rows and columns. Note that .csv files can be created by Excel.

Take a look at Figure 1.70, "Tab-delimited Text". The first row represents the event of a person saying 'so from here'. The first value (as well as the first column of the complete file) represents the tier name, the second and third represent begin time in different formats, the fourth and fifth represent the end time, the sixth an seventh represent the duration and the last value represents the annotation.

Figure 1.70. Tab-delimited Text

```
K-Spch  00:00:00.280    0.28    00:00:01.250    1.25    00:00:00.970    0.97    so from here.
K-Spch  00:00:04.040    4.04    00:00:04.400    4.4 00:00:00.360    0.36    yeah
K-Spch  00:00:10.060    10.06   00:00:10.460    10.46   00:00:00.400    0.4 ja
K-Spch  00:00:14.620    14.62   00:00:14.990    14.99   00:00:00.370    0.37    ja
```

You are able to import CSV or Tab-delimited Text files in ELAN: File > Import > CSV / Tab-delimited Text File.... In the dialog window browse to and select a file that contains CSV or Tab-delimited data and click Open.

The second dialog window contains two sections (see Figure 1.71, "Import CSV / Tab-delimited Text"). The upper section shows a sample table containing data from the selected file. Both rows and columns are numbered. The lower section enables you to specify which columns to include and what data type they represent. This means that the format of the files is flexible: it is not prescribed what data is expected nor how it is formatted. The numbers of the columns in the Import Options section correspond to the numbers of the columns in the sample table. The data types you can select are:

- Annotation

- Tier

- Begin time

- End time

- Duration

Select at least one column with data type 'Annotation'. If you select a column for begin time, end time and duration, the latter will be ignored in the import process.

Figure 1.71. Import CSV / Tab-delimited Text

The option `Specify first row of data` enables you to exclude a header by excluding the first few lines. The option `Specify delimiter` lets you specify the delimiter if ELAN did not guess the correct delimiter. The delimiters supported by ELAN are comma, tab, colon, semi-colon and the vertical line (vertical bar).

If you enable the option `Default annotation duration` ELAN creates all annotations from the selected file with durations equal to the number of milliseconds specified. This option works only if there is no time data or only the begin or end times.

`Default annotation duration` will create annotation units with the specified duration.

`Skip empty cells` will leave out the cells in the csv that are empty. Different tiers can be imported with different segmentations with this option.

`Combine with template (.etf)` allows to import annotations into tiers defined in a template, described in more detail below (Figure 1.74, "Import tab-delimited text in combination with a template")

Finally click OK to import the data. If a transcription document was open when starting the import, the imported tiers and annotations will be added to the already open document, otherwise a new transcription document is created with the imported annotations as its contents.

Another example

To demonstrate that the format of the imported file can be flexible, take a look at the following tab-delimited text:

Figure 1.72. Tab-delimited text, different orientation

In this example each column represents a tier with the tier names in the first row and the annotation in the other rows. This file can be imported by selecting the following import options:

Figure 1.73. Import CSV / Tab-delimited Text

Note that the `Specify first row of data` option is set to 2. As a consequence ELAN starts importing annotations from row 2 instead of row 1. Furthermore, ELAN tries to extract tier names from the first line of the file if the column they are part of is specified as 'annotation'. This results in this example in two tiers: K-Spch and W-Spch.

To merge a CSV file with an existing `*.eaf` file, open the `*.eaf` file first and then choose Import CSV/Tab-delimited Text File. For information on merging a CSV file that has been imorted into a new document with an existing *.eaf file, please see the section called "Merging transcriptions".

Import with a template:

Figure 1.74. Import tab-delimited text in combination with a template

When the `Combine with template (.etf)` checkbox is selected, the Select... button allows to browse to an ELAN template file (`*.etf`). The import function will then use the tiers and tier types etc. from the template as the basis for the new transcription and add the imported annotations to those tiers. The matching is based on tier names being exactly the same in the template and in the tier column (as in the first screenshot) or column headers in the delimited text file (as in row 1 of the sample tables in the second and third screenshot).

The import function will try to apply constraints as defined by the tiers and types in the template, but success is not guaranteed. Especially if the template defines many levels of tier dependencies, proper import might fail. Depending on the structure of the delimited text file, the Skip empty cells option may have to be selected or deselected for a successful import.

📑 Note

> This option is still experimental.

Subtitle / Audacity Label file

It is possible to import subtitles that are stored in the SubRip `*.srt` format: File > Import > Subtitle / Audacity Label File.... HTML and similar formatting tags are filtered out and multiple speakers are merged into one. The correct encoding of the file has to be specified in the import window.

Audacity Label files are a specific kind of tab-delimited text (`*.txt`) files. They can be imported here without the configuration step that is part of the general Import CSV/Tab-delimited Text File import.

If this import is started when a document is already open, the imported contents is added to that transcription. Otherwise a new transcription document is created.

Praat TextGrid file

ELAN offers the possibility to import a Praat `.TextGrid` file: click on File > Import > Praat TextGrid File.... In the dialog window that now appears, you can browse to the file you wish to import. You

are also able to Include Praat PointTiers. When selecting this option, specify the Default PointTiers annotation duration in milliseconds. Check Skip empty intervals / annotations if you want to do so. Finally, if you tick the Attempt to auto detect .wav file checkbox, an audio file with the same name as the TextGrid file will be added to the linked files, if it is found.

If there is already a annotation document opened in ELAN, the imported TextGrid is added to the document in one or more new tiers. If there is no annotation document opened, a new document consisting of the TextGrid data is generated.

In addition to TextGrid files in the default encoding for the operating system, ELAN supports Praat TextGrid files with UTF-8 and UTF-16 encoding.

WebAnnotation JSON file

It is possible to import a WebAnnotation JSON file via File > Import > WebAnnotation JSON File..., the file extension is `.json` or `.jsonld`. There are no configuration options. The contents of the file should comply with the W3C Web Annotation Data Model [https://www.w3.org/TR/2017/REC-annotation-model-20170223/] specifications, even though the import function only supports a subset of those specifications (those elements that map quite naturally to ELAN elements).

Tiers from recognizer

Importing Tiers from recognizers will import the tiers in a new file if there is no file currently open in elan. But if a file is open, the tiers will be in the currently open file. To import the tiers from recognizers, go to File > Import > Tiers from Recognizer.... Selecting this option, first will prompt for the import file. If there is no file is open, the tiers are directly imported to the new file. But if a file is already open, then a **'Create tiers from segments'** dialog appears. For more information about this dialog see Figure 2.14, "Silence Recognizer".

Shoebox file

Importing a document from Shoebox is very much the same as importing a document from Toolbox (see the section called "Toolbox file"). As with the Toolbox import, information about the tier relations can be provided by means of a `.typ` file or by using a marker file.

When reconstructing the vertical alignment of words on interlinearized markers, the position is recalculated based on the number of bytes per character. But in some files this leads to incorrect alignment, therefore this recalculation can be turned off by unchecking Correct alignment based on the number of bytes per character. This import also tries to take non-spacing characters into account.

Screen display and navigation through a document

Whenever you open, create or import a document, the ELAN window appears. This section introduces you to the setup of the ELAN window (the section called "Basic Information: The ELAN window"), and explains the navigation through it (the section called "How to navigate through a document" and the section called "How to play a document").

Basic Information: The ELAN window

The ELAN window displays the Menu bar, the Media Player options, and a number of Viewers, e.g.:

Figure 1.75. The ELAN window

All Viewers are synchronized and thus display the same point(s) in time. I.e., whenever you access a point in time in one of the Viewers, all the other Viewers will immediately jump to the corresponding point in time.

In all Viewers, color coding is used to facilitate the orientation in the document.

This section introduces the setup of the Viewers, the Menu bar, the Media Player options and the color coding. Detailed information about how to navigate through the ELAN window follows in the subsequent sections.

The Video Viewer

The Video Viewer displays up to four video images (of the linked *.mp4, *.mpg etc. video files). However it is possible to link more than four video files. To make videos (in)visible click View > Media Player > in the ELAN menu and (un)check a video file name. A video file will not be unlinked when it is made invisible in this way.

Figure 1.76. The video viewer

Note that you can right click on the video viewer to *detach* it, i.e. create a separate window for the video. To re-attach the video window, right click on it and select *attach*.

To change the size of the video viewer you don't need to detach the video viewer. Instead, you can drag the vertical divider on the right side of the window up and down to make the video viewer respectively smaller and bigger (see also Figure 1.76, "The video viewer").

📄 Note

If you encounter problems while playing video files, change the media framework via Edit > Preferences > Edit Preferences.... Select Platform/OS and toggle Media Framework appropriate for your operating system.

Right clicking in the video window and selecting Player Info… will display a dialog with information about the video file, e.g.:

Figure 1.77. Video information dialog window

📄 Note

For users of a one-button mouse on macOS computers: press the **CTRL** key and click.

A static picture containing the currently displayed frame can be stored using the context menu of the video window (right click > Save Current Frame as Image…)

Clicking anywhere in the video viewer copies the coordinates of the mouse cursor relative to the upper left corner of the video to the clipboard. The coordinates can have different formats depending on the modifier key used:

- no modifier key: `x,y [original width, original height]`, where x and y are coordinates in the original coordinate system.

- with **ALT** key: `x,y` where x and y are between 0 and 1 (0.000, 1.000) identifying a relative position in the (0,0,width,height) image space.

- with **SHIFT** key: `x,y` where x and y are coordinates in the original coordinate system (not bothering about original dimension or aspect ratio)

- with **ALT**+**SHIFT** `x,y [current width, current height]` where x and y are coordinates based on the current width and height of the video viewer.

Normally the aspect ratio of the video as detected by the media framework is correct, but sometimes it is not. In those cases the aspect ratio of the video viewer can be set by right clicking the video, selecting Use Aspect Ratio... and choosing one of the aspect ratios offered in the sub menu.

You can zoom in on the video by right-clicking the viewer and selecting Zoom. You can choose from a number of zoom percentages.

To copy the Non-adjusted media time, right-click the video and select Copy Non-adjusted Media Time This function will disregard any offset that may be applied to sync the video with another video and will copy the actual timecode to the clipboard.

To place the video viewer in the center, see the section called "Editing preferences" - Media option

The Annotation Density Viewer

The Annotation Density Viewer is a kind of a timeline which allows you to:

- Navigate through the whole media file. The length of this viewer always corresponds to the whole media file, so e.g. by clicking in the middle you will always go to the middle of the media file. The selection is represented as a small grey bar.

- See how many annotations are concentrated at a particular moment of the time (the *Annotation Density*). The more annotations available for a particular moment, the more the Annotation Density bar is filled. This can be useful to track places in the media file that still have to be annotated.

By default the Annotation Density Viewer shows the annotation density of all tiers. It is also possible to view the annotation density of a selection of either tiers, types, participants or annotators. To do so, right click the Annotation Density Viewer and select one or more tiers, types, participants or annotators.
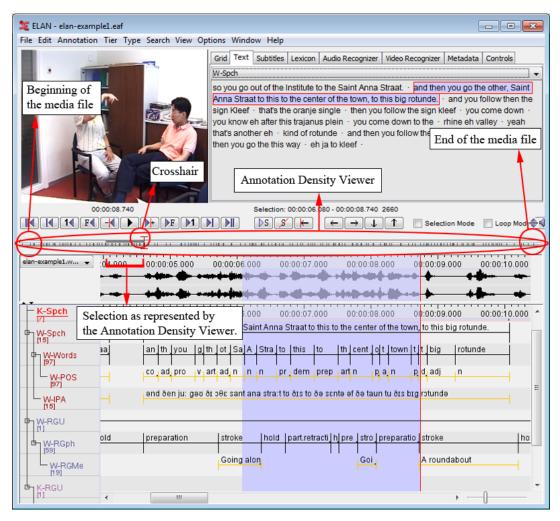
Figure 1.78. Annotation density viewer

The Waveform Viewer

The Waveform Viewer (in the user interface also referred to as Signal Viewer) displays the waveform of the audio file(`*.wav`) or of the audio track of a video file.

📄 Note

Different resolutions of `*.wav` files are supported: 8 bits (mono and stereo), 16 bits (mono and stereo) and 24 bits (mono and stereo). Both PCM and A-law encoded wave files can be loaded. There is basic support for RF64 wav files (which can exceed the file size of 4GB). Supported video files for waveform visualisation depend on the framework of the operating system (Windows or macOS)

Above the waveform, time code information is displayed. This time code information can be hidden by right clicking in the Waveform Viewer and clicking on Time Ruler Visible in the context menu. In the same manner it can be made visible again.

While listening to the sound, a red vertical bar, the crosshair, moves through the waveform and indicates which part of the waveform corresponds to the current point in time. Furthermore, whenever you have selected a time interval, the corresponding part of the waveform will be highlighted in light blue color.
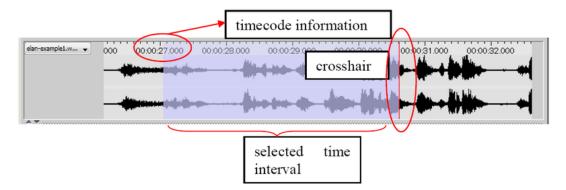
At any time, you can press **ALT** and drag the time axis for a panning effect (i.e. go to the left to go back in the time or to the right to go further). In the case of video files, the waveform is only displayed

if there exists an additional `*.wav` file (see the section called "Basic Information"). If this is not the case, the Waveform Viewer will not be available.

📄 Note

On slower machines, the Waveform Viewer may not always update properly when moving to the next page.

Figure 1.79. The waveform viewer
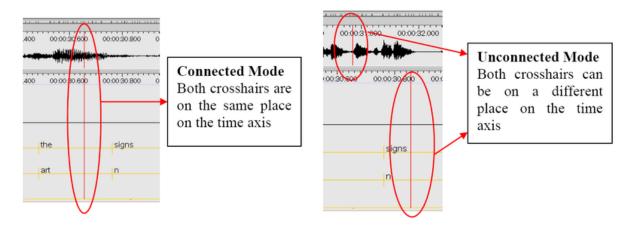
The Waveform Viewer supports 3 modes. You can select the active mode by a right click on the Waveform Viewer. In the menu Stereo Channels, the following options are available:

1. Separate. 2 Waveforms are shown, one for each channel.

2. Merged. The 2 channels are merged and the result, one waveform is displayed.

3. Blended. Both channels are displayed on 1 waveform, differences are designated with colors.

Another option in the context-menu (right click) of the Waveform Viewer is *connected*. If this option is checked, the time scale of the Waveform Viewer and the Timeline Viewer are connected:

Figure 1.80. Waveform viewer modes

The context menu also contains options for to open the file or the selection in Praat or to clip the selection with Praat (see the section called "Opening a wave file in Praat" and the section called "Exporting a selection to a wave file with Praat") or with Java Sound (built-in).
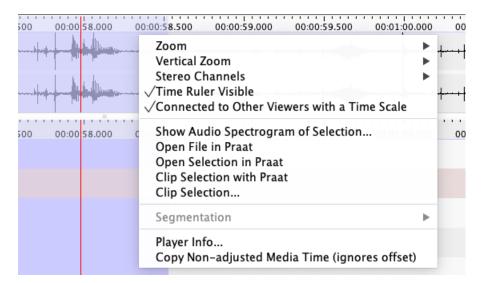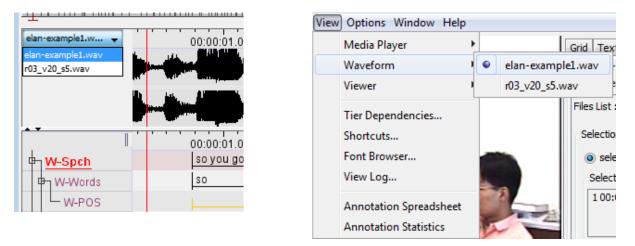
Figure 1.81. Waveform context menu

You can load multiple waveforms into your project. Only one will be visible. There are 2 ways to switch between waveforms: from the View>Waveform option in the main menu, or from the drop-down menu on the left side of the waveform viewer. The chosen waveform will be displayed in the waveform viewer:

Figure 1.82. Waveform visualisation switching

The Audio Spectrogram detail window

In addition to audio visualization in the Waveform Viewer, it is possible to visualize audio sources in the Spectrogram Viewer (see the section called "The Spectrogram Viewer"). The creation of a spectrogram image, and updating it during playback and while navigating backward and forward in the recording, requires a lot of computing power. This can slow down the application on average desktop and laptop computers, to the extend that it is best to switch the Spectrogram Viewer off.

As an alternative, it is possible to create a static audio spectrogram based on the current selection in the Waveorm Viewer and show it in a popup window. This window can be resized, shows the spectrogram image with more detail and can be used e.g. to finetune the selection. From the context menu, choose the option Show Audio Spectrogram of Selection.... A window like the following will appear .
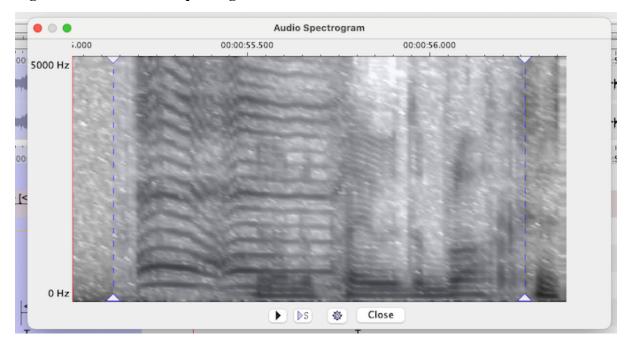
Figure 1.83. The Audio Spectrogram window

A spectrogram of the selection has been created, with some extra samples (roughly 10%) to the left and right of the selection. There is a time ruler at the top, frequency labels to the left and a few buttons below the image. The time selection is marked by a light blue, semi-transparent color and dashed lines with triangular knobs to mark the begin and end time boundary. The boundaries can be changed by dragging the knobs to the left or right. Or by using the keyboard:

• **ARROW LEFT** and **ARROW RIGHT** move the left boundary (begin time) with one pixel to the left or right

• **SHIFT+ARROW LEFT** and **SHIFT+ARROW RIGHT** move the right boundary (end time) with one pixel to the left or right

The selection is connected to the main window, i.e. changing the selection here, changes the selection in the main window (in case of dragging with the mouse, the selection is updated after releasing the mouse button).

The buttons have the following functions:

• Play plays the entire interval visible in the window, starting from where the crosshair happens to be

• Play Selection plays the selection from begin to end

• Configure opens the Spectrogram Settings window, as described in section the section called "The Spectrogram Viewer"

• Close closes the window

> 📄 **Note**
>
> The spectrogram detail window is not connected to the media player in the main window. Instead, the audio samples which are provided by the Waveform Viewer for the visualization of the spectrogram, are used to create an in memory JavaSound player. This means that the audio that is produced, is taken from the media file which is visible in the Waveform Viewer.

> 📄 **Note**
>
> There may be platform specific issues with audio playback in case the source is a video file and the audio has been extracted with the FFmpeg option.

The right mouse button context menu of the spectrogram image allows to select which channel(s) to display (in case the source has more than one channel). It is possible to show only channel 1 **or** 2, or 1 **and** 2 merged or separate.

The Spectrogram Viewer

The Spectrogram Viewer displays a spectrum of frequencies of an audio file (`*.wav`) or of the audio track of a video file. The horizontal axis represents time (though the viewer currently does not show a time ruler), the vertical axis frequency and color intensity represents the amplitude of a particular frequency range at a particular time.
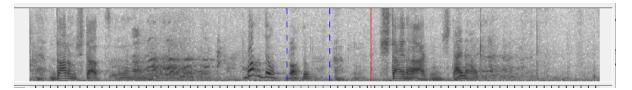
For the creation of the spectrogram the data of the audio signal have to be transformed to a spectrum of frequencies. The transform applied in ELAN is a Fourier transform (a fast Fourier transform, FFT [https://en.wikipedia.org/wiki/Fast_Fourier_transform]). The transform is applied successively to small segments of the audio signal ("sliding windows") and creates an array of *frequency bins* (each *frequency bin* representing a range of frequencies) for each segment. To reduce the effects of discontinuity resulting from this segmentation, a window function (see the window function options) can be applied to the segments before the transformation.

The viewer shows the usual red vertical bar, the crosshair, to indicate the current point in time. It also marks the boundaries of the selected time interval, currently by means of blue vertical lines (instead of a light blue overlay on the image).

📝 Note
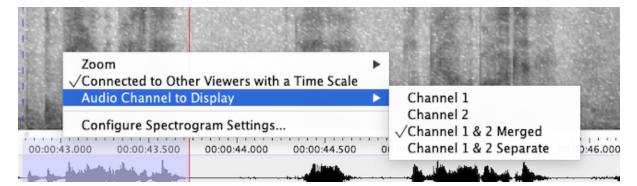
Generation of the spectrogram image requires a considerable amount of processing, therefore it might be useful to switch off this viewer (see the section called "Show / Hide Specific Viewers") if it doesn't serve your purposes (this is true for other viewers as well).

Figure 1.84. The spectrogram viewer

The viewer provides a few options to customize the appearance of the image. The context menu that is shown with a right click on the viewer contains the following items:

Figure 1.85. The spectrogram context menu

• the Zoom menu allows to zoom in or out like in other viewers with a timeline

• the Connected to Other Viewers with a Time Scale option determines whether or not the visible interval remains in sync with e.g. the Waveform and the Timeline viewers

- the Audio Channel to Display menu items allow to select the channel(s) to display. Currently only files with a maximum of two channels are supported and in most cases using both is a good option (although it requires slightly more data processing).

- the Configure Spectrogram Settings... menu creates a dialog in which parameters that control the processing and visualization of the audio data can be configured.

📑 **Note**

> At the moment the settings are global (application wide) settings, but already open windows are not updated immediately when settings are changed.

The audio Spectrogram settings window contains two panels, one dedicated to visualization parameters and one to data transformation parameters.

Figure 1.86. Spectrogram display settings panel

The following parameters determine the visualization of the spectrogram:

- The range of displayed frequencies (in Hz.) can be set by entering the minimum and maximum in two textfields. The minimum value is displayed at the bottom of the image, the maximum at the top.

- The default color scheme (Gray) for the spectrogram image is grayscale with higher color intensities (darker parts) corresponding to higher amplitudes. The Reversed gray option makes a grayscale image with the lighter parts corresponding to higher amplitudes. The Color gradient option produces a color image for which a foreground color (higher amplitudes) and a background color can be specified. (The resulting image can contain more than two colors; the specified colors are taken as two points in HSV color space with possibly multiple colors in between.)

- The brightness and contrast of the image can be adjusted in two ways.

    - When the Adaptive contrast checkbox is selected, the intensity of the color adapts to the actual values in the current visible interval. The highest amplitude value is black, the lowest white (in

the default color scheme). As a result, when scrolling to the left or right, the darkness of that part of the image that was already in view, might change (because of the new data in the interval). For performance reasons this setting may be temporariliy ignored e.g. when the media player is playing. By default this option is not selected and the appearance of each section ("window" of the audio remains the same regardless of what else is in the visible area.

- Otherwise (if the Adaptive contrast checkbox is not selected) the brightness and contrast of the image can be adjusted by specifying a correction for the foreground and/or background color, making them darker or lighter. The value entered for Foreground brightness correction increases (negative value) or decreases (positive value) the intensity of the darker parts (in the default color scheme). The entered values are interpreted as percentages of the original range. Similarly the Background brightness correction changes the lighter parts of the image (a positive value makes the lighter parts lighter).

Figure 1.87. Spectrogram transform settings panel

The following parameters determine how the data of the audio signal are transformed to a spectrum of frequencies:

- The Window function (window shape) drop-down list provides a number of Window functions [https://en.wikipedia.org/wiki/Window_function] to choose from. The *Rectangular (none)* option means that no window function is applied.

- The Window length determines the duration of the segments that are passed to the transform function. The actual duration of the windows may differ slightly from the value entered here, because the number of audio samples passed to the function is adjusted to a power of two. A smaller window (shorter duration) leads to higher time resolution and lower frequency resolution (i.e. less frequency bins), a bigger window leads to lower time resolution and higher frequency resolution.

- The Stride length determines the size of the step with which the sliding window is moved over the audio signal and therefore the amount of overlap between successive segments. The stride size can not be larger than the window size.

Apart from the Apply and Cancel buttons, there is Restore Defaults button which resets all settings to their default values (after clicking Apply).

The Subtitle Viewer

Figure 1.88. The subtitle viewer

You can turn on the subtitle viewer for a tier by selecting that tier from the pull down menu in the tab Subtitle Viewer. During playback, the Subtitle Viewer displays the annotations of the selected tiers at the current media time, both during playback and in static situations.

By default, the Subtitle Viewer can display up to four tiers as a subtitle:

1. Select the Subtitle Viewer tab in the right upper corner of the ELAN window

2. In the pull-down menu, click on the tier you want to be displayed.

Figure 1.89. Display multiple tiers as a subtitle

The number of tiers to display as subtitle can be between 1 and 8. To set this number, click Edit > Preferences > Edit Preferences... from the main menu and select Viewers in the Preferences dialog. Change the number of viewers to the desired value in the pull down menu and click Apply.

The Grid Viewer

The Grid Viewer displays the content as well as the begin time, end time and duration of all annotations from a single tier. You have to activate this Viewer through selecting the Grid tab next to the video window.

Figure 1.90. The grid viewer

It is possible to select annotations within the Grid Viewer (by clicking on them), or to edit them (by double-clicking on them).

The time format of the begin time, end time and duration can be changed. Right click on the Grid Viewer, select Time Format and select one of the available formats: hh:mm:ss:ms (hours:minutes:seconds.milliseconds), PAL (hours:minutes:seconds.frames), NTSC (drop frame) (hours:minutes:seconds.frames) and msec (milliseconds).

By default the Grid Viewer works in single tier mode. To switch to other multi tier modes with symbolic association tiers or with symbolic subdivision tiers, click on this dropdown menu button which is indicated in the figure that is above. In that case all the annotations of the selected tier will be shown in the grid, together with all symbolic associated/symbolic subdivision tiers (see the section called "Basic Information: Annotations, tiers and tier types"). Empty cells of dependent tiers can also be filled in this way.

Figure 1.91. Multiple tiers with symbolic association in the grid viewer

Exporting from the Grid Viewer

The results as shown in the Grid Viewer can be exported to a (tab-separated) text file, for later manipulation with e.g. a spreadsheet. Follow these steps:

1. First make a selection of the tiers you want to be displayed (and thus exported)

2. Right click on the Grid Viewer and select Export Table as tab-delimited text…

3. Enter the name of the file to be created and choose Save

Figure 1.92. Export from the grid viewer

The Text Viewer

The text viewer shows all values of all annotations on a selected tier as ongoing text. Within this viewer, there are 3 kinds of highlighting shown:

• Text inside a red box: the value of the annotation on the selected annotation tier that matches the current point of time.

• Text inside a dark blue box: the active annotation (see also the section called "The Timeline Viewer and the Interlinear Viewer")

• Text with a light blue background: the selected time interval.

Optionally, you can make the annotation boundaries visible in the text viewer. Right click in the text viewer and select Toggle visualization to enable this. The boundaries are marked by a dot.

A selection of the text in the Text Viewer can be copied to the clipboard. To do so, first select (part of) the text using your mouse. The selection you make in the Text Viewer is enlarged to include the whole of each annotation your selection spans. However, only your exact selection will be copied. Right click in the Text Viewer and select Copy.

Figure 1.93. The text viewer

The Lexicon Viewer

See Figure 2.52, "The Lexicon Entry tab"

The Comments Viewer

See the section called "Adding comments to .eaf files "

The Metadata Viewer

On the Metadata tab in the main window the IMDI or CMDI metadata that belong to the media can be displayed. Click Select Metadata Source..., select an IMDI or CMDI file and click Select. See also the section called "Changing the linked media files"

The default metadata keys are now displayed (see also the section called "Editing preferences") in either a table (for IMDI)or tree (for CMDI/IMDI). To change the view, right click the table and select Tree View. Right click and select Table View to return to the table view. If you want to change which keys are displayed, click Configure... and (de)select the keys. For CMDI metadata, you can collapse or expand all nodes, or just the top-nodes.

Figure 1.94. The metadata viewer displaying a CMDI file

Audio And Video Recognizers

The Recognizers tab provides a user interface to recognizer components that have been installed as extensions. These recognizer components generally apply pattern recognition algorithms to automatically detect events of interest (pauses, speaker turns, utterances, gesture phases etc.) in an audio or a video file. Each recognizer has it's own icon representing the kind of recognizer.

A few recognizers are included in the ELAN distribution, one of them, the Silence Recognizer, is described in more detail here (seethe section called "Creating a tier using the Silence Recognizer"). The screenshot below shows an example of the list of installed recognizers.

Figure 1.95. The Recognizer Tab

The latest interface specification for audio and video recognizers is available in this document: Avatech-interface-spec-2014-03-06.pdf [https://www.mpi.nl/tools/elan/docs/Avatech-interface-spec-2014-03-06.pdf]. The old AVATecH and AUVIS project pages are no longer online. The recognizers developed in those projects are likewise not available as web services anymore and their extensions have been removed from the recognizer tab.

The Timeline Viewer and the Interlinear Viewer

All annotations can be displayed in the Timeline Viewer or the Interlinear Viewer. Only one of the two Viewers can be switched on at a time.

The Timeline Viewer is always shown when a document is opened in ELAN. It displays the tiers and their annotations, whereby each annotation corresponds to a specific time interval. Because the display of an annotation is limited to this time interval, an annotation does not always fit in the annotation frame. A small grey square is the bottom right corner of the upper part of a annotation frame indicates that an annotation is truncated.

The height of the tiers can be reduced to make more tiers visible. To do so, open Edit > Preference..., select Viewers, check `Reduced Tier Height` and click OK. Above the tiers, a time scale is displayed. This time-scale can be hidden by right clicking in the Timeline Viewer and clicking on Time Ruler Visible in the context menu. In the same manner it can be made visible again.

The horizontal resolution of the timeline can be changed by zooming in or out. There are several ways to do this, as described in the following section (the section called "Changing the time zoom").

During playback, a red vertical bar, the crosshair, moves through the annotations and indicates the current point in time. Normally the crosshair will start from the left if it reaches the right side of the viewer. If you right click in the Timeline Viewer and select Ticker Mode, the crosshair will stop when it reaches the center of the viewer, while the viewer itself scrolls to the left.

Whenever you have selected a time interval, it will be highlighted in light blue; and whenever you have selected an annotation, this becomes the active annotation and will be highlighted in a dark blue frame.

If desired the latter can also be indicated with a bold line. To activate this, right click on an annotation somewhere in the timeline viewer and check the Active Annotation Bold box in the context menu.
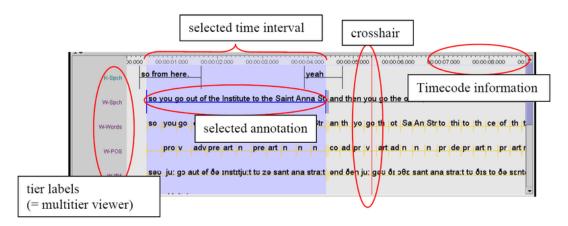
Figure 1.96. The timeline viewer

In the Timeline Viewer you can (a) select and modify time intervals (see the section called "How to make a selection") and (b) enter annotations (see the section called "How to enter and edit annotations").

The Interlinear Viewer offers an alternative perspective on the tiers and their annotations. It shows parent-child relations between annotations using vertical text alignment (interlinearization). You can enable it selecting the Show Interlinear Viewer radio button when you right-click on the tier name's panel and select Viewer. Switching it on, will automatically switch off the Timeline Viewer.

The following screenshots compare how information is displayed in the two Viewers.

Figure 1.97. The ELAN window with timeline/interlinear viewer

Whenever the Interlinear Viewer is switched on, it displays an annotation block (i.e., an independent, time-alignable parent annotation together with its referring annotations, see the section called "Basic Information: Annotations, tiers and tier types"). To move forward/backward to the next block, click on the arrow icons at the top of the Viewer. During playback, the Viewer automatically moves forward to the next annotation block.

The Interlinear Viewer differs from the Timeline Viewer in that it does not allow to modify the time interval or to enter new annotations. It is similar to the Timeline Viewer in that it allows to edit existing annotations. A more versatile variant of the Interlinear Viewer is available as the the section called "Interlinearization mode" working mode.

The Tier Name panel

The Tier Name Panel gives an overview of the different existent tiers. Both tier names (see also the section called "Basic Information: Annotations, tiers and tier types" and the section called "How to define a tier and its attributes") and tier types are defined by the user (see also the section called "How to define a tier type"). One of the tiers in the Tier Name Panel is the active tier (indicated by its underlining and red color), which means that new annotations will be added to this tier (when pressing ALT+ N). Hovering over a tier name will show more information about the tier. (see screenshot below)
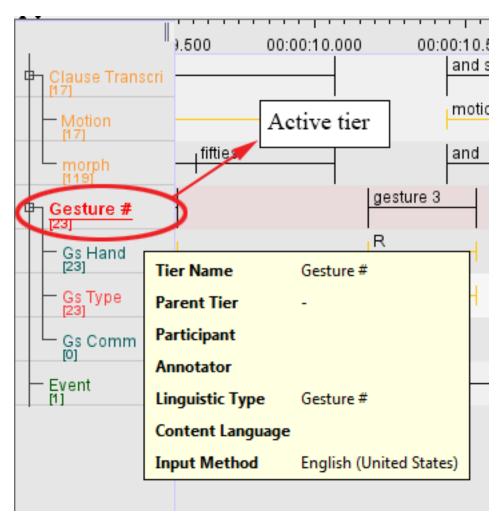
Figure 1.98. The tier name panel

To make a tier the active tier, choose one of the following actions:

• Double click on a tier label.

• Right click on the desired tier in the Tier Name Panel and check `Activate (tiername)`

• Select the active tier with the keyboard shortcut **CTRL + ARROW UP/DOWN**

To deactivate a tier, either do one of the following actions:

- Hold **Shift** and double click on a tier label.

- Right click tie active tier in the Tier Name Panel and uncheck `Activate (tiername)`

- Select the active tier with the keyboard shortcut **CTRL + ARROW UP/DOWN**

To select the tiers to display (and their order) see the section called "Switching tiers on/off" and the section called "Sorting tiers"

It is possible for ELAN to show the number of annotations per tier. Right click on the Tier Name Panel and select Show Number of Annotations.

The Timeseries Viewer

The Timeseries Viewer can display time series data as line graphs. Like the Timeline and Waveform viewer, it has a horizontal time-scale bar, a red vertical crosshair indicating the media time and a light blue rectangle to highlight the selected time interval. It has also the same zoom and pan options.

It can display multiple "track panels" and each track panel can display multiple "tracks". Track panels and tracks can be added and removed via a popup menu. Each track panel derives its value range (vertical axis) from one of the tracks. The viewer has a facility to transfer data from a track to annotation values. Based on the time intervals of the annotations on a chosen (time-alignable) tier, the minimum, maximum or average of the data within these intervals of the selected track will be copied to annotations on a dependent, symbolically associated tier.

The Timeseries Viewer will be created after at least one supported timeseries data file has been associated to the transcription via menu Edit > Linked Files and then the tab "Linked Secondary Files". These data files can be synchronized to the media files in the "Media Synchronization Mode".

📄 Note

Currently supported file formats are CSV/Tab delimited text files, Praat .PitchTier and .IntensityTier files, a proprietary .log file produced by MPI CyberGlove software and a special kind of plain text (.txt) file, containing a time-value pair on each line. Software developers can add support for other formats by implementing a Service Provider Interface (more information can be found in the source code release notes).

Figure 1.99. Linking timeseries data files

Displaying data from an already linked CSV/Tab delimited text file in the Timeseries Viewer is done as follows:

1. Right click in the Timeseries Viewer and select Configure Tracks....
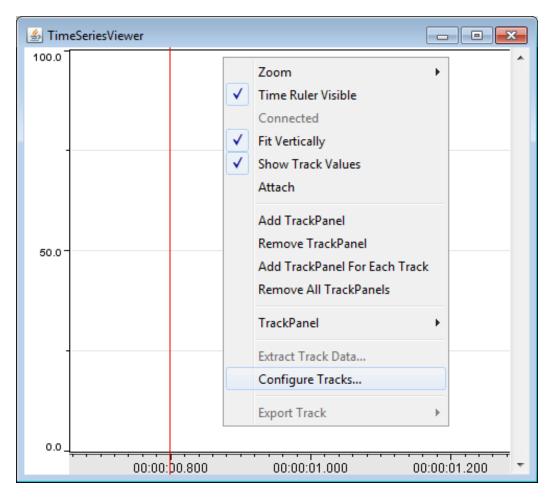
Figure 1.100. Timeseries Viewer popup menu

2. If you have more than one file linked as secondary file, choose the file you wish to use from the pull down menu that is now displayed and click OK.

3. In the next window you see a sample table with several lines and columns of the chosen file. At least one of the columns must contain time data. Select that column by selecting the appropriate column number at `Time Column Index`. If the time codes have a fixed interval, you can check the option `Continuous Rate`. Its underlying purpose is to speed up the calculations for displaying a data track.

Figure 1.101. Timeseries Viewer: Configure tracks

4. After you have selected a column as the time column, you can begin creating tracks. On the Add tab, enter a `Track Name` and optionally a `Track Description`. Select the number of the column in the data that you want to use for this track and specify the range for the vertical axis. This can be

automatically calculated by selecting `Calculate Range From Data` or it can be set manually by selecting `Manual Setting` and entering the `Minimum Value` and `Maximum Value`.

The `Derivative` option allows you to display the first, second or third derivative of your data. Derivatives are useful if we are, for example, dealing with data that represent the position of an object, but we wish to see the velocity of that object. Because velocity is the first derivative of position, we would select 1. In this example, 2 would represent the acceleration and 3 the rate of change of acceleration, also called *jerk* or *jolt*.

Enter the units of your data, for instance *meters* for position or *Pascal (Pa)* for pressure at the `Units (String)` option. Select a color by clicking the colored box at `Track Color`.

Finally click the Add button. The track is now added to the list of Current Tracks which is above the Add tab. Continue adding tracks for each column of data you wish to display. After adding tracks, click on the Close button.

5. To display the track right click on the Timeseries Viewer again. Select Add TrackPanel to add a new track panel. Right click the new track panel and select *TrackPanel* > Add Track. A list of not yet displayed tracks is displayed. Click one to add it to the track panel.

The other options from the popup menu are:

• Zoom: zoom in and out horizontally.

• Time Ruler Visible: hides or shows the time scale bar.

• Connected:

• Fit Vertically: fit the track panel(s) vertically to the Timeseries Viewer window.

• Attach: attaches of detaches the Timeseries Viewer to the main window.
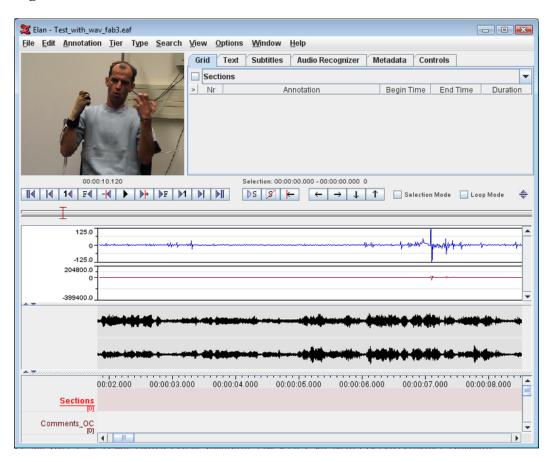
Figure 1.102. Timeseries viewer in the main window

- Add TrackPanel: create a track panel

- Remove TrackPanel: remove current track panel.

- Add TrackPanel For Each Track: create a track panel for each of the existing tracks.

- Remove All TrackPanels: remove all track panels form the Timeseries Viewer window.

- TrackPanel > Set Range For Panel: set the vertical range to the range specified for a track.

- TrackPanel > Remove Track: remove a track from the current track panel.

- TrackPanel > Add All Tracks: add all tracks to the current track panel.

- TrackPanel > Remove All Tracks: remove all tracks from the current track panel.

- Extract Track Data: Extract data from a track and add it to a tier. This process consists of two steps:

  1. Selection of a source and a destination tier. The annotations of the source tier provide the segments for which to extract data from a track. The destination tier should be a (Symbolic Association) dependent tier of the source tier. The numerical values that have been extracted per segment will be stored in dependent annotations on the destination tier.

  2. Selection of the track to extract data from. There are a number of options for what has to be extracted (calculated):

Figure 1.103. The Extract Track Data window

- Minimum - the smallest value found in the interval

- Maximum - the largest value in the interval

- Range - the maximum minus the minimum value in the interval

- Average (arithmetic mean) - the average of all values in the interval

- Sum - the total of all values in the interval

- Median - the median of all values in the interval

- The value at the begin of the segment

- The value at the end of the segment

The Allow existing values to be overwritten checkbox allows to determine whether existing annotations on the destination tier may be overwritten or not.

The Menu options

The following Menu options are available at the top of the ELAN window:

- File: use this menu to open, create, save, im-/export or exit a document (see the section called "Screen display and navigation through a document") and to configure automatic backups.

- Edit: use this menu to define, modify and delete annotations, tiers and tier types (see Chapter 2, *Annotations*).

- Annotation: use this menu to define, modify, copy, paste and delete annotations (see Chapter 2, *Annotations*).

- Tier: use this menu to define, modify and delete tiers. You can also create tiers based on annotations (see Chapter 2, *Annotations*).

- Type: use this menu to define, modify, delete and import tier types (see Chapter 2, *Annotations*).

- Search: use this menu to search for text (see the section called "Tokenize a tier").

- View: use this menu to get an overview of the tier dependencies (see the section called "Basic Information: Annotations, tiers and tier types"), the videos (see the section called "The Video Viewer") and waveforms (see the section called "The Waveform Viewer") that are active and the shortcut keys.

- Options: use this menu to (de)activate the Bulldozer mode (see the section called "Activating and deactivating the Bulldozer mode or Shift mode"), to choose between annotation mode and synchronization mode and to select a language and video standard.

- Window: this menu shows you a list of projects that are currently open and you can switch between these (see the section called "Switching between multiple files").

- Help: use this menu to read information about ELAN.

The Media Player options

With the Media Player options, you can control the playback of the file. The following options are available at the bottom and at the left side of the ELAN window:

Figure 1.104. The media player options

Table 1.1. Media Controls

Icon	Meaning	Shortcut
	Go to the beginning of the video/ audio fragment	**CTRL+B**
	Go to the previous scroll view (make the beginning point of the current timeline view the end point)	**CTRL+PAGE UP**
	Go back one second	**SHIFT+LEFT**

Icon	Meaning	Shortcut
	Go back one frame (= 40 ms for PAL, 33.4 ms for NTSC)	**CTRL+LEFT**
	Go back one "pixel" on the timeline viewer (smallest unit, depends on the zoom factor of the timeline viewer, default value 10 ms)	**CTRL+SHIFT+LEFT**
	Start / Pause the playback	**CTRL+SPACE**
	Go to the next "pixel" on the timeline viewer (smallest unit)	**CTRL+SHIFT+RIGHT**
	Go to the next frame	**CTRL+RIGHT**
	Go to the next second	**SHIFT+RIGHT**
	Go to the next scroll view	**CTRL+PAGE DOWN**
	Go to the end of the media fragment	**CTRL+B**

Table 1.2. Selection Controls

Icon	Meaning	Shortcut
	Play the selected interval.	**SHIFT+SPACE**
	Clear the selection.	**CTRL+C**
	Move the crosshair to the begin / end of selection	**CTRL+/** or **CTRL+SHIFT+K**

Table 1.3. Annotation Controls
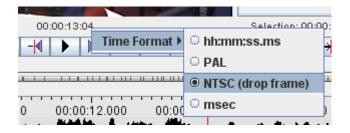
Icon	Meaning	Shortcut
	Go to the previous annotation on the active annotation tier	**ALT+LEFT**
	Go to the next annotation on the active annotation tier	**ALT+RIGHT**
	Go to the annotation above.	**ALT+UP**
	Go to the annotation below.	**ALT+DOWN**

Table 1.4. Selection Mode

Icon	Meaning	Shortcut
Selection Mode	While playing, select an interval automatically	**CTRL+K**
Loop Mode	Keep playing the selected interval (if used together with play selection ▷S)	**CTRL+L**

- The time code can be displayed in the following formats:hh:mm:ss:ms (hours:minutes:seconds.milliseconds), PAL (hours:minutes:seconds.frames), NTSC (drop frame) (hours:minutes:seconds.frames) and msec (milliseconds). It can be used to access points in time (by clicking on it and entering a value in "hh:mm:ss:ms" format). To change the display format, right click on it and select a format from the menu. The SMPTE time code formats PAL and NTSC only indicate the way time values are converted to frame numbers; this is independent of the actual video standard of the associated video(s).

Figure 1.105. Timecode formats

- The sliders available when the Controls tab is selected allow you to control the playback rate and the volume.
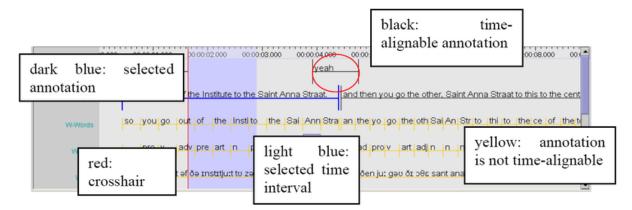
The color coding

In all its displays, ELAN makes use of recurring colors in order to facilitate the orientation in the document. The following colors are used:

- Red: Position of the crosshair (i.e., current point in time);

- Light Blue: Selected time interval;

- Dark Blue: Active annotation.

- Black with long segment boundaries: Annotations that can be aligned to the time axis.

- Yellow with short segment boundaries: Annotations that cannot be aligned to the time axis.

For example:

Figure 1.106. Color coding

How to change the display of the ELAN window

The ELAN window setup as described and illustrated in the section called "Basic Information: The ELAN window" above is the default display. But you can easily change the display according to your needs. The following options are available:

- increasing/decreasing the size of the ELAN window (the section called "Increasing/decreasing the size of the ELAN window");

- switching Viewers on/off (the section called "Switching Viewers on/off");

- increasing/decreasing the size of Viewers (the section called "Increasing/decreasing the size of Viewers");

- switching tiers on/off (the section called "Switching tiers on/off");

- rearranging the order of tiers (the section called "Rearranging the order of tiers" and the section called "Sorting tiers");

- displaying a tier in the any of the tab panes viewers (the section called "The Subtitle Viewer" and further);

- changing the time resolution (the section called "Changing the time zoom");

- changing the font size (the section called "Changing the vertical (intensity) zoom");

- de-attach/re-attach the video window (the section called "The Video Viewer").

Increasing/decreasing the size of the ELAN window

The size of the ELAN window can be increased or decreased. Do one of the following:

1. Go with the mouse to the borders of the ELAN window. The mouse will turn into a double-headed arrow. Click and move it to increase/decrease the size of the window.

2. In the top right corner of the ELAN window, click on the Maximize icon to activate the full-screen modus; click on the Restore Down icon to return to the previous size.

Switching Viewers on/off

Depending on the type of media file, ELAN automatically displays three Viewers (Video, Waveform and Timeline Viewer). Furthermore one can choose an additional viewer in the tabs on the right / left(For more info on how to choose the viewers in left / right of video see Figure 1.32, "Edit Viewers

preferences") of the video Viewer: a text viewer, a grid viewer, a subtitle viewer, a lexicon viewer , a audio recognizer, a video recognizer or a metadata viewer. Note that it is not possible to activate both the Timeline and the Interlinear Viewer at the same time. To select which viewers you want and which viewers should be visible see the section called "Show / Hide Specific Viewers".

📄 Note

> If a media file is not available (e.g., the `*.mpg`/`*.mov` file in case of audio data, or the `*.wav` file in case of some video data), the corresponding Viewer is not available either.

Show / Hide Specific Viewers

To show or hide the specific viewers like grid viewer, text viewer, subtitle viewer or lexicon viewer and the audio and video recognizer, select View > Viewer. A list of the viewers and recognizers that can be shown or hidden is displayed. Now select or deselect the viewers by clicking them. Only the selected viewers are displayed in the pane next to the video. If the video in the center, to choose which viewers are to be shown in the left and right pane of the video, and to sort the order of the viewers and recognizers see Figure 1.32, "Edit Viewers preferences" .

Increasing/decreasing the size of Viewers

The size of all Viewers (except for the Video Viewer) can be increased and decreased relative to the size of other Viewers. Do one of the following:
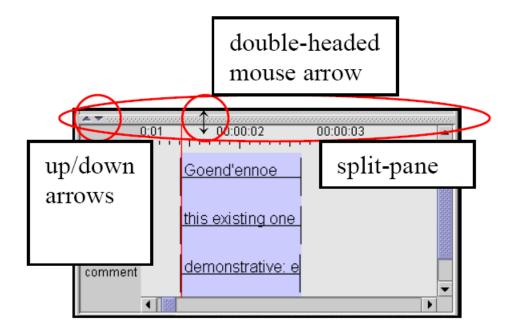
1. Use the up/down-arrows of the split-pane.

   Click on the up/down-arrow to increase/decrease the size of the corresponding Viewer.

2. Use the mouse.

   Go with the mouse to the split-pane. The mouse will turn into a double-headed arrow. Click and move it up/down to increase/decrease the size of the corresponding Viewer.

Figure 1.107. Changing size of the timeline viewer

The width of the tier label panel left of the timeline viewer can also be changed. Put your mouse cursor on the arrows in the top right corner of this panel. When the appearance of the mouse cursor changes

you can drag the right border to the left or to the right and by doing so decrease or increase the size of the tier label panel.
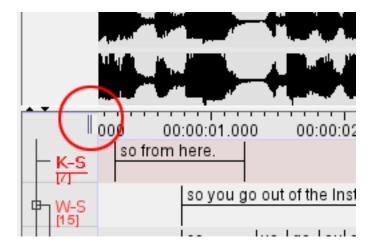
Figure 1.108. Changing the size of the tier label panel

Switching tiers on/off

By default, ELAN automatically displays all available tiers, but each tier can be switched on or off manually, allowing you to focus only on the tiers of interest for the task at hand.

To switch tiers on/off, do the following:

1. In the Timeline Viewer, right-click in the tier name panel.

2. In the pull-down menu select the sub menu Visible Tiers and (un)check the tier name

Switching off a tier can be done directly by right clicking on its name and selecting hide <tier name> from the pull down menu. Alternatively you can open a window containing all tier names by selecting Show/Hide More…(see Figure 1.111, "Visible Tiers") in the popup menu.
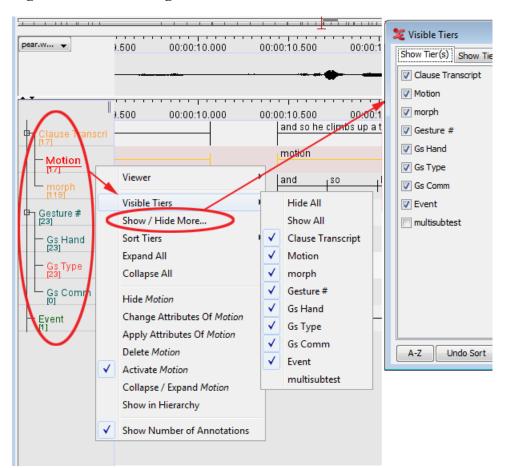
Figure 1.109. Switching tiers on/off
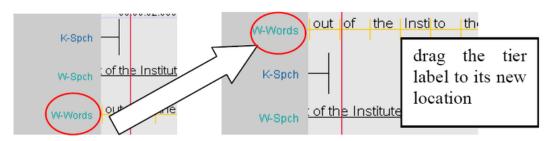
If you switch a tier on, it will be put on the place where you clicked.

- Show / Hide More... : (see Figure 1.111, "Visible Tiers")

- Sort Tiers : (see the section called "Sorting tiers")

Rearranging the order of tiers

Within the Timeline or Interlinear Viewer, you can rearrange the order in which the tiers are displayed. Just drag the tier label to its new location. The tiers will be displayed in the new order.

Figure 1.110. Rearranging the order of tiers

If you exit the document, ELAN will save the order of tiers in the following way: first, all activated tiers (in the order as they appear in the Timeline or Interlinear Viewer), followed by all non-activated tiers in alphabetical order.

View tiers by Type/Participant/Annotator

The tiers can also be viewed by its tier type or by participant or by annotator. In order to do that, right click in the tier name panel and select Show / Hide More.... This dialog window appears:

Figure 1.111. Visible Tiers

- Show Tier(s)

  Displays a list of all tiers in the transcription and the selected tiers are the visible tiers

- Show Tier Type(s)

  This shows a list of all the tier types in the transcription. Select the all the types you want to view. The tiers of the selected types are selected automatically in the Show Tier(s) list.

- Show Participant(s)

  This shows a list of all the participants in the transcription. Select the all the participants you want to view. The tiers of the selected participants are selected automatically in the Show Tier(s) list.

- Show Annotator(s)

  This shows a list of all the annotators in the transcription. Select the all the annotators you want to view. The tiers of the selected annotators are selected automatically in the Show Tier(s) list.

- A-Z

  This button is used for sorting the list of tiers shown alphabetically.
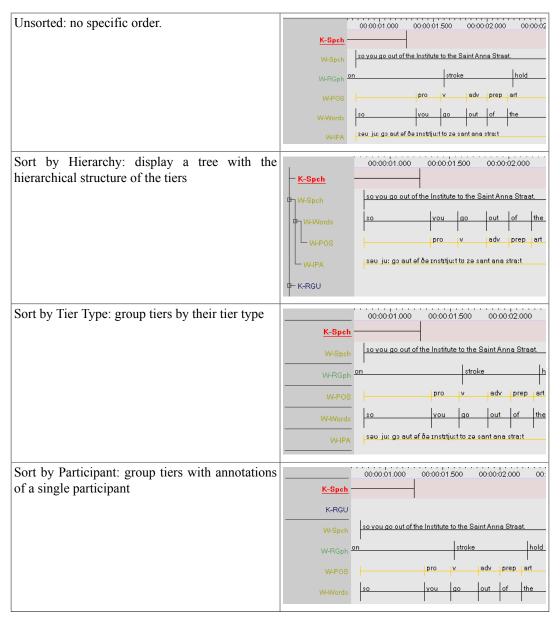
- Undo Sort

This is to undo the sort and restore the order back.

Sorting tiers

The order of the visible tiers in the timeline and interlinear viewer can be altered. To achieve this, right click in the tier name panel and select the sub menu Sort Tiers. Then choose one of the following options:

Table 1.5. Sort tier options

Unsorted: no specific order.	
Sort by Hierarchy: display a tree with the hierarchical structure of the tiers	
Sort by Tier Type: group tiers by their tier type	
Sort by Participant: group tiers with annotations of a single participant	

It also possibly to sort the tiers alphabetically along with the anyone of the sorting options before. To do this, right click on the tier name panel and select Sort Tiers > Sort Alphabetically.

Creating and applying tier sets

In some projects the number of tiers per transcription can amount to several dozens, depending on the number of phenomena that are studied, the number of levels of the analysis, the number of participants and/or the number of annotators contributing to the document. The space available for the Timeline
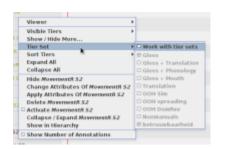
Viewer is not large enough to display all tiers and even if it was, it would be very difficult to have a good overview of all tiers and to discern annotations and the tiers they belong to. The sections on how to show or hide tiers (see the section called "Switching tiers on/off" and the section called "View tiers by Type/Participant/Annotator") demonstrate how to hide tiers from the view , either individually or as a group based on a common attribute (type, participant etc.). But even with the Show/Hide More option it can take several or many clicks to assemble a custom group of tiers to be visible. Moreover, after switching to a different set of tiers and then back again, the same selection steps have to be performed again.

To make switching between groups of tiers and combining groups of tiers easier and more reliable Tier Sets have been introduced. A Tier Set is a custom group of tiers that can combine tiers independent of hierarchical relations and independent of tier attributes (like Type or Participant etc.). A tier set is identified by a unique name by which the whole set can be made visible or hidden. The user can define multiple tier sets, the configuration of these sets is stored in a preference file. This file can be shared with colleagues so that they have the same setup available. Making a set visible in the Timeline Viewer also makes those tiers available in the tier list for e.g. the Grid Viewer, while hiding a set removes its tiers from that tier list. The main advantage of tier sets is in working with a corpus with consistent tier names.

A screencast is available on Vimeo to help you setting up a tierset: `https://vimeo.com/183124013`

The tier set feature has to be activated explicitly in the Preferences (see Preferences)or by clicking Tier Sets > Work with Tier Sets... in the context menu of the tier list in the timeline viewer.

Figure 1.112. Work with tier sets

A side-effect of activating this feature can namely be that no tiers are visible at all or that some tiers are missing after opening a file that adheres to a different tier naming convention. Therefore this should be used with care.

Tier sets can be managed via the menu Edit > Edit Tier Sets.... No document needs to be open for this, tier sets can be managed on the basis of files in a domain or a set of files selected from the file system.

Figure 1.113. Creating and editing tier sets

The left side of the window shows the list of currently defined tier sets.

• The Visible checkbox determines whether the tiers in that set are visible or hidden, whether the set is selected in the list of tier sets.

• The list can be sorted alphabetically by the A-Z button or in custom order using the Up and Down buttons.

• The - button deletes the selected tier set(s) (without warning!), the + button produces the Create new tier set dialog that allows to specify the name and contents of a new set.

The right side of the window shows the attributes and contents of the set selected in the list of tier sets to the left.
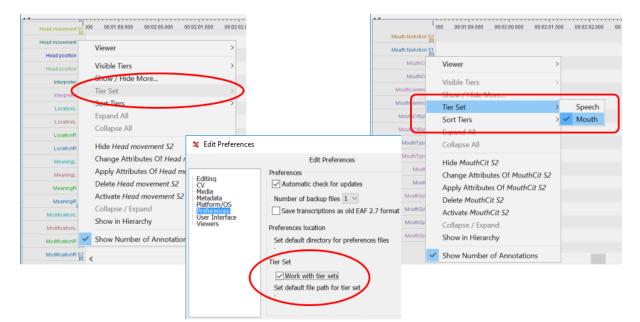
• Each set has to have a unique name and can have a description.

• The tiers that belong to this set are listed below the description. Each tier can individually be marked as Visible or hidden.

• The Edit button creates the Edit tier set dialog which is essentially the same window as the one in which to create a tier set. The name and description can be edited and the list of tiers can be updated and re-ordered. Adding tiers is done in the usual extended tier selection panel.

The Apply button stores the tier sets and closes the window. By default the tier set configurations are stored in a file `TierSet.xml` in the ELAN data folder. This file can be shared with colleagues so that they have the same sets available.

Even if tier sets have been defined they cannot be applied immediately. As shown in the screenshot below the Tier Set is greyed out initially in the context menu of the tier name panel. This feature can be enabled in the Preferences tab of Edit > Preferences > Edit Preferences. After enabling it here the

Tier Set menu will be enabled and the available sets are shown and can be selected or deselected. The choices made here also update the list of tiers that can be selected in e.g. the Grid Viewer.

Figure 1.114. The Tier Set context menu item before and after enabling tier sets in the preferences.

Changing the time zoom

The default zoom for the Waveform and the Timeline Viewer is 100%, corresponding to 10 milliseconds per pixel. The zoom can be changed simultaneously for both Viewers. Do either of the following:
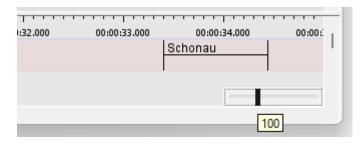
1. a. Use the keyboard combinations **CTRL+=** to zoom in (Ctrl +), **CTRL+-** to zoom out (Ctrl -) or **CTRL+0** to zoom to the default level (Ctrl 0)

   b. Click with the right mouse button on either the Waveform Viewer or the Timeline Viewer.

   c. Go to Zoom. The following menu appears:

Figure 1.115. Changing time zoom

d. Click on a zoom rate to select it. A check mark appears next to the selected zoom rate.

- Click on a lower percentage to get a finer zoom.

- Click on a higher percentage to get a wider zoom.

2. a. Place mouse cursor on the Waveform Viewer or the Timeline Viewer.

b. Press the **CTRL** key and keep it pressed.

c. Move the scroll wheel of your mouse. Moving down is zooming out and moving up is zooming in.

3. Use the zoom slider in the right bottom corner of the Timeline Viewer.

Figure 1.116. Timeline zoom slider

This slider can be hidden via a setting in the Viewers preferences panel (see Figure 1.32, "Edit Viewers preferences").

There is another zoom option called Zoom to Selection (see Figure 1.115, "Changing time zoom"). To use it, first make a selection (see the section called "Making a selection"). Then right click on the Waveform Viewer or Timeline Viewer and select Zoom > Zoom to Selection. The selection is now displayed almost as wide as the Waveform and Timeline Viewer. In the context menu beneath Zoom to Selection the option Custom is selected and the zoom factor is displayed.

Changing the vertical (intensity) zoom

Sometimes it can be handy to zoom in on the intensity of the signal displayed in the waveform viewer. This way you can more easily make the distinction between parts where someone is speaking and those where there is a silence. Such a visual amplification is available through the right-click context menu in the waveform viewer:
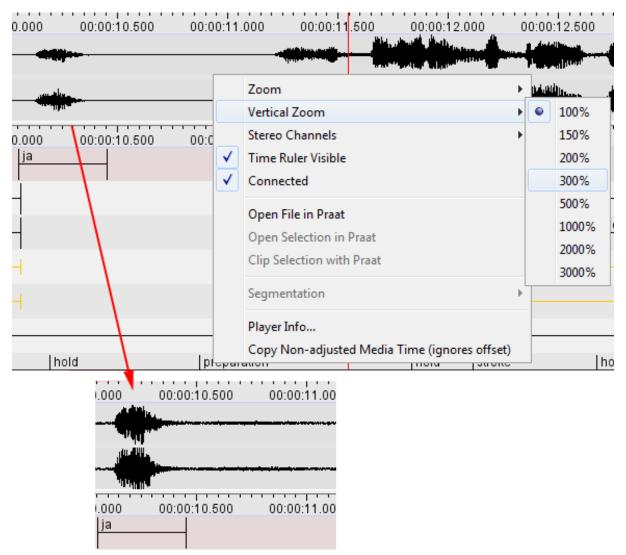
Figure 1.117. Changing vertical (intensity) zoom

Please note that this vertical zoom does not change the audio characteristics in any way.

Changing the font size

The default font size is 12 pt., but it can be changed separately for the different annotation viewers in ELAN. Do the following:

1. Right click on one of the viewers (Grid, Subtitle, Text, Timeline, Interlinear Viewer).

2. Go to Font size. The following menu appears:
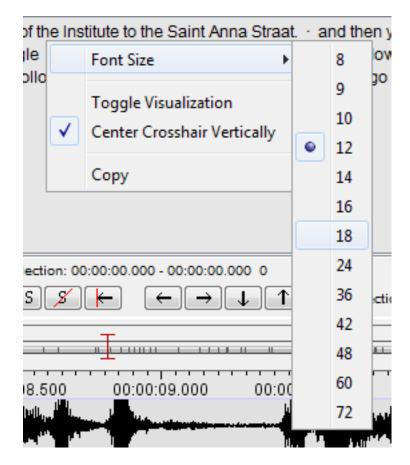
Figure 1.118. Changing the font size

3. Click on a font size to select it. A check mark appears next to the selected font size.

Font Browser

If you are not sure that the font you want to use can display all the (special) characters of an annotation (for instance IPA characters), you can check this by using the Font Browser utility ELAN offers. Click on View > Font Browser... to open the Unicode Font Finder-Explorer (see Figure 1.119, "Font Finder-Explorer"). In the first list of the explorer you can select a system font for which you want to know what Unicode subsets it can display. These subsets are displayed in the list below the list of system fonts. If you click on a Unicode subset, this subset is displayed in a new window (Font Browser for Codepage).

Another way of checking whether your special characters can be displayed in the desired font, is to enter text in the bottom text box of the Font Finder-Explorer and click on Check. Now the lists on the right of the Font Finder-Explorer will display the fonts and Unicode subsets that can display the text in the text box. Clicking on a Unicode subset will display that subset in the Font Browser for Codepage-window.

Clicking the Clear button will clear the lists, except for the list of system fonts.

Figure 1.119. Font Finder-Explorer

Importing and exporting preferences

The following display preferences can be imported and exported:

• Font

• Font size

• Tier name color

• Selected tier per viewer

• Visible/hidden tiers

Importing and exporting these preferences make it possible to apply preferences to another document. To export preferences click Edit > Export Preferences..., select a destination folder, enter a file name and click on Save. To import preferences click Edit > Import Preferences..., look up the preference file and click on Select.

Editing shortcuts

ELAN has shortcuts for many of it's functions. The default shortcuts, which are mentioned throughout this manual (for an overview see the section called "The shortcut keys"), can be changed via Edit > Preferences > Edit Shortcuts.... This dialog window appears.

Figure 1.120. Edit Keyboard Shortcuts

- The shortcuts in the table can be sorted by clicking a column header (e.g. Description or Category); the items will be sorted alphabetically, ascending or descending.

- To change a shortcut, select it and click on Edit Shortcut. Press the desired shortcut on your keyboard and click OK

  - click Apply to change the shortcut only in the selected mode.

  - click Apply in all modes to change the shortcut for this action in all the other modes if applicable.

    If the shortcut was already assigned to a function, you are asked whether the shortcut should be reassigned.

- After changing one or more shortcuts click Save to save the changes

- To restore the default shortcuts

  - click Reload Default to restore the shortcuts in the currently selected mode in this dialog.

  - click Reload All Default to restore the shortcuts in all the modes.

  Clicking those buttons will only update default shortcuts for the current instance of ELAN. Click Save to override the current shortcuts with the default shortcuts.

- Click Cancel to discard the changes

How to copy current time of the media

To copy the current time from the media, go to the Edit menu and select Copy Current Time or use the shortcutkey **Ctrl+Alt+G**

How to navigate through a document

ELAN supports the following options for navigating through a document:

- accessing points in time (the section called "Accessing points in time");

- going to a selection (the section called "Going to the borders of a selection");

- jumping step by step through a document (the section called "Jumping step by step through a document");

- navigating using the Timeline Viewer (the section called "The Media Player options")

- navigating using the Grid Viewer (the section called "Viewing a list of annotations within one tier (Grid Viewer)");

- navigating using the Interlinear Viewer (the section called "Jumping from annotation to annotation (Interlinear Viewer)").

All Viewers are synchronized in time, i.e., when you navigate to a specific point or selection in one Viewer, all other Viewers will immediately jump to the corresponding point or selection:

- The Video Viewer will display the corresponding video frame

- The Waveform Viewer will display a crosshair at the corresponding location in the waveform.

- The Subtitle, Timeline and Interlinear Viewers will display the corresponding annotation(s).

Accessing points in time

To access a point in time, do one of the following:

1. Use the time code box.

   a. Click on the time code above the media playback controls (left side of the ELAN window). The Goto dialog window appears.
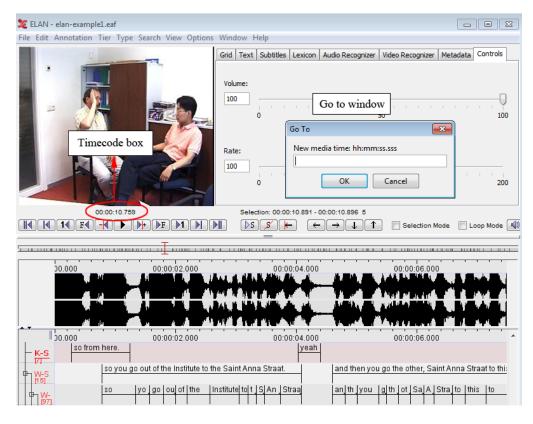
Figure 1.121. Click the timecode box

b. Enter the time code in the following format: "hh:mm:ss.sss" (hours:minutes:seconds.milliseconds). The crosshair jumps immediately to the selected time code.

You can omit part of the time code information:

- If the digits are "00", you can omit "hours:", "hours:minutes:", or "hours:minutes:seconds."

- It is optional to specify ".milliseconds".

📑 Note

The character preceding the milliseconds is a dot, not a colon.

2. Use the crosshair

Figure 1.122. Use the crosshair

Click somewhere in the Text, Subtitle, Timeline, Waveform or Grid Viewer. The crosshair will jump to that point. By holding the **ALT** button and dragging the time axis to the left or to the right you can scroll through the annotations.

3. Use the media control buttons

   See the section called "The Media Player options"

4. Use the 'Shift' + Scrollwheel function. When pressing and holding 'Shift', you can scroll horizontally with the scrollwheel on your mouse. On a laptop or macbook, you can use two finger scrolling to achieve the same effect.

Going to the borders of a selection

If you have made a selection (see the section called "How to make a selection"), you can move the crosshair to the begin or end of that selection from anywhere within the file.

To jump to the begin, click on the [←] button, which is part of the Selection Controls. Then the button will show an arrow in the other direction [→], which brings the crosshair to the end of the selection.

Jumping step by step through a document

It is possible to move the crosshair back and forth step by step. To achieve this, use one of the media controls as described in the section called "The Media Player options".

If the media framework ELAN is using cannot determine the video format, you can alter the step size when using the next/previous frame control. This is useful in order to work with a "natural" frame duration, depending on the video format that is used (i.e. 25 frames/second for PAL or +/- 30 frames/second for NTSC).

Do the following:

1. Click on Options menu.

2. Go to Set Frame Length. The following menu appears:

Figure 1.123. Set Frame Length menu

Select an option from the menu:

- PAL: The 1 frame step size for video data corresponds to one PAL frame (40 ms)

- NTSC: The 1 frame step size for video data corresponds to one NTSC frame (33 ms)

3. Click on a Frame Length mode to select it. A radio bullet appears next to the selected step mode.

Viewing a list of annotations within one tier (Grid Viewer)

ELAN allows you to list all annotations from a single tier, and then select one annotation and jump to it.
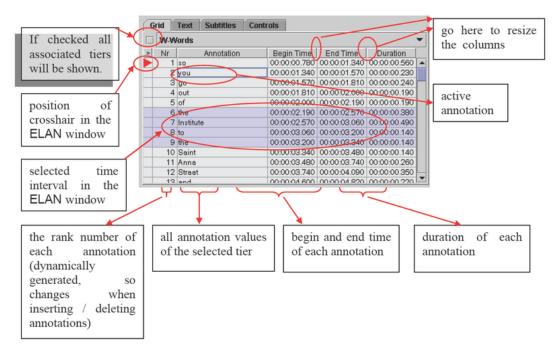
To activate the Grid Viewer, click on the Grid tab:

Figure 1.124. The grid viewer

The structure of the Grid View window is as follows:

Figure 1.125. Structure of the grid viewer

You can change the Grid Viewers interface at any moment by right clicking in the Grid Viewer. A context menu will appear:

Figure 1.126. Grid viewer with context menu

In this context menu, you can choose between the following options:

• choose the columns displayed in the Grid Viewer

• changing the Grid Viewers font size

• toggle between the time code format (hh:mm:ss.mmm) and milliseconds

You can use the Grid Viewer window to navigate to an annotation in the ELAN window. You have the following two options:

1. Move the crosshair to the begin of an annotation.

   Do the following:

   a. In the Grid Viewer window, click with the mouse button on an annotation.

      A red triangle appears next to the annotation in the Grid Viewer window, and the crosshair moves to the beginning of that annotation in the ELAN window.

2. Select a time interval.

   Do the following:

   a. In the Grid Viewer window, click with the mouse button on the first annotation that you want to select.

   b. Keep the mouse button down and drag the mouse to another annotation.

      In the Grid Viewer window, all selected annotations are highlighted in light blue color. In all other windows, the corresponding time interval is selected and highlighted in light blue color (starting with the beginning of the first annotation and ending with the endpoint of the last).

📄 **Note**

   Selecting a time interval also changes the current time. This happens implicitly by moving the crosshair to the begin of the annotation.
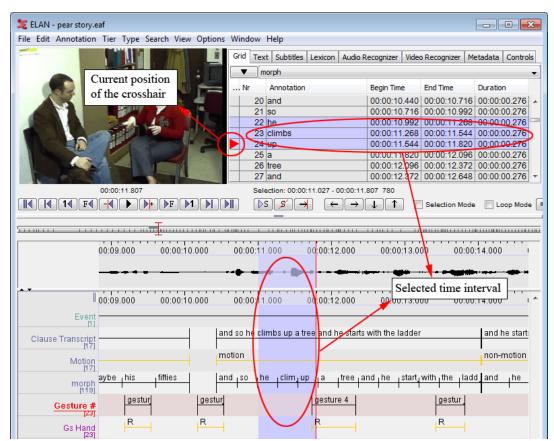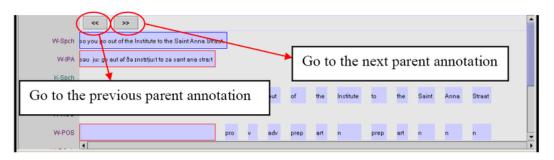
Figure 1.127. Grid viewer with selected annotations

Jumping from annotation to annotation (Interlinear Viewer)

In the Interlinear Viewer, you can jump from one annotation block forward/backward to the next block. Do one of the following:

1. Click on the left arrow button at the top of the Interlinear Viewer to move to the previous annotation block.

2. Click on the right arrow button at the top of the Interlinear Viewer to move to the next annotation block.

Figure 1.128. Interlinear viewer

How to play a document

ELAN supports the following options for playing a document:

- playing a document (the section called "Playing a document");

- playing a selection (the section called "Playing a selection");

- playing around the crosshair/selection (the section called "Playing around a selection");

- playing in slow motion (the section called "Playing in slow motion").

Playing a document

During playback, the red crosshair moves through the Waveform, Timeline, Interlinear, Annotation Density, Text and Grid Viewers, indicating the point in time that is being replayed. If you stop the playback, the crosshair will stop at this point in time. The playback will resume from this point and will continue until the end of the document, or until the playback is paused again.

Make use of either one of the following three options to start/pause the playback:

1. Use the Play/Pause button located under the Video window:

    a. Click on the Play icon to start playback. After the playback starts, the Play icon turns into a Pause icon.

    Figure 1.129. Play/pause button

    b. Click the Pause icon to pause the playback again.

2. Use the Shortcut key **CTRL+SPACE** to start the playback. Use it again to pause the playback.

📑 **Note**

If two or more ELAN documents are open at the same time, the sound may not work properly. Should this happen, close all documents except for one.

Playing a selection

It is possible to play only a selected part of the document. Do the following:

1. Make a selection (see the section called "How to make a selection"), or click on an annotation.

2. To play the selection, do one of the following:

    a.
    Click on the Play Selection ⟨S⟩ icon.

b. Or use the shortcut key SHIFT + SPACE.
The selected part is played immediately. To listen to the selection again, repeat step 2.

📝 Note

If the crosshair is positioned somewhere within the selection (i.e., if it had been manually moved forward or backward, see the section called "Jumping step by step through a document"), playback will start from that position and stop at the end of the selection. Otherwise, the whole selection will be played.
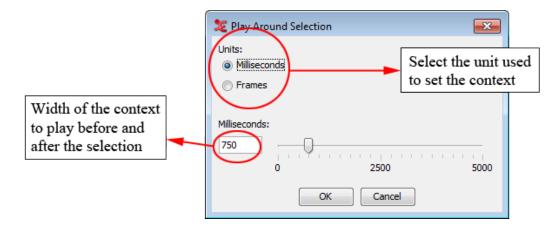
If you want to loop over the selection, be sure to check the Loop Mode box (next to the Selection Mode checkbox).

Playing around a selection

When playing a selection it is possible to extend the playback interval with a context up to a few seconds before and after the selection. This can e.g. be useful to get an idea about the preparation and the finishing of a gesture.

1. Click on Options menu.

2. Go to Play around selection. The following dialog appears:

Figure 1.130. Play around selection
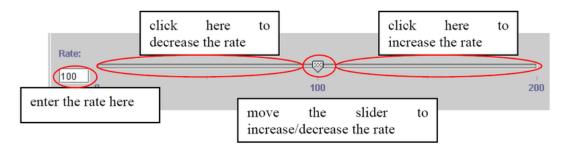
3. Choose how many (milli)seconds or frames there should be played before and after the selection. Click on the OK button.

4. Now make a selection and press **CTRL-SHIFT-SPACE**. This keyboard shortcut is the only way to play around a selection.

Playing in slow motion

To change the playback rate, use the Rate slider to the right of the video window. The following options are available:

• Click somewhere above or below the slider to increase or decrease the playback rate by 1%.

• Move the slider up or down to increase or decrease the playback rate.

• Enter the desired playback rate in the box at the left of the slider and press the key ENTER. ELAN accepts rates between 1% and 200%.

Figure 1.131. Playback rate

Once you have selected a playback rate, the document will be played at the selected rate. If you want it to be played at a different rate, you have to manually change the playback rate, repeating the steps above.

📑 **Note**

On slower machines, the slow motion playback may not work properly.

Changing playback rate and volume via the keyboard

In the Options > Rate and volume toggle… menu, a pre-set value can be specified for the playback rate and volume which can be activated through a keyboard shortcut:

- **CTRL+SHIFT+R** : alternate the current volume with the pre-set value

- **CTRL+ALT+R** : alternate the current playback rate with the pre-set value

Adding comments to .eaf files

This section will describe how to add comments to an annotation project. ELAN supports annotation and display of time-based resources such as audio and video with a referencing and note exchanging system. Persistent references to time slices and tiers are possible by making and exchanging comments referring to these. Different strategies for exchanging notes based on these links are available, allowing for collaboration among users. There is support for sharing of comment collections in files stored on shared directories (for instance using a cloud directory sharing service) and for exchanging of single or multiple comments with their metadata via ordinary email messages (the DWAN server system that used to be an option as well, is no longer online). Downloading an eaf file from the IMDI-browser will also provide the user with the comments file.

A screencast is available on Vimeo explain the comments function: `https://vimeo.com/126680719`

Each eaf has a specific id and all shared comments regarding that eaf will be shown to you.

Figure 1.132. Comments-tab in ELAN

The Comments tab is divided in two panels; the left-hand side shows all comments associated with the .eaf file you are working on, the right-hand panel is used to enter your comments. Furthermore, there are six buttons which have different functions:

- Add Comment the section called "Adding, deleting & changing comments"

  After creating a selection on the timeline viewer, enter your desired comment on the right-hand side. After clicking this button, the selection gets added to the left-hand side of the comments dialog.

- Change Comment the section called "Adding, deleting & changing comments"

  This button is only active when comments have been entered. You can change a comment using this button.

- Delete Comment(s) the section called "Adding, deleting & changing comments"

  This allows you to delete a comment from the comments-list.

- Filter...the section called "Filter comments"

  The filter button allows you to filter comments based on a regular expression.

- Log in... the section called "Comments setup and settings"

  Clicking this button allows you to log into a comments server. The DWAN-server that used to be maintained by the MPI is no longer available, therefore this function is only useful if a compatible server is hosted by a third party.

- Other... the section called "Other functions "

  This button has a group of functions: To Mail. To Clipboard, From Clipboard, Search & Settings. These functions are also available by right-clicking anywhere in the right-hand side of the comments tab.

Comments setup and settings

To take full advantage of the comments-system, you can adjust some settings. Go to the Comments tab and click Other... and select Settings from the drop-down menu. In the new dialog the following settings can be modified:

Figure 1.133. Directory settings and default comment attributes

The first options under Comments settings allow to specify folders that the system will use for storing and searching files.

• Location of shared directory. This can be a local folder, in which you can save your eaf and comment files. You can also point this to e.g. a Dropbox folder. This way, you can share the files via a cloud network.

• Use shared directory for this transcription. When a shared folder is set, you can check this option so your current comment file will be saved to this folder. The eaf file itself will be saved to a location of choice, not in this shared directory.

• Search Comments in. To search for comments, a folder needs to be set that holds the comments-files. Once set, the search will be recursive, so sub-folders containing comment-files will also be searched.

• Search EAF files in. The EAF files can be stored in a separate location from the comments files. When you search for comments and want to open the EAF file associated with that, the system needs to know where the EAF file is located. So it is advised to point this to the location where you normally store your EAF files.

The Default comment field values options allow to set some default attributes for commments.

• Sender email address. When comments are shared by email, this will be used as the default `From` address (your email address).

• Recipient email address. When comments are shared by email, this will be used as the default `To` address.

• Initials. Your initials, code or name as the creator of new comments.

• Thread ID. The Thread ID can be used to identify individual "conversations", if you want to use the comment system in that way. To do that, choose some meaningful short name for each separate conversation.

Note that you can edit all these fields in each of your comments at any time.

Under Stored Comments several server related properties can be set.

• Service URL. The base URL of a comments web service (no default service provided anymore)

• Server login name. The login name you registered with that service.

• Update check time. This ranges from 2 minutes to over 50 minutes. The system will check the server every X minutes, to see if there are any changes made or added in the project you are working on. If a change was made, a warning will display with information regarding the change. It will also prompt you for an action.

Figure 1.134. Conflicting comments dialog

• Postpone will postpone the prompt to take action for now.

• Keep local comment as new will keep the external comment and also save your local comment in place as a new comment.

- Keep local version (and export it) will keep the local version and export it to the server.

- Replace local version will replace the local version you have with the external comment.

When done, click Apply to apply the changes.
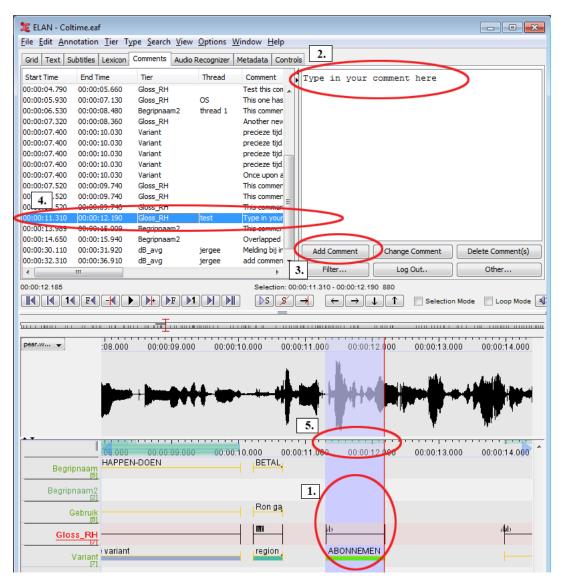
Adding, deleting & changing comments

Adding comments can be done either when you are logged in, but also when you are not logged into the server. When not logged in, the comments will be saved locally and will only be synced with the server once you log in.

To add a comment, create a selection on the timeline that you wish to comment on (1.). Make sure you select the tier you want to refer to by double-clicking it, so it becomes highlighted in red.

Next, type in the comment in the right-hand field of the comments-tab (2), and finally click Add Comment to add the comment to the left-hand field of the comments-tab (3). Your comment will appear in the list of comments on the left (4). A visual indicator of the position of the current comment will also be added to the Timeline viewer (5).

Figure 1.135. Adding comments

Deleting a comment is only possible for the comments that you have added, comments added by other users cannot be deleted. They have a yellow background, indicating it is a comment from another user.

Figure 1.136. A comment added by another user.

To delete a comment that you added previously, select it in the left-hand screen and click Delete Comment(s). To delete multiple comments at once, press and hold Ctrl(Windows)/Cmd(macOS) and click the comments to delete. When done, click Delete Comment(s).

Changing a comment is done in a similar way to deleting a comment. First, select the comment you want to change from the list of comments (only comments made by you can be altered). Next, click inside the text-edit field on the right and alter your comment. The comment will change to a red color, indicating that it is not yet saved on the server. Finally, click Change Comment to save the changes you made.

Figure 1.137. Changing a comment.

📒 **Note**

There are keyboard shortcuts available for adding or changing a comment:

- When adding a new comment, press and hold Ctrl(Windows)/Cmd(macOS) and press enter to add a comment.

- When changing a comment press and hold Ctrl(Windows)/Cmd(macOS) and press enter to change the comment.

Viewing comments & display options

While you can only edit your own comments, you can always view all comments. Clicking any comment in the left dialog once will display the comment on the right. Double-clicking the comment will highlight the selection on the timeline viewer where the comment refers to. If a tier is linked to the comment, it will be activated as well.

The left-hand side of the comments-tab displays various columns regarding a comment. These provide certain information regarding the comment and the Time line. The available columns are: *Start Time, End Time, Tier, Initials, Comment, Thread, Sender, Recipient, Creation Date & Modification Date.* To display or hide any of these columns, right-click anywhere in the comments overview display and check or uncheck a column. You can also sort the order of the columns and resize them.

Figure 1.138. The comments display overview

You can also alter the sort order of the comments, by clicking the header of the column you want the comments to be sorted by. You can sort by Start Time ascending, for instance, but also by Comment, alphabetically. The columns will remain hidden or displayed until you alter these settings again.

Filter comments

The Filter... button allows you to filter out comments displayed on the screen. This can be helpful when multiple users have submitted comments and you only want to see comments made by a certain user, or by a certain thread name for instance. To filter the comments, click the Filter... button. A dialog appears, in which you can type in a regular expression you wish to use as a filter.

Figure 1.139. Changing a comment.

Once you apply your filter, you will see the results in the main screen. Note that the background will turn blue, indicating that a filter is active. The Filter... button font will also be blue as an indicator that a filter is active. To remove your filter, simply click the Filter... again and remove the regular expression you entered, then click apply. The filter will then be removed.

Other functions

The button Other... incorporates multiple functions. You can access these by just clicking on the button, the options will appear below it.

Figure 1.140. Other functions

The functions available are:

1. To Mail... This allows you to send a mail containing the selected comment in XML format to another user. (you must have a native mail application installed for this to work).

2. To Clipboard. This sends the current comment-info in XML-format to the clipboard so you can paste it in another application, or re-use it later in ELAN.

3. From Clipboard. This allows you to paste data from the clipboard into the comments field. This needs to be in XML-format to work. If you have received an e-mail containing comment-data, you can copy that data to the clipboard and use this function to add the comment to the list.

4. Search... This function allows you to search for a comment in the list of current comments. This is done by entering a regular expression, and can be set to search all columns or a single column.

Clipboard and e-mail comment function

Copying data to the clipboard or sending it via e-mail will be done in XML format. When you click the To Mail... option, your local e-mail application will open and create a new message. It will also fill out the recipient address that is associated with the comment you selected. The recipient can be filled out in the settings, see the section called "Comments setup and settings". The data in the e-mail will look like this:

Figure 1.141. E-mail a comment

When you receive an e-mail, just select all the xml-data and copy that to the clipboard. Next, in ELAN choose From Clipboard and the comment will be added to the list of comments.

In a similar fashion, you can also copy data to and from the clipboard. This can be helpful if you use a web-based e-mail client or for if you want to copy multiple comments to a text-document, for instance.

Multiple File Operations

How to select multiple files

In the dialog window below (which pops up as the first dialog for most of the multiple file operation) do one of the following:

Figure 1.142. Multiple file selection dialog

- Select an existing domain from the list and click Load. (Click Delete if you want to delete the domain.)

- Create a new domain:

  1. Click New Domain...

  2. Click in the new dialog on the Look in pull down box and browse to the directory that contains the annotation files.

  3. Double-click an annotation file (`*.eaf`) to select it. It now appears in the rightmost box. Alternatively, you can click on the annotation file name and click the >> button.

     Repeat this for every annotation file you want to include.

     It is also possible to select a complete directory. All `.eaf` files in a selected directory will be included.

  4. Click OK to continue the exporting process; otherwise click Cancel to exit the dialog window without exporting.

  5. If you clicked OK you can save this domain: enter a name and click OK. If you do not want to save the domain click Cancel.

- Create a new domain from an IMDI search:

  1. Click New Domain from IMDI Search...

2. Browse to and select an IMDI file that has been exported from a metadata search in the standalone IMDI Browser.

3. Click Open.

4. You can save this domain: enter a name and click OK. If you do not want to save the domain click Cancel.

Editing multiple files and analysis of multiple files

Normally ELAN allows to edit only a single file at a time. There are situations in which it is convenient to edit multiple files at once. The menu item File > Multiple File Processing gives a number of options to do just this. When selecting either of them, you are warned that you should have copies of the files you are going to work on in case you want to restore the files (there is no **Undo** for multiple file edits).

Create transcription files for media files

When you choose this option Create Transcription Files for Media Files..., you see the following dialog.

Figure 1.143. Create transcription files for multiple media files Dialog

Options :

- Select folder containing media files : click Browse button to select the folder containing media files.

- check this option process sub-folders as well, recursive to include all the media files in the recursive sub-folders of the selected folder.

- To apply a template (only if required) for the new transcription files, check this option Use a template for the new transcription files and click on Browse button to select the template file.

- Next option allows you to select a location for the new transcription files.

- To put the transcription files in the same folder as the media files, select in the same folder as the media files.

- To put them in a different folder, select in other folder and click the Browse button to select the destination folder.

- You could always have more than one media file in a transcription. In order to group the media files for a transcription, check this option Combine videos based on. In order to define how the media should be grouped, select one of the following:

  - different suffix : to combine the media file with a different suffix and has the same file name.

  - different prefix : to combine the media file different prefix value and has the same file name.

  To specify a separator in the file name to identify the suffix or prefix, check this option Specify custom affix separator character ('-' and '_' are built in).

Click on Start to create the transcriptions based on the options set.

Edit Multiple Files

The option Edit Multiple Files... shows, after clicking Yes in the warning dialog mentioned above, the Multi File Editor. The first thing to do here is to load a domain by clicking Load domain. Loading a domain is the same as for the Scrub Transcriptions... option. To be able to load a domain you must of course have created one beforehand (see the section called "How to select multiple files"). After loading a domain, the data is shown in the table. In this table you can edit tiers on the Tiers tab and tier types on the Tier Types tab.

To edit a name, annotator or participant of a tier, double click the corresponding table cell or select it and start typing. To change the tier type of a tier, select one from the drop down menu. You can add a tier by clicking Add tier, add a depending tier by clicking Add dependent tierand remove one by clicking Remove tier.

📄 Note

If there are hierarchy inconsistencies (e.g. if a tier in one file does have a parent while a tier with the same name in another file does not) removing tiers is not possible. The button Remove tier is therefore greyed out.

Figure 1.144. Multiple File Editor

On the Tier Types tab, the name of a tier type can be changed by double clicking the corresponding table cell in the Type Name column.

Changes made in the **Tiers** and **Types** tabs are applied to all the files in the domain after clicking the **Save changes to domain files** button.

Scrub Transcriptions

When you choose Scrub Transcriptions..., you first need to specify a new domain or select an existing domain. This option helps you to "clean" the annotation files (*.eaf) of possible tabs or white space characters which are often overlooked by the user but are still saved in the file. To select, create or delete a domain see the section called "How to select multiple files".

In the next dialog, you can specify what characters to delete, new line characters, tab characters and/or white space characters, and in what position these characters have to be. Click Start to start the scrubbing process. The progress of the scrubbing is shown in the progress bar.

Figure 1.145. Transcription Scrubber

Annotations from overlaps

The option Annotations From Overlaps... for multiple files is the same function as annotations from overlaps in the current open file (see the section called "Creating annotations from overlaps"), but applied to a selection of files. The first step allows to select a custom set of files in a file browser or to load a set of files that have been stored as a domain. For loading or creating new domain see the section called "How to select multiple files" The list of tiers is the sum of all tier names encountered in the selected files. The options in the next steps are the same, clicking the **Finish** button in the last step the new tier is created and populated with annotations in all files of the domain.

Calculate inter-annotator reliability

In some projects two or more annotators annotate the recordings and the need exists to assess the level of agreement, in order to be able to improve the training of the annotators and in the end the annotation quality. Several algorithms have been implemented and are being applied in research projects but there doesn't seem to be consensus on what the best approach is for this type of data (time aligned annotations). The option Calculate inter-annotator reliability... allows to calculate an agreement measure for annotations on multiple tiers in multiple files, providing three different methods based on existing algorithms. Even though multiple tiers can be selected, the comparison is always performed on pairs of tiers. These methods are provided "as is". Implementation of other algorithms might be added later, if time allows. The three methods are:

• calculation of (modified) Cohen's kappa according to Holle & Rein

• calculation of (modified) Fleiss' kappa

• calculation of the ratio of the overlap and the total extent of two annotations

• calculation of the degree of organization using Staccato

Using the Calculate inter-annotator reliability... function requires a few steps to be taken. Some of the steps differ depending on the choices made; other steps are common to all methods. The steps and their availability depending on the choices made are described below.

- 1. Method selection - selection of the algorithm that will be applied

- 2. Customization of the selected method (*if applicable*)

- 3. Document and tier configuration

- 4. Select files and matching (*not available when current document & manual selection is chosen*)

- 5. Tier selection

- 6. Execute the calculation and save the output file

The following sections will describe each step of the process in more detail.

Step 1. Method selection

Figure 1.146. 1. Method selection

The first step in the process allows you to select a method of comparing. Depending on what method is selected, the steps after that will differ. The options to choose from are:

- By calculating modified Cohen's kappa:

  This option implements (part of) the Holle & Rein algorithm as described in this publication: Holle, H., & Rein, R. (2014). EasyDIAg: A tool for easy determination of interrater agreement. Behavior Research Methods, August 2014. The manual of EasyDIAg [http://sourceforge.net/projects/easydiag/] can be consulted for a detailed description and explanation of the algorithm.

- By calculating modified Fleiss' kappa:

  This option provides a modified implementation of Fleiss' kappa [https://en.wikipedia.org/wiki/Fleiss'_kappa]. The modification concerns (as is the case for the modified Cohen's kappa implementation) the matching of annotations to determine the "subjects" or "units" and the

introduction of the "Unmatched" or "Void" category for annotations/events that are not identified by all raters. If the raters only have to apply labels to pre-existing segments, the problem of matching annotations does not exist. Fleiss' kappa works for two or more raters (the other options are limited to two raters).

• by calculating the ratio of overlap and total extent:

This is a simplified version of the function that used to be under Tier > Compare annotators.... It calculates a raw agreement value for the segmentation, it doesn't take into account chance agreement and it doesn't compare annotation values. The current implementation only includes in the output the average agreement value for all annotation pairs of each set of tiers (whereas previously the ratio per annotation pair was listed as well).

• by applying the Staccato algorithm:

This will compare the annotations (the segmentations) of 2 annotators using the Staccato algorithm. See this article for more information on the Staccato algorithm: Luecking, A., Ptock, S., & Bergmann, K. (2011). Staccato: Segmentation Agreement Calculator according to Thomann. In E. Efthimiou G. & Kouroupetroglou (Eds.), Proceedings of the 9th International Gesture Workshop: Gestures in Embodied Communication and Human-Computer Interaction (pp. 50-53) [https://pub.uni-bielefeld.de/publication/2392895].

Once you choose a method, click Next to continue. Note that when the Kappa or Staccato are chosen, the next step will be 'Customize compare method'. Otherwise the next step is 'Document & tier configuration'.

Step 2. Customize compare method

Figure 1.147. a. Customize compare method for modified Cohen's kappa

When the modified Cohen's kappa method is chosen, this step allows you to specify the minimal required percentage of overlap. This is the amount of overlap as a percentage of the duration of the

longest of the two annotations. The higher the percentage, the more the annotations have to overlap to match.

You can choose to generate and export agreement values per pair of tiers, in addition to the overall values. Since this algorithm compares annotation values as well, it is best to select tiers that share the same (controlled) vocabulary. When done, click Next.

Figure 1.148. b. Customize compare method for modified Fleiss' kappa

The options for Fleiss' kappa are slightly different; the slider here allows to specify a percentage between 1 and 100. Since there can be any number of raters, the annotation matching algorithm tries to create clusters of as many overlapping annotation as possible, taking into account the required average of the percentages of the overlap and each involved annotation's duration. The figures below try to illustrate the problem.

Figure 1.149. Six raters, two possible clusters of four and six matching annotations, the overlap in light blue
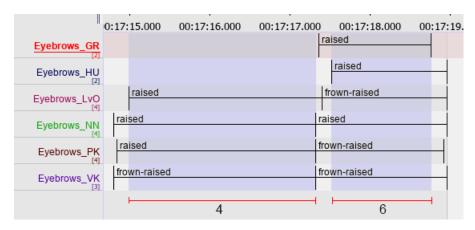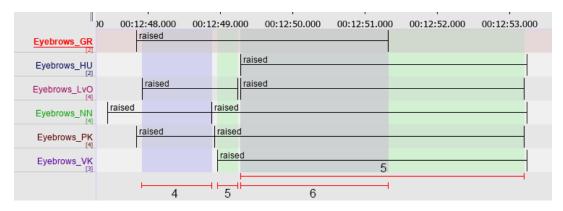
Figure 1.150. Six raters, four of the possible clusters of matching annotations are marked in blue and green

The algorithm gives preference to clusters with more annotations, as long as the required average percentage of overlap is met. If not, a cluster with less annotations is selected. Each annotation can only be part of one cluster.

The Also export matrices checkbox allows to not save the tables of values (see the worked example at Wikipedia), but it is recommended to accept the default.

Figure 1.151. c. Customize compare method for Staccato

When you've selected the Staccato algorithm as the compare method, the settings as shown above will appear. You can customize the settings for the Staccato algorithm here. This algorithm takes chance into account by comparing the segmentation with a series of randomly generated segmentations, the Monte Carlo simulation. The nomination length granularity determines how many memory slots for segments of different length will be used internally. For more in-depth information regarding these settings, please see the reference article mentioned before. When done, click Next.

Step 3. Document & tier configuration

Figure 1.152. 3. Document and tier configuration

The next step is to configure where the tiers that you want to use for comparison are located and how they should be paired. In the upper part of the dialog, you select the location of the tiers:

- *In the current document* (available when you have an .eaf file open, otherwise this option will be greyed out)

- *In the same file* (choose a single file in the next step)

- *In different files* (choose multiple files in the next step)

In the lower part of the dialog, you can select in what way the pairing of tiers to compare is done:

- *Based on manual selection* (select tiers from a list in the next step)

- *Based on prefix or suffix* (customize in the next step)

- *Based on same tier name* (only available when the option 'in different files' is chosen)

When done, click Next to continue.

Step 4. Select files & matching

Next, you will select which files you want to use for comparison and how to match tiers.

Figure 1.153. 4. Select files & matching

The screen above shows all possible options. The options available for this step will differ depending on the configuration you made in the previous step. It will not be available when the option *'in current document'* together with *'based on manual selection'* was chosen in the previous step.

- *Select files from file browser* Browse for one or more files that you wish to use.

- *Select files from domain* Choose or create a domain of files to use, see the section called "How to select multiple files".

- *Combine files based on different suffix/prefix* When using multiple files, choose how they should be combined. E.g. if there is a certain naming convention for the files and the annotations of the first annotator are in files like "Recording4_R1.eaf" and those of the second annotator in files like "Recording4_R2.eaf" and suffix has been selected, these files will be combined automatically.

- *Combine tiers based on different suffix/prefix* Similar as with files, when a certain naming convention has been used for tiers of different annotators, they can be combined on the basis of a prefix or suffix, e.g. A_LeftHand and B_LeftHand in case of prefix-based matching.

- If Fleiss' kappa was selected and the tiers to compare are in different files, an option is available to create and store new EAF files containing the matching tiers (experimental).

Step 5. Tier selection

Figure 1.154. 5. Tier selection (current document / manual selection)

In this last dialog, you will select the tiers used for comparison. The layout will be different, based on what you selected in previous steps. The screen above displays the dialog when you've chosen the option '*In the current document*' & '*based on manual selection*' in step 3. You can manually select which tiers to compare.

Figure 1.155. Tier selection (in the same file / based on suffix)

The dialog above will appear when '*based on pre/suffix*' in step 3 is chosen. Marking a tier from an annotator will result in a highlighted corresponding tier in the lower part of the dialog. When '*based on same tier name*' was chosen, you can only select tier names, corresponding tiers will not be visible in the dialog.

Finally, click **Next** or **Finish** to save the output text file to a location on your computer.

Annotations from subtraction

The option Annotations From Subtraction... for multiple files is the same function as annotations from subtraction in the current open file (see the section called "Create annotation by subtraction"), but applied to a selection of files. The first step allows to select a custom set of files in a file browser or to load a set of files that have been stored as a domain. For loading or creating new domain see the section called "How to select multiple files" The list of tiers is the sum of all tier names encountered in the selected files. The options in the next steps are the same, clicking the **Finish** button in the last step the new tier is created and populated with annotations in all files of the domain.

Modify boundaries of all annotations

The multiple file function Modify Boundaries of All Annotations is similar to the variant for the current document (see the section called "How to modify the boundaries of all annotations of selected tiers"). Except that there is no Undo and good backup copies of the files should be there. The files to process can be selected in a file browser or from a domain (see the section called "How to select multiple files"). The options in the second step are the same and the Finish button starts processing of all files. At the end of the process a short report lists which tiers in which files have been altered.

Statistics for multiple files

This function Statistics for Multiple Files is similar to annotation statistics for the current file (see the section called "Annotations Statistics"). The main difference after selecting the files in the domain is

that it is possible to select which tiers to include in the calculations. The tables in the tabs do not have the column showing the total annotation duration as a percentage of the media duration but most do have a column for the number of files a certain value (tier or type name etc.) has been encountered in. After changes in the selection of files or in the selection of tiers the **Update Statistics** button needs to be clicked before the new calculations are started.

N-Gram analysis

Since ELAN 4.7, you are able to do an N-Gram analysis over multiple eaf files. This functionality has been developed by Larwan Berke, you can find an extensive PDF document about this implementation on the third-party resources page of ELAN: https://archive.mpi.nl/tla/elan/thirdparty [https://archive.mpi.nl/tla/elan/thirdparty#documentation]

When you first open the N-gram analysis, a new dialog window will pop up that contains the various options for the search and the resulting table showcasing a few statistics.

Figure 1.156. N-Gram Analysis dialog
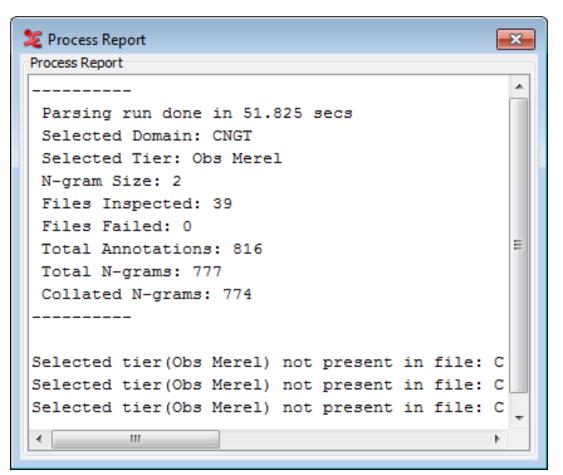
The first step is to select the search domain, see the section called "How to select multiple files". Once that is done a list of tiers seen in the domain will be shown. A note of caution: the code assumes that all files in the domain will contain the same tiers. It then loads the first file in the domain to extract the tiers and display it in the window. Check the tiers you want to analyse.

Figure 1.157. N-Gram Analysis dialog 2

Next, define the N-gram size in the text box. The software can handle any positive size greater than 1. When set, clicking the "Update Statistics" button will start the search and will calculate the statistics. The annotations are extracted from the files, N-grams created from them, and finally collated into groupings of same N-grams for statistical analysis. When done, you will see a pop-up window with a process report. If there were any errors, they will also be displayed here.

Figure 1.158. N-Gram Analysis process report


```
----------
 Parsing run done in 51.825 secs
 Selected Domain: CNGT
 Selected Tier: Obs Merel
 N-gram Size: 2
 Files Inspected: 39
 Files Failed: 0
 Total Annotations: 816
 Total N-grams: 777
 Collated N-grams: 774
----------

Selected tier(Obs Merel) not present in file: C
Selected tier(Obs Merel) not present in file: C
Selected tier(Obs Merel) not present in file: C
```

When the search is done, the result table will be displayed in the main window. Some of the columns from the data are visible here, however only a small subset is displayed simultaneously to avoid overcrowding the GUI. The visible columns are: N-gram, Occurrences, Average Duration, Minimal Duration, Maximal Duration, Average Annotation Time, and Average Interval Time.

Figure 1.159. N-Gram Analysis dialog 3

The first column shows the N-gram. The vertical marker "|" separates the annotations contained in the bigram. For example, if a trigram was selected it would show something similar to "FINISH|READ| BOOK" and so on for larger N-gram sizes.

Finally, in order to see the entire data that was produced it is necessary to export the results into a text file for further processing. This is done by clicking on the "Save" button and a dialog will pop up asking the user where to save the data. It is exported in a CSV-like format (Comma-Separated Values). The CSV file uses tabs "\t" as the delimiters and newlines "\n" as the record separators to avoid ambiguity with the values. A sample row is: "HOLD|IX-1p\t7.9934\t0.348\t0.13754 ..." and contains numerous columns.

Furthermore, it is possible to export the N-grams individually in order to process it separately from ELAN. The data is exported by clicking the "Raw Data" button in the GUI. After supplying the file the data will be exported in the same CSV format as discussed above. For more in-depth information about the N-Gram analysis function and the resulting data, please consult the PDF mentioned earlier on the ELAN website, or consult it here: https://parasol.tamu.edu/dreu2013/Berke/ images/DREU_Final_Report.pdf (This link may become outdated at some point).

Create multiple media clips

Apart from exporting part of a clip in a project (see the section called "Media clip using script"), you can also export multiple clips from the same (or multiple) projects. The clips will be clipped based on the annotation-times on a specified tier.

As with the exporting of a part of a clip, Windows users will need to put a copy of ffmpeg.exe or ffmbc.exe in the program folder of ELAN. (see the section called "Media clip using script") for more info.

To utilize this function, you will need to create a tab-delimited text file first. Go to File > Export Multiple files as... > Tab-delimited text.... (note, this will not work for the single file tab-delimited export, as there is no option to include the videofile-path). Choose a domain, or create a new one (if you want to create multiple clips from only one .eaf, create that .eaf as a new domain). Select the tiers you want to include in the Tab-delimited text file (each annotation on a tier will result in a clip).

Make sure you check the following options:

- Exclude tier names from output

- Exclude participant names from output

- Include file path column

Under the time column and format options, you will need to check:

- Begin Time

- End Time

- Duration

- ss. msec

The other options have to remain unchecked. Next, click OK and the file will be exported to a text file.

Go to File > Multiple File Processing... > Create Multiple Media Clips.... Choose the exported tab delimited text file you just created, and specify a folder to save the clipped videos to. Click OK to start the process. When done, a process report dialog will appear with information about the clipping process.

Merge tiers

Similar to merging tiers within a single project (see the section called "Merging tiers"), this function allows you to merge tiers in multiple projects. This means the merged tiers will be added to each project you select in the process.

To start, click File > Multiple File Processing... > Merge Tiers.... You will be presented with a dialog in which you either select the eaf files from the file browser, or select files from a domain. When done, select the tiers to use for the merging process.

Figure 1.160. Merge Tiers Multiple Files 1

Next, select the merge criteria, either regardless of the annotation values or according to specified constraints within the annotations of a chosen tier. When checking the option Only process overlapping annotations , ELAN only merges annotations that have the same value. In this case, the values of both annotations are not concatenated, so the created annotation contains the value only once.

Figure 1.161. Merge Tiers Multiple Files 2

In the next step, set a name for the destination tier and decide whether this tier will be a root tier or a child tier. Also select or create a tier type for the new tier.

Figure 1.162. Merge Tiers Multiple Files 3

Lastly, specify the value for the destination tier. You can set a value in a time format, which will put in the specified time values inside the annotation units on the new tier. You can also choose to set a specific value to be filled out into the annotation units. The final choice is to concatenate the values of the annotations from the tiers you have selected for merging.

Figure 1.163. Merge Tiers Multiple Files 4

After clicking Finish, the tiers will be merged and inserted into each eaf you chose at the start. A process report will show an overview of what has been done.

Update transcriptions after changes in ECV's

If one or more of the tier types of a transcription are linked to ECV's (External Controlled Vocabularies, see the section called "Using an External CV") and annotations have been created using those ECV's, it might be necessary to update existing transcriptions after changes have been made to those ECV's. This function allows to update an entire corpus or a selection of files.

If you choose the option Update Transcriptions for ECVs, you'll see the following dialog:

Figure 1.164. Update transcriptions based on ECV's

First a source folder containing the transcription files should be selected, specifying whether or not sub-folders should be processed too. For the destination it is possible to choose to overwrite the existing source files (this should only be done if there are recent back up copies of the files) or to select a folder where the updated transcriptions should be stored. Furthermore it is possible to specify which content language to use; this is only useful if the ECV's are multilingual.

The default behavior of this update action is to change annotation values after changes in an external controlled vocabulary, based on the reference of the annotation to (the id of) an entry in the controlled vocabulary. The Don't change the annotation value... option allows to change this default behavior; instead of updating an annotation value based on a reference to a CV entry, it checks if the annotation value is still in the controlled vocabulary and, if so, updates the reference or, otherwise, removes the reference.

Update Transcriptions with elements from a Template

When the files of a corpus are based on a template, the need can arise during the project, to update all files with new tiers, tier types and/or new controlled vocabularies. To achieve this, the template can be updated with the new elements and the changes can then be applied to the files. It is also possible to create a new template with only the new tiers and types etc. and use that one for the updating, because new elements in the template will be added but existing tiers, types and CV's etc. in the .eaf files *will not* be deleted if they are not (or no longer) in the template.

To start the process of updating a set of files, choose Update Transcriptions with Template... A message will be shown, warning that there is no **undo** for the changes that are going to be made to the files. Then this dialog will be shown:

Figure 1.165. Update transcriptions with a template

The following options are available:

- The Load Domain... button allows to select the files to update.

- The Select Template... button allows to browse to the template file containing the new elements.

- The Dry-run checkbox ensures, if selected, that the selected files will not be changed. After clicking the Start button the files will be compared with the template file and a report will be shown, listing the detected differences.

- The Replace CV's with the same name checkbox determines that existing CV's will be replaced by the CV with the same name in the template. If this checkbox is not ticked, new entries in the CV with the same name in the template will be added to the CV in the file (while new entries in the CV in the file will not be removed if they are not in the CV in the template).

- The Start button starts the update process. After the last file has been processed a report will be shown with a listing per file of the differences and of the changes applied. The report can be saved as a `.txt` file

Apart from adding new elements to the existing files, this process also allows to update some properties of existing elements, as long as these changes can't result in data loss. E.g. the **Annotator** and **Participant** properties of a **Tier** can be updated, but not the **Parent Tier**. The **Tier Type** of a **Tier** can only be changed to a **Type** with the same overall constraints (**Stereotype**). The **Controlled Vocabulary** property of a **Tier Type** can be changed, but not its **Stereotype**. **Controlled Vocabularies** can be converted from *internal* to *external* or the other way round. Etc.

Multiple file export options

ELAN offers the possibility to export multiple annotation files as one file. To do so click on File > Export Multiple Files As... and one of the options.

- Toolbox File(UTF-8) (see the section called "Toolbox file(UTF-8)")

- FLEx File... (see the section called "FLEx file")

- Praat TextGrid... (see the section called " Praat TextGrid")

- Tab-delimited Text... (see the section called "Tab-delimited Text")

- List of Annotations... (see the section called "List of annotations")

- List of Words... (see the section called "List of Words")

- Selected Tiers as EAF... (see the section called "Export tiers as EAF")

- Annotation Overlaps Information... (see the section called "Annotation Overlaps Information")

- Theme Data Files... (see the section called "Theme Data Files")

- Annotated Part Overview... (see)

Toolbox file(UTF-8)

To export multiple files as toolbox files, click on File > Export Multiple Files As... > Toolbox file(UTF-8).... This process involves 3 steps.

1. **Step 1/3: File and Tier Selection**

Figure 1.166. Export as Toolbox file step 1

a. First you have to select the files that are to be exported. You can select multiple files you can choose any one of the below options

- Select files from file browser : This will option a multiple file selection dialog which allows you to select multiple files and you can also choose a directory to export all the files in the directory.

- Select files from domain : the section called "How to select multiple files"

b. Next select the tiers which are to be used for the export process. Using the arrow buttons, you can sort the order of the tiers.

c. Insert blank line after this marker (see the section called "Toolbox file(UTF-8)")

From the drop down list select the tiers to use in the overlaps computation. You can select all the tiers displayed in the list if you click on Select All, or deselect them if you click on Select None.

Once you have made your choice for the tiers for which the overlaps should be found, you can select next, this will bring you to the next step.

2. **Step 2/3: Export Settings**

Figure 1.167. Export as Toolbox file step 2

In this step you can define output settings and the Toolbox options. The option are more clearly defined in the section called "Toolbox file(UTF-8)"

3. **Step 3/3: Save as Settings**

Figure 1.168. Save as Settings

FLEx file

To export multiple files as FLEx files, click on File > Export Multiple Files As... > FLEx File.... This process involves 4 steps.

1. **Step 1/4: File selection and element mapping**

Figure 1.169. Export as FLEx file step 1

a. First you have to select the files that are to be exported. You can select multiple files, choose any one of the below options

   • Select files from file browser : This will option a multiple file selection dialog which allows you to select multiple files and you can also choose a directory to export all the files in the directory.

   • Select files from domain : the section called "How to select multiple files"

b. Next select if you want to export the interlinear-text and paragraph tier. You can set the correct tier type to use as element type and paragraph type in de dialog below that.

2. **Step 2/4: Export Settings**

Figure 1.170. Export as FLEx file step 2

In this step you can select a tier type to use for the 'morph-type' tiers. It's also possible to uncheck this, if not needed. From the dialog, you can also map the tier types to the different items, which are listed on top.

3. **Step 3/4: Element-item 'type' and language attribute**

Figure 1.171. Element-item type and language attribute.

In the next dialog, you can specify the element-item tier type and set a language for it. ELAN can try to extract that from a tier name, (if the box is checked) but it is also possible to add (or remove) a value for a language or type. To do so, enter a value ('en' in this example) and click Add. Then, you can select the added value from the drop-down menu under 'language'. You need to set a type and language for every Tier Type Name in order to be able to go to the final step. For more information on the structure of FLEx, see Figure 1.69, "FLEx to ELAN structure".

4. **Step 4/4: Save as settings**

Figure 1.172. Save as Settings

Praat TextGrid

To export multiple files as Praat TextGrid, click on File > Export Multiple Files As... > Praat TextGrid .... This process involves 2 steps. See the section called "Praat TextGrid file" for more details.

1. **Step 1/2: File and Tier Selection** (see 1)

2. **Step 2/2: Export Settings** (see xref3

Tab-delimited Text

To export multiple files as Tab-delimited Text, click File > Export Multiple Files As... > Tab-delimited Text....

1. A multiple file selection dialog appears(see the section called "How to select multiple files" . Select or create a domain and click on OK to continue.

2. In the next dialog that appears, select tiers and options as you would do when exporting a single Tab-delimited Text file (see the section called "Tab-delimited text file").

3. You can also choose to include a column for the file name and file path to the exported text file. To do so, check or uncheck the appropriate boxes. Instead of adding a column containing the file

name and/or path, you can also choose to put the name and path in a row preceding the annotations of each file. When both file name and path are unchecked, but the row option is checked, the file path will be exported in a row.

Figure 1.173. File name & path options for Multiple Tab-delimited text export.

List of annotations

To export multiple files as List of annotations, click File > Export Multiple Files As... > List of annotations....

1. A multiple file selection dialog appears(see the section called "How to select multiple files" . Select or create a domain and click on OK to continue.

2. Then in the next dialog that appears, select the tiers (see the section called "How to select tiers") from which the annotations are to be exported. Note that the annotations are not separated into words. Check Count occurrences if you want the list to include the number of occurrences for each annotation.

List of Words

To export multiple files as Tab-delimited Text, click File > Export Multiple Files As... > List of Words....

1. A multiple file selection dialog appears(see the section called "How to select multiple files" . Select or create a domain and click on OK to continue.

2. Then in the next dialog that appears, select the tiers and other options as you would do when exporting a single Tab-delimited Text file (see the section called "Alphabetical list of words").

Export tiers as EAF

There could be situations in which you want to discard or select tiers from multiple `.eaf` files, for instance if you want to present a third party with a limited number of tiers. To do so, select File > Export Multiple Files As... >Selected Tiers as EAF.... In the first dialog(see the section called "How to select multiple files") you can select the files from which you want to export a selection of tiers.

Once you have selected your files, **Export Tiers from Multiple Files** dialog appears.

Figure 1.174. Exporting by selecting tiers

To export, do the following:

1. Select tiers (see the section called "How to select tiers") for the export.

2. In the Output Options section you can specify...

   a. whether to `Export parent tiers of the selected dependent tiers automatically` or to `Only export dependent tiers if their parent tiers are selected`.

   b. whether to `Save files with original names` of to `Make use of suffixes`. In case of the latter, you can specify whether to save the files with their original name followed by a suffix or to save the files with a new base name and followed by a suffix number.

   c. whether the files should be saved in the original directory, in a (possibly new) directory which is local for each files, or together in the same directory.

   d. whether of not ELAN should export files that result in having no tiers.

3. Finally, click Export to export the .eaf files containing only the selected tiers.

Annotation Overlaps Information

This function is available via menu File->Export Multiple Files As-> Annotation Overlaps Information....

This function allows the user to select one "reference" tier and multiple other tiers that will be compared (sequentially) with the reference tier. The comparison is done on the level of the annotations.

The following information will be present in the resulting tab-delimited text file:

1. **The Header line with column names**

Column 1- 4:

Begin time	End time	Duration	Reference Tier Name

These columns will contain information for all annotations of the reference tier. The annotation values are in the column with the tier name as the header, the time info in the first 3 columns.

Next for each "comparing" tier there will be 11 columns, the header of which consists of the tier name and a suffix and the column contains the following information:

1	Name-ov	0 or 1, whether there is an overlapping annotation or not (0=no, 1=yes)
2	Name-same	0 or 1, whether the overlapping annotation has the same value. If there are more than one overlapping annotations the value is 0 (0=no, 1=yes)
3	Name-ov-dur	The duration of the overlap, the total overlap duration in case of more than one overlapping annotation
4	Name-no-ann	The number of overlapping annotations
5	Name-value	The value of the overlapping annotation, concatenated, comma separated, in case of multiple overlaps
6	Name-bt-to-bt-After	The amount of time from the beginning of the reference annotation to the beginning of the first non
7	Name-et-to-bt-After	The amount of time from the end of the reference annotation to the beginning of the first non overlapping annotation
8	Name-et-to-et-After	The amount of time from the end of the reference annotation to the end of the first non overlapping annotation
9	Name-bt-After	The begin time of the annotation on the comparing

		tier after the reference annotation
10	Name-et-After	The end time of the annotation on the comparing tier after the reference annotation
11	Name-value-After	The value of the first annotation on the comparing tier after the reference annotation

2. **The content per file**

After the header, for each file there will be the following information/data:

- one line containing the file name

- a number of rows equal to the number of annotations of the reference tier

- each cell filled with the information corresponding to the header description (above)

All time values are in milliseconds.

Theme Data Files

Theme© (www.patternvision.com [http://www.patternvision.com]) is an application for detection and analysis of hidden patterns in time (so called t-patterns) in behavioral data. It is possible to export annotations of selected tiers such that they can be imported into a Theme project. A Theme project requires at least two text files:

- a Category or Variable-Value-Table File (a file named vvt.vvt) containing a listing of the coding Classes and the Items of each class. This is similar to the idea of Controlled Vocabularies and their Entries in ELAN.

- One or more Raw Data Files, text files in which each line represents a record of a time-stamped event. Each line contains a time value, a label for the Actor, a 'b' or 'e' to indicate whether it is the begin or end of the event and the "code" or value of the event.

When exporting, ELAN creates the "vvt.vvt" file and for each transcription file it creates a raw data file, converting each annotation into two records for the raw data, one for the begin time and one for the end time of the annotation.

To start, click File->Export Multiple Files As-> Theme Data Files....

1. A file and tier selection dialog appears. Select the appropriate files or a domain, then choose the tiers you would like to use for the export.

   📑 **Note**

   If a file from the selected eaf files is currently opened, you will be presented with a warning. Make sure you have saved your current transcriptions before starting this export process as local changes will be overwritten.

2. In the next dialog that appears, set the various export options. There are two specific configuration options for the export:

   - When tiers are connected to a Controlled Vocabulary there is an option to include the entire CV in the .vvt file, otherwise ELAN will just add all values that are actually present in the annotations.

   - For the Actor it is possible to use either the tier name or the participant label of a tier (if it is there).

3. When done, click Finish to start the export. Afterwards, you can find the exported vtt.vtt file and the text file(s) in the directory you specified in step 2.

Annotated Part Overview (preliminary)

This export function is available via File->Export Multiple Files As-> Annotated Part Overview.... It creates a tabular text file containing information on how much of a recording has been annotated on at least one of the selected tiers. The export window allows to select a domain or a custom set of files and lists the tiers found in the files. After selection of the tiers to include in the processing, clicking the Next or Finish button starts the processing and shows a Save As dialog. The output can be opened a spreadsheet application, it contains a row for each file and the following columns (summarized):

- **File name**

- **Total duration of non-empty annotations**, annotations with textual contents

- **Total duration of empty annotations**, to be able to distinguish segmenting-and-labeling from segmenting alone

- **Total annotated part**, the total duration of the 'virtually' merged segments represented by these (empty and non-empty) annotations

- **Last annotation time**, the end time of the farmost annotation (on any tier), instead of the duration of the recording, which is currently not stored in the `.eaf` file and therefore not to available to this multiple file function

- **Total as a percentage of last time**, the annotated part as percentage of the last annotation time

- **Number of annotations**, the total of empty and non-empty annotations on the selected tiers

- **Number of empty annotations**, on the selected tiers

- **Number of processed tiers**, the number of tiers actually processed in this file

- **Number of empty tiers**, the number of tiers present in the file but without annotation

- **Number of tiers not processed**, either not present in the file or not selected

Finally there is a row with the totals of all rows, giving an overview of all the files in the domain.

Multiple file import options

ELAN offers the possibility to import multiple files at once and save them as *.eaf files. To do so click on File > Import Multiple Files As... and one of the options.

- Toolbox File(UTF-8) (see the section called " Toolbox file (UTF-8)")

- Praat TextGrid... (see the section called " Praat TextGrid")

- FLEx File... (see the section called "Flex files")

Toolbox file (UTF-8)

To import multiple toolbox files for conversion to *.eaf, click File > Import Multiple Files As... > Toolbox file(UTF-8).... This operation consists of 3 steps.

Figure 1.175. File selection

1. Select the files you would like to import by clicking browse and adding the files to the list.

2. Next you will have to select what settings to import. Either select a *.typ file or use the 'Set field markers' option. See the section called "Shoebox file" for working with *.typ files and field markers.

3. Lastly, you will have to configure the Save as settings.

Figure 1.176. Save as Settings

When the operation has completed, you will be presented with a process report. The multiple *.eaf files are now ready to be used in ELAN.

Praat TextGrid

Multiple TextGrid files created in Praat can be imported and converted to *.eaf files. This process involves 3 steps.

1. Choose the `*.TextGrid` files that will be imported for conversion to `*.eaf`. You can also set the encoding (default, UTF 8, UT 16).

2. In the next dialog, you can define the settings to be used for importing:

Figure 1.177. Import Settings

In this dialog, you can choose to include Praat PointTiers, if empty annotations or intervals should be skipped or not and whether or not for each converted `.TextGrid` file a corresponding `.wav` file should be added to the linked files.

3. Lastly, you will be asked how the files should be saved and in what location.

Figure 1.178. Save as Settings

When the operation has completed, you will be presented with a process report. The multiple *.eaf files are now ready to be used in ELAN.

Flex files

To import multiple FLExtext files for conversion to *.eaf, click File > Import Multiple Files As... > FLEx file.... This operation consists of 3 steps.

1. Select the FLEx files you want to use for conversion to *.eaf. Do so by clicking the Browse... button in the dialog and choose the proper files.

2. In the next dialog, you can define the settings to be used for importing:

Figure 1.179. Import Settings

You can select whether to use the 'interlinear-text and 'paragraph' element in FLEx, import the participant information and what the smallest time-alignable element should be: 'phrase' or 'word'. Choose on what level you want to create tier types and set a duration per phrase element (required).

3. Finally, you can choose how to save the files and in what directory to save them:

Figure 1.180. Save as Settings

Finally, configure how and where to save your files. You can choose to save with an .XML or .flextext extension, and you can skip files that would result in having no tiers.

Interact with an Annotation Server (experimental)

It is now possible to upload annotations to a Web Annotation server that is compliant with the W3C Web Annotation data model and protocol (currently supporting Elucidate Server [https://github.com/dlcs/elucidate-server] and AnnoRepo [https://github.com/knaw-huc/annorepo]). This can become another way of sharing your annotations with other researchers using ELAN or any other tool supporting the web annotation format and protocol. Once analysis and presentation tools become available for collections in a Web Annotation server, this features provides alternative ways of exploring your data. The format of this export is similar to the file based export to the section called "WebAnnotation JSON"

Exporting to Annotation Server

To export annotations to the Annotation Server go to File > Annotation Server > Export Annotations to Server. The following window appears:

Figure 1.181. Export to Annotation Server

1. Select the tiers you wish to export to the server

2. Click the Export to Server button to start the export.

3. At the end of the process a message window will show how many tiers and annotations have been exported.

Figure 1.182. Export results

📒 Note

Authentication may be required in order to upload annotations to a server. In that case an additional field will allow to paste a so-called API-key, which has to be obtained from the server administrator.

Listing the Collection ID's from the server

Tiers that are exported to the Annotation Server are stored as `AnnotationCollection`'s and are given a collection ID by the server. ELAN stores the mapping of tier name to collection id in a local database, per annotation document. With an ELAN document open, it is possible to see a listing of previously exported tiers by choosing File > Annotation Server > List Collection IDs from Server. A window appears showing the mapping:

Figure 1.183. List of Collection ID's

Each tier can occur multiple times in the table, if it has been exported more than once. The collection ID and the creation date and time should be different in that case. The collection ID is a unique identifier which eventually will be one of the properties that can be used to retrieve the web annotations from the Annotation Server.

Importing collections from an Annotation Server

If you have the collection ID of an annotation collection (i.e. a tier) stored in an Annotation Server, it is possiblt to import it into the current or a new document. Choosing File > Annotation Server > Import Collection from Server shows the following window:

Figure 1.184. Import a collection with its ID

A collection ID (which is a full URL), can be pasted or entered in the text field and after clicking OK the tier with its annotations will be imported and added to the open or a new document.

📋 Note

Authentication may be required in order to import annotations from a server. In that case an additional field will allow to paste a so-called API-key, which has to be obtained from the server administrator.

Chapter 2. Annotations

You can use the ELAN program for annotating your data. This annotation process involves three steps: defining tier types and tiers (see the section called "Adding new tier types" and the section called "How to define a tier and its attributes"), selecting time intervals (see the section called "How to make a selection"), and entering annotations (see the section called "How to enter and edit annotations").

Basic Information: Annotations, tiers and tier types

The following illustration shows an example of an annotation document:
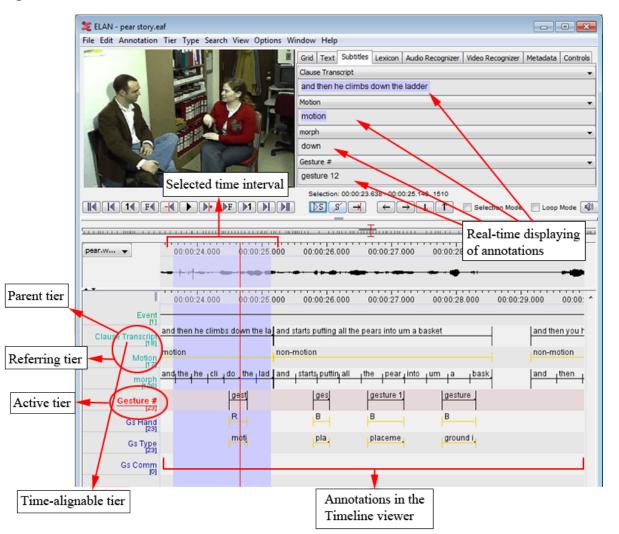
Figure 2.1. Annotation basics

Each annotation is entered on a tier and assigned to a time interval (either directly or to the time interval of another annotation).

All tiers can be displayed simultaneously in the Timeline or Interlinear Viewer (the section called "The Timeline Viewer and the Interlinear Viewer"), but four of them can be displayed additionally in the Subtitle Viewer. It is useful to select the tier you are currently working on in a Subtitle Viewer because this viewer is bigger and supports line wrapping (which makes it easier to read along during playback).

It is also possible to select one tier as the *active tier*. This can be done by double clicking on the tier name in the Timeline or Interlinear Viewer. When a tier is active, its name is underlined and displayed in red. Adding a new annotation to a tier by the keyboard shortcut **ALT+N** is only possible when that tier is active (see the section called "How to enter and edit annotations").

A tier is a set of annotations that share the same characteristics, e.g., one tier containing the orthographic transcription of the speakers utterances, and another tier containing the free translation of these utterances.

The following two types of tiers exist:

• Independent tiers, which contain annotations that are linked directly to a time interval, i.e., they are "time-alignable".

• Referring tiers, which contain annotations that are linked to annotations on another tier (i.e., to annotations on their so-called "parent tier"). They are usually not linked directly to the time axis. (Some of them may be linked – but only within the time interval determined by their independent parent tier, see below.)

One example: a transcription tier could be independent and time-alignable, as it is linked directly to the time intervals of the speakers utterances. A translation tier, by contrast, would be referring and not time-alignable: it refers to the transcription tier – not directly to the time axis. By definition, it inherits its time alignment from the transcription tier, i.e., from its parent tier.
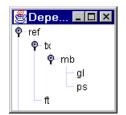
In the Timeline and Interlinear Viewers, the label of a referring tier is assigned the same color as the label of its independent parent tier.

It is possible to build up nested hierarchies, i.e., tier A can be the parent tier of tier B, and tier B can be the parent tier of tier C, etc.

For example:

Table 2.1. Nested tier dependencies

tier:	type:	hierarchical relation:
ref (referent)	independent	parent of tx and ft
tx (text)	referring	parent of mb
mb (morpheme break)	referring	parent of gl and ps
gl (gloss), ps (part of speech)	referring	-

Figure 2.2. Tier dependencies

Figure 2.3. Tier dependencies in the timeline viewer

📋 Note

> Parent and child tiers are linked in such a way that some changes made on a parent tier will also affect its child tiers (but not vice versa):

• If you delete a parent tier, all its child tiers are automatically deleted as well. Similarly, when you delete an annotation on a parent tier, all corresponding annotations on its child tiers are deleted as well.

• If you change the time interval of an annotation on a parent tier, the time interval of the corresponding annotation on all its child tiers are changed accordingly. The time interval of a child tier cannot be changed independently.

You can view the existing dependency relations by clicking on View menu, and then on Tier Dependencies.

Each tier is assigned to a tier type (see also the section called "How to define a tier and its attributes"). A tier type denotes the linguistic data that is contained in the referring tier. Examples of names for tier types are *utterances*, *words*, *orthography*, *phonetic transcription*, *PoS* (part of speech), but any name can be used. Each tier type specifies a number of constraints that hold for all tiers assigned to that type. Such constraints are bundled into so-called 'stereotypes'. The following five stereotypes are currently available:

Table 2.2. Tier type stereotypes
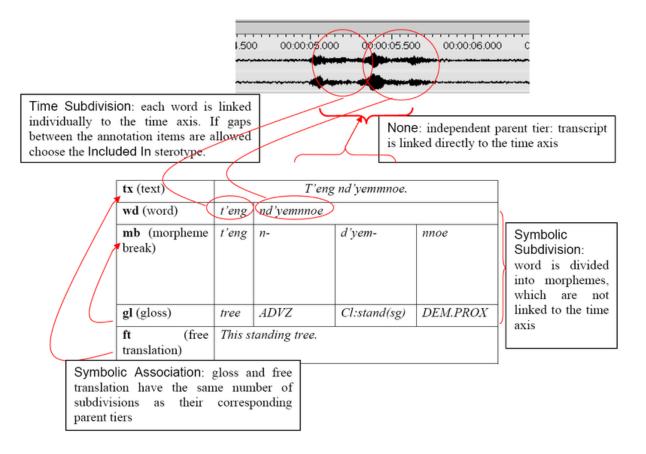
• None	The annotation on the tier is linked directly to the time axis, i.e., the annotation is entered on an independent tier. Two annotations cannot overlap.
	Tiers of this type are independent tiers (cannot have a parent tier).
• Time Subdivision	The annotation on the parent tier can be sub-divided into smaller units, which, in turn, can be linked to time intervals. Note that there are no time gaps allowed, i.e., the smaller units have to immediately follow each other.
	Tiers of this type are dependent tiers (must have a parent tier).
	E.g., an utterance transcribed on a parent tier can be sub-divided into words – each of which is then linked to its corresponding time interval.

	📝 **Note**
	Annotations on such tiers are time-alignable. They differ from annotations on independent tiers in that they are assigned to an interval that is contained within the interval of their parent annotation.
• Symbolic Subdivision	Similar to Time Subdivision, except that the smaller units cannot be linked to a time interval. The smaller units form a chain (an ordered sequence) of units. Tiers of this type are dependent tiers (must have a parent tier). E.g., a word on a parent tier can be sub-divided into individual morphemes (which are not linked to a time interval).
• Included In	All annotations on a tier of this type are linked to the time axis and are enclosed within the boundaries of an annotation on the parent tier. However, there can be gaps between the child annotations (unlike Time Subdivision). Tiers of this type are dependent tiers (must have a parent tier). E.g., a sentence with a silence can be split up into words while the silence corresponds to a gap in the child annotations (i.e. the separate words).
• Symbolic Association	The annotation on the parent tier cannot be sub-divided further, i.e., there is a one-to-one correspondence between the parent annotation and its referring annotation on this tier. Tiers of this type are dependent tiers (must have a parent tier). E.g., one sentence on a parent tier has exactly one free translation. Or one word has exactly one gloss.

The following example illustrates (four of) the different stereotypes (see also Figure 2.3, "Tier dependencies in the timeline viewer"):

Figure 2.4. Examples of (four of) the stereotypes
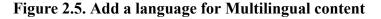
You can define an unlimited number of tiers. It is useful to make decisions about the type of information that you want to enter (and consequently about the type of tiers that you need) at a relatively early stage in the annotation process. However, it is always possible at a later stage to change the parent of a dependent tier (see the section called "Changing the parent of a tier") or to copy a tier (the section called "Copying tiers") and to alter the copy.

Defining languages for your content

ELAN supports the option to specify the content language of annotations and Controlled Vocabularies. The default list of language identifiers provided in the user interface, is taken from the ISO 639-3 standard. A tier can have one language reference, whereas a CV can incorporate multiple languages. For more info how to setu p a tier, please see the section called "Adding new tiers". How to setup a CV with one or more languages, please see the section called "Setting language(s) for a CV".

A screencast is available on Vimeo about setting up multilingual content: `https://vimeo.com/126786931`

To edit the list of languages you can choose from for a CV or a tier, click Edit from the main menu, and select Edit list of languages.... In the dialog that appears, you will see the current list of languages in the upper pull-down menu. To add a language, choose one from the available languages in the pull-down menu on the bottom of the dialog, and click Add.

Figure 2.5. Add a language for Multilingual content

To change a language, select the language you want to change from the List of languages and choose the new language from the All available languages list. Next, click Change to change the language.

To delete a language from the List, select it in the list of languages and click Delete.

In addition to adding a language by selecting it from the All languages list, it is possible to define custom language identifiers by typing in values in the three editable textboxes of the All languages pull-down menu. E.g. when ISO 639-1 two letter codes are preferred or required (these are currently not provided as a list) or if subtags for region, variant or script etc. need to be added.

When done, click Close to close the dialog.

How to define a tier type

ELAN supports the following options:

• adding a new tier type (the section called "Adding new tier types");

• changing the attributes of a tier type (the section called "Changing tier types");

• deleting a tier type (the section called "Deleting tier types");

• importing a tier type (the section called "Importing tier types").

Adding new tier types

Every tier is assigned to a tier type. The tier type specifies the stereotypical constraints (and, as a consequence, whether or not the tiers are time-alignable).

Information about the relationship between tiers is given in two different places: for each individual tier it is given in the Add tier dialog window (see the section called "Adding new tiers"), and for all tiers belonging to one tier type it is given in the Add type window (this section), i.e.:

• Add tier attributes window: specify the parent tier of the individual tier.

- Add type window: specify the stereotypical constraints of tiers belonging to one type.

To create a tier type, do the following:

1. Click on the Type > Add new tier type... menu. The Add Type dialog window appears:

Figure 2.6. Add a tier type

2. Enter the following information for the new type:

   a. Go to Type name. Enter the name or label for the type.

   b. Go to Stereotype. Select the stereotypical constraint that apply to the tiers based on this type (see the section called "Basic Information: Annotations, tiers and tier types").

   c. Optionally go to Use Controlled Vocabulary. Select a controlled vocabulary or leave this to **None** (see the section called "Associating a tier type with Controlled Vocabularies").

   d. Optionally go to Lexicon Connection. Click the + button to select a lexicon and a lexicon entry field to be associated with this tier type (see the section called "Consulting a Lexicon", the section called "Set the Lexicon Service and entry field information for a Tier Type" and the section called "Lexicons").

   e. Optionally go to ISO Data Category to associate this type with a data category via the + button (see the section called "Associating a tier type with a Data Category".

   f. The Time-alignable checkbox cannot be edited, its value depends on and changes with the selected Stereotype.

3. Click Add to save the settings; otherwise click Close to exit the window without saving them.

Changing tier types

Do the following to change the attributes of an already existing tier type:

1. Click on Edit > Change tier type… The Change type dialog window appears.

2. The labels of all available tier types are displayed in the Current Types table, e.g.:

Figure 2.7. Change a tier type

3. Click on the tier type whose attributes you want to change.

4. Change the settings. The Stereotype drop down box will be disabled if there is at least one tier based on the selected type. For changing the *Stereotype* of a tier consider the options to copy (see the section called "Copying tiers") or to reparent a tier (see the section called "Changing the parent of a tier").

   Associations with a Lexicon field or a ISO Data Category can be changed via their + button or removed by clicking their - button.

5. Click Change to save the changes. Otherwise click Close to exit the window without saving the changes.

Deleting tier types

To delete a tier type, do the following:

Click on Edit menu.

1. Go to Delete tier type…

2. A dialog window appears. The names of all available tier types are displayed in the Current Types table, e.g.:

Figure 2.8. Delete a tier type

3. Click on the type you want to delete.

4. Click Delete to delete the type; otherwise click Close to close the dialog window.

You can only delete a tier type if it is not used by any of the tiers. If it is used, the following error message appears:

Figure 2.9. Tier type is used

Importing tier types

It is possible to import tier types from `.eaf` or `.etf` files.

Click on Edit > Import types.... Here you can select a .eaf or .etf file by clicking on Browse. Select the file you want to use in the next window and click Select. Finally click Import to import the tier types from the selected file.

Associating a tier type with a Data Category

📄 **Note**

> This functionality is currently out of order, please see the note at the section called "Data categories"

In ELAN it is possible to associate a tier type with a data category of the ISO Data Category Repository (DCR). To do so click on Type in the ELAN main menu and select Change Tier Type.... In the window that now appears, click the Browse... button to the right of the `ISO Data Category` option. In the next window(Figure 2.10, "Local Data Category Selection"), select a category from the list and click Apply to associate the selected data category to the tier type.

Figure 2.10. Local Data Category Selection

More information about the ISO DCR and how to use it can be found in the section called "Data categories".

How to define a tier and its attributes

ELAN offers the following options:

• adding a new tier (the section called "Adding new tiers");

• changing the tier attributes (the section called "Changing tier attributes");

• deleting a tier (the section called "Deleting tiers").

Adding new tiers

Do the following to add a new tier:

1. Click on Tier > Add New Tier.... The Add Tier dialog window appears.

2. Define the tier attributes:

Figure 2.11. Tier attributes

Enter the following information:

1. Go to Tier name. Enter the tier name.

   The tier name is the name that is displayed in the Timeline, Interlinear and Subtitle Viewer.

2. Go to Participant. Enter the name of the participant whose utterance or other behavior is being transcribed (optional).

3. Go to Annotator. Enter the name or code of the creator of the annotations on this tier (optional).

4. Go to Parent tier. Select a parent tier from the pull-down menu.

   • If the tier has no parent tier, select **none**.

   • Otherwise select the appropriate parent tier.

5. Go to Tier type. Select a tier type from the predefined list in the pull-down menu (see the section called "How to define a tier type").

> 📝 **Note**
>
> The list of possible tier types is dependent on the parent tier that is chosen. E.g., if there is no parent tier ("none" in the pull down menu), the tier types to choose from are of the stereotype "none" (see the section called "How to define a tier type").

6. Go to Input Method. Select what language to use as input method from the pull-down menu or select None. If None is selected, it means that the system's default language will always be used as input method for this tier.

   Whenever you enter or change annotations on that tier, the text entry box is automatically preconfigured for the default character set.

7. Set the Content Language. This refers to the annotated language in the tier. This can be different for each tier, useful if you have multi-lingual content. Also see the section called "Defining languages for your content").

8. Click the More Options... button if you want to change the color of the tier name and the font of the annotations. In the new dialog window you can change them by clicking Browse..., selecting a color or font and clicking OK. The color chooser has four tabs. The last three contain different ways to choose a color, which is subsequently displayed in the lower part of the window. In the first tab you can add or insert the color displayed below and you can copy, paste and delete the selected color. The list of favourite colors is saved and used the next time you start ELAN. To apply the new color and font click Apply. The following dialog window will then appear:

Figure 2.12. Change Tier Attributes Settings for Multiple Tiers

In the upper part of the window you can select the attribute settings you wish to apply, i.e. Tier, Color, Tier, Highlight and Tier Font. In the bottom part of the window you can decide to change the preferred attribute settings for multiple tiers in one action, i.e. by selecting all tiers of the same type, or all depending tiers, or all tiers with the same participant. Finally click Add to save the tier and its attributes. Otherwise click Cancel to exit the window without saving.

Importing tiers

You can reuse the tiers (together with their tier types) you created before in other `.eaf` or `.etf` (templates, see the section called "Saving a template") files:

1. Select Tier > Add new tier…

2. Click on the Import tab

3. Click on Browse…

4. Select the eaf or etf file from which you want to import a tier and confirm your choice by clicking on Select

5. Finally click on Import.

Figure 2.13. Import tiers

> **📒 Note**
>
> If you import a tier type that already exists, a postfix like –cp1 or –cp2 will be added to the imported version.

Creating a tier using the Silence Recognizer

ELAN offers the possibility to create tiers based on the output of an audio recognizer. A simple silence recognizer that comes with ELAN will show how to use an audio recognizer in ELAN. The recognizer

can determine segments of silence in an audio signal based on examples given by the user. To use it for segmenting the audio first open the tab Recognizers in the main ELAN window. From there, you can load or save a (previous) parameter setup, detach the tab from the main screen and configure if a recognizer should prompt for file in- and output locations (if applicable). Some recognizers include a separate help file, which you also can consult.

Figure 2.14. Silence Recognizer

Next, choose the appropriate recognizer from the pull down menu at the top of the tab, in this case Silence Recognizer MPI-PL. In the parameters section you can choose the appropriate waveform the recognizer should use. The Selections Panel allows you either create selections from silent parts of the waveform or it can analyse a specified tier, if the chosen recognizer allows it.
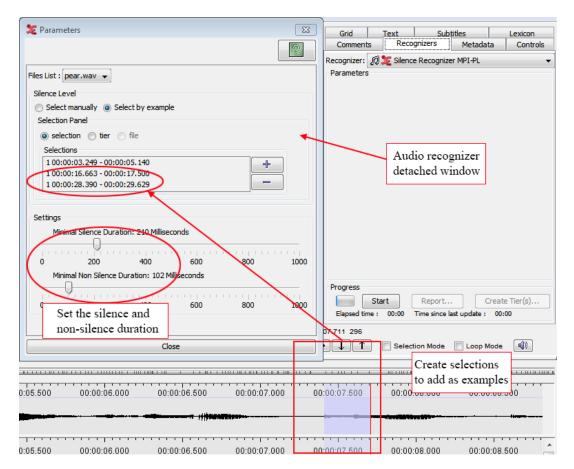
Figure 2.15. Audio Recognizer detached

The Silence Recognizer uses examples to determine what is silence and what is not. To give the recognizer an example first select a part of the audio that is silence (see also the section called "Making a selection"). Then click on + in the Selection Panel. The begin and end times of the example are shown in the list beneath the Add Channel buttons. By selecting a line in the list and clicking - an example can be removed. By double clicking a line in the list the associated time interval is selected in ELAN.

After giving sufficient examples, click on the Start button to start the recognition. During the recognition you can click Cancel to stop the recognition.

The result of the recognition is a segmentation in the Waveform Viewer for each channel for which an example is given. In the case of the silence recognizer the segments are either labelled 's' for the beginning of a silent segment or 'x' for the beginning of a non-silent segment. If you are not satisfied with the segmentation, you can change the examples or the duration parameters and start a new recognition.

🗒 Note

The second and subsequent runs of the audio recognizer can be several times faster than the first run. This is caused by the buffering the audio recognizer applies.
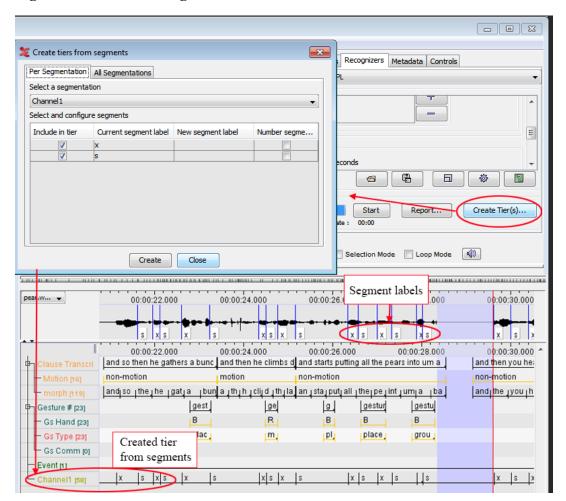
Figure 2.16. Audio Recognizer results

If the labelling is correct, you can create a tier with annotations reflecting the labelling in the Waveform Viewer. Click on Create Tier(s)... in the Recognizers tab. On the tab Per Segmentation of the dialog window select the channel that has the segmentation you want to use from the pull down menu. In the table Select and configure segments first select the labels that must be included in the tier. If necessary, change the label by clicking in the third column of a label and enter a new label. Check the Number segments column if you want to number each annotation with a particular label. The number will be appended to the label. Finally, click the Create button to create the tier.

If all segment labels are to be used, open the All Segmentations tab instead of the Per Segmentation tab. On the All Segmentations tab you are only asked to select the channels for which a tier must be created. Again, clicking the Create button will make ELAN create the tier.

Each recognizer will have its specific controls. These controls can be found in the parameters section of the Recognizers tab. In the case of the silence recognizer there are two sliders: Minimal Silence Duration and Minimal Non Silence Duration. When using another recognizer, these sliders are replaced by the controls implemented by that recognizer.

To learn more about creating and adding other recognizers, some information can be found in this specification document (Avatech-interface-spec-2014-03-06.pdf [https://www.mpi.nl/tools/elan/docs/Avatech-interface-spec-2014-03-06.pdf]) and the recognizer API example set [https://www.mpi.nl/tools/elan/recognizer_api_V4.1.zip] .

Changing tier attributes

Do the following to change the attributes of an already existing tier:

1. Click on Tier menu.

2. Go to Change tier attributes…

3. Click on the drop down box and select the tier which attributes you want to change. The Change tier attributes dialog window for that tier appears.

4. After making the changes, click on Change to save them. Otherwise click Cancel to exit the window without saving.

Figure 2.17. Tier menu

Figure 2.18. Change tier attributes

> 📝 **Note**
>
> Changing the Parent Tier in this dialog is only possible if there are no annotations on the tier (because of possible data loss). To change the Parent Tier in case there are already annotations, use the Reparent Tier option (see the section called "Changing the parent of a tier"). Similar limitations apply to changing the Type of the tier, for the same reason. To change the Type of a tier in a safe way, use the Copy Tier option (see the section called "Copying tiers").

Deleting tiers

Do the following to delete a tier:
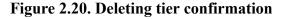
Use the Tier menu.

1. Click on Tier menu.

2. Go to Delete tier… A dialog window appears.

3. The labels of all available tiers are displayed a pull down box, e.g.:

Figure 2.19. Delete tier

4. Click on the tier that you want to delete. To select multiple tiers and delete them in one action, select the tiers you want to delete (either by browsing through them with the mouse, or using the control or shift key), in the pull down box in the delete tier dialog window. A warning dialog appears asking you to confirm the deletion., e.g.:

Figure 2.20. Deleting tier confirmation

5. Click Yes to delete the tier/s and all its child tiers; click No to not delete them.

📄 **Note**

> If you delete a tier, you will lose all its annotations.
>
> If you delete a parent tier, all its child tiers will be automatically deleted as well. Please make sure that you do not accidentally delete a child tier.

To delete a parent tier without deleting its child tiers, you have to assign the child tier to another parent or make it an independent tier. Afterwards you can safely remove the parent tier. For instructions on how to change a tiers parent, see the section called "Changing the parent of a tier".

Merging tiers

ELAN offers the possibility to merge two tiers. There are several reasons why you would want to merge tiers. For instance, one could think of a situation where two people are transcribing different parts of the same media file. If both people have finished, the transcriptions can be merged via File > Merge Transcriptions.... You can also merge tiers across multiple eaf files, using the File...Multiple File Processing... option. See the section called "Merge tiers" for more information about that.

To get the annotations of both tiers onto one tier, use Tier > Merge Tiers.... In the dialog window select the two tiers to merge and click Next. Set the criteria for merging the annotations. When checking

the option Only process overlapping annotations, ELAN only merges annotations that have the same value. In this case, the values of both annotations are not concatenated, so the created annotation contains the value only once.

Enter a name for the new tier and select the desired tier type or add a new type. Next, select `Concatenate the values of overlapping annotations` and click Finish to create the new tier. Now all annotations of the original tiers are on the new tier. Overlapping annotations are merged to one annotation. The merged annotations begins where the first of the overlapping annotations begins and ends where the last one ends. The values of the overlapping annotations are concatenated. Optionally you can specify the value the merging annotations should have.

Merging tiers can also be used to get some time statistics of the combination of two tiers. Again, select Tier > Merge Tiers..., select the two tiers to merge and click Next. After setting the criteria, entering a name for the new tier and selecting the desired tier type, select Value in the following time format and the desired time format. Finally click Finish. Overlapping annotations are merged and the annotation's value is the total duration of the overlapping annotations. (More about annotations statistics can be found in the section called "Annotations Statistics".)

As a final example, consider an audio recognizer (see the section called "Creating a tier using the Silence Recognizer") creating not one but multiple tiers. If you want to put the annotations of those tiers on one tier, you could use the Merge Tiers... option to achieve this.

Merging tiers (Classic)

Another way to get the annotations of both tiers onto one tier, is to use Tier > Merge Tiers...Classic. This method is somewhat simplified.

In the dialog window select the two tiers to merge and click Next. Enter a name for the new tier and select the desired tier type. Now select `Concatenate the values of overlapping annotations` and click Finish to create the new tier. Now all annotations of the original tiers are on the new tier. Overlapping annotations are merged to one annotation. The merged annotations begins where the first of the overlapping annotations begins and ends where the last one ends. The values of the overlapping annotations are concatenated.

When checking the option Only process if the overlapping annotations have the same value, ELAN only merges annotations that have the same value. In this case, the values of both annotations are not concatenated, so the created annotation contains the value only once. Optionally you can specify the value the merging annotations should have.

Merging tiers can also be used to get some time statistics of the combination of two tiers. Again, select Tier > Merge Tiers..., select the two tiers to merge and click Next. After entering a name for the new tier and selecting the desired tier type, select Set the duration over the overlap as the annotation's value and the desired time format. Finally click Finish. Overlapping annotations are merged and the annotation's value is the total duration of the overlapping annotations. (More about annotations statistics can be found in the section called "Annotations Statistics".)

As a final example, consider an audio recognizer (see the section called "Creating a tier using the Silence Recognizer") creating not one but multiple tiers. If you want to put the annotations of those tiers on one tier, you could use the Merge Tiers... option to achieve this.

Merging tier groups

In ELAN it is possible to merge a tier group (an independent tier and all of its dependent tiers) with another tier group and put the result in a new tier group. The names of the tiers in the new group are the names of the tiers in the first group with a suffix entered by the user. The mapping of tiers is based on name patterns. For instance, the tier *W-RGph* is mapped onto *K-RGph*, but not onto *K-Spch*.

To merge two tier groups, click in the main menu on Tier > Merge Tier Groups... and select the two independent tiers to be merged. After clicking Next, enter a suffix that is to be concatenated to the tier names of the first tier group for the naming of the new tier group and click Finish.

Changing the parent of a tier

Previous versions of ELAN offered an option to change the parent of a tier. As this operation could cause data loss when not performed carefully, this function has been disabled. However, what can be done now as an alternative is making a copy of a tier and altering that copy. This prevents the possible loss of data, as the original tier stays the same. At the same time the tier type of the copy can be changed as well. In order to maintain the overview during this potentially complicated operation a wizard guides you through this process, which can be started via the Tier > Change parent of tier… menu.

> 📄 **Note**
>
> Though an undo option is available it still is a good idea to make a backup of your files before proceeding.

Table 2.3. Change parent of tier wizard

As the result of this process the selected tier (and its children) will be copied and they will become dependent upon the newly chosen parent tier. In our example the W-Words tier, previously a child of W-Spch, became an independent tier:

Figure 2.21. Dependent to independent tier

Note that as the tier is not moved but copied the names have been changed: a postfix "-cp" has been added to the copies. The original can be deleted afterwards if you are satisfied with the result of the operation, while the copies can be renamed to reflect the original tier names.

If you decide to assign a tier to a different parent tier, ELAN will automatically align its annotations with that of the new parent tier (based on overlapping time intervals). In this case, if there is an annotation on the referring tier, but no overlapping annotation on the parent tier, ELAN will delete this annotation. Be very careful that you do not lose such annotations accidentally. A referring tier can be turned into an independent time-alignable tier without any problem.

Managing tiers for Participants

Add New Participant

ELAN offers the possibility to create a new set of tiers for a new participant based on an already existing set of tiers. To add a new participant click on Tier > Participant > Add New Participant.... The following *Add new participant* dialog window appears:

Figure 2.22. Add New Participant

If you have not defined a tier structure, or set a first participant yet, this warning will appear:

Figure 2.23. Warning dialog box

This means you need to create a tier structure first, which is a tier with at least one child tier, or create a new tier with a participant set in the tier attributes. See the section called "Adding new tiers" on how to create a tier structure and how to define a tier with the participant attribute set.

The new set of tiers for the extra participant can either be based on an existing tier group with all its depending tiers, or on all tiers with a specific participant attribute.

Figure 2.24. Add New Participant Example 1

In this example, the option "tier structure" is selected (1). This means only one tier structure will be copied and used with a new participant. "W-Spch" is the selected tier structure. The name of the new participant will be "Participant 3" (2).

In this case, the prefix is changed to distinguish the new tier structure, with the value "W" being changed to "X" (3). The value to be replaced can also be left empty. In that case, the replacement value will be added to the name of the structure. By clicking "OK", the process will be started and the new tier structure for the new participant will be added to the timeline viewer (4). When this is done, you can close the dialog box.

It is also possible to add a new participant based on an existing participant. This method will copy **all** of the tier attributes to the new participant:

Figure 2.25. Add New Participant Example 2

In this example, the participant option is checked (1). This means all tiers and tier structures associated with a certain participant will be copied to the new participant. In this case, the participant of whom the attributes will be copied is "Participant 2". The tiers that will be copied are the "K-RGU" structure and the "K-Spch" tier.

Next, specify a name for the new participant (2). The name will be in the tier attributes after the adding of the participant has been done. Finally, you will need to add or change the prefix or the suffix for the new tiers (3). The value to be replaced can be left empty; the value for replacement cannot. In the example, the prefix "K" is changed to "X". When everything is set, click "OK", the new tiers will appear in the timeline (4). After that, you can close the dialog box.

Hovering over the tier in the timeline window will show the tooltip, displaying tier info with the associated participant:

Figure 2.26. Add New Participant tooltip result

> ### 📄 Note
>
> - The new prefix/suffix value for replacement cannot be empty.
>
> - The value to be replaced can be left empty. If you do not enter any value to be replaced, the new value for the replacement will be added either as suffix or prefix (depending on your choice) to the selected tier.
>
> - Only the tier structures are copied, annotations on the source tiers will not be copied.

Adding depending tiers to participants' tier structures

In case of transcriptions with consistent tier sets for several participants, as described in the previous section, it is possible to add new depending tiers to each selected participants' tier hierarchies. The prerequisite is that a systematic tier naming scheme is in place. To use this function choose Tier > Participant > Add Dependent Tiers to Participants' Tier Structures..., which will show the following window.

Figure 2.27. Add Dependent Tiers to Participants' Tier Structures window

The window is divided into two main panels. The Select Participants table lists all participants, extracted from the Participant attribute of tiers. By default all are selected, but it is possible to deselect participants. The Tier Options panel allows to configure the properties of the tiers to be created. Here is a list of the options the panel provides:

- Specify whether the suffix or prefix of tier names can be replaced to create the new tier names. If e.g. there are tiers named "A_mb", "B_mb" etc., where the first part of the name denotes the participant and the last part the "type" of the tier, the suffix is the part to be replaced when constructing the new tier names (which could be e.g. "A_ps", "B_ps" etc.).

- the New Prefix or Suffix Value textfield can be used to specify the new part of the tier name that comes before or after the participant indicator. (E.g. "ps" in the previous example.)

- in the Separator Character field, the character that is used to separate the affix from the rest of the name, can be selected (e.g. the "_" in the previous example). It is also possible to enter a custom value by selecting Custom... in the drop-down list.

- the Example Parent Tier field should be filled with the full name of one of the parent tiers to which a new tier should be added as dependent tier (e.g. "A_mb" in the example above). This tells the function what the old value of the prefix/suffix is, where in the tier hierarchies the new tiers should be added an which tier types are acceptable for the new tiers.

- the Annotator field allows to enter the optional Annotator attribute of the new tiers

- in the Tier Type drop-down list the tier type of the new tiers can be selected. This list is populated and updated after an example parent tier has been entered.

- in the Input Method list an optional input method can be selected

- the Content Language list allows to select a language for the annotations on the new tiers

While filling in the fields, a tier name is shown in blue at the bottom of the panel, as an illustration of the result with current settings.

Clicking the OK button creates the tiers, unless required fields are not sufficiently filled in. In that case a warning message will be shown, pointing at the field to correct or fill out. Otherwise, if all went well, a report will be shown listing the tiers that have been created.

🗐 Note

> This function can be applied to individual files. If many existing transcription files need to be enriched with new tiers, it is sometimes also possible to use the update multiple files with a template function (see the section called "Update Transcriptions with elements from a Template")

Delete Participant

It is possible to delete all tiers of one or more participants via the Tier > Participant > Delete Participant... menu.

Figure 2.28. The Delete Participant window

The window that appears shows a list of all participants, extracted from the Participant attributes of the tiers. Here all participants can be selected whose tiers are to be deleted. After clicking the Delete

button a confirmation dialog will be shown, listing all tiers that will be deleted. The usual rules apply, e.g. if a tier is deleted, all dependent tiers are deleted as well. After clicking the Yes button the tiers are actually deleted, while No closes the dialog to return to the participant selection window.

Data categories

📄 Note

The ISO 12620 data category registry (database and web service) moved several times and is currently accessible via DatCatInfo [https://datcatinfo.net/], linking to the web interface to the actual data at https://datcatinfo.termweb.eu [https://datcatinfo.termweb.eu/search/terms]. The old web service ELAN relied on to list profiles and data categories is no longer there. The new system provides a REST API which is quite different from the old one, therefore ELAN is currently not able to connect to the service. It is not clear yet whether the ELAN code will be updated to support the new API and be able to list the profiles and data categories again or if this functionality will be removed.

In ELAN users are free to invent their own tier setup and labelling method. This flexibility is often necessary due to the nature of the data that is to be transcribed. Moreover, people that are involved in the transcription process may not be fluent in English and as a result an international (English) annotation scheme is not applicable. In those cases a controlled vocabulary (see the section called "Controlled Vocabularies") and templates (see the section called "Saving a template") are convenient tools to help annotators.

The downside of all this flexibility is the amount of work involved to make language resources interoperable. When dealing with only a few resources, data can be manageable, but with an increasing number of resources a convenient way to make them interoperable becomes more important. For this purpose the ISO Data Category Registry is developed.

The Data Category Registry (or DCR) is an list of linguistic concepts covering a range of linguistic domains. The concepts in the DCR can be referenced to from all sorts of tools and resources. Therefore, the DCR acts as a intermediate between those tools and resources.

Referencing to a Data Category is implemented in ELAN as follows. Depending on the type of data you are referencing from (tier type (the section called "Associating a tier type with a Data Category"), controlled vocabulary entry (the section called "Associating a CV entry with a Data Category") or annotation (the section called "Associating an annotation with a Data Category")), the following or a similar window is displayed.

Figure 2.29. Local Data Category Selection

The left panel shows the categories stored on your local system. Since there are none in the left panel, the right panel does not display any name or description. To add categories, click on Add Categories. The following window appears:

Figure 2.30. Remote DCR

This window displays the DCR on a remote server. It includes all profiles and the data categories of those profiles. To select one or more data categories for local storage first click a profile in the left panel. All data categories of the selected profile are displayed in the middle panel ordered by alphabet, ID or Broader Concept. If you select a data category, information of the category is displayed in the right panel. For instance, the data category *partOfSpeech* has Id 1345 as can be seen below. Holding the **CTRL** key while clicking multiple lines in the middle panel enables you to select more than one data category. The same holds for using the **SHIFT** key for selecting a range and using **CTRL+A** for selecting all categories from the list. Click on Apply to storing the selected data categories on you local system.

Figure 2.31. Remote Data Category Selection: Result

In the same way as described above more data categories, also from other profiles, can be selected and stored on your local system. Afterwards, you can highlight a category and associate it to a CV or tier type by clicking Apply:

Figure 2.32. Local Data Category Selection: Result

The original purpose of this system is to associate (parts of) your data to a common labelling system to improve interoperability between resources and tools. To do so, select a data category and click on Apply. This will associate the selected data category to an annotation, entry of a controlled vocabulary or tier type, depending on the point from which you entered the Local Data Category Selection.

Controlled Vocabularies

The creation of a CV

When you frequently use a certain tier type with a limited number of annotation values, it might be a good idea to associate a Controlled Vocabulary (CV) with it. Such a CV consists of a number of predefined values that a user can choose from when editing an annotation, thus making the task of the annotator less error-prone. E.g. one can choose to create a controlled vocabulary for part of speech tagging, as the tags that are used often reoccur. In order to do this select Edit > Edit Controlled Vocabularies. The following dialog will appear:

Figure 2.33. Edit controlled vocabulary

1. The CV being edited , the entries and the set language.

2. Import & export options, link to an external CV.

3. Annotation value as it is going to be filled in.

4. A description of the entered value.

5. Association with an ISO Data Category.

6. Buttons to move the selected entry up/down, top/bottom. Undo/redo changes to the CV.

To create a new CV, do the following:

1. Enter a CV Name and a description. (Each language within a CV can have a different description)

2. Click on the Add button

3. Choose a language from the pull-down menu (see the section called "Setting language(s) for a CV" for more info on setting a language.)

4. Now enter each CV entry and its description.

5. Confirm every entry addition by clicking on the Add button or by hitting Enter. When adding entries for an additional language, click Add to add a new entry, or click Change to add the entry next to an existing entry.

📝 Note

The undo function in the CV dialog window only works as long as that window is active. Once it is closed changes cannot be undone any more.

📝 Note

Also see the notes concerning the section called "Associating an annotation with an external resource".

By clicking More Options... (not yet shown in the figure above) you can choose a color that will fill the lower part of every annotation frame containing the selected CV entry. Moreover, you can choose a shortcut key to edit an annotation with a single key stroke.

The color chooser has four tabs. The last three contain different ways to choose a color, which is subsequently displayed in the lower part of the window. In the first tab you can add or insert the color displayed below and you can copy, paste and delete the selected color. The list of favourite colors is saved and used the next time you start ELAN.

Setting language(s) for a CV

New in version 4.7.0 is the option to set one or more languages for a CV. This allows you to have one CV that contains a certain vocabulary in both English and German, for instance. You can choose from nearly 7700 ISO specified languages.

◆ Caution

This change in 4.7.0 introduced an incompatibility with earlier versions of ELAN; the structure of the `.eaf` changed to some extent. As a result, when opening a new `.eaf` in an older ELAN version, the entries of CV's will be missing!

By default, the language for a CV is set to 'undetermined (und)'. To change this, click the drop-down menu and select Edit Languages...

Figure 2.34. Edit CV Languages

From the dialog that appears, select the desired language and click Change. The undetermined language in the upper drop-down menu will now be replaced with the chosen language. see the section called "Defining languages for your content" to edit the list of available languages.

Figure 2.35. Edit CV Languages 2

If you would like to add an additional language, select the desired language from the lower drop-down menu, and click Add afterwards. The chosen language will be added to the CV languages, and will be visible in the upper drop-down menu. When you are done adding or changing languages, click Close to close the dialog.

Figure 2.36. Edit CV Languages 3

Back in the main CV dialog, you will now find the list of languages under the current CV (1). There is also a second column with the language label under 'Entries' (2). Select the language from the list (1) you want to add, change or delete entries for.

Figure 2.37. Edit CV Languages 4

1. The list of current languages linked to the CV

2. Entries for the linked languages

3. Annotation value and description as it is going to be filled in.

4. Add/change/delete entries within the chosen language.

By clicking an empty field in the entries (2), you can now enter the required values (3). To add a new entry, enter the values and click Add (4). This will yield a new line in the entries. To add a value to an existing entry, as shown in the screenshot, click Change. Lastly, you can delete entries as before with the Delete button.

If you use a lot of multi-lingual CV's, you can also set the preferred default language to work with. More info about this can be found here: see the section called "Using CV's"

Associating a CV entry with a Data Category

> ### 📋 Note
>
> This functionality is currently out of order, please see the note at the section called "Data categories"

In ELAN it is possible to associate an entry of a controlled vocabulary with a data category of the ISO Data Category Repository (DCR). To do so click on Edit in the ELAN main menu and select Edit Controlled Vocabularies.... In the window that now appears, click the Browse... button to the right of the `ISO Data Category` option. In the next window, select a category from the list and click Apply. Or you can add a category by clicking Add Categories; first select a profile in the left panel and then select a data category in the middle panel. Next, click on Apply to put the selected data category in the list of selected categories. Then click to highlight it and click Apply to associate it to the CV entry.
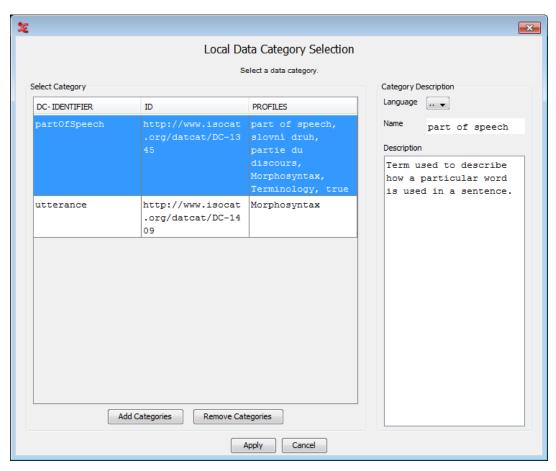
Figure 2.38. Local Data Category Selection

More information about the ISO DCR and how to use it can be found in the section called "Data categories".

Importing an existing CV

Instead of creating a Controlled Vocabulary again every time ELAN also supports the possibility to reuse existing CV's. A first way to achieve this is to use a template file that contains a CV when creating a new `*.eaf` file. (See the section called "Creating a new document") Alternatively you can import a CV for already existing files:

1. select Edit > Edit Controlled Vocabularies

2. Click on Import CV

3. Select the template (.etf), .csv or .txt file from which you want to import a CV

4. Choose Open

5. Now all CV's that are stored in the selected template file will be imported

If you try to import a CV with the same name as an already existing CV a dialog will pop up asking what to do:

Figure 2.39. Controlled vocabulary already exists warning

- Skip CV: ignore the CV from the template

- Replace Existing CV: overwrite the existing CV with that from the template

- Rename CV: opens a dialog asking you to give a new name for the imported CV

- Merge CV's: entries from the imported CV that are not in an existing CV are imported.

Using an External CV

A created or imported controlled vocabulary is part of the current transcription and is therefore part of the `.eaf` file when the transcription is saved. However, there are situations in which a controlled vocabulary is shared between two or more transcriptions and change to that CV should be propagated to all transcription that use that CV. For this purpose, ELAN enables you to link to a controlled vocabulary that is either on your local or attached hard disk or somewhere on the network or the Internet:

1. select Edit > Edit Controlled Vocabularies

2. click External CV

3. Do one of the following:

   - click on Browse... and browse to and select the file containing the External CV

   - enter a URL that point to an External CV somewhere on the Internet

4. Click OK to make the link

Similar to the Import CV process, if you try to import a CV with the same name as an already existing CV a dialog will pop up asking what to do (see Figure 2.39, "Controlled vocabulary already exists warning").

The entries of an External CV cannot be edited nor their ISO Data Category. Also the order of the entries cannot be change. The possibility to add a color and shortcut key to an entry via the More Options... button is still there.

Exporting a CV

A CV can be exported by clicking the export .ecv button in the Edit Controlled Vocabularies window. This will open another window, in which you can select the CV's to be exported:

Figure 2.40. Export Controlled Vocabulary

Finally, select a location to save the CV. The file will have the .ecv extension.

Associating a tier type with Controlled Vocabularies

Open the Change the Tier type dialog (via Type > Change tier type) and make a choice from the Use Controlled Vocabulary dropdown box:

Figure 2.41. Change tier type

Using CV's

When editing an annotation that belongs to a tier type associated with a CV, a drop down box appears, the suggest panel. Just select the value to be filled in:

Figure 2.42. Controlled vocabulary drop down box

📝 Note

- It is possible to by-pass the controlled vocabulary constraints by holding shift and double clicking on the active annotation (right clicking and selecting "Modify annotation value" while holding shift does the same).

- If a CV entry is associated with a data category of the ISO DCR (see the section called "Associating a CV entry with a Data Category"), the annotation is also associated with that data category.

If you have a created a multi-lingual CV, you can set the preferred language to use by going to Options > Language for multilingual content. From there you will see a list of languages that you have used with a CV. Select the one preferred and afterwards, values you enter are in the CV-language chosen. When the default language changes, all Annotations which are associated with a CV entry will be adapted (if the selected language actually exists in the CV, and if the entry is not empty in that language).

Figure 2.43. Select language for multi-lingual CV

More information about entering annotations from a CV and working with the Suggest Panel, see the section called "Entering annotations from a Controlled Vocabulary"

Suggesting CV entries based on the parent annotation value

The CV entries for the child annotation value can be filtered/suggested based on the parent annotation value. To enable this,a preference setting has to be enabled. By default, this preference is unselected. Edit -> Preferences -> Edit Preferences -> CV -> Suggest child annotation CV entries based on the parent annotation value

The child tier should have a tier type associated with a CV and a symbolic association stereotype. The automatically filtered/suggested values depend on the two options:

1. Suggest CV entries based on any of the words in the parent annotation value

   If this option is selected, the suggestions are based on all individual words of the parent annotation value. The result also is in combination with Suggest Panel options selected in the CV preference.

Figure 2.44. Look for CV entries that contain instead of start with the search string unselected

Figure 2.45. Look for CV entries that contain instead of start with the search string selected

The suggestion also considers Also look in the descriptions of CV entries and Ignore case

2. Suggest CV entries based on the exact sequence of words of the parent annotation value

If this option is selected, as the name suggests, suggestions are based on the exact sequence of words of the parent annotation value.

Figure 2.46. Suggestions based on the exact sequence of words

The suggestion also considers Also look in the descriptions of CV entries and Ignore case

Consulting a Lexicon

> 📝 **Note**
>
> In addition to the web based lexicon services described here, ELAN (version 5 and higher) also contains a built-in lexicon component, available as part of the Interlinearization mode. See section the section called "Lexicons".

To improve consistency an annotator can use a controlled vocabulary (CV). From a CV an entry can be selected that serves as annotation value. Sometimes, however, it is not immediately clear what CV entry should be selected given a certain media fragment. In that case looking up a proposed annotation in a relevant lexicon could help to make a decision. ELAN enables the user to perform lookups in lexicons through the following steps:

1. Set up a new Lexicon Service (see the section called "Setting up a Lexicon Service").

2. Enhance a Tier Type so that it's tiers can do a lexicon lookup (sec the section called "Set the Lexicon Service and entry field information for a Tier Type").

3. Perform the lookup (see the section called "Perform a Lexicon Lookup").

Setting up a Lexicon Service

A Lexicon Service is in essence a link to a certain lexicon that is hosted on a certain lexicon server. That means that for a user to be able to connect to a lexicon, he needs to connect to the server and subsequently select one of the available lexicons. You can either add a new lexicon or Import an existing one. To add a new lexicon do the following:

1. In the main menu select Edit > Lexicon Services...

2. In the Edit Lexicon Services click Add

Figure 2.47. Edit Lexicon Services

> 📝 **Note**
>
> If the Add button is disabled, it means that there is no ELAN extension loaded that can handle Lexicon Services. To install an extension, please consult the release notes of ELAN at https://archive.mpi.nl/tla/elan/release-notes.

3. On the first page of the dialog that now appears..

   a. Enter a name that is later used to refer to this Lexicon Service.

   b. Select the Lexicon Server Type.

   c. Enter the URL of the lexicon server.

d. Enter your username and password for the lexicon server.

Figure 2.48. Name and Lexicon Service Info

4. Click Next

5. On the second page of the dialog select the lexicon you wish to connect to. The bottom half of the page will contain the description of the lexicon.

Figure 2.49. Lexicon Selection

6. Click Finish

A new Lexicon Service will appear in the drop down list of the Edit Lexicon Service dialog. Click Close to close this dialog.

To import a lexicon do the following:

1. In the Edit Lexicon Services click Import

2. The following dialog window will appear:

Figure 2.50. Previously defined services

Once you have selected a service name from the list, you can either click Delete to delete it, or Import to import it. The service name will now be displayed in the Edit Lexicon Service window and you can now add it (see steps 2-6 above).

Set the Lexicon Service and entry field information for a Tier Type

Before an annotation can be used for a lexicon lookup, the tier type of the tier must have a reference to a Lexicon Service. In addition, there must be a reference to the field of an entry that is of interest. To set this information for a tier type, do the following:

1. In the main menu select Type > Change Tier Type...

2. Select the a tier type in the pull down menu Select Type.

3. Next to Lexicon Connection click Select...

4. Select a Lexicon Service in the drop down list at the top of the dialog.

5. A list of fields that compose a lexical entry is requested from the lexicon server and shown in the table. Select one.

6. Click OK

Figure 2.51. Select a Lexicon Service and Entry Field

In the Change Type dialog the Lexicon Service name and Lexical Entry Field name are shown. Click Change to commit to the new Lexicon settings.

Perform a Lexicon Lookup

If the tier type of a tier is enhanced with lexicon service and entry field information, lexicon lookups using the annotations of that tier can be performed:

1. Open the Lexicon Entry tab in main ELAN window.

2. Select an annotation on a tier of which the tier type is enhanced with a lexicon service and entry field information.

3. In the Lexicon Entry tab the annotation is entered in the field Annotation and the Get Lexicon Entries button is enabled to indicate a lookup is possible.

4. Select a constraint from the drop down list.

5. Click the Get Lexicon Entries button.

6. A lookup is performed and the results are presented on the right side of the tab in the form of a tree structure.

7. Open an entry and subsequent entry nodes by clicking the open icon in front of a node (if there is one).

8. If a node value consist of a URL, selecting the node will open the URL in your default browser.

9. If you select the top node of an entry, the Change annotation button is enabled to indicate that you can use the value of the entry field as value of the active annotation. Click this button to do so. You can also change the value of the dependent tiers by clicking Change annotation + dependents.

Figure 2.52. The Lexicon Entry tab

Signbank lexicon service

The Signbank lexicon service is a lexicon service plugin that makes a connection between ELAN and an instance of Signbank. Signbank is a webapplication for constructing sign language lexicons (the original Signbank: http://www.auslan.org.au/ [http://www.auslan.org.au], several spin offs: https://github.com/Signbank). When ELAN is connected to a Signbank instance that is equipped with this functionality, ELAN is able to show videos of Signbank entries.

Setting up the connection

For setting up the connection, the section called "Consulting a Lexicon". Note that when your are editing the lexicon service, you need to select the type Signbank. After editing a new lexicon service, the initial data from the Signbank is downloaded immediately. This may take a while because all available media is also downloaded.The reason for this is that ELAN needs quick access to the media and this is not guaranteed if the data is requested on demand. As the download is going on, now a progress dialog appears to indicate the progress. See below images of the progress dialog.

Figure 2.53. Signbank progress dialog: in progress

The dialog mentions the service name, progress of the package, video and image files download. The Last updated on label indicates the last time the package was downloaded. It also gives feedback to the user to be patient until the download is complete.

Figure 2.54. Signbank progress dialog: completed

Once the download is complete, it indicates that the cache is created and now the lexicon service can be used. It also updates the last updated time to the current time. Now the user can start using this lexicon.

Changing the location of cached data

It may prove necessary to change the location of cached data (download data from signbank) from the special ELAN data folder (see 1.1.2. Special ELAN data folder) to elsewhere, because this data may take a great amount of space. See below for how to do this.

It can be set in the Preferences tab of Edit > Preferences>Edit Preferences. At the item 'Location of cached data' click Browse... and navigate to a new folder in a different location than the default and click Select. The X restores the default file location.

Using the Signbank lexicon service

On top of the default functionality of a lexicon service the Signbank lexicon service comes with a few extras. For those extras to work tiers with a tier type that is connected to the Signbank lexicon must also have a connection to an ECV (see the section called "Using an External CV") from that same Signbank. If so the following functionality is available.

For a start, when right clicking an annotation of a tier with a tier type that is connected to the Signbank lexicon, there is an option Show in Signbank in the menu. If clicked the Signbank entry corresponding to the annotation value (which is an ECV entry) is opened in a browser.

Figure 2.55. Show in Signbank

Also, when hovering over an annotation of a tier with a tier type that is connected to the Signbank lexicon, while **CTRL** is pressed, a video of Signbank entry corresponding to the annotation value is displayed next to annotation box.

Figure 2.56. Hovering over annotation while pressing CTRL

A videos is also displayed when opening a annotation for editing, in this case selecting an ECV entry, and hovering over an ECV entry in the list.

Figure 2.57. ECVentry hovering

Finally, when searching for lexicon entries in the Lexicon tab, each entry in the result has a link to the lexicon entry in the Signbank webapplication. If clicked the Signbank entry is opened in a browser.

Figure 2.58. Lexicon entry link

The fields to be shown in the Lexicon result tab can be configured. When Filter Result Fields is clicked, below shown dialog appears where the fields to be seen in the result tab can be selected and saved. Then when Get Lexicon Entries is clicked, instead of all the fields, only the saved fields are shown in the result tab. If no fields are selected, all the fields will be listed in the result tab.

Figure 2.59. Lexicon filter result fields

Figure 2.60. Lexicon filtered result

Media Displayer

Apart fron the linked media (audio and/or video) that is the basis of the transcription, it is now possible to display media files that are linked to annotations and controlled vocabulary entries. This may come in handy e.g. when choosing an entry from an controlled vocabulary and you want audio and/or visual information to help you choose the correct entry.

Setting up the connection

To accomplish this a special lexicon service extension that also functions as a media provider should be used. First set up the lexicon service (see the section called "Setting up a Lexicon Service"). Next, connect the lexicon service to a tier type (see the section called "Set the Lexicon Service and entry field information for a Tier Type") that is also connected to a external controlled vocabulary with entry IDs that can be used by the lexicon service (see the section called "Associating a tier type with Controlled Vocabularies").

An example of a lexicon service extension that does all this is the Signbank extension. It connects to a Signbank web service (the original Signbank: http://www.auslan.org.au/, several spin offs: https://github.com/Signbank) that also provides a corresponding ECV containing the necessary data.

📋 Note

The media displayer now also works for external references of annotations, in case the referent is video (see the section called "Associating an annotation with an external resource".

Using the media displayer

When everything is set up correctly, there are several ways to display the media. First, when holding the CTRL button while hovering over the annotation a media file is displayed next to the annotation, if there is a media file linked to the CV entry of the annotation.

Figure 2.61. Hovering over annotation while pressing CTRL

Secondly, when opening an annotation for editing, a list of CV entries is shown. Hover over a CV entry in the list to display the corresponding media.

Figure 2.62. ECVentry hovering

How to make a selection

In order to make an annotation, you first have to select a time interval. ELAN supports the following options:

- making and saving a selection on an independent tier (the section called "Making a selection on an independent tier");

- making and saving a selection on an independent tier while playing (the section called "Using the selection controls");

- making and saving a selection on a referring tier (the section called "Making and saving a selection on a referring tier");

- deselecting a selection (the section called "Deselecting a selection");

- changing the time alignment of an existing selection (the section called "Changing the boundaries of an existing selection and annotation");

- activating and deactivating the Bulldozer mode (the section called "Activating and deactivating the Bulldozer mode or Shift mode").

Making a selection on an independent tier

There are several ways to make a selection. If you wish to only use the mouse, do the following:

1. Go either to the Waveform or the Timeline Viewer.

2. Go with the mouse to the beginning of the time interval you want to select.

3. Click the mouse button, keep it clicked and drag it to the endpoint of the time interval you want to select.
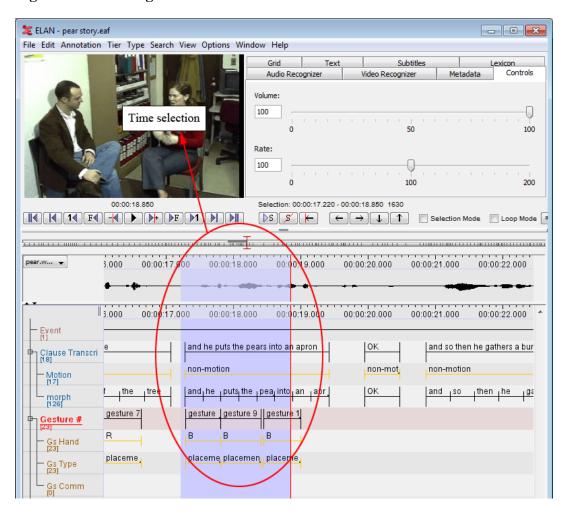
Figure 2.63. Making a selection

The video image will be continuously updated. The selected part is highlighted in light blue color. You can use the shortcut **SHIFT+A** to put the selection in the center of the Timeline Viewer.

The selection can be extended beyond the size of the current window. The display in all Viewers will automatically move along.

You can change the beginning and endpoints of the selection. Choose one of the following options:

1. Either use the mouse: press the **SHIFT** key, keep it pressed and click with the mouse to the left/right of the selected part. The selection will be extended to include this point.

2. Or enable the *Selection Mode* by selecting the Selection Mode checkbox. When selection mode is enabled, you can use the media controls to edit the selected part. When moving the crosshair

in Selection Mode, the current selection is narrowed or broadened, depending on the direction the crosshair is moved to. For a complete overview of the use of the media controls, see the section called "The Media Player options".
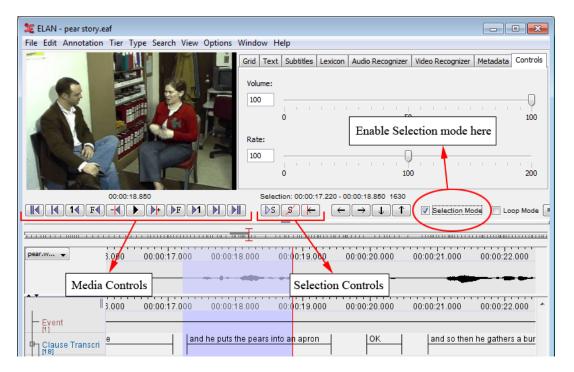
Figure 2.64. Change selection

If there is not already a selection, there is another way to make a selection. First put the cross hair at a position where you want the beginning or the end of the selection to be. Then press the **SHIFT** key and keep it pressed while clicking with the mouse at the position where you want the other end of the selection to be. A selection between the cross hair and the click position is created.

Selecting multiple annotations

It is also possible to select an annotation as a whole by clicking on the annotation unit. The annotation will then be selected. This way, it's also possible to make a selection spanning multiple independent annotations. To do so, press and hold crtl+alt (windows) or alt (osx) and click on the annotations you want to select. The selection will expand accordingly. To remove an annotation from the selection, just click it again and it will be deselected.

Using the selection controls

The selection controls allow you to navigate through or to change the active selection. For their use, see again the section called "The Media Player options".

Example: extending the begin of a selection with one second

Assume that you have made a selection and that you want to add a time interval of 1 second to the begin of it. In that case you should perform the following actions:

1. Move the crosshair to the begin of the active selection

Figure 2.65. Crosshair on begin of selection

2. Enable Selection Mode

Figure 2.66. Selection mode

3. Go back one second by clicking the corresponding button from the media controls.

Figure 2.67. Back one second

4. Turn off the selection mode and enter an annotation for the selection.

Figure 2.68. Selection mode off

Creating annotation units without gaps

It is possible to start a new selection immediately after a previous selection has ended, i.e., the endpoint of one selection will be taken as the starting point for the next selection. Do the following:

1. Make a selection on an independent tier (see the section called "Making a selection on an independent tier").

2. Double-click on the selection.

3. Optionally enter the content of the annotation unit. Press the keys **CTRL+ENTER**. The selection is saved.
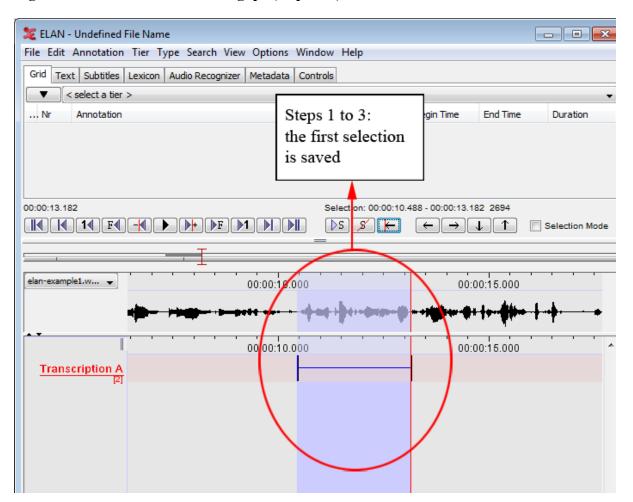
Figure 2.69. Annotation without gaps (steps 1-3)

4. Press the keys **ALT+SHIFT+C** (or **ALT+C**) or click on the clear selection icon to deselect the selection (see the section called "Deselecting a selection" for deselecting a selection).
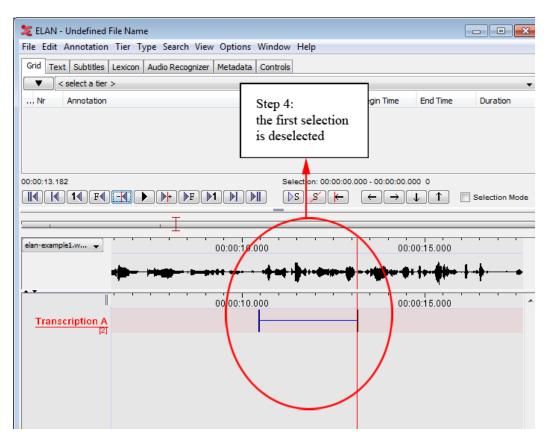
Figure 2.70. Annotation without gaps (step 4)

5. Enable the selection mode. Then, play the video or sound file until the playback stops. The new selection extends from the endpoint of the previous selection until the point when the playback was stopped.

Repeat steps 2 and 3 to save the new selection.

Repeat steps 3 to 5 to add another selection.
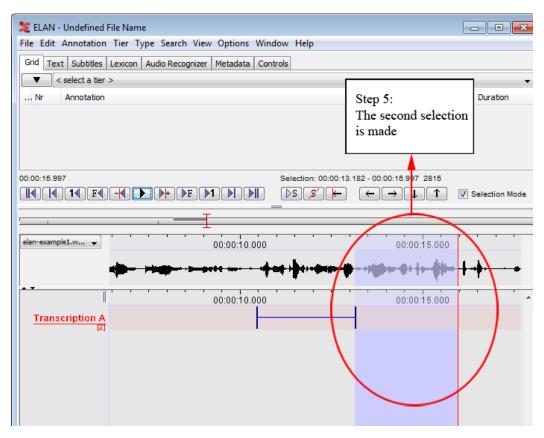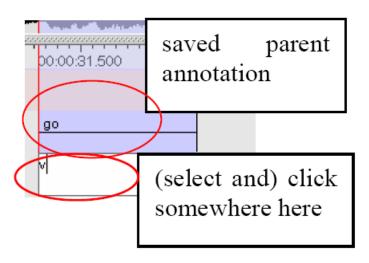
Figure 2.71. Annotation without gaps (step 5)

Making and saving a selection on a referring tier

A referring tier inherits all its time alignments from its parent tier. To make and save a selection on a referring tier, do the following:

1. Select and save a time interval on the corresponding parent tier (see the section called "Making a selection on an independent tier" and the section called "How to enter and edit annotations").

2. Double-click somewhere within the time interval of the parent annotation at about the height of the referring tier. The Inline Edit box appears.

Figure 2.72. A selection on referring tier

3. Do one of the following:

    a. Enter an annotation (see the section called "How to enter and edit annotations"), and then press the keys **CTRL+ENTER** to save the selection.

    b. Press the keys **CTRL+ENTER** (without entering an annotation) to save the selection.

Deselecting a selection

To deselect a selection, do one of the following:

1.

    Use the Deselection icon from the selection controls:

2. Use the shortcut key **ALT+SHIFT+C** or **ALT+C**.

3. Use the shortcut key **CTRL+SHIFT+Z**. This shortcut also cancels selecting mode (see the section called "Making a selection on an independent tier").

📄 Note

Whenever you select another time interval, the old selection is automatically deselected, unless you enabled Selection Mode.
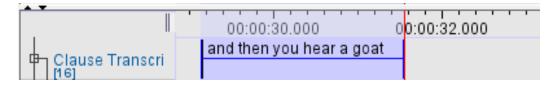
Changing the boundaries of an existing selection and annotation

By selection

In the timeline viewer, go through these steps:
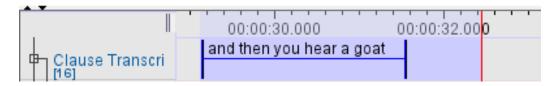
1. Click on an annotation unit to select it.

Figure 2.73. Select annotation unit

2. Select the region where you want the modified annotation to be placed.

Figure 2.74. Region for modified annotation
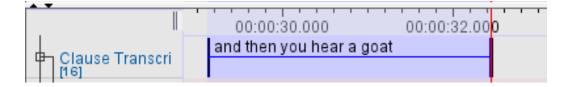
3. Right click on the original annotation and select Modify annotation time or press **CTRL+ENTER**

Figure 2.75. Modify annotation time

4. Now the length of the annotation becomes that of the selection from the second step.

Figure 2.76. Annotation length changed
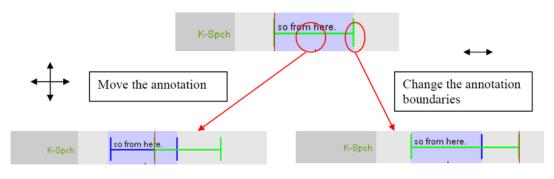
By dragging with the mouse

If you press **ALT**, the active annotation is indicated in green and becomes adjustable with the mouse:

• drag in the middle of the annotation and drop it somewhere else to move it

• drag and drop the borders to change the boundaries of the annotation unit

Figure 2.77. Changing selection boundaries with the mouse

Note

• Only the time-alignment of annotations on the following types of tiers can be modified: annotations on independent tiers, and annotations on referring tiers that fall under the Time Subdivision stereotype (but note that in the latter case, the alignment cannot be extended beyond

the boundaries of its parent annotation, see the section called "Basic Information: Annotations, tiers and tier types").

- To modify the time alignment of annotations on all other tiers, change the time alignment on the corresponding parent tier (following the steps above). The time alignment on all referring tiers is automatically updated. The annotations on the referring tier that are no longer within the borders of the annotation on the parent tier are discarded. If you want to shift the annotations on a referring tier in the same way as the annotation on the parent, use the methods described in the section called "Shifting annotations".

- If two annotations are adjacent you can snap them by specifying the maximum close-value in ms.

By entering a new begin and/or end time

The precise begin and end time of an annotation can be entered in the hh:mm:ss:ms format in a window that can be invoked via the Annotation > Modify Annotation Time... menu or the corresponding keyboard shortcut (by default **CTRL+SHIFT+M**).

Figure 2.78. Modify annotation boundaries by typing

By entering a number of milliseconds for shifting

Annotations can be shifted to the left or the right by a number of milliseconds. See the section called "Shifting annotations".

Changing selection boundaries with the shortcut keys

Use the shortcut key **CTRL+J** and **CTRL+U** to change the left boundary of a selection to the left and to the right, respectively. To change the right boundary use the shortcut key **CTRL+SHIFT+J** and **CTRL+SHIFT+U**.

Shifting annotations

To shift a specific annotation, go through these steps in the timeline viewer:

1. Click on an annotation unit to select it.

2. Right click the annotation and select Shift Active Annotation or press **CTRL+SHIFT+ENTER**.

3. Enter a number of ms/ss.ms/mm:ss.ms/hh:mm:ss.ms (between -510 ms and 1080 ms) by which the annotation should shifted. If the number is greater that zero, the annotation is shifted to the right. If it is less that zero, is shifted to the left.

Figure 2.79. Shift annotations

4. Click on OK.

📑 Note

> The number of milliseconds you can enter is limited by the end of the annotation to the left and the begin of the annotation to the right, or by the begin or end of the timescale.

It is also possible to shift more than one annotations at once. To do so, first decide which annotations you would like to shift:

• The annotations within a selection. In this case, select the annotations to shift.

• The annotations to the left or to the right of a point on the timeline. In this case, place the crosshair on that point.

Then click Annotation in the main menu and select Shift >. This sub menu has the following options:

• Active Annotation: same as described above.

• Annotations on the Active Tier in the Selected Time Interval.

• Annotations on the Active Tier, Left of Crosshair.

• Annotations on the Active Tier, Right of Crosshair.

• Annotations on All Tiers, Left of Crosshair.

• Annotations on All Tiers, Right of Crosshair.

All these options result in a window as in Figure 2.79, "Shift annotations". Enter a number of millisecond and click OK. All annotations referred to in the Annotation > Shift > menu option are now shifted by the number of millisecond you entered.

A final option is to shift all annotations on all tiers. To do so, click Annotation > Shift All Annotations.

Activating and deactivating the Bulldozer mode or Shift mode

ELAN supports three editing modes: Overwrite mode, Bulldozer mode and Shift mode. These modes are somehow comparable to the Overwrite and Insert modes that are supported by many text editor programs such as, e.g., Microsoft Word.

• Normal (i.e. overwrite) mode: if you extend a selection into a time interval that is already occupied by an annotation, that annotation is (partly or wholly) overwritten.

Figure 2.80. Normal mode

- Bulldozer mode: if you extend a selection into a time interval that is already occupied by an annotation, that annotation is moved to the right/left. Think about it as a bulldozer which pulls all annotations together, discarding the spaces in between.

Figure 2.81. Bulldozer mode

- Shift Mode: like Bulldozer Mode, but the spaces between annotations are preserved too. This resembles most to the insert mode of text editors (see also the section called "Synchronizing video files").

Figure 2.82. Shift mode

The following conventions apply for the Bulldozer mode:

- Annotations are moved to the right if you extend your selection from left to right. They are moved to the left if you extend your selection from right to left.

- If a moved annotation extends into the time-interval of yet another annotation, that other annotation is moved accordingly. If it extends into empty space, no other annotations are affected.

🗒 Note

Moving annotations may thus affect the whole document, and may thereby destroy previous time alignments. Please make sure that the Bulldozer Mode is not accidentally switched on.

The Overwrite mode is the default mode. To switch to another mode, do the following:

1. Click on Options > Propagate Time Changes.

2. Click on Normal Mode, Bulldozer Mode or Shift Mode. A check mark appears next to it. To switch back to the Overwrite mode, repeat steps 1 and 2 above. The check mark disappears.

How to enter and edit annotations

All annotations are entered either into the Inline Edit box or into the Edit Annotation box, which are accessed via the Timeline, Interlinear, Grid, Text and Subtitle Viewers. ELAN supports the following options:
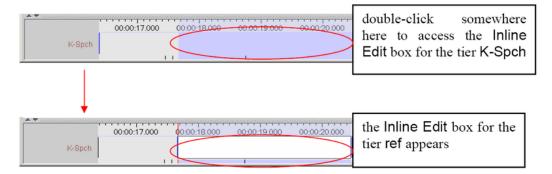
- entering annotations (section the section called "Entering annotations (into the Inline Edit box)" and the section called "Entering annotations (into the Edit Annotation box)");

- entering annotations before/after other annotations (the section called "Entering annotations before/after other annotations");

- modifying the content of annotations (the section called "Modifying the content of annotations");

- deleting annotations (the section called "Deleting annotations");

- entering annotations in different character sets (the section called "Entering annotations in different character sets").

Entering annotations (into the Inline Edit box)

An Inline Edit box from a selection

1. Select the time span in which you want to place the annotation.

2. Access the Inline Edit box by doing one of the following:

   a. Either double-click in the Timeline Viewer on the selection at about the height of the tier where you want to enter the annotation.

   b. Or click on Edit menu, then click on New annotation here (active tier only).

   c. Or use the key equivalent **ALT+N** (active tier only).
   In all cases, the Inline Edit box appears, e.g.:

   Figure 2.83. Inline Edit box

3. Do one of the following:

   a. Press the keys **CTRL+ENTER** (without entering an annotation) to create an empty annotation.

   b. Enter an annotation and then press the keys **CTRL+ENTER** to save the selection.

      It is possible to enter text that contains line breaks. The text entry box automatically displays a scrollbar if necessary.

      The Inline Edit box is automatically preconfigured for the default character set of the tier. If you want to use a different character set, do the following:

i. Right-click in the Inline Edit box. A pull-down menu appears that displays the available character sets.

Figure 2.84. Select character set

ii. Click on the appropriate character set. From now on, the characters are entered in the selected set.

📄 **Note**

If you are using a third-party keyboarding solution like Keyman, make sure to select the default system language as input language for the tier to be edited (e.g. Dutch if your system language is set to Dutch).

iii. To switch back to the default character set, repeat the steps above and select the default set from the pull-down menu.

📄 **Note**

Only selections on time-alignable tiers can be saved in this way. To save a selection on a referring tier, see the section called "Making and saving a selection on a referring tier".

4. Save the annotation by doing one of the following:

a. Use the shortcut keys **CTRL+ENTER**, you can change this shortcut to ENTER in Preferences, see the section called "Editing preferences"

b. Or right-click in the Inline Edit box and click on Commit Changes in the pull-down menu.

To exit the Inline Edit box without saving, do one of the following:

1. Use the shortcut key **ESC**.

2. Or right-click in the Inline Edit box and click on Cancel Changes in the pull-down menu.

> **📝 Note**
>
> When annotations are created, they can be aligned with the video frames by setting it as a preference (see the section called "Editing preferences")

An Inline Edit box from begin and end time

An alternative method to get an Inline Edit box is:

1. Click a time that should become the begintime of the annotation.

2. Press **SHIFT+ENTER**.

3. Click a time that should become the endtime of the annotation.

4. Again press **SHIFT+ENTER**.

An Inline Edit box appears on the selected tier. You can now enter an annotation and save it in the way explained above.

A new annotation from manually entered begin and end time

It is possible to create one or more annotations by manually entering the begin and end time for the new annotation(s).

1. Choose Annotation > New Annotation from Begin-End Time or use the shortcut key **CTRL+ALT +SHIFT+N**

2. A dialog window appears as shown below:

Figure 2.85. New annotation from begin and end time

3. a. Enter the begin time in the first textfield and the end time in the second field

   b. Optionally enter an annotation value in the third text field. If more than one tier are selected, the value will be applied to all new annotations.

   c. In the table in the lower part of the window, select the target tier or tiers.

   d. Click OK to create the annotation(s)

An Inline Edit box for a active annotation

To open the inline edit box for a active annotation, either double-click in the Timeline Viewer on the active annotation or use the shortcut key **CTRL+SHIFT+E**.

Entering annotations (into the Edit Annotation box)

The Edit Annotation box differs from the Inline Edit box in that it has a bigger, resizeable, display and supports line wrapping. It is therefore better for entering longer texts.

To enter an annotation into the Edit Annotation box, do the following:

1. Either make a selection in the Timeline Viewer (see the section called "Making a selection on an independent tier"), or click on an existing annotation in the Timeline or Interlinear Viewer.

2. Access the Inline Edit box (see the section called "Entering annotations (into the Inline Edit box)").

3. Do one of the following:

   a. Either use the shortcut keys **SHIFT+ENTER**.

   b. Or right-click in the Inline Edit box. A pull-down menu appears. Click on Detach Editor.

4. The Edit Annotation box appears:

Figure 2.86. Edit annotation

The Edit Annotation box is automatically preconfigured for the default character set of the tier (see the section called "Changing tier attributes"). If you want to use a different character set, do the following:

a. Click on Select Language. A pull-down menu appears that displays the available character sets, e.g.:

Figure 2.87. Select language

b. Click on the appropriate character set. From now on, the characters are entered in the selected set. (For an overview of the input methods for the character sets see the section called "Entering annotations in different character sets").

c. To switch back to the default character set, repeat the steps above and select the default set from the pull-down menu.

5. Edit the annotation.

6. Save the annotation by doing one of the following:

a. Use the shortcut keys **CTRL+ENTER**.

b. In the Edit Annotation box, click on Editor and then click on Commit Changes in the pull-down menu.

To exit the Edit Annotation box without saving, do one of the following:

1. Use the shortcut key **ESC**.

2. In the Edit Annotation box, click on Edit and then click on Cancel Changes in the pull-down menu.

To return to the Inline Edit box, do one of the following:

1. Use the shortcut keys **SHIFT+ENTER**.

2. In the Edit Annotation box, click on Attach Editor in the pull-down menu.

Entering annotations from a Controlled Vocabulary

If the tier type of the selected tier has a controlled vocabulary both the Inline Edit box and the Edit Annotation box show a list of the CV entries from which an entry can be selected. To commit to a selected entry, use **Enter** or **Ctrl+Enter**. Double clicking an entry has the same effect.

The use of just a list works well when the number of entries is limited. For larger CVs another method of selecting the correct entry can be used. If you are either in the Inline Edit box or Edit Annotation box you can find and select a CV entry by reducing the selection list as you type the first characters of an entry. To enable this method, do the following:

1. Right-click in the Inline Edit box or Edit Annotation box.

2. In the pull-down menu select Toggle Suggest Panel....

3. The box now changes to two parts: a text field on the top and a list on the bottom.

4. Start typing the first few characters of the entry you want to select in the text field.

5. As you type, the list is updated to only containing those entries that start with the characters you typed so far.

6. Using the arrow up and down keys or just by clicking an entry you can select an entry.

7. **Enter** or **Ctrl+Enter** or double clicking an entry commits the selected entry and changes the annotation to the value of the selected entry.

Instead of right clickling in the Inline Edit box, you can also use the tabs Suggest Pane and List Pane to switch between the suggest panel and the list panel.

Figure 2.88. The suggest panel
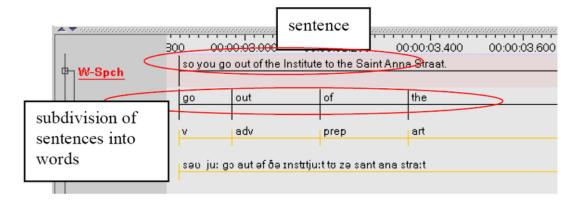
There are some options you can set for the suggest panel, which can help in searching the CV entries. These options can be set in Preferences (See the section called "Editing preferences")

Entering annotations before/after other annotations

Annotations on some tiers can be subdivided into smaller units. In this way, you can for example break up a sentence into different words (or words into morphemes, etc.), as in the following illustration:

Figure 2.89. Subdividing annotation
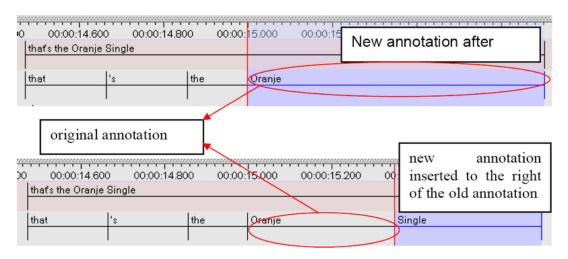
To divide an annotation into smaller units, do the following:

1. In the Timeline or Interlinear Viewer, click on the annotation that you want to subdivide. It appears in a dark blue frame.

2. Do one of the following:

   a. Right-click on the selected annotation. A pull-down menu appears. Click on either New Annotation before or on New Annotation after to subdivide the annotation.

   b. Or click on Edit menu. Then click on either New Annotation before or on New Annotation after to subdivide the annotation.
   If you click on New annotation before, the original annotation is divided and the new annotation is inserted to its left (as in the illustration below). If you click on New annotation after, it is inserted to its right.

Figure 2.90. New annotation after

📄 Note

This option is only available for those tiers that are assigned to the stereotypes Time Subdivision and Symbolic Subdivision (see the section called "Basic Information: Annotations, tiers and tier types").

An annotation is always subdivided into two units. If you need further subdivisions, repeat the steps above.

Creating depending annotations for a active annotation

To create depending annotations for a current active annotation, select Annotation > Create depending annotations. This will create dependant annotations for the current active annotations on all its child tiers.

Creating annotations on dependent tiers

1. In the main menu, select Tier > Create Annotations on Dependent Tiers.... A dialog appears with a list of independent(parent) tiers.

2. Select all the parent tiers of the dependent tiers, on which the annotations are to be created.

3. Click on Next, will bring up this dialog box:

Figure 2.91. Create annotations on dependent tiers

4. Select all the dependent tiers on which the annotation are to be created.

5. Select Empty Annotations on a dependent tier to create empty dependent annotations of the parent annotations or select Annotation With Value of Parent to create dependent annotations with the value on the parent annotations.

6. Check Overwrite the annotation values to overwrite the values of the dependent annotations(if any) with the values of the parent annotation.

7. Finally click on Finish to create annotations.

Creating annotations from gaps

Gaps between annotations on a tier can be transformed into annotations on the same tier or on a new tier. To do so, click Tier > Create Annotations from Gaps.... In the dialog window, select the (independent)

tier for which you wish to transform the gaps into annotations. Then select whether you wish to put the new annotation on the same tier or on a new tier and specify a tier name in the latter case. Also specify the contents of the new annotations: either a specific value, the duration of a gap or no contents.

It is possible to select multiple tiers when creating annotations from gaps. Selecting multiple tiers can be done by holding the **CTRL** key while clicking other tiers than the one already selected. The **SHIFT** key can be used in a similar way to select the range of tiers from the one that is selected to the one that is clicked. The gaps created from multiple tiers are periods where each of the selected tiers has no annotation.

Creating regular annotations

When there is a time interval you wish to divide in a number of adjacent annotations with the same duration you can do that by clicking on Tier > Create Regular annotations. In the dialog window (see Figure 2.92, "Regular Annotations") select the tiers that must receive the new annotations. Then do one of the following:

1. Enter Start time and Duration (End time is calculated automatically).

2. Enter Start time and End time (Duration is calculated automatically).

Finally, enter a Annotation size and click Apply.

Figure 2.92. Regular Annotations

The total duration of the new annotations does not exceed the interval between start and end time. So if start time is 1.000 seconds and, end time is 4.000 seconds (interval of 3 seconds) and the annotation size is 2.000 seconds, than only one new annotations is created because two would make a duration of 4 seconds which exceeds the interval defined by the start and end time.

> 📑 **Note**

> If there is an overlap between the new annotations and one or more existing annotations, these existing annotations will be removed.

Duplicating annotation values

If two annotations are adjacent and have the same length (for instance after creating regular annotations, see the section called "Creating regular annotations") and if the second annotation is empty, you can copy the value of the first annotation to the second. This is done by selecting the first annotation and pressing **CTRL+SHIFT+D**. Now the second annotation is selected and you can repeat the process to get a string of adjacent annotations with the same value.

An additional result of **CTRL+SHIFT+D** is that if there are annotations on another tier that have the same begin and end time as the annotations you are working on, and the second of those annotations is empty, then the value of the first annotation on that tier is also copied to the second annotation on that tier.

Snap Annotations

ELAN offers the possibility to create the so-called snap-annotations. If two annotations are adjacent you can snap them by specifying the maximum close-value in ms. You can set the value to snap annotations in the edit preferences (see the section called "Editing preferences").

Automatic labelling and numbering

Sometimes you don't want annotations to contain transcriptions of the media you are working on, but you rather want to label and number them for later processing. In ELAN this is accomplished by clicking Tier > Label And Number Annotations.... In the dialog as shown in Figure 2.93, "Label and Numbering annotations" you can select one or more tiers to provide labels and numbering.

Figure 2.93. Label and Numbering annotations

The formatting options of the label and numbering include:

- Include label part: the text to act as label (or prefix) for each annotation.

- Insert delimiter: a delimiter between the label in front and the number.

- Include number part:

  - Integer: the number of each annotation is an integer and the increment value is also an integer.

  - Decimal: the number of each annotation is a decimal and the increment value can also be a decimal.

- Prepend leading zeros: leading zeros for easy sorting in post-processing (e.g. 001, 002, 003 etc).

- Start value: the value the numbering must start with.

- Increment: the value with which the number in the next annotation is incremented.

The result of the options is shown below the options in a blue box. This result is updated as you change the options.

Modifying the content of annotations

To modify an annotation, do one of the following:

1. In the Timeline or Interlinear Viewer:

   a. Click on the annotation that you want to modify. It appears in a dark blue frame.

   b. Do one of the following:

      i. Right-click on the selected annotation. A pull-down menu appears. Click on Modify annotation value.

      ii. Or click on Annotation > Modify annotation value.

      iii. Or use the key equivalent **ALT+M**.
      The Inline Edit box appears (see the section called "Entering annotations (into the Inline Edit box)").

2. In the Grid Viewer:

   a. Double-click on the annotation that you want to modify.
   The Inline Edit box appears (see the section called "Entering annotations (into the Inline Edit box)").

Modifying the boundaries of annotations

The boundaries or alignment of annotations can be changed at any time, as described in the the section called "Changing the boundaries of an existing selection and annotation" section.

Moving annotations to another tier

It can occur that you put an annotation on the wrong tier. If the start and end time of the annotation unit are correct, you can move the complete annotation to a different tier by pressing and holding **ALT** and clicking and dragging the annotation with the mouse. This feature currently only works with time based annotations.

Dependent or child annotations will only be moved along with the parent annotation unit if it is clear from the tier names to which tier they should be moved. (from tx@A to tx@B for example). If this is not clear, the dependent annotations might be lost. A safer way to move annotations is to use the copy and paste annotation groups. See the section called "Copy and paste annotation groups"

Changing the case of Annotations

To change the case of letters of all annotation on a tier, click Tier > Change Case of Annotations.... Then select the tier(s) you wish to change, choose the case the annotations should be converted to and click OK. When changing the text to Upper-case it is possible to change only the first character (initial capital). Similarly when changing to Lower-case, it is possible to specify that each annotation should begin with a capital.

Deleting annotation values

To delete the value of annotation, do the following:

1. In the Timeline or Interlinear Viewer, click on the annotation from which you want to delete the value. It appears in a dark blue frame.

2. To delete it, do one of the following:

   a. Right-click on the selected annotation. A pull-down menu appears. Click on Remove Annotation Value.

   b. Or click on Edit menu. Then click on Remove Annotation Value.

   c. Or use the key equivalent **ALT+Delete**.
   The annotation value is deleted without further warning.

Deleting annotations

To delete one annotation, do the following:

1. In the Timeline or Interlinear Viewer, click on the annotation that you want to delete. It appears in a dark blue frame.

2. To delete it, do one of the following:

   a. Right-click on the selected annotation. A pull-down menu appears. Click on Delete annotation.

   b. Or click on Annotation in the main menu. Then click on Delete annotation.

   c. Or use the key equivalent **ALT+D**.

ELAN also gives the possibility to delete multiple annotations: click on Annotation > Delete in the main menu. Now click one of the five menu items:

• Annotations on the Active Tier in the Selected Time interval...

• Annotations on the Active Tier, Left of the Crosshair...

• Annotations on the Active Tier, Right of the Crosshair...

• Annotations on All Tiers, Left of the Crosshair...

• Annotations on All Tiers, Right of the Crosshair...

To delete a number of specific annotations on more than one tier, select those annotations by holding **ALT** while clicking them. The annotations get a purple border. Then right click in the Timeline Viewer and select Delete Selected Annotations.

📄 Note

Annotations are deleted without further warning.

If you delete an annotation on a parent tier, the corresponding annotations on all its child tiers will be automatically deleted as well. Please make sure that you do not accidentally delete a child annotation. An annotation on a child tier can be deleted without consequences for the annotation on its parent tier.

Deleting annotations and annotation values on multiple tiers

To delete annotations or annotation values on multiple tiers, select Tier > Remove Annotations or Values.... This will bring you this dialog:

Figure 2.94. Delete annotations or annotations values

Select the tiers on which the annotations are to deleted. First select whether to delete Annotations or Annotation Values on the selected tiers and then select All Annotations to delete all the annotations/ values, Empty Annotations for removing annotation units with no values in them or select Annotations where value is... to delete only annotations/ values where annotation value matches the given value and finally click on OK.

Split Annotations

Annotations that are on a tier without a parent tier (top level tier) and that have no or only depending annotations on symbolically associated tiers can be split in two ways:

• To split a annotation exactly in the center, select a annotation and click on Annotation > Split Annotation.

This will split your annotation exactly in the center and both the annotations will have the same values.

• To split a annotation at a specific point, select a annotation and right click on the point where you want to split the annotation and select Split Annotation.

This will split the annotation at the point where the right click is made.

You can also split an annotation (in one of the ways described above) which has only time aligned depending annotations (symbolically associated). In this case the annotation will be split together with its depending annotations.

Merging annotations

It is possible to merge two annotations on an independent tier into one annotation. To do so, select the first annotation of the two you want to merge. Then either right click in the Timeline Viewer or click Annotation in the ELAN main menu and select Merge with Next Annotation. The resulting annotation starts where the first had started and it ends where the second had ended. Note that the annotations on dependent tiers are also involved in the merging process. The result is that the last annotation on a dependent tier that is part of the first annotation on the independent tier is stretched so that it ends where the second annotation on the independent tier begins. Figure 2.95, "Merge annotations" illustrates this.
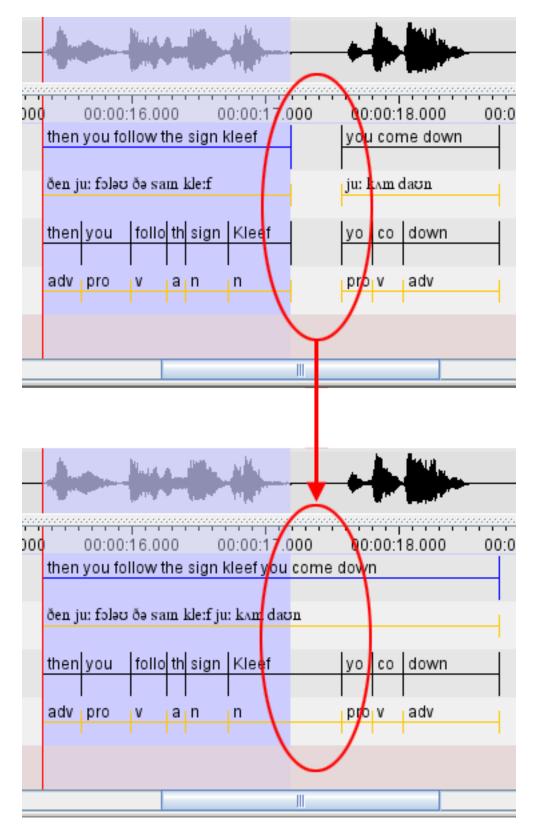
Figure 2.95. Merge annotations

It is also possible to merge with the annotation before the selected annotation. It works in the same way as "Merge with next annotation. To do so select Merge with Annotation Before instead of "Merge with next annotation" in the above mentioned options.

Entering annotations in different character sets

ELAN supports different character sets, making use of the following input methods:

1. GATE Unicode Kit (GUK)

   Characters can be entered by using a different keyboard mapping. This method is implemented using the GATE Unicode Kit developed at Sheffield University, Department of Computer Science.

   If you select a character set that is based on GUK, a visual representation of a keyboard appears on the screen, which informs you about the implemented keyboard mapping. The following illustrations show the mappings of "ipa-96 (SAMPA)" and "Arabic (WINDOWS)":

   Figure 2.96. ipa-96 keyboard map

   Figure 2.97. Arabic keyboard map

   To enter a character, do one the following:

   a. Either press the corresponding key on your keyboard.
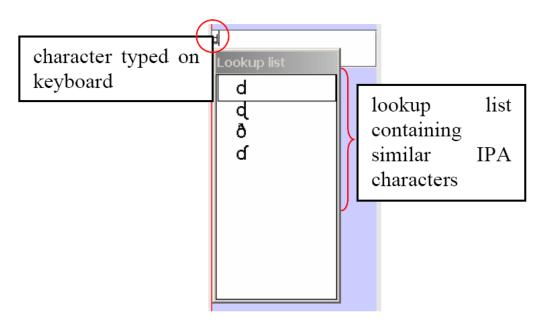
   b. Or click on the corresponding key of the keyboard displayed on your screen.
   The visual representation has the layout of the standard UK keyboard. If you do not have a UK keyboard, there may be mismatches between the characters and their visual representation.

For example, the IPA character "ə;" is matched to the key "@", i.e., in order to get "ə", you have to type "@". On a standard UK keyboard, the key "@" is located to the left of the key "enter" (see the illustration above). On other keyboards, however, "@" may be located on a different key. In such cases, if you press the key to the left of "enter", you will *not* get the character "ə". To get "ə", you have to search for the location of "@" on your keyboard, and then press that key. (Note that these mismatches only arise if you use the physical keyboard, but not if you use the visual representation on the screen.)

The character set "ipa-96 (SAMPA)" can be used to enter IPA characters. However, the current version of ELAN only supports SAMPA, but not X-SAMPA. As a consequence, some of the characters that you require may not be available yet (see http://www.phon.ucl.ac.uk/home/sampa for further information).

2. Roman Typographic Root (RTR)

   IPA characters can be entered using the RTR input method. This method is based on the following principle: whenever you type a character, all typographically similar characters are displayed in a lookup list, as shown in the following illustration:

   Figure 2.98. Roman typographic root

   To select a character from the lookup list, do the following:

   a. Use the UP and DOWN arrow keys to navigate to the desired character.

   b. Press ENTER or SPACE to confirm the selection.

📄 Note

   Do *not* use the mouse within the lookup window. If you do, the window will disappear.
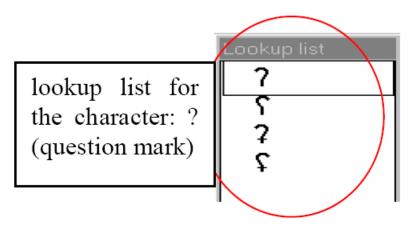
   The input of IPA characters is restricted to the official IPA-96 character set. Withdrawn or superseded characters are not supported.

   The RTR-mapping is generally obvious, but please pay attention to the following features:

   • Use lower case letters only.

- IPA characters can map to non-letter symbols, e.g.:

Figure 2.99. Non-letter symbols
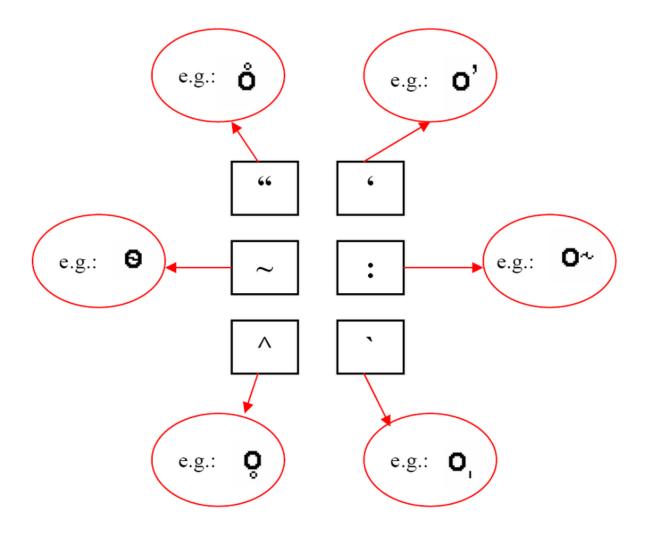
- Diacritics are entered with the help of the following keys:

Table 2.4. Diacritics

key	position of diacritic in relation to character
"	(quotation mark) above
'	(apostrophe) in the upper right
:	(colon) to the right
`	(grave accent) in the lower right
^	(circumflex accent) under
'	(tilde) on

For example:

Figure 2.100. Diacritics

> 📑 **Note**
>
> On Windows 2000, if you use an international keyboard, you have to type SPACE after typing the quotation mark (") or one of the accents (', `, ^).
>
> It is possible to stack diacritics, but there may be problems displaying them.

3. Chinese characters

Chinese characters, both traditional and simplified, are entered using the Pinyin method. Characters are selected by starting to type Roman characters. Candidates are shown in a lookup window while the user types along. The desired character is selected with the UP and DOWN arrow keys, e.g.:

Figure 2.101. Chinese characters

Use the following procedure:

- Enter the pinyin word with the keyboard. For each pinyin word, a list of Han symbols is shown in a popup window.

- Navigate to a Han symbol with the UP and DOWN arrow keys.

- Page through the list with the PAGE UP and PAGE DOWN keys.

- Select a Han symbol with the SPACE or the ENTER key.

📑 **Note**

Do *not* use the mouse within the lookup window. If you do, the window will disappear.

On Windows 98, you cannot display both Chinese and IPA characters.

4. Keyboard tools

If the options above don't fulfil your needs (e.g. the character set is not supported or you don't want to use the on-screen display for a large amount of annotations) you might want to look for a third-party solution. Such a tool provides a mean to remap your keyboard to the desired input character set. For details, we refer to the following programs:

Keyman (Windows), http://www.tavultesoft.com/

Figure 2.102. Keyman

Ukelele (Mac), http://scripts.sil.org/ukelele

Figure 2.103. Ukelele

📄 **Note**

> If you are using a third-party keyboarding solution like Keyman, make sure to select the default system language as input language for the tier to be edited (e.g. Dutch if your system language is set to Dutch).

Associating an annotation with a Data Category

📄 **Note**

> This functionality is currently out of order, please see the note at the section called "Data categories"

In ELAN it is possible to associate an individual annotation with a data category of the ISO Data Category Repository (DCR). To do so first select an annotation. Then click Annotation in the main menu of ELAN and select Modify Annotation Data Category (alternatively press **ALT+SHIFT+M**). In the next window (Figure 2.104, "Local Data Category Selection: Annotation"), select a data category from the list and click on Apply to associate the selected data category to the annotation.
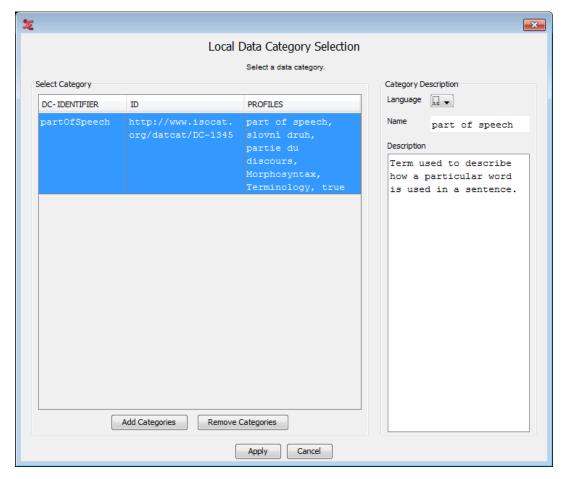
Figure 2.104. Local Data Category Selection: Annotation

More information about the ISO DCR and how to use it can be found in the section called "Data categories".

Associating an annotation with an external resource

An annotation can be linked to an external resource (e.g. a video, an image, a pdf file etc.) identified by a remote URL or a local file path. The external resource can e.g. be an illustration, example or definition of the contents of the annotation.

The Timeline viewer adds an additional menu item to the context menu of the active annotation, in case a reference to an external resource is detected. If it is an online resource, the menu's label is Show in Browser, if it is a local resource it is Open External Resource in Default Application (the system's or user's default application for that file type).

Figure 2.105. Open external resource context menu item

To create or modify the link from an annotation to an external resource, click Annotation > Modify Annotation External Resource URL.... In the small window that appears, a local path or remote URL can be pasted or otherwise entered.

Figure 2.106. Open external resource context menu item

> **Note**
>
> To provide compatibility with existing practice, a URL as the value of an annotation is treated, as much as possible, in the same way as linked external resources. The same applies to a URL in a CV entry's description or value field.

> **Note**
>
> A **CTRL+Click** (or **CMD+Click**) on an annotation with an external link, may work as a shorcut to the **Show/Open** menu actions (this is system dependent).

> **Note**
>
> Hovering the mouse pointer over the active annotation (in the Timeline viewer) while holding down the **CTRL** key, may show the inline media displayer (see the section called "Media Displayer") if the referenced resource is a video.

Spell checking

In the annotation edit box, either attached or detached, and in the Text tab, you can let Elan spell check the text. To do so, first make sure the tiers that you want to spell check have its content language set to the same language as the spell checking service (see the section called "Changing tier attributes") and then set up a spell checking service.

Setting up a spell checking service

To set up a spell checking service, go to Edit > Edit Spell Checking Services... (see figure below). In the dialog click Add. Choose the Type and the language for which this spell checking service is set up. Each spell checker type requires its own specific information.

Figure 2.107. Spell Checking menu

The type Hunspell (see also http://hunspell.github.io/) requires only a path to a dictionary file and its corresponding affix file. Click on Browse and browse to and select a Hunspell dictionary file (with extension .dic, download from https://cgit.freedesktop.org/libreoffice/dictionaries/tree/). Make sure that in the directory of the dictionary the corresponding affix file (with extension .aff) is also there. Also make sure that the files are saved as plain text and not embedded in HTML (this can be verified in a text editor), by following the `plain` links. After selecting the dictionary file click Open. Then click OK.

Figure 2.108. Hunspell setup

The type Gecco is a web service for spell checkers developed at Radboud University Nijmegen. Currently only one implementation is available, Valkuil for Dutch (see http://valkuil.net/info). Select Dutch as the language and click OK.

Figure 2.109. Gecco setup

The new spell checking service is identified in the dropdown menu by a the language identification, the type and a service URL or a path. When selecting a service a more elaborate description is displayed in the text box. Click Close to finish.

Spell checking

In the timeline viewer open an annotation editor, either attached or detached, on a tier that has its content language set to the language of the spell checking service you have set up. On opening and on editing the text will be spell checked. Spell errors are indicated by a red wavy line under the incorrect word (see figure annotation_editor.png).

In the Text tab, when selecting a tier that has its content language set to the language of the spell checking service you have set up, the annotation text is spell checked. Again, spell errors are indicated by a red wavy line under the incorrect word.

Figure 2.110. Spell checking

In the Transcription mode, when you click on an annotation (cell) within a tier that has its content language aligned with your configured spell-checking service, the annotation text undergoes spell checking. Any spelling errors are highlighted with a red wavy line beneath the incorrect word.

Figure 2.111. Spell checking in transcription mode

How to undo an action

Most actions in ELAN can be undone via the Edit > Undo (or **CTRL+Z**) menu. A redo (go back to the state before the undo action) is available to, via Edit > Redo (or **CTRL+Y**) .

The following commands support undo/redo:

- Add new tier / tier type

- Copy tier / change parent tier

- Insert annotation before / after

- Change tier attributes / tier type

- Clear selection (through the icon or keyboard shortcut)

- Delete annotation / tier / tier type

- Modify annotation / annotation time

- Create new annotation

- Copy / Paste / Duplicate annotation

- Play around selection time setting

- The tokenizer

- Add / Change / Edit controlled vocabularies

Tokenize a tier

One of the main advantages of using a hierarchical structured tiers is the possibility to split the content of an annotation unit on a parent tier automatically into smaller parts on another tier. This can be a child tier, but also an independent tier, outside of a hierarchy. E.g. the words on an utterance tier could be split into separate words. This is called tokenizing in ELAN. These steps will guide you through this process:

1. Go to Tier > Tokenize tier

2. Select a Source (= parent) and Destination tier

3. Optionally create a new destination tier by selecting Create New Tier…

4. Select a delimiter. The default is a space, but other choices are possible (e.g. "-" for morpheme breaks).

5. If you would like punctuation symbols tokenized as well, please specify them in the field next to 'Punctuation tokens'.

6. If the destination tier already contains annotation units, choose between overwriting or preserving them. If its still empty you can ignore this option.

7. Select Create destination annotation for empty source annotation if you want to create for every source annotation a destination annotation, even if its empty.

Figure 2.112. Tokenize tier
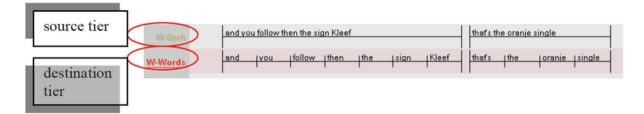
8. Click on Start, the tokenization will begin:

Figure 2.113. Tokenization

9. When it is finished, you will see that every annotation unit from the source tier has been tokenized on the destination tier:

Figure 2.114. Source and destination tier

All tokens (words in this example) on the destination tier have the same size (i.e. duration), even when tokenizing to a tier from the type time subdivision. You can adjust their length, as described in the section called "Changing the boundaries of an existing selection and annotation".

For symbolic associations, there is no need to use the tokenizer. Instead, go to the grid viewer and make sure the checkbox next to the dropdown menu is selected. Now you can fill in the annotations of the symbolic associations in their column of the grid. See also the section called "The Grid Viewer". If you want to copy or filter the contents from 1 tier to another symbolic associated tier, have a look at the section called "How to filter and copy tiers".

Tabs and newline characters are always treated as delimiters.

How to filter and copy tiers

Filtering tiers

Filtering a tier (Tier > Filter Tier…) works in a very similar way as the previously described tokenizer. The main difference is that filtering is meant for transfer of annotation information between tiers that are symbolic associations.

A step-by-step approach to filtering:

1. Select a source tier (from which the information will be copied)

2. Choose a destination tier. If necessary create a new tier (with the Create new tier… button)

3. Optionally specify a filter. If a filter expression is found, it will be removed from the destination annotation. Without any filter, the complete source tier is copied to the destination tier.

4. If the destination tier already contains annotation units, choose between overwriting or preserving them. If its still empty you can ignore this option.

5. Select Create destination annotation for empty source annotation if you want to create for every source annotation a destination annotation, even if its empty.

6. Click on Start to begin the filter operation or Close to go back to ELAN's main screen.

Figure 2.115. Filter tier

Copying tiers

As easy as you can change the parent of a tier, it also is possible to copy the complete content of a tier. This option differs in that it does not put any constraints on the designated parent tier and it leaves you the choice whether you want to copy the tiers children.

This process can be started via the Tier > Copy tier menu. Follow the steps below:

1. Choose a tier to copy. If you also want to create a copy its dependent tiers, check the Copy dependent tiers as well box.

2. Specify the parent tier for the copy. To make it independent, select Transcription (no parent)

3. By default, the tier type will be kept. If you want to change it, select another one from the dialog window and click on Finish.

4. Now the tier (and optionally its children) will be copied. "-cp" will be added to the names in order to prevent confusion with the original tier. It is also possible to check the option to rename the original tiers, so that the copies can use the original names. This way, "-orig" will be added to the original tier(s).

Figure 2.116. Copy tier

Note that this is similar to the change parent tier functionality (see the section called "Changing the parent of a tier"). However it differs in 2 aspects:

- It is not mandatory to copy the child tiers.

- The parent for the copy can be any tier in the transcription, including the tier itself, or no parent.

Copying annotations from one tier to another tier

This variant copies annotations from one tier to another, already existing, tier (so the tier itself is not copied).

This process consists of three steps and can be started via the Tier > Copy Annotations from Tier to Tier menu.

1. In the first step select one source tier from which to copy the annotations.

2. In the second step one destination tier can be selected. Listed are dependent tiers of the source tier and all independent, top-level tiers.

3. The third step allows to specify which annotations to copy and whether or not existing annotations on the destination tier can be overwritten.

Figure 2.117. Copy annotations from one to another tier

The following options are provided:

- All annotations copies all annotations regardless of their value.

- Annotations where value is..., the value can be entered in the text field.

- Treat as regular expression. This option can be checked if the value entered in the text field should be used as a regular expression to match annotations on the source tier. By default annotations are matched using case-sensitive, exact matching. The matching always concerns the entire value (no substring matching).

- Allow existing annotations to be overwritten allows to protect existing annotations from being changed. This concerns both the value and the alignment.

How to modify the boundaries of all annotations of selected tiers

It can occur that you want to modify the start and/or end times of all annotations of one or more tiers with the same amount of milliseconds. If e.g. the start of annotations is such that playing the interval of an annotation often misses the very first part of the audio of the utterance, it is best to move the start time of each annotation a bit to the left. Such annotations could e.g. have been created by a recognizer which made the boundaries of segments a bit too "tight".

This function can be started via Tier>Modify Boundaries of All Annotations... and involves two steps.

1. File and Tier Selection

   If an annotation document is open, the single file option Use currently opened file is active. In the multiple file situation (see the section called "Modify boundaries of all annotations" Select files from browser and Select files from domain are active. When tiers are loaded in the table, select one or more tiers to process and click Next.

2. Specification of start and/or end time modification

   There are two main options in this step:

   • Enter left and/or right boundary shift in milliseconds: the two textfields allow to enter the amount of modification of the begin and end time, in number of milliseconds. A negative value, e.g. -200, moves the boundary to the left, a positive value to the right. If the entered values are not the same, the duration of the annotations will change as well as a result of the modification. (If the values are the same, it might be better to use the Shift Annotations option, in the single file case).

   • Overwrite options: here it is possible to specify whether the modification of an annotation may or may not change the boundaries of the neighbouring annotations. If e.g. the Overwrite annotations to the left checkbox is selected, a change of the begin time of one annotation (with a negative value, moving the begin boundary to the left) is allowed to change the end time of the annotation left of it or before it. If the checkbox is not selected, a boundary is only allowed to change as much as there is "empty space" before it.

   Annotations are processed from left-to-right and, depending on the selected options, the processing of one annotation may alter the next annotation before it is processed itself. For predictable results, in some cases it is best to perform the modifications in two rounds, once for the begin times, once for the end times.

   Figure 2.118. Modify boundaries of all annotations of selected tiers

   Click Finish to apply the modifications.

How to detect annotation overlaps

In some cases you might want to have more information about annotation overlaps and their respective duration. There are two methods to detect overlaps. The first creates annotations from overlaps (see the section called "Creating annotations from overlaps") and the second specifically aimed at comparing annotators (see the section called "Comparing annotators").

Creating annotations from overlaps

Creating annotations from overlaps uses a wizard that has been developed to automatically detect annotation overlaps and optionally adds the length of the overlaps to the newly created annotations. This function can be started from Tier > Create Annotations from Overlaps…. This will open the *Create Annotations From Overlaps* window that is based on 4 steps.

1. Step 1/4: File and Tier Selection

   Select the tiers to use in the overlaps computation. You can select all the tiers displayed in the list if you click on Select All, or deselect them if you click on Select None. Once you have made your choice for the tiers for which the overlaps should be found, you can select next, this will bring you to the next step.

   Figure 2.119. Create annotations from overlaps

> **📄 Note**
>
> a. At least two tiers has to be selected in order to reach the second step.
>
> b. To check the option Select files from file browser or Select files from domain, see the section called "Annotations from overlaps"

2. Step 2/4: Overlaps Computation Criteria

In this step you can define the overlaps computation criteria in the *Overlaps Computation Criteria* window

Figure 2.120. Overlaps Computation Criteria

You can choose within 4:

• Regardless of their annotation values. If this option is selected, all the possible overlaps will be computed.

• And their annotation values are equal. If this option is selected, only the overlaps with the very same annotation values will be computed.

- And their annotation values are different. If this option is selected, only the overlaps with different annotation values will be computed.

- According to specified constraints. Select this option and click on Constraints... the *Annotation Value Constraints* dialogue window will be displayed:

Figure 2.121. Annotation Value Constraint

Here you can specify for which tiers you want the constraints to be applied by selecting the tiers from the drop down list and inserting the value they should contains. Then click Add: the constraints table will display the tiers you have selected together with the value they have to contain. Once you have made your selections, you can click OK to go back to the *Overlaps Computation Criteria* window and click Next to go to the third step.

3. Step 3/4: Destination Tier Name Specification

In order to create the annotation from overlaps, you have to define its destination tier. You first enter the name for the destination tier, and then select if you want it to be either a root tier, or a child of a parent tier. In the former case, you can select the Tier Type Name and Stereotype for the destination tier from the table. In the latter case, specify from the drop down list which tier you want to be the parent tier. In both cases, if there is no correct tier type available, you can create a new one which matches the destination tier by clicking the Add new type... button. You can now go to the next and last step, by selecting Next.

Figure 2.122. Destination Tier Name Specification

4. Step 4/4: Destination Tier Value Specification

   Here you can specify the value for the destination tier.

Figure 2.123. Destination Tier Value Specification

You have four options:

- `Specify the value for the destination tier`. If selected, created annotations will be filled with the overlap duration. You can choose one of the following time format:

  a. Msec

  b. ss.msec

  c. hh:mm:ss.ms

  d. SMPTE Time code hh:mm:ss:ff (either PAL or NTSC drop frame)

- `A specific value` If selected, you can enter the value for all annotations that are created.

- `Value from a specific tier`. Here you can specify the tier (you can select it from the drop down list) whose annotation values will be used for the created annotations.

- `Concatenate the values of overlapping annotations.` If selected, created annotations will be filled with the concatenated values of the overlapping annotations

  - `Compute values by annotation time` If selected, the values of the annotations are concatenated based on the begintime of the annotation.

  - `Compute values from the tier in the selected order` If selected, the values of the annotations are concatenated based on the selected tier order. The tiers can be moved up and down within the list using the buttons below the list.

Finally, you can click on Finish. The new tier will be created and populated.

Comparing annotators

The Tier > Compare Annotators... function is no longer there. A simplified version of this function is now one of the options in the the section called "Calculate inter-annotator reliability" function (the "ratio of overlap and total extent" option).

Create annotation by subtraction

New annotations are created by subtracting annotations from annotations on other tiers. This function can be started from Tier > Create Annotations from Subtraction…. This will open the Create Annotations From Subtraction window that is based on 4 steps.

1. Step 1/4: File and Tier Selection (see Figure 2.119, "Create annotations from overlaps")

2. Step 2/4: Subtract Computation Criteria

   In this step you can choose the subtraction computation criteria in the *Subtract computation criteria* window

Figure 2.124. Subtract Computation Criteria

You can choose one of the following:

- `Subtraction based on 'exclusive or' logic` If this option is selected, the subtracts will be computed based on the 'Exclusive-or" logic.

- `Subtraction` Selecting this option, will allow you to select a reference tier from which the annotations of the other selected tiers will be subtracted from.

Click on the icon to know more about the differences between the above mentioned options.

3. Step 3/4: Destination Tier Name Specification (see Figure 2.122, "Destination Tier Name Specification")

4. Step 4/4: Destination Tier Value Specification

Here you can specify the value for the destination tier.

Figure 2.125. Destination Tier Value Specification

For detailed explanation on the options see Figure 2.123, "Destination Tier Value Specification".

Finally, you can click on Finish. The new tier will be created and populated.

Copy and paste annotations

Copy annotation

Just like in many other applications, you can reuse elements you are editing with the well-known copy and paste operations. To copy an annotation do the following:

1. Select the annotation

2. Right click on it and select Copy Annotation, select the same option in the Annotation drop-down menu or press **CTRL+C**

It is possible to customize what is placed to the clipboard when copying an annotation in the the section called "Editing preferences" window (`Editing` panel, bullet point 9). This is especially relevant when pasting annotations in other applications than ELAN. It is e.g. possible to copy the annotation value only or to copy a citation form that includes the media file name and tier name etc.

Paste annotation

In a similar way you can paste these annotations from the clipboard into ELAN:

1. Go to the tier where you want to paste the annotation

2. Select Annotation > Paste Annotation or press **CTRL+V**

📝 Note

Annotations can only be pasted onto tiers of the same tier type! If you want to copy the annotation to a tier of a different type, use *Duplicate Annotation* (see the section called "Duplicate annotation").

Please note that pasting an annotation can result in different behavior according to the context:

• By default the annotation is pasted onto the tier where it originates from.

• If that is impossible (i.e. there is no tier with the same name as the originating tier, e.g. in another file) the copy of the annotation will be placed on the active tier (see the section called "Make a tier the active tier" on how to activate a tier)

However, pasting an annotation will never change the time alignment of that annotation unit. This means that the annotation will be placed on exactly the same time as it was found when the copy operation was performed. If you want to change its timing, move it afterwards or use the *Paste annotation here* option (see the section called "Paste annotation here").

Paste annotation here

Using this function (Annotation > Paste annotation here or press **CTRL+SHIFT+V**) you can copy an existing annotation to any tier and time you want. (Of course the tier still has to belong to the same tier type). The copy will be placed on the position of the crosshair on the active tier.

Right click somewhere in the timeline viewer and select Paste annotation here from the context menu to copy the annotation to the position of the mouse cursor – both the tier and the time position.

Copy and paste annotation groups

When copying and pasting annotations, only the selected annotation will be copied and pasted. If you want the annotation including all its child annotations (those on the dependent tiers) to be transferred, select

• Copy Annotation Group, accessed via the main menu item Annotation or in the context menu (opened by right clicking), and

• Paste Annotation Group, via the main menu item Annotation or Paste Annotation Group here via the main menu or the context menu.

Pasting all annotations in a copied annotation group only works if the tier structure (the dependent tiers and their tier types) of the source is the same as the tier structure of the destination. Or alternatively, the source and destination should follow the same naming convention. An example in which the destination follows the same naming convention is shown in

Figure 2.126. Copying a annotation group

The annotations from the tiers labelled W-RGU, W-RGph and W-RGMe were copied by Copy Annotation Group. Then they were pasted on the tier labelled M-RGU. Since that tier and the tiers labelled M-RGph and M-RGMe follow the same naming convention as the source tiers (the suffixes match) the annotation group could be pasted on the three target tiers M-RGU, M-RGph and M-RGMe. Note that the tier structure of the target tiers is different from that of the source tiers.

Duplicate annotation

Select an annotation and make the tier where you want it to be duplicated the active tier. Now select Annotation > Duplicate Annotation or press **CTRL+D** to create an exact copy of the annotation onto the active tier. This function also works for tiers that don't share the tier type.

Copy all annotations from one tier to another.

The process of copying multiple annotations from one tier to another is described in the the section called "Copying annotations from one tier to another tier" section.

Synopsis

Figure 2.127. Synopsis copy and paste annotations

Web Services

ELAN offers several ways to interact with web services. These web services are tools or applications that run on a web server, accept some resource(s) as input, apply an algorithm and return the result as output. Some of the Recognizers in the the section called "Audio And Video Recognizers" tab are available as a web service, in which case audio or video or any other type of sequential data are uploaded to be processed online. The web services available in this menu work with text rather than multimedia content. These tools can automate certain parts of the annotation process, such as tokenizing and part of speech tagging. These web services can be found by clicking Options > Web Services . The next section describes the web service WebLicht.

WebLicht

WebLicht (Web-Based Linguistic Chaining Tool [https://weblicht.sfs.uni-tuebingen.de/weblichtwiki/index.php/Main_Page]) is an execution environment developed at Tuebingen University as part of the CLARIN infrastructure. Most of the tools in this environment perform NLP (Natural Language Processing) type of tasks on textual data and most of them are tailored to work with language data in one of the well-described and well-resourced languages. The tools are encapsulated as web services and can be combined into processing chains.

To make use of the WebLicht service, go to Options > Web Services > Weblicht. In the dialog that opens, you can specify how to use WebLicht

- by selecting the type of input, either plain text or one or more selected tiers

- and by selecting the type of service, either a preconfigured tool chain (recommended) or a single tool service (deprecated)
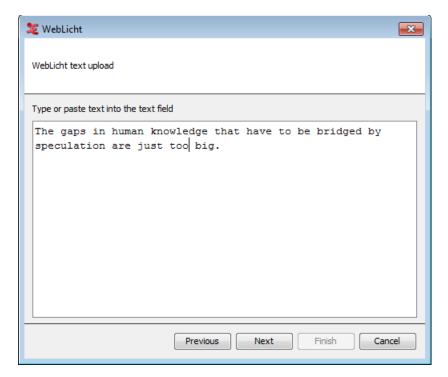
Figure 2.128. WebLicht service: upload plain text or the contents of a tier

Then click the Next button.

The interface of the second step depends on the choice between plain text and tier(s) in the first step.

- In case of plain text a text area will be visible in which you can paste or type plain text.

Figure 2.129. WebLicht service: text input

> WebLicht
>
> WebLicht text upload
>
> Type or paste text into the text field
>
> The gaps in human knowledge that have to be bridged by speculation are just too big.
>
> Previous | Next | Finish | Cancel

- In case of tier selection a table will be shown in which one tier can be selected. Below the table, specify the content type (`Sentence` if the annotations on the selected tier contain sentences or `Word/Token` if the annotations contain single words). There are some limitations on the tiers you can select for each type; `Sentence` tiers are expected to be a toplevel tier or a symbolically associated dependent tier thereof, `Token` tiers are expected to be on a symbolic subdivision tier.

Figure 2.130. WebLicht service: select a tier to upload

> WebLicht
>
> Select a tier and a content type
>
> Select the tier to upload

Tier Name	Parent Tier	Tier Type	Participant	Annotator	Input Met...	Content ...
K-Spch	-	utterance			English (Un...	-
W-Spch	-	utterance			English (Un...	-
W-Words	W-Spch	words			English (Un...	-
W-POS	W-Words	part_of_sp...			English (Un...	-
W-IPA	W-Spch	phonetic_t...			English (Un...	-
W-RGU	-	gestures			English (Un...	-
W-RGph	W-RGU	gesture_p...			English (Un...	-
W-RGMe	W-RGph	gesture_m...			English (Un...	-
K-RGU	-	gestures			English (Un...	-
K-RGph	K-RGU	gesture_p...			English (Un...	-
K-RGMe	K-RGph	gesture_m...			English (Un...	-
multisubtest	-	utterance			English (Un...	-

> Specify the content type of the tier to upload [Sentence ▼]
>
> Previous | Next | Finish | Cancel

Then click Next.

The third step depends on the choice between a tool chain and a single tool in the first step.

- In case of a tool chain (see section the section called "WebLicht with tool chain file"), the stored chain file can be specified here. Click the Select button and browse to the saved file. It is necessary to specify the type of content the first tool in the chain expects, plain text or TCF. In the lower text field the access key can be pasted.

Figure 2.131. WebLicht service: select the tool chain file and paste the access key

- In case of a single tool, a list of available tool services will be shown. There is no guarantee a listed tool service (still) works this way (direct access without authentication).

For plain text in the first step, the list contains several services that detect sentence boundaries and then tokenize these sentences. Select one of the tokenize servicese. In case of successful processing the result will be two tiers, for sentences and tokens. If you want to add Part of Speech and/or Lemma annotations, you can use the tiers produced in this step as the input for such services (part of speech taggers) in a second run. There is an option to specify the duration (in ms) of each sentence

Figure 2.132. WebLicht service: select a tokenizer
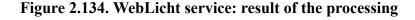
For tier input different services are available which can parse text, tag Parts of Speech, etc. Each service has a short description that specifies its function. Hovering over a service with the mouse will show a tooltip containing more information of the service. If the service you are looking for is not listed, you can manually specify its URL.

Figure 2.133. WebLicht service: select a service

After configuration with a tool chain file or a single tool, click Finish to start processing. When the processing was successful, you will see a dialog stating the operation is complete. Depending on the

service you selected for processing, the tokenized sentence and/or part of speech tags will be added as children of the tier you selected for processing.

Figure 2.134. WebLicht service: result of the processing

Parser used: Berkeley NLP Parser

WebLicht with tool chain file

The recommended way to use WebLicht from within ELAN is by supplying a preconfigured tool chain file. This file can be created as follows:

- go to the WebLicht web interface [https://weblicht.sfs.uni-tuebingen.de/weblichtwiki/index.php/Main_Page] and click the Start WebLicht button. You'll be prompted to log in to WebLicht via the CLARIN Service Provider Federation. If your home organisation is not in the list, you can request a CLARIN account.

- once logged in, create a tool chain, preferably using some of your own data, and run the tools to check the results

- click the Download button to save the tool chain configuration. The file will have a name like `chain_randomnumber_.xml`, it might be practical to rename it to something more intuitive, especially if you plan to have more than one tool chain

- make sure to make a note of the type of input of the first tool in the chain, it should be Plain Text or TCF

- to be able to run the tool chain from within ELAN, you'll need to generate a so-called **API Key**. ELAN will run the tool chain via WebLicht as a Service (WaaS) [https://weblicht.sfs.uni-tuebingen.de/WaaS/apikey], and that service requires an API key to authenticate the request. The API key can be generated on the WaaS home page and then stored locally. The key will be valid for 3 months, after that a new key will have to be generated.

- in the third step of running WebLicht in ELAN, all this information has to be provided: the tool chain file, the type of input and the API key (see Figure 2.131, "WebLicht service: select the tool chain file and paste the access key")

Information about annotations

Annotations Spreadsheet

A very basic view of the annotations in the independent tiers can be obtained by clicking on View > Annotations Spreadsheet. This option simply lists all annotations of all independent tiers ordered by time.

Annotations Statistics

ELAN offers a possibility to compute some statistics of the tiers. Click on View > Annotations Statistics and select a tier in the dialog window:

Figure 2.135. Annotations Statistics

If you wish the Annotations tab to show the statistics of a dependent tier, uncheck `Show only root tiers` and select it. Uncheck the next option if you want ELAN to count all contiguous annotations with the same value as 1.

The *observation period* is the interval between the beginning of the first annotation of all tiers and the end of the last annotation of all tiers. If you want ELAN to use the total media duration, just check `Use media duration as observation period`.

The Annotations tab shows the following information:

• Annotation

• Occurrences: the number of occurrences (contiguous annotations containing the same value count as only one occurrence if the relevant option is checked).

• Frequency: the frequency defined as the number of occurrences divided by the observation period.

• Average Duration: the average duration defined as the total duration of the annotations with the same value divided by the number of occurrences.

• Time Ratio: the time ratio defined as the total duration of the annotations containing the same value divided by the observation period.

• Latency: the latency defined as the time interval between the beginning of the observation period and the first occurrence of an annotation.

The Annotations II tab shows:

• Annotation

• Occurrences: the number of occurrences.

• Minimal Duration: the duration of the shortest (in time) annotation.

• Maximal Duration: the duration of the longest (in time) annotation.

- Average Duration: the average duration defined as the total duration of the annotations with the same value divided by the number of occurrences.

- Median Duration: the median duration defined as the duration that separates the lower half of the annotation durations from the higher half.

- Annotation Duration Percentage: the ratio between the total duration of the annotations and the total media duration expresses as a percentage.

- Latency: the latency defined as the time interval between the beginning of the observation period and the first occurrence of an annotation.

In the Annotations II tab, contiguous annotations with the same value are *not* counted as 1. The observation period is the same as the media duration.

The Tiers tab shows:

Figure 2.136. Tiers statistics

- Tier: the name of a tier.

- Number of Annotations: the number of annotation on that tier.

- Minimal Duration: the minimal duration of an annotation on that tier.

- Maximal Duration: the maximal duration of an annotation on that tier.

- Average Duration: the average duration of the annotations on that tier.

- Median Duration: the median duration of the annotations on that tier.

- Total Annotation Duration: the total duration of all annotations on that tier.

- Annotation Duration Percentage: the percentage of the total annotation duration of the media duration on that tier.

- Latency: the time interval between the beginning of the observation period and the first annotation on that tier.

The *observation period* on the Tier tab is the interval between the begin and the end of the media.

The statistics on the tabs Tier Type, Participant, Annotator, Language are similar to the statistics on the Tiers tab except that they apply to tier type, participant, annotator and content language respectively.

The information in the columns displayed on every tab can be sorted in different ways. By clicking the header of the preferred column, the information below gets sorted. This is done in an ascending order first, and subsequently toggling between ascending and descending order. Saving the statistics will also output the selected sort order.

Figure 2.137. Sort order statistics

Annotation Density Plot

An annotation density plot visualizes the distribution of annotations (the segmentation) over the media duration or over a selected time interval. The textual annotation values are ignored.

Figure 2.138. An annotation density plot

There are a number of options to customize the visualization and the result can be saved as an image. Apart from the usual tier selection options the following settings are available:

- Limit to current selection the plot will be restricted to the current time selection

- Tier column width the width in pixels available for the tier names. Set to 0 to switch off tier names.

- Image width the width of the image in pixels

- Image height the height of the image in pixels

- Tier height the height of a tier layer in pixels

- Tier height - Fill when this box is checked the tier height is calculated such that the entire height of the image is occupied by the selected tier layers. A minimum height is applied, a message will be shown if the tiers don't fit.

- Margin height the white space above and below each colored band in pixels

- Include outlines when checked, grid lines will be drawn to mark the boundaries between tiers

After making changes click the Update button to see the new settings in effect. Click the Export button to show a Save As window and save the result as a PNG image.

Figure 2.139. The effect of unticking the Fill checkbox

Chapter 3. Working modes

Different working modes are available, some of which are optimized for a specific task. The modes are accessible via the **Options** menu and are described below.

Annotation mode

The Annotation mode is the generic (and default) mode for working with annotations in several ways. The window in this mode offers most viewing, editing and searching options that are described in this manual (Chapter 1, *ELAN documents*, Chapter 2, *Annotations* and Chapter 4, *Searching*).

Synchronization mode

The Synchronization mode is described in the section called "Synchronizing video files", where the emphasis is on the synchronization of video files. The synchronisation mode can also be used for synchronising audio files and timeseries files. For each audio file a simplified waveform view is added that allows for dragging and playing a selection. The waveform is active and responsive when the associated wave player is activated. There will be at most one timeseries player that can be used for synchronizing multiple timeseries files.
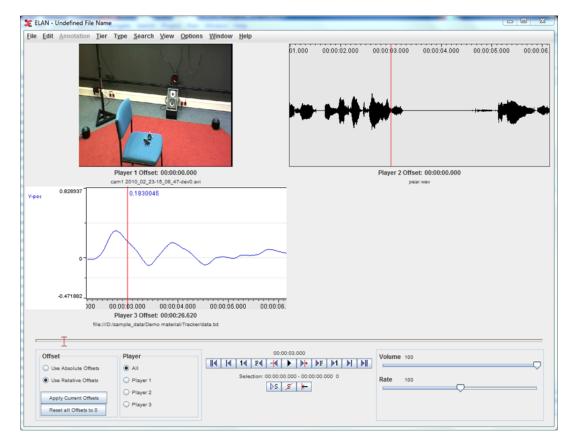
Figure 3.1. Synchronization of video, audio and timeseries files

Transcription mode

Transcription Mode is a mode designed to increase the speed and efficiency of transcription work. The interface is keyboard-driven and minimizes UI actions. All annotations of a certain tier type are

displayed in a vertical list for easy visual access. Transcription mode brings down the transcription work to the bare essentials: listen, type, listen, type, listen, type. To go open transcription mode, select Options > Transcription mode . If you go to Transcription mode for the first time, a Settings dialog will come up (see Figure 3.2, "Settings Dialog") else transcription window is opened with the last used settings (see Figure 3.10, "Transcription Window").

📒 Note

Transcription mode presupposes that the initial segmentation of the recording is already done. The rationale for this is that the most efficient work flow for transcribing large amounts of linguistic data is a two-step process: first segmenting the recording into turns —also attributing turns to the appropriate speaker— (this can be done in Annotation mode(see the section called "Annotation mode")or in the special purpose Segmentation mode(see the section called "Segmentation mode")), and then transcribing and translating these turns.

Selecting the tier types for transcription

If you go to Transcription mode for the first time, a Settings dialog will come up else click on the configure... button in the settings panel (see Figure 3.10, "Transcription Window"). Here you can select the tier types and the number of columns.

Figure 3.2. Settings Dialog

Options in this dialog :

• Font size : Specifies the font size for the table in the Transcription window.

• Number of columns : Specifies the number of columns table in the Transcription window Use the + to add a column and - to remove the last column.

• Settings Table

    • Column : Specifies the column index in the transcription table
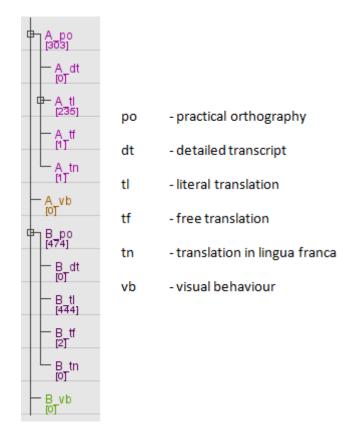
- Select type for column : Specifies the selected tier type for that column. To select a type click on the <select a type> cell in the table to get the list for available types for that column. Click on the type to select it.

- Select tiers... button see the section called "Selecting tiers for transcription".

📄 **Note**

You select *tier types*, not individual tiers. This is because Transcription mode displays all annotations on all tiers of a certain type in a vertical column.

For the purposes of this description we will assume that the user is working with a file that has six main tier types:po (practical orthography), dt (detailed transcript), tl (literal translation), tf (free translation), tn (translation in lingua franca) and vb (visible behaviour). Our example file contains tiers of these types for two participants, and the overall tier structure looks like this :

Figure 3.3. Tier Structure

In our example, we choose the type po (practical orthography) as the first column. We can leave it at that if we just want to work on the transcript. Or we can display any number of columns next to the primary one, the number of columns depends on the available linked tier types — for instance the free translations and/or a literal translations and/or detailed transcript and/or translations in lingua franca.

For the other columns we can only select tier types that are time-aligned with the first using the stereotype "Symbolic Association" (see Table 2.2, "Tier type stereotypes"). In our example, we can have maximum four columns with tier types dt (detailed transcript), tl (literal translation), tf (free translation) and tn (translation in lingua franca). We *cannot* choose the tier type vb (visible behaviour)here, because it is not time-aligned with our primary column. Thus the primary column type can be of any type, but the types for other columns should be somehow symbolic associated.

Having selected the tier types , click "Apply". Now the chosen tier types are displayed in vertical columns (see Figure 3.10, "Transcription Window"), and the two largest differences from the default Annotation mode become visible: (i) all annotations are displayed vertically (top to bottom) rather than horizontally (left to right), and (ii) columns display all annotations of a certain type. For instance, the po (practical orthography) column displays turns from both speakers A and B.

Note
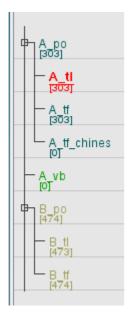
> Transcription mode presupposes that you use tier types to differentiate the types of information in your tiers. Thus the tier type of your free translation tier should be different from the tier type of your main transcription.

Selecting tiers for transcription

Once you have selected the types for transcribing, it is possible to select the tiers if the transcription has more than one tier of a type which is selected for other columns. To select the tiers, click the select tiers... button in the Settings Dialog(see Figure 3.2, "Settings Dialog". This dialog comes up.

Figure 3.4. Select Tiers Dialog

Show / Hide ...	Top Level Ti...	Type 1 : po (...	Type 2 : dt (d...	Type 3 : tl (lit...	Type 4 : tf (fre...	Type 5 : tn (tr...
✔	A_po	A_po	A_dt	A_tl	▼	A_tn
✔	B_po	B_po	B_dt	B_tl	A_tf	B_tn
☐	C_po	C_po	C_dt	C_tl	A_tf_Chines	C_tn
☐	D_po	D_po	D_dt	D_tl	D_tf	D_tn
☐	E_po	E_po	E_dt	E_tl	E_tf	E_tn
☐	F_po	F_po	F_dt	F_tl	F_tf	F_tn
☐	G_po	G_po	G_dt	G_tl	G_tf	G_tn
☐	H_po	H_po	H_dt	H_tl	H_tf	H_tn

Apply   Cancel

In this tier structure, we have two tiers (A_tf and A_tf_chines) of the same type 'free translation'. So its possible to select which tier you want to edit from the "*Select Tiers Dialog*".

Using this dialog you can choose which tiers should be used on which column in the transcription table. This dialog is also used to show or hide tiers from the transcription table. To hide tiers, uncheck the box in the first column.

Show or Hide tiers

To hide a group of linked tiers, any one of the following options could be done.

1. **If tier names are shown** : right click on the tier name will popup this menu. Select hide all tiers linked with this tier from it.

Figure 3.5. Context Menu

2. **If tier names are not shown** : right click on the row no in the "No" column. This will popup the same menu as before. Select hide all ties linked with this tier from it.

To hide or show more number of tiers,

• select show / hide more tiers from the above popup menu. This will bring this dialog : the section called "Selecting tiers for transcription" . Check all the tiers to show up in the table and uncheck the tiers that are to be hidden.

• Another way to open the above dialog is to click Configure... button from the settings panel in the transcription window (see :Figure 3.10, "Transcription Window"). This will popup the transcription mode settings dialog (see : Figure 3.2, "Settings Dialog" and click on the Select tiers... button. This will bring the same dialog (see : the section called "Selecting tiers for transcription").

Change color of tiers

To change the color of any tiers, do the following steps.

1. **If tier names are shown** : right click on the tier name will popup this menu. Select change color for this tier.
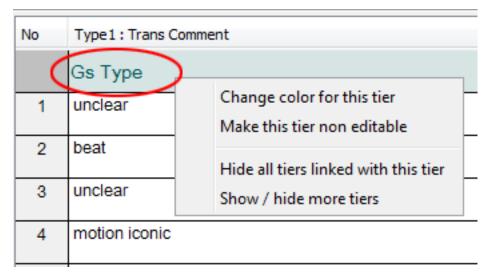
Figure 3.6. Context Menu

2. **If tier names are not shown** : right click on an annotation in the desired column. A similar context menu will be shown. Select change color for this tier from it.

From the context menu that shows up, you can set the tier background color, the tier highlight color and the font. To do so, click on one of the Browse buttons.

Figure 3.7. Context Menu Tier

The next context menu will let you select the color. Click on one of the tabs: 'Swatches', 'HSB' or 'RGB' and choose a color. A preview is shown on the bottom of the dialog. When done, click OK or, if you want to add the color to your favorites, go to the tab Favorites and click the 'add' button. The color will be added to your favorites.

Figure 3.8. Context Menu Color

The Tier Attributes context dialog will now show the chosen color next to 'Tier Color'. If you are satisfied, click Apply. Alternatively, you can also set the tier highlight color in a similar way.

In the context dialog that pops up, you can select what attributes must get the specified color. Also you can select what kind of tiers should get the specified colors. When done, click Apply and the changes will be made.

Figure 3.9. Context Menu Apply

Making tiers non-editable

Making a tier non-editable is to avoid putting that tier in edit mode while navigating through the table . For instance if some of the tiers of the same type are transcribed and few are not, to avoid the transcribed ones to get in to edit mode the tiers can be locked(made non-editable). To freeze/unfreeze the tiers, right click on the tier name (if shown) or right click on any of the locked or unlocked annotation of the tier will popup either this menu - see : Figure 3.5, "Context Menu " or this menu - see :Figure 3.12, "Context Menu for a Annotation". In the popup menu select make this tier non editable to freeze a tier or select make this tier editable to unfreeze the tier.

Entering text for a annotation

Transcription mode is built for high-speed transcription work. Once you have selected the tier types for the columns , the table is loaded with the tiers of the selected types. Clicking on Apply button in the settings dialog(see Figure 3.2, "Settings Dialog") will bring this window:

Figure 3.10. Transcription Window

Options available in this window :

- : Press this button to toggle the video/settings panel from left to right and vice versa. The video can also be detached for viewing independent of the main window. Do to this, right click on the video and click Detach.

- : Play/pause button for the media.

🗐 Note

The other media player options like, go to next of previous second/pixel/frame are only available as shortcut actions. To see the shortcuts used for this action, go to menu View -> Shortcuts... and select transcription mode.

- : Play/pause the selection made.

- : Clears the selection.

- Loop mode : If checked, plays the media of the selected annotation constantly in a loop until a new annotation is selected. Default: unchecked.

- Options in Settings Panel

  - Automatic playback of media : If checked, the annotation is automatically played when it enters into edit mode, else you can manually play the media using **TAB** or **SHIFT+TAB** (for details see the section called "Shortcut keys used in this mode"). Default: checked.

  - Create missing annotations : If checked, you will be able to create new annotations by double-clicking an empty column. If unchecked, you can only alter the existing annotations.

  - Show tier names :If checked, the tier names are shown in the table and hovering over an annotation will give you the time interval of that (see Figure 3.10, "Transcription Window"). If unchecked, colour coding distinguishes different tiers/participants, and hovering over an annotation will give you the tier/participant name (see Figure 3.11, "Transcription Window without showing the tier names"). Default: checked.

  - Colors only on the "No" column : If the show tier names is unchecked, you can choose whether the color distinction should be displayed all over the table or only in the "No" column. If checked, color difference made only in the "No" column, else the whole table has the color differentiation. Default: unchecked.
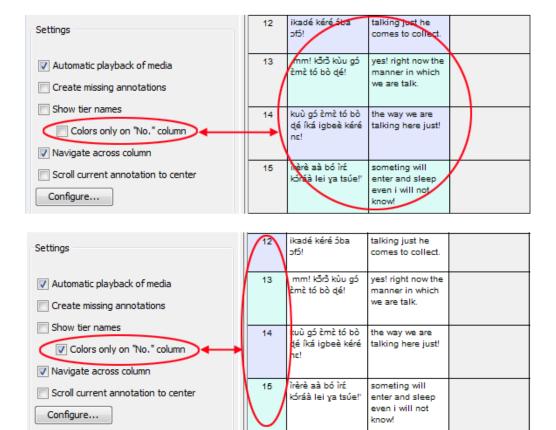
Figure 3.11. Transcription Window without showing the tier names

- Navigate across column : This controls the behaviour of the **ENTER** i.e. decides whether to move to the next annotation in the same column or to move across the columns. If checked, you move across columns (from left to right). If unchecked, you move only within a column (from top to bottom). Default: unchecked.

- Scroll current annotation to center : If checked, this mode keeps your current annotation always in the middle of the screen. Default: unchecked.

- Configure... : Press this button to get the settings dialog (see Figure 3.2, "Settings Dialog") to select new types or to change the font size or to show/hide tiers.

- ↕ : This is used to resize the video/settings panel and the transcription table. Click and drag to resize.

Each cell in the table is the representation of the annotation on a tier. The annotations are sorted/aligned based on the begin time of it. Clicking on any annotation(cell) activates it for editing. You can directly start typing the text for the annotation. After entering the text, press **ENTER**to save the changes made in the current annotation and to put the next annotation to edit mode. The Navigate across columns setting controls whether you go down within a column or you move across columns (from left to right). There are several ways to put a annotation to edit mode. A single mouse click on an annotation or a right click made on a annotation will put the annotation in edit mode and pops up the context menu below and for other keys see the section called "Shortcut keys used in this mode".

Other options available in this mode :

- right clicking on annotation will popup this menu.

Figure 3.12. Context Menu for a Annotation

- Commit Changes : Saves the changes made to the annotation

- Cancel Changes : Cancels the changes made to the annotation

- Edit in Annotation mode : see the section called "Modifying annotations"

- Merge with Annotation Before : see the section called "Merging annotations"

- Merge with Next Annotation : see the section called "Merging annotations"

- Delete Annotation : Deletes the current annotation from the tier

- make this tier non editable : Freezes the tier in the table thus making it non editable

- hide all tiers linked with this tier : Hides the group of tiers that are linked with the current editing tier

- show/hide more tiers : see the section called "Show or Hide tiers "

- All the columns are resizeable: just mouse click and drag the boundaries to fit your desired widths.

- You can re-order columns on the screen simply by dragging them to the desired location.

- The video can also be detached for viewing independent of the main window. Do to this, right click on the video and click Detach.

- The wav form can also be hidden if needed. You can do this by unchecking the Signal Viewer from View > Viewer > Signal Viewer. To show the wav form check the Signal Viewer option from the same menu.

Merging annotations

To merge two annotations, first make a right click on the annotation. This will pop up a context menu(see Figure 3.12, "Context Menu for a Annotation"). Then select either Merge with annotation before or Merge with next annotation . For detailed information see the section called "Merging annotations".

Modifying annotations

This mode only allows to edit the text of the annotation. In order to do some modifications on the annotation, i.e. to finely manipulate annotation boundaries or to split a annotation, right click on the annotation. This will popup a menu (see Figure 3.12, "Context Menu for a Annotation") . Select edit in annotation mode to jump to this annotation in the annotation mode which allows you do modify the annotations. For detailed information on changes the annotation boundaries see the section called "Changing the boundaries of an existing selection and annotation" and to split annotations see the section called "Split Annotations"

Shortcut keys used in this mode

The following shortcut keys are the mainly used default shortcuts in this mode.

- **ENTER** saves the current annotation, moves to the next annotation, and plays the new annotation if the Automatic playback option is selected. The Navigate across columns setting controls whether you go down within a column or you move across columns (from left to right).

- **ALT+UP** arrow moves up to the next cell in the same column and works same as **ENTER**.

- **ALT+DOWN** arrow moves down to the next cell in the same column and works same as **ENTER**.

- **ALT+LEFT** and **ALT+RIGHT** arrows move left and right across columns and also behaves like the **ENTER**.

- **TAB** plays the current annotation. It acts as a play/pause key, so press it again to pause playback, and press again to continue playing. If a selection is made in the waveform or in the timeline, then **TAB** will play/pause the selected interval.

- **SHIFT+TAB** plays back the current annotation from the start.

- **CTRL+A** merges the current annotation with the next annotation on tier.

- **CTRL+B** merges the current annotation with the annotation before it on the tier.

- **SHIFT+DELETE** deletes the current annotation.

- **ESCAPE** cancel the changes before committing the annotation.

To see a full list of shortcuts used in this mode, go to View -> Shortcuts.... To edit/customize the shortcuts see the section called "Editing shortcuts"

Segmentation mode

In earlier versions of ELAN **on-the-fly segmentation** was offered in a separate window in which segments could only be created on one tier at a time. This functionality has now been moved to the Segmentation mode in the main window. The layout of the window looks much like that of the Annotation mode but the segmentation mode is designed for rapid creation of segmentations only: it does not provide the means to enter text into annotations. Annotations are not created based on a time selection (like in Annotation mode) but are created by marking begin and/or end time of segments by pressing the Enter key. This can be done both when the media player is playing and when the media player is paused.

There are two main elements specific to the segmentation window, a configuration panel in the tab pane and the timeline based segmentation area. The on the fly segmentation can be performed on root tiers (i.e. tiers with a tier type with stereotype "None") and on depending tiers of type "Included In". Only tiers of these types can be displayed in the segmentation area. There is only one tier editable at a time and it is displayed at the top, just beneath the timeline, decorated quite distinctly from the other tiers. The Up and Down Arrow keys can be used to navigate through the list of tiers and quickly change the tier that is editable. The editable tier remains in the list, with a red marker indicating its position. This makes it easy to see which tier is the previous (up) or next (down). The media player doesn't have to be paused for changing the editable tier.

In general the segmentation process consists of the following steps:

1. Select the candidate tiers for segmentation via the right-mouse context menu in the tier name area.

2. Configure the behavior of the keystroke in the Segmentation panel.

3. Use the media controls to play the media file or use the keyboard shortcuts (see the section called "Playing a document & making selections").

4. Do either of following:

   a. Press **ENTER** to start/end an annotation unit, while the movie or sound file is playing

   b. If the tier type of the selected tier has a controlled vocabulary and the entries of this CV have shortcut keys, press the shortcut key of the desired CV entry instead of **ENTER** to start/end an annotation unit. See also the section called "The creation of a CV".

5. Annotations are created immediately and every new annotation is a separate item in the undo/redo list (in contrast to the old segmentation implementation).

Figure 3.13. Segmentation mode

Configuring the behavior of the segmentation keystroke

The following options are available for the segmentation keystroke (Enter):

- two keystrokes per annotation; the first keystroke marks the begin, the second the end of the annotation

- one keystroke per annotation; the end time of one annotation is the begin time of the next. This creates a chain of adjacent annotations

- one keystroke per annotation, each annotation has a user definable, fixed duration. The keystroke marks either the begin or the end of the annotation

An additional option, the delayed mode, allows for compensating for the time lag between the observation of an event and pressing the key. The value (in milliseconds) is subtracted from the time value of the keystroke.

Editing annotations

The alignment of annotations on the editable tier can be modified by dragging with the mouse. When the mouse hovers over an annotation it is highlighted in green. Depending on the position of the mouse either the whole annotation can be dragged, the left boundary or the right boundary can be modified. When the mouse is over an annotation, pressing the Backspace key or the Delete key deletes that annotation.

Clicking on the annotation that is highlighted, or altering the boundaries, not only activates the annotation, but also creates a selection, painted in blue. This way you can play the selection quickly

(see the section called "Playing a selection") and correct the annotation unit if needed. You can also create a selection yourself by clicking and dragging the mouse. See the section called "How to make a selection") for more information on creating selections.
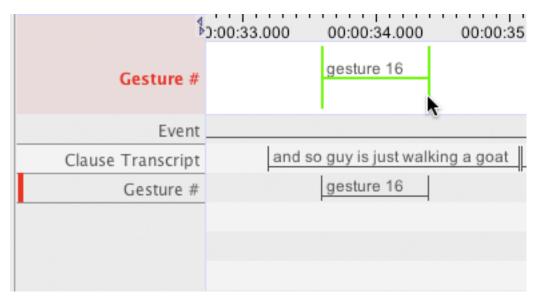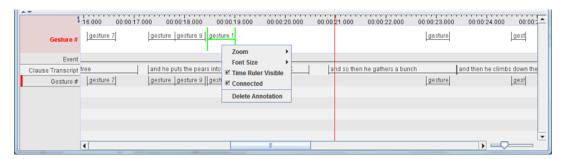
Figure 3.14. The mouse over an annotation

The right mouse button context menu contains items for zooming, for changing the font size and for deleting the annotation at the position of the mouse click. Zooming can also be done with the zoom slider in the lower right corner of the segmentation view.

Figure 3.15. The context menu

Merging annotations and splitting annotations

To merge two annotations, make a right click on the annotation and from the menu that pop-up select either Merge with annotation before or Merge with next annotation . You can also use shortcut keys **CTRL+A** to *merge with next annotation* and **CTRL+B** to *merge with annotation before*. For detailed explanation on merging see the section called "Merging annotations".
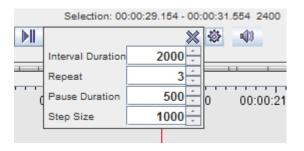
To split an annotation, make a right click on the annotation at the point where the annotation is to be split into and select Split Annotation from the menu that pops up. For more information see the section called "Split Annotations".

Step-and-Repeat playback mode

In addition to the standard play/pause buttons and the play selection button, there is a button for starting and stopping playback in step-and-repeat mode. In this mode the player starts at the current position, plays a segment of *n* milliseconds *m* times and then moves the cursor *t* milliseconds forward. Then

the next iteration starts. This mode has to be stopped explicitly, it is not stopped by e.g. clicking in the viewer.

Figure 3.16. Step and Repeat configuration

A configure button next to the step-and-repeat button shows a little settings panel in which the user can set the duration of the interval, the number of times it should be played, the pause between successive runs and the step size (number of milliseconds) for moving the interval forward.

This playback mode is new and still experimental.

Keyboard shortcuts

The following keyboard shortcuts are available in this mode:

• **ENTER** marks the begin and/or end of an annotation. The annotation is created on the active, editable tier.

• **UP** select the tier that is above the current tier in the list of tiers

• **DOWN** select the tier that is below the current tier in the list of tiers

• **DELETE** deletes the annotation under the mouse pointer (highlighted in green)

• **BACKSPACE** deletes the annotation under the mouse pointer (highlighted in green)

• **CTRL+SPACE** start/stop the step-and-repeat playback mode

Interlinearization mode

Interlinearization mode is a text oriented mode designed for parsing and glossing annotations to one or more lines of interlinearized text. This can be done manually or with the use of one or more so-called Analyzers. The segmentation and (typically) the transcription of speech events need to be done in one or more of the other modes before interlinearization can be added in this mode.

Analyzers are software modules that accept an annotation as input and produce suggestions for one or more annotations, on one or more tiers, as output. Examples of the type of processing analyzers can perform are tokenization, morphological parsing and lookup of glosses. The behavior of some analyzers can be configured in a settings panel. Some analyzers need a connection to a lexicon, others can perform their task based on the input alone. Analyzers are implemented as extensions so that third party users and developers can create and add their own analyzers. At least eventually: the LEXAN API, as it is called, still has to be finalized, documented and published.

Part of the user interface of this mode is a Lexicon panel, the front-end of a Lexicon Component module. It allows to create, import and edit a lexicon and its entries. Lexicons are stored separately from annotation data in a new data format. These are the lexicons that analyzers can get access to.

To start the Interlinearization mode, click Options > Interlinearization Mode from the main window.
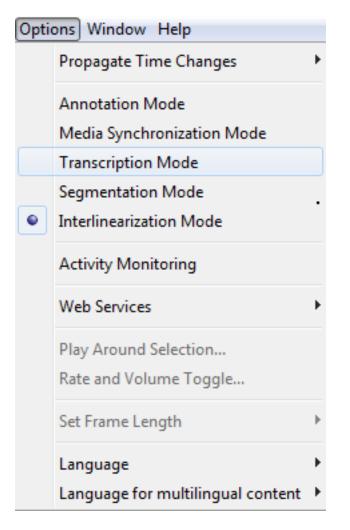
Figure 3.17. Select Interlinearization Mode

The main screen is split in two, the left side containing 2 panels, the right side consisting of a single panel.

- Top-left panel: contains a single button which gives access to several configuration windows, both for analyzer settings proper and for configuration of source and target tier types for analyzers.

- Bottom-left panel: Lexicon editor. Lexicons may be used by an analyzer and are user-defined.

- Right panel: Viewer and editor showing the annotations in an interlinear text style.

Figure 3.18. Interlinearization mode main view

To start working in Interlinearization Mode, you need to have already set up a tier structure and have to have some segmentations (annotations on a top level tier). The values of annotations can be edited in this mode and annotations on dependent tiers, including subdivisions, can be created, but not primary segmentations on top-level, independent tiers. This can be done in Annotation mode and/ or Segmentation mode. It is still possible to add new tier types and tiers in this mode (please refer to the section called "How to define a tier type" and the section called "How to define a tier and its attributes" for more information about tier structures).

If you want to use an analyzer that requires a connection to a lexicon, you should first create or import a lexicon and link one or more tier types to specific fields in a lexical entry (see the section called "Adding new tier types" and the section called "Set the Lexicon Service and entry field information for a Tier Type").

Analyzer setup and configuration

In order to analyze/interlinearize annotations with the assistance of an analyzer (i.e. other than strictly manual annotation), some configuring is required first. The top-left panel, titled Configure Analyzer Settings, shows the number of current configurations below the Analyzer & Source-Target Configuration button which opens a separate window when clicked, see below.

Figure 3.19. Analyzer and source-target configuration dialog

Configuration of analyzer settings consists of two parts:

- Required: specification of source and target tier types. Analyzers need input text and the system needs to know what the input source is. Analyzers will produce output and the system must know what the target(s) is (are). Source and target(s) specifications are based on **tier types**; the types will resolve to combinations of tiers based on those types.

- Optional: an analyzer might provide a way to modify settings and thus to change its behavior. But it should work with default settings too.

By ticking the Show tier mapping checkbox the table shows to which tiers each "analyzer-tier type" configuration applies. E.g. in case there are three speakers and the speech tier for all speakers use the same tier type, three tiers will be listed if that type is selected as source of an analyzer etc.

Figure 3.20. The configuration dialog with resolved tier mappings visible

To add new configurations or edit existing ones click the Edit configuration... button to the right. The Remove configuration... removes the selected configuration, if any.

In the dialog that appears, you can, working from left to right, choose the analyzers you would like to use and set the source and target tiers for each analyzer.

Figure 3.21. Analyzer - Tier Type configuration screen, choosing an analyzer

First, you choose a certain analyzer, as described in the section called "Types of analyzers and their settings". You can configure multiple analyzers, one per line.

Figure 3.22. Choosing the source tier type

Each chosen analyzer will need a source and at least one target tier type for it to function. The source and target tier should not be the same. By default, the user interface tries to assist in the setup by only listing types in the source and target columns if certain constraints are met. E.g. in the column for the source a tier type is only listed if there is at least one tier based on that tier type and if that tier has at least one dependent tier (which can then be selected as target).

Figure 3.23. Choosing the target tier type(s)

In the columns of the target tier types only tier types are listed for which there is at least one tier created as dependent tier of a tier of the source type. If the analyzer supports two target tiers the rightmost column will allow selection of the type for the second target, otherwise this column will be disabled.

When the List all available types... checkbox is ticked, the check on source and target types is not performed and all tier types of the transcription are listed. After selecting the target type(s) a warning message might still be shown that there are possible issues with the configuration, but the user can choose to ignore this.

Figure 3.24. A warning concerning a missing link to a lexicon field or the absence of suitable tiers

Some constraints are not checked:

• if an analyzer potentially produces multiple (suggestions for) annotations, the type of the target should be one of the subdivision types. This is not enforced.

Note

If an analyzer needs access to a certain field in a lexicon the selected type for the source and/or the target should be linked to the proper field in the right lexicon, see the section called "Adding new tier types". This way the analyzer knows which lexicon and which field to query.

Note

As of ELAN 6.9 (Nov. 2024), the interlinearization mode tries to infer additional targets for analyzers connected to a lexicon. Based on links of sibling- and/or depending tiers (through their tier type) to fields in lexical entries and possibly on the language specified for the tiers and the entry fields. Therefore it is possible to create annotations on more than the one or two target tiers configured in the Analyzer - Tier Type window.
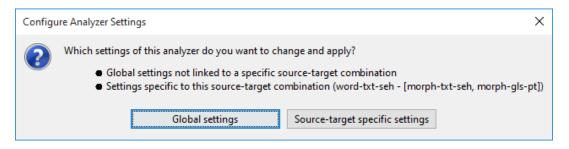
When you are done with the configuration, click Apply to finish and go back to the main dialog with the table listing the current configurations.

Note

Sometimes, especially after changing an existing configuration, it is necessary to save the file and open it again to see the effect of the changes.

A selected configuration can be removed here too by clicking the Remove configuration button. If the selected analyzer supports customization of settings, the Configure <analyzer name> button will be active and clicking it will show a **Configure Analyzer Settings** window (double clicking the analyzer in the table has the same effect). But before the actual analyzer settings window opens, you can choose whether global settings or configuration specific settings are going to be updated. This allows for different settings for an analyzer for different source-target combinations, e.g. depending on the language of involved tiers. In case of doubt choose Global Settings.

Figure 3.25. The choice to edit global or specific settings

The next section describes some analyzers and their settings.

Types of analyzers and their settings

The following analyzers are distributed with ELAN:

• Parse Analyzer

• Gloss Analyzer

• Lexicon Analyzer (a combination of the Parse and Gloss analyzers)

• Whitespace Analyzer

Figure 3.26. Analyzer settings configuration panel

The names are somewhat misleading; all of the Parse, Gloss and Lexicon analyzers require access to a lexicon. The Parse analyzer morphologically parses annotations from a word (or token) level tier, based on lexical units (prefixes, stems, suffixes etc.) available in the lexicon (internally the parser is implemented as a state machine with a stack). The results are shown as parse suggestions in a suggestion window from which the user can select one. This analyzer requires one source tier and one target tier, where the target is of a subdivision tier type.

The Gloss analyzer looks up the source annotation in the lexicon and lists all glosses found in the matched entries. The results are again presented as suggestions from which the user can select one. This analyzer requires one source tier and one target tier, where the target is of a symbolic association tier type.

The Lexicon analyzer is a combination of the parse and the gloss analyzer. By configuring the lexicon analyzer, the source tier containing the annotations will both be parsed and glossed in one action. This analyzer requires one source tier and two target tiers.

As mentioned in a note in the previous section, for the above three analyzers it is possible that more than the one or two target tiers will receive new annotations, if this can be inferred from tier, tier type and lexical entry field configurations.

The Whitespace analyzer splits the selected source annotation at white spaces and places the result on the target tier. It does not need any user confirmation. This analyzer requires one source tier and one target tier, where the target is of a subdivision tier type. Currently the behavior of this analyzer can not be configured (e.g. with respect to treatment of punctuation marks), this might be added in the future.

When configuring analyzers and their source and target tiers, it is possible that the target tier from one analyzer, is the source tier for the next analyzer. The configuration of the tiers is based on *tier types* rather than on individual tiers.

📒 Note

Configuration on the basis of individual tiers might be added later as an option as well.

The Lexicon analyzer

The Lexicon analyzer is a combination of the parse and the gloss analyzer. When a lexical entry matches a part of the input token during the matching process (and thus becomes part of one of the suggestions), the glosses of that entry are added to the suggestions too (these "glosses" can be the from

any field of the entry, depending on the tier-typ configuration), By configuring the lexicon analyzer, the source tier containing the annotations will both be parsed and glossed in one action. This analyzer requires one source tier and two target tiers. (The LEXAN API currently limits the number of target tiers to two, this might be too restrictive and may need to be reconsidered in a future release.)

Figure 3.27. Lexicon Analyzer settings configuration window

The Lexicon analyzer supports the following configurable settings:

- Include variants in the parsing process if this option is checked the parser will also look at the variant field in the process of matching morphemes from the lexicon with parts of the word or token it has received as input

- Match longer prefixes/suffixes first by default the parser tries to match shorter prefixes before longer ones. This has an effect on the order of the suggestions

- Exclude aborted parses from results if the parser hasn't finished (one iteration of) the matching process within the maximum number of steps, it adds an "++ABORT++" label at the position in the suggestions where it stopped. This option allows to filter them out of the presented results.

- Case sensitive matching tells the analyzer whether or not to ignore case in the matching process

- Match entry field language against tier content language this option allows the analyzer to only include fields with a language string equal to the content language of the target tier (the short id, e.g. "nld" for Dutch). If a lexical entry contains e.g. glosses in multiple languages, only the gloss(es) with the same langauge as the target tier will be suggested.

- Only suggest parses with same category constituents when this option is selected the analyzer only includes suggestions where the parts have the same grammatical category (based on exact matching, no support for regular expressions yet).

- Use the citation form of the lexical entry in the output with this option selected the analyzer will use the citation form field in the output (if it exists).

- Don't parse the input if it is found in the lexicon as a whole tells the analyzer to first check if the full input text is an entry in the lexicon and, if so, return that entry (or those entries) without further attempts to parse it into affixes and stem etc.

- Maximum number of parse steps this option determines when the parser should stop the matching process to prevent an unusable number of suggestions

- Affix marker character by default the analyzer assumes the character that is used to mark a lexical entry as a prefix (a-) or suffix (-a) is a hyphen. This can be changed here (ideally this information should be an accessible property of the lexicon). Apart from this marker, the analyzer has hard coded, built-in support for the morpheme types "prefix", "suffix", "root", "stem" to determine what to try to match in the parsing process.

- Clitic marker character this field allows to specify the character used to mark clitics in the lexicon. Clitics are treated the same as affixes in the parsing process.

- String for missing values sets the text the analyzer should use to indicate that a part (e.g. a gloss) is missing in the lexicon

- "Replace" field in the lexicon this analyzer supports replacement of a matched morph by one or more characters to make the next parse step (more) successful. This replacement text should be in the lexical entry and by default the analyzer looks for a (custom) field "replace". If it is in another field, it can be specified here.

- Maximum number of prefixes to detect allows to specify a maximum number of prefixes to match by the analyzer. When this maximum is reached, the remaining part of the input will be checked for matching stems and/or suffixes.

Changes in these settings will only be passed to the analyzer after clicking Apply Settings!

The Parse analyzer

This analyzer is largely the same as the parser part of the Lexicon analyzer, with the same configurable settings. It does not support the Match entry field language against tier content language option, which belongs to the glossing part of the process.

The Gloss analyzer

This analyzer performs a look-up of the input token in the lexicon and returns all values of the lexical entry field it is configured for (via the tier type). This doesn't have to be the "gloss" field of the lexical entries, but can be any field.

Figure 3.28. Gloss Analyzer settings configuration window

This analyzer supports the following configurable setting:

- Also match against variant fields in the look up process if this option is checked the glosser will also look at the variant field in the process of matching the input

- Match gloss language against tier content language this option allows the analyzer to only include glosses with a language string equal to the content language of the target tier (the short id, e.g. "nld" for Dutch). If a lexical entry contains e.g. glosses in multiple languages, only the gloss(es) with the same langauge as the target tier will be suggested.

- Also match against the citation field in the look up process with this option selected the analyzer will use the citation form field in the output (if it exists).

- String for missing values sets the text the analyzer should use if the specified field is not found in matched entries in the lexicon

The Whitespace Text analyzer

This analyzer splits the input text it receives into multiple tokens based on white spaces. It allows to configure how e.g. punctuation marks should be treated.

Figure 3.29. Whitespace analyzer configuration panel

The + (Add) and - (Remove) buttons can be used to add or remove a category of characters, represented by a row in the table. A category can contain one or more characters; if there are more than one, each character is separately treated according to the setting for that category. The table has two columns, one labelled **Marks**, where the special characters or marks can be entered, and one labelled **Action**,

specifying the way those characters should be handled in the tokenization process. When clicked on, the second column shows a dropdown list with predefined actions:

- **Treat as white space** means that the input will be split at the position of this character and that this character itself will not be in the output

- **Create separate annotation** means that the input will be split at this position and that this character will become a separate token (annotation) in the output

- **Keep with preceding token** means that this character will become part of the same annotation as the characters to the left of it

- **Keep with following token** means that this character will become part of the same annotation as the characters to the right of it

- **Remove** means that this character will be removed from the input string without causing a split of the input (i.e. it is filtered out)

The Apply button has to be clicked to inform the analyzer of the changes and to put them into effect.
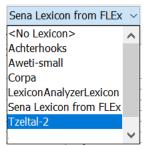
Lexicons

The main purpose of the Lexicon Component in ELAN is to support the (semi-automated) interlinearization process. It is not intended as a full-fledged lexicon tool, though the data model supports a bit more than strictly necessary for its main purpose. The data model and XML-based data format are similar to the LIFT format (Lexicon Interchange Format)), but simplified. These are the main fields of a lexical entry:

- **lexical-unit** (1)

- **morph-type** (0 or 1)

- **citation** (0 or 1)

- **variant** (0 or more)

- **phonetic** (0 or more)

- **sense** (1 or more)

  - **grammatical-category** (1)

  - **gloss** (1 or more)

The main field is **lexical-unit** (equivalent to lemma, headword, the primary lexical form). **morph-type** indicates the word part (e.g. stem, prefix, suffix), analyzers can use this information when processing the input text. The **grammatical-category** field is the category of the lexical item, the part of speech. The Edit Entry window shows which other fields can be added at the moment. (The data model defines more fields than visible in the entry window, but support and documentation hereof is still pending.) The user can add custom fields at the level of the **entry** and at the level of **sense**, these will be visible as **field: name**.

In the lexicon panel of the interlinearization mode the contents of lexicons can be displayed, one lexicon at the time. But it is possible to store and manage multiple lexicons on disc and choose which one to display. The leftmost drop-down box above the table lists all lexicons that have created or imported.

Figure 3.30. The list of available lexicons

If a lexicon has been selected and is displayed in the panel, lexical entries can be added, edited or removed.

Figure 3.31. Lexicon overview

The lexicon overview can be adjusted to display or hide certain columns. To do so, you have to right-click on a lexical entry to display the context-menu. From there, you can show or hide columns of the lexicon. Some fields can occur more than once in an entry (like **variant** or **gloss**), these will be displayed in a single cell, each value surrounded by square brackets. The context menu also has Add, Remove and Edit items, which have the same function as the buttons below the table (see the section called "Editing lexical entries").

Figure 3.32. Lexicon context-menu

The order of the entries in the table can be changed by clicking on any of the column headers, a little arrow indicates whether the items are sorted in ascending or descending order.

Creating & editing lexicons

Actions for creating and editing lexicons are available in the Lexicon Actions drop-down menu. The actions always apply to the lexicon that is selected in the lexicons list and visible in the table. The actions will be discussed below.

Figure 3.33. The Lexicon Actions

• The Open Lexicon Editor Window... option opens a separate window for editing lexicons and lexical entries. Especially when multiple modifications are intended, this is more convenient than the one-time use entry edit dialogs described in the section called "Editing lexical entries". The separate Lexicon Editor is described in section the section called "The Lexicon Editor".

• To create a new lexicon from scratch choose Create New Lexicon.... A new dialog will appear, allowing you to fill out details about your new lexicon.

Figure 3.34. Create New Lexicon dialog

When the details have been filled out (a Name and a Language are required), click Apply to create the lexicon. The lexicon file will be stored in a predefined folder labeled **LexanLexicons** inside the the section called "Special ELAN data folder". The file name is based on the name of the lexicon and the file extension is `.xml`.

- The Open Lexicon... option allows to open a lexicon that is already in the native format but is in a different folder. The lexicon file is then copied to the **LexanLexicons** folder. So, this action behaves like an import function for lexicons that don't need to be converted (e.g. a file that has been shared by a colleague).

- The Close Lexicon option closes the current lexicon, the table will be cleared and **No Lexicon** will be selected.

- The Save Lexicon action updates the corresponding file in the default location, if there have been any changes in the lexicon.

- The Save Lexicon as... option allows to save the lexicon somewhere else in the same file format, e.g. for the purpose of creating a back up file. This action is similar to the export functions but then in the native format.

- The Save Lexicon with Current Entry Order option saves the lexicon with the entries sorted as they are visible in the table of lexical entries (the lexicon framework doesn't allow to specify a custom sort order yet).

- The Edit Lexicon Properties... option opens a window similar to the **Create New Lexicon** dialog, the properties that have been entered before can be updated or completed here. But there are two additional tabs, Custom Fields and Sort Order, described below. Click the Apply button to apply the changes to the lexicon.

Figure 3.35. The Edit Lexicon Properties window

Figure 3.36. The Custom Fields tab

The Custom Fields tab allows to specify the names of custom fields (at the level of **entry** or **sense** or both) to be used in lexical entries of this lexicon.

Figure 3.37. The Sort Order tab

In the Sort Order tab a preferred sort order can be specified by entering an ordered list of tokens consisting of one or more characters. The sort order will be applied to the **lexical-unit** field but also for the **variant** and **citation** fields. After changing the sort order it might be necessary to click the header of the lexical-unit column once to enforce re-sorting of the table.

🗒 Note

> It is advised to use the Save Lexicon with Current Entry Order option to apply the new sort order to the underlying data structure as well (not only to the view). After that, new lexical entries will be inserted according to the custom sort order.

- The Import Lexicon... option opens a file browser window that allows to select either a Toolbox lexicon (`.dic, .db, .txt`), a lexicon in LIFT format (`.lift`) or a CorpAfroAs lexicon (`.eafl`). In case a Toolbox file has been selected, a configuration window will be shown (see below), otherwise a converter will import as much as possible from the original lexicon data into an ELAN lexicon and will add the new lexicon to the list of lexicons.

Figure 3.38. Import Toolbox dictionary

This window allows to specify mappings from Toolbox field markers to ELAN lexicon entry fields. The main element in the window is a tabel with in the first column the list of markers that have been found in the Toolbox dictionary file. In the second column the corresponding field name can be selected from a list or a custom field name can be entered (either *custom-field-name* or *sense/ custom-field-name*). For some fields it will be possible to enter a language code in the third column, depending on the value in the second column. Markers that don't have a mapping in the second column will be ignored during the import process.

Other elements in this configuration window are:

- The Lexicon Name text field allows to enter a name for the lexicon. A name is required, it is also used for the file name of the lexicon.

- The Show raw field names in table checkbox only has an effect on how field names appear in the table, as *friendly* name or as their *raw* equivalents from the lexicon XML file.

Figure 3.39. Import Toolbox dictionary showing "raw" field names

- The Split semicolon separated fields into multiple fields option requests the import function to produce multiple fields from a single input field, if the input contains one or more semicolon characters and if the target field is allowed to appear more than once in an entry (e.g. the input has multiple gloss values separated by semicolons).

After clicking the OK button several warning messages might be shown, e.g. if required information is missing, if a required field has not been selected in the second column (e.g. **lexical-unit**) or if a field has been selected more often than allowed.

- The Export Lexicon... action only supports conversion to LIFT format.

📑 **Note**

There is currently no Undo/Redo mechanism for lexicon edit actions!

📑 **Note**

If multiple transcription windows are open, using the same lexicon, modifying the lexicon should preferably be done in only one window. On Windows there is no guarantee that all windows are updated correctly after a change in a lexicon made in a different window!

📑 **Note**

More documentation about the structure of lexical entries, about which fields are required, what is hard-coded etc. will follow. The same for native, import and export formats etc.

Editing lexical entries

After creating or importing a lexicon, lexical entries be added, edited or removed. Adding a new entry can be started by clicking the Add button on the bottom of the Lexicon main panel or by choosing

the Add menu item from the context menu, if there are already entries in the table. A new dialog will appear, allowing to create a new entry by entering values for, at least, the required fields.

Figure 3.40. Add a Lexical Entry

When you are finished, click Apply to add the entry to the lexicon. Required fields will be highlighted if you click Apply while not all required fields are filled in.

Editing a lexical entry can be done by either clicking the Edit button on the bottom of the panel, by right-clicking the entry and choosing Edit from the context-menu or by double-clicking the entry in a cell that can not be edited directly in the table (i.e. the **lexical-unit** field or any field of which there can be more than one). A dialog will open, displaying the chosen lexical unit as a tree structure. Some general information is shown, such as ID and date of creation.

Figure 3.41. Edit a lexical entry

- To edit the entry, click on the value that you would like to add or change or press **TAB** to select the first editable field. It will become active and ready for editing, a dark blue border will highlight it. Hitting the **TAB** key accepts the value of the current field and activates the next field.

- The fields that are displayed in a dark grey color are so-called placeholders: these fields are not yet there, but can be created. When activating such field the + button can be clicked to add the field to the entry and to make the text field editable.

- In general when activating a field + or - buttons can be present and can be enabled or greyed out, depending on whether more fields of that type can be added or if that field can be deleted. Pressing the + or = keyboard key triggers the **Add Field** action, the - key the **Remove Field** action.

- When a field is removed it will still be shown in the entry, but in a red color.

- Changes will only be applied to the entry in the lexicon after clicking the Apply button or pressing the **ENTER** key.

- Changes can be discarded by clicking the Cancel button or pressing the **ESC** key or the **CTRL +W** combination.

Removal of a lexical entry is done by highlighting the entry in the lexical entries table and then either clicking the Remove button on the bottom of the Lexicon panel, or by right-clicking the entry and choosing Remove from the context-menu.

The Lexicon Editor

The Lexicon Editor combines the editing actions described in the previous sections in a separate, blocking window. The window is split in two with on the left side the same lexical entry table as in the main window and on the right side the lexical entry editor.
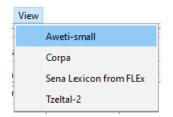
Figure 3.42. The Lexicon Editor window

The menu bar has three menus:

- The File menu holds most of the lexicon **Open**, **Save** and **Import/Export** actions that are described in the section called "Creating & editing lexicons"

- The Edit menu currently only contains the **Edit Lexicon Properties** item.

- The View menu lists the available lexicons. Selecting a lexicon here will load it and fill the Lexical Entries table with its entries.

Figure 3.43. The Lexicon View menu

The left panel shows the name of the current lexicon at the top, with visual highlighting when there are unsaved changes to the lexicon. The lexical entry table covers most of the space, it has the same column showing/hiding and sorting mechanisms as mentioned above. A single click on an entry loads the entry in the entry edit panel to the right. The **UP** and **DOWN** keys select the entry in the previous or the next row and loads it. The Add button creates a new entry and loads it. The required fields in an entry are filled with template text, to be modified by the user.

Figure 3.44. A new lexical entry

Navigation from one field to the next is again performed with the **TAB** key, the **ENTER** key applies the current changes to the entry. If another entry is selected and loaded, changes in the current entry will be applied without a prompt.

📄 **Note**

If there are multiple ELAN windows open with the same lexicon visible, it might be necessary to reload the lexicon in those windows to see the modifications made in the Lexicon Editor.

Figure 3.45. Filtering the lexical entries

It is possible to filter the entries in the table by entering a search string in the Filter Entries text field and pressing **Enter**. The input is treated as a regular expression. The filter can either be applied to all visible columns or to a single column, selected in the Column drop down box. This allows, for example, to only show the entries with a citation form starting with a 'c'. The filter can be removed again by clicking the Reset button and all entries will be listed again.

The interlinearization panel

The interlinearization panel shows the tier structures and the annotations that were created in e.g. the annotation mode. The annotations are displayed in cells, like in a table or a list, each cell containing a top level annotation and its dependent annotations. The annotations of all visible top level tiers are sorted on time and then added to the cells. This is the main editing area in this mode, for manual and for assisted annotation. In this panel the configured analyzers can be invoked, either for an individual annotation or for a sequence of annotations, via the Analyze/Interlinearize button. The result(s) produced by the analyzer will either be shown here in a Suggestions Window (if there are suggestions to choose from) or immediately applied and added to the designated tier(s) as determined by the analyzer configuration.

Figure 3.46. Interlinearization Panel

The panel has a small "toolbar" at the top with the following options:

- The Play Selection panel allows to play back the audio of the interval corresponding to the selected cell (light blue background). A small progress bar shows the playback progress within the interval. The source of the audio is the first media file in the list (the "master media").

- The Analyze/Interlinearize button starts an automatic, sequential processing of annotations, starting from the active annotation (or the first one of the right type in the view) and then continuing to the right and to the bottom. If the analyzer produces multiple suggestions and user interaction is required, it will pause when the suggestions window is on screen, until the user selects one of the suggestions and the analyzer will continue with the next annotation. When the suggestion window is closed with the **ESC** key or by clicking the window Close button, this ends the automatic processing. Clicking inside the suggestions window but outside of any of the suggestions, will ignore these suggestions but continue with the automatic processing.

- The Recursive option only has an effect if there are more than one analyzer source-target configurations. When this checkbox is selected, annotations created by an analyzer will in turn be analyzed, if the first analyzer's target tier is configured as a source tier for another analyzer.

- The Font Size - and + buttons decrease or increase the size of the font(s) in the panel. The keyboard equivalents are **CTRL+-** and **CTRL++** (**CTRL+=**).

- The **divider** to the left of the Analyze/Interlinearize button can be dragged to change the width of the area of the tier labels.

- The Configure button creates a dialog in which the user can customize margins and colors used in the interlinearization panel.

Figure 3.47. Interlinear editor configuration window

The following properties can be set:

- Two background colors for alternating rows.

- Background and border color of the annotation boxes (if any).

- Horizontal and vertical margins in pixels between the text and the border of the annotation boxes.

- The width in pixels representing a white space between annotations (1 - 20).

Figure 3.48. Customized Interlinearization Panel

The tiers that are visible in the editor can be configured via the right-click context menu of the tier names area.

Figure 3.49. Visible tiers context-menu

The Show / Hide More... option opens the same window as described in the section called "Switching tiers on/off" and in the section called "View tiers by Type/Participant/Annotator". The Speaker tier is not a real tier but it shows the Participant attribute of the top level tier in this cell. The TC (time code) tier shows the begin and end time of the top level annotation in this cell.

📒 Note

Hiding the top level tier(s) hides all depending tiers, effectively removing the corresponding cells.

A context-menu will also be shown when right-clicking an annotation. Depending on the annotation there can be different options. This allows you to start interlinearization of an annotation, delete an annotation, or add new annotations.

Figure 3.50. Interlinearization context-menu with Analyze option

If the active annotation is on a tier that is the source tier for any of the analyzers, there will be the Analyze / Interlinearize option, which invokes the analyzer with this annotation as input. The option the section called "Deleting annotations" will always be there, while the section called "Entering annotations before/after other annotations" and the section called "Creating depending annotations for a active annotation" are available depending on the type of tier the annotation is on.

Another option is also only available for some annotations.

Figure 3.51. Context-menu with Add to Lexicon option

The Add to Lexicon option is only available for annotations that are on a tier that is linked to a field in a lexicon (via its tier type). This action opens the new entry window (Figure 3.40, "Add a Lexical Entry") and adds the value of the annotation to the corresponding lexical entry field.

The Suggestion View

When you click the Analyze / Interlinearize button or right-click and choose Analyze / Interlinearize from the context-menu, the process of analyzing or interlinearization will commence. If the analyzer produces multiple, alternative suggestions that need to be disambiguated, a window will appear showing the suggestions in a layout similar to that of the interlinearization panel. This is called the Suggestion View. The window is positioned just below the source annotation but it can be moved and resized. If you tick the Remember window position checkbox, the Suggestion View will show up with the size and location of the last time it was shown.

Figure 3.52. Suggestion View

The image above displays the Suggestion View with suggestions produced by the Lexicon analyzer, which has two target tiers. It both suggests possible parsings of the input annotation as well as glosses, based on lexical entries found in the current lexicon. You can select the suggestion that best matches your expectations by clicking on it. This is recorded by the analyzer and the next time the same input occurs, it will move the suggestions that have been selected most often to the top of the list. A little header saying "*chosen x times before*" will appear in those suggestions. If the width of a suggestion makes part of the header invisible then hovering the mouse over the header will show the text in a tooltip. Hovering the mouse over a part of the suggestions will show relevant fields of the lexical entry that part is based on.

Figure 3.53. Tooltip showing some fields of the originating lexical entry (for suggestion S6)

There are several keyboard shortcuts for mouse-less interaction with the Suggestion View.

Figure 3.54. Keyboard shortcuts in the Suggestion View

- Digits 1-9 and 0 select suggestion number 1-10

- Digits while pressing **SHIFT** will select suggestions 11-20

- **SPACE** or **ENTER** selects the first suggestion

- **TAB**, **BACKSPACE** or **DELETE** ignore this suggestion (and move to the next one in case of sequential processing)

- **ESC** or **W** closes the Suggestion View

- the **UP**, **DOWN**, **LEFT**, **RIGHT** arrow keys scroll the view area, if there are scroll bars

- **PAGE UP** and **PAGE DOWN** scroll through the pages in the scroll area

- the **R** key reverses the order of the suggestions

- the **I** key switches between incremental and normal selection mode (see Figure 3.55, "Incremental, step-by-step selection of suggestions, left to right")

It can happen that there are so many suggestions that they are (too) hard to overview. There may be many similar-looking options, e.g. similar looking parses. To get some visual aid, you can press **Shift** and hover the mouse pointer over the fragments of a suggestion. This will trigger a colouring effect: all suggestions with the same value at that position will be displayed with the same background colour (i.e. there will be as many different colours as there are different values in that position). By then clicking one of the suggestions (**Shift** still pressed), only the suggestions with the same colour will remain in view, all others are removed. This can be a way to narrow down the available choices.

Alternatively it is possible to switch to Incremental selection mode. In that mode disambiguation of the fragments is supported by showing only the alternatives for one specific position, starting with the first one from the left. After choosing the best option, the remaining alternatives for the second fragment are shown, and so on. When filtering for alternatives for a fragment, only the surface form is taken into account (different entries with the same value in the relevant field are shown as one).

Figure 3.55. Incremental, step-by-step selection of suggestions, left to right

The checkbox Only show suggestions chosen before removes all suggestions from the view that have never been chosen before. In other words, if an input word has been disambiguated before, it is possible to only show the previous disambiguation(s). The number of currently shown and hidden suggestions is displayed near the checkbox. When the input hasn't been seen before the window will be empty with this option. the **C** key switches this mode on or off. This option can not be used in combination with the incremental selection workflow.

Figure 3.56. Only the suggestions chosen before are visible

The right mouse button context menu of a suggestion contains one option Don't show this suggestion again, which, if the analyzer supports this, will expel this output or this suggestion (i.e. this combination of elements in the suggestion) for this input from future suggestions.

Figure 3.57. The context menu of a suggestion

Chapter 4. Searching

The ELAN tool allows you:

- to search annotations within one document (see the section called "Searching in a single annotation file")

- to display the results (see the section called "Displaying search results")

- to jump to the corresponding annotation the ELAN window (see the section called "Jumping from the search results to annotations")

- to search through multiple annotation files (see the section called "Searching through multiple annotation files")

- to do a structured search through multiple annotations files (see the section called "Structured search through multiple annotation files")

Searching in a single annotation file

To search for text, do the following:

1. Click on Search menu.

2. Go to Find (And Replace)... (alternatively you can press **CTRL+F**) The following dialog window is displayed:

Figure 4.1. Search dialog

3. The following search options are available:

   a. Go to Annotation on tier and, from the pull-down menu, select the tier to be searched.

   b. Go to matches and type in the item to be searched. If the tier type of the selected tier has a Controlled Vocabulary (see the section called "Changing the parent of a tier"), this field is a pull down menu containing the entries of the controlled vocabulary. Note that it is still possible to enter a string that is not in the list.

You can always make use of regular expressions to conduct your searches when "regular expression" is checked. (see Appendix A, *REGULAR EXPRESSION SEARCH* for the regular expression syntax).

By default the search is not case sensitive. To change this, select the "case sensitive" checkbox.

Optionally, specify the interval to search in (from ... s …ms to … s … ms). Make a choice between searching within a time interval and finding annotations that overlap with a certain interval. Click on Add new constraint to add a second tier and search item. Up to 10 constraints can be used. There exist 2 kinds of them:

1. Constraints based on structural distance. (Annotation units "around" a certain annotation entity). This option is only available for tiers that are symbolically associated to (or are a symbolical subdivision of) the tier mentioned in the first search box.

   For example: annotations contained in a structural distance of −1 to 2 tx-annotations from *trees* on the tier tx are *sees*, *trees*, and, *flowers*.

Table 4.1. Annotation example

tier	annotations							
st (sentence)	He sees trees and flowers.							
tx (word)	he	sees		trees		and	flowers	
mb (morpheme break)	…	see	-s	tree	-s	…	flower	-s
ps (part of speech)	…	V	SUF	N	SUF	…	N	SUF

2. Constraints based on temporal distance. This means search results are restricted on the basis of the temporal relation between two intervals:

   • *is inside*: the annotation is completely contained within a given interval
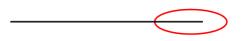
   • *overlaps*: at least a part of the annotation is contained within the given interval
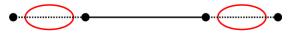
   • *overlaps only begin time of*: the annotation only has its end part in common with the given interval

   • *overlaps only end time of*: the annotation only has its begin part in common with the given interval

   • *is within … around*: the annotation is contained in an interval around either the begin time or the end time

- *is within ... around begin time of*: the annotation is contained in an interval around the begin time

- *is within ... around end time of*: the annotation is contained in an interval around the begin time

It is possible to search on different tiers within one annotation. For example, the search parameters illustrated below search for all annotations on the tier tx, which contain "-s" in one of their morpheme breaks and "N" in one of their parts of speech. (Both "-s" and "N" are in distance of "0 words", i.e., they occur within the same word as specified on the tier tx.) I.e., these parameters would find "trees" and "flowers" in the above example, but not "sees".

Figure 4.2. Find with constraints

Another option is searching for sequences of utterances, words or other annotations on the same tier, e.g.:

Figure 4.3. Find sequences

on the tier tx, the annotation "flowers" follows 1 to 3 annotations after the annotation "trees"

You can delete the second (or third) search item. Click on Delete last constraint to delete it.

3. Right click in a text box to change the input character set and select the suitable language from the pull-down menu.

   You only need this option if you want to select a non-default character set. The box automatically displays the default set of the selected tier (see the section called "Changing tier attributes").

   After you have specified your search parameters, click OK to start the search process.

📑 Note

Make sure the box next to *Regular Expression* is checked when you search for "special" characters (i.e. all characters that are not plain letters or digits) like diacritic characters.

Advanced searching: an example

Suppose we are investigating turn taking and we want to find all switches from speaker W to speaker K that don't overlap, with gaps of at most 2 seconds. In order to find this, we fill in the search form as follows:

Figure 4.4. Advanced searching example

Displaying search results

After the Search is completed, ELAN displays the following search results:

- The number of annotations where the search item was found.

- The full content of each annotation where the search item was found along with full content of its parent annotation and all its child annotation.

- The begin, end time and duration of each annotation where the search item was found. To see more information per match, right click in the results and check the other desired fields (file – tier – before – after).

- Clicking the header of a column sorts the result table based on that column. Depending on the column, sorting is by alphabet or by time value. Clicking the header again changes the sorting from descending to ascending or vice versa.

Figure 4.5. Search results

The search history

When performing multiple search actions one after another, a search history is created. With this feature it becomes possible to browse through all the entered queries:

1. Enter a query and click OK to activate it.

2. Click on the New Query icon

3. Enter another query and choose OK

4. Now you can use the back and forward icons to browse through the search history.

Figure 4.6. Search history

> 📒 **Note**
>
> When closing the search dialog, the query history is removed. If nonetheless you want to save a certain search command, have a look at the section called "Saving and loading queries".

Refining a search

After a search operation, one can perform another search within the results of the previous action. To achieve this, select Search on current result from the Query menu in the Search dialog window. This will get you the following window:

Figure 4.7. Refining a search

By specifying extra search constraints, you can narrow down the results. This is similar to the addition of an extra search constraint.

Saving and loading queries

When using complex queries, or one is repeating the same search on annotation files, it is convenient to reuse a query. This is possible in ELAN thanks to the save and load query functionality.

Saving a query

1. Specify a query in the search dialog

2. Either choose Query > Save or click on the save icon in the toolbar.

3. Enter a file name for the active query.

4. Choose save.

Figure 4.8. Saving a query

Loading a query

1. Either choose Query > Open or click on the Open icon in the toolbar.

2. Select an existing query from the file dialog.

3. Choose Open.
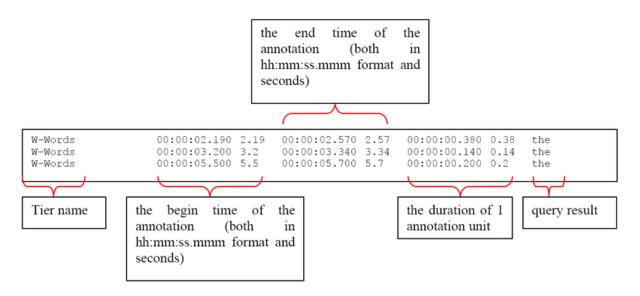
Exporting the search results

To export the results of a query to a tab-separated file, go through these steps:

1. In the query dialog, select Query > Export matches

2. Specify a file name

3. Click on save

An exported file looks as follows:

Figure 4.9. Exported search results


```
                                    the    end    time    of    the
                                    annotation       (both      in
                                    hh:mm:ss.mmm  format  and
                                    seconds)


W-Words              00:00:02.190 2.19   00:00:02.570 2.57   00:00:00.380 0.38   the
W-Words              00:00:03.200 3.2    00:00:03.340 3.34   00:00:00.140 0.14   the
W-Words              00:00:05.500 5.5    00:00:05.700 5.7    00:00:00.200 0.2    the


  Tier name          the   begin   time   of   the                the duration of 1      query result
                     annotation       (both      in               annotation unit
                     hh:mm:ss.mmm  format  and
                     seconds)
```

If you right click in the table containing the search results, a popup menu appears. Tick the checkboxes to show or hide columns that are related to the found annotations. In the same popup menu you will find the option Export Table as tab-delimited text, which literally saves the displayed result table to a text file, as shown in the example below.

Figure 4.10. Export search results

Search and replace

ELAN offers the possibility to replace all of the found search results. Choose Edit > Replace in the Search Dialog:

Figure 4.11. Search and replace

Then enter the text that should replace the found results and choose OK.

Jumping from the search results to annotations

You can jump from the annotations displayed in the Search-Dialog window to the corresponding annotations in the ELAN window. This option allows you to access other types of information that are linked to the annotations found: to listen to the audio, to watch the video, to view the waveform or to read the annotations on different tiers.

Do the following: In the Search-Dialog window, click on the annotation that you want to jump to. It will be highlighted in blue color. In the ELAN window, the corresponding annotation is automatically accessed.

The screen display should look similar to the following illustration:

Figure 4.12. From search results to annotations

Searching through multiple annotation files

A screencast regarding this section is available on Vimeo: `https://vimeo.com/67896572`

ELAN offers an option to search for an expression through multiple files. This search function searches for (whole) words within annotations to match the given query. To access it, go to Search > Search multiple eaf… This will open the following dialog box:

Figure 4.13. Searching through multiple annotation files

In order to perform a multiple file search, go through the following steps:

1. Click on `Define search domain` and do one of the following:

   - Select an existing domain from the list and click Load. (Click Delete if you want to delete the domain.)

   - Create a new domain:

     a. Click New Domain...

     b. Click in the new dialog on the Look in pull down box and browse to the directory that contains the annotation files.

     c. Double-click an annotation file (`*.eaf`) to select it. It now appears in the rightmost box. Alternatively, you can click on the annotation file name and click the >> button.

        Repeat this for every annotation file you want to include.

        It is also possible to select a complete directory. All `.eaf` files in a selected directory will be included.

     d. Click OK to continue the exporting process; otherwise click Cancel to exit the dialog window without exporting.

     e. If you clicked OK you can save this domain: enter a name and click OK. If you do not want to save the domain click Cancel.

   - Create a new domain from an IMDI search:

     a. Click New Domain from IMDI Search...

b. Browse to and select an IMDI file that has been exported from a metadata search in the standalone IMDI Browser.

c. Click Open.

d. You can save this domain: enter a name and click OK. If you do not want to save the domain click Cancel.

2. Enter a search expression and optionally enable a regular expression and/or case sensitive search.

3. Click on Search. The result screen will appear:

Figure 4.14. Searching through multiple annotation files: result

Nr	File	Tier	Before	Annotation	After	Begin Time	End Time	Duration
13	elan-example1	W-Words	of	the	town	00:00:08.130	00:00:08.260	00:00:00.130
14	elan-example1	W-Words	then	the	sign	00:00:12.200	00:00:12.340	00:00:00.140
15	elan-example1	W-Words	's	the	Oranje	00:00:13.694	00:00:13.836	00:00:00.142
16	elan-example1	W-Words	follow	the	sign	00:00:16.263	00:00:16.404	00:00:00.141
17	elan-example1	W-Words	to	the	Rhine	00:00:22.635	00:00:22.930	00:00:00.295
18	elan-example1	W-Words	follow	the	signs	00:00:30.305	00:00:30.465	00:00:00.160
19	elan-example1	W-Words	go	the	this	00:00:33.220	00:00:33.620	00:00:00.400
20	elan-example1	W-RGMe		Going out of	Going out of M...	00:00:01.600	00:00:02.240	00:00:00.640
21	elan-example1	W-RGMe	A roundabout	Going to the	Going along Or...	00:00:12.120	00:00:12.680	00:00:00.560
22	elan-example1	W-RGMe	Going along Or...	Going to the	Going down	00:00:17.000	00:00:17.200	00:00:00.200
23	pear story	Clause Tra...	and then you s...	and so he	and he starts pi...	00:00:10.440	00:00:14.040	00:00:03.600
24	pear story	Clause Tra...	and so he clim...	and he starts	and he puts the...	00:00:14.050	00:00:17.070	00:00:03.020
25	pear story	Clause Tra...	and he starts pi...	and he puts	OK	00:00:17.283	00:00:19.423	00:00:02.140
26	pear story	Clause Tra...	and so then he ...	and then he	and starts puttin...	00:00:23.413	00:00:25.173	00:00:01.760
27	pear story	Clause Tra...	and then he cli...	and starts	and then you h...	00:00:25.183	00:00:28.553	00:00:03.370
28	pear story	Clause Tra...	and then he go...	and **the** man	and starts picki...	00:00:38.943	00:00:40.663	00:00:01.720
29	pear story	Clause Tra...	and the man go...	and starts		00:00:40.673	00:00:43.843	00:00:03.170
30	pear story	morph	with	the	ladder	00:00:13.476	00:00:13.752	00:00:00.276
31	pear story	morph	off	the	tree	00:00:16.312	00:00:16.689	00:00:00.377
32	pear story	morph	puts	the	pears	00:00:18.084	00:00:18.351	00:00:00.267
33	pear story	morph	down	the	ladder	00:00:24.668	00:00:24.919	00:00:00.251
34	pear story	morph	all	the	pears	00:00:26.531	00:00:26.868	00:00:00.337
35	pear story	morph	and	the	man	00:00:39.158	00:00:39.373	00:00:00.215

The result window contains the following fields for every found annotation:

• Nr: the order in which the results appear

• File: the file that contains the result

• Tier: the tier that contains the result

• Before / After: the annotation as found before and after the annotation that matches the search expression

• Annotation: the search result itself

• Begin time, end time, duration: of the annotation unit that was found

📋 Note

It is not possible to restrict the search results to a certain tier or to specify extra structural or temporal constraints.

The displayed search results can be exported to a tab-separated text file as well. The exported files are very similar to that described in the section called "Exporting the search results".

- Just click on Export next to the Search button.

- Enter a file name.

- Select the Save button.

If you click on one of the listed annotations, a new ELAN window will be opened, and that annotation unit will be selected. When clicking on another result, the newly opened window is reused.

If you right click somewhere in the list ELAN shows a context menu with the following options:

- File, Tier, Before, After, Begin Time, End Time, Duration: uncheck a columns if you do not want it to be in the result table

- Font Size > :change the font size of the result table

- Toggle Time Format: change time format to milliseconds or back to hh:mm:ss.ms.

- Export Table as Tab-Delimited Text...: export the table as tab-delimited text without the hidden columns

FASTSearch

Since version 4.7.0, you can now also search via FASTSearch. This method of searching is the same as the search function described in the section called "Searching through multiple annotation files". There is a difference on the back-end of the search function, which makes it a bit faster than the normal search function, especially when you have to search through a large amount of files.

Find & replace in multiple files

You can also do a find & replace in multiple eaf files. To do so, click Search > Find & Replace in Multiple Files... from the main window of ELAN. A warning message will appear, telling you find & replace actions cannot be undone. So make sure you have backups of the files you will work on, and click Yes to continue. A dialog will appear, looking like this:

Figure 4.15. Find & replace in multiple files dialog

First, select a domain to work with by clicking Define search domain. This will open the Search Domain dialog (see the section called "How to select multiple files" for more info about setting up a domain). Select a domain containing the eaf files you want to do the search & replace action on, or create a new domain. Click Load to load the chosen domain.

Back in the multiple file find & replace dialog, the eaf files and their location are now visible in the top-half of the dialog. You can select if all tiers must be searched, or a selection. When you choose the latter, a dialog with all the tiers from the files will pop-up:

Figure 4.16. Tier selection dialog

Check the tiers you want to include in the find and replace function and select OK. Next, fill out the query you want to find and replace. You can also search by means of regular expressions (see Appendix A, *REGULAR EXPRESSION SEARCH*), and/or do a case sensitive search, by checking either or both options.

In the Replace by field, fill out the desired text that will be put in place of the found results. Finally, click Find and Replace to start the process. When done, a process report will be shown, with information regarding the inspected files, the number of hits and the number of files that have changed.

Figure 4.17. Process report dialog

```
Process Report                                          ✕

Multiple file find and replace report

Number of files to process:   4
Tiers to search: all tiers
Search string:  TR
Replace string:  T
Regular expression:  false
Case sensitive:  false


Number of hits in C:\Users\jergee\Desktop\ELAN\
Saved file successfully: C:\Users\jergee\Deskto
Number of hits in C:\Users\jergee\Desktop\ELAN\
Saved file successfully: C:\Users\jergee\Deskto
Number of hits in C:\Users\jergee\Desktop\ELAN\
Saved file successfully: C:\Users\jergee\Deskto
Number of hits in C:\Users\jergee\Desktop\ELAN\
Saved file successfully: C:\Users\jergee\Deskto
Find and replace completed.


Summary:
Number of files in domain:   4
Number of files inspected:   4
Number of files changed:   4
Number of files failed:   0
```

Structured search through multiple annotation files

If you want to perform a detailed search over multiple EAF-files, but the options offered by *Search multiple EAF* (see the section called "Searching through multiple annotation files") are not comprehensive enough, you can use yet another search mode. This allows you to restrict the search domain to certain tiers, to use regular expressions, etc. while examining multiple annotation files at once. This search function will search for (whole) words to match the given query, but also will match parts of words that match the query.

The function can be reached via Search > Structured search multiple eaf.... When you click on this option for the first time, you will be asked to define a search domain in the form of one or more .eaf files. The next time you open the Structured search, it uses the last defined search domain. The search

window offers the possibility to define a new search domain: click on Define Domain and do one of the following:

- Select an existing domain from the list and click Load. (Click Delete if you want to delete the domain.)

- Create a new domain:

  1. Click New Domain...

  2. Click in the new dialog on the Look in pull down box and browse to the directory that contains the annotation files.

  3. Double-click an annotation file (`*.eaf`) to select it. It now appears in the rightmost box. Alternatively, you can click on the annotation file name and click the >> button.

     Repeat this for every annotation file you want to include.

     It is also possible to select a complete directory. All `.eaf` files in a selected directory will be included.

  4. Click OK to continue the exporting process; otherwise click Cancel to exit the dialog window without exporting.

  5. If you clicked OK you can save this domain: enter a name and click OK. If you do not want to save the domain click Cancel.

- Create a new domain from an IMDI search:

  1. Click New Domain from IMDI Search...

  2. Browse to and select an IMDI file that has been exported from a metadata search in the standalone IMDI Browser.

  3. Click Open.

  4. You can save this domain: enter a name and click OK. If you do not want to save the domain click Cancel.

After defining a search domain for the first time or when you open the Structured search with a search domain from the previous usage, the following window will open:

Figure 4.18. Search eaf files

As you can see there are three tabs offering different kinds of search:

• Substring Search: finds all annotations in which the search string occurs (see the section called "Substring Search Tab").

• Single Layer Search: finds all annotations or N-grams in which the search string or regular expression occurs, both case sensitive and insensitive and possibly restricted to one (type of) tier (see the section called "Single Layer Search tab").

• Multiple Layer Search: finds annotations in three related tiers. You can use multiple search strings and regular expression and make constraints on duration and time slot as well as constraints on how to search strings are to be combined (see the section called "Multiple Layer Search tab").

Substring Search Tab

This tab offers the simplest search. It just asks for a search string. After entering the search string you can click on Find (or press Enter) to start the search process. This will result in a screen like the one below:

Figure 4.19. Substring Search Results

It shows tokens that contain the search string and some tokens in the context printed in italic typeface. The default number of tokens in the context is three on both sides. When the number of hits exceeds the maximum number the window can contain, you can view the rest of the hits by clicking the < and > button that appear above the list of hits to go back or forward one page. To view an annotation in the timeline view of the main window simply double click it:

Figure 4.20. Hit in transcription

For further investigation of the results the search window offers a context menu that enables you to view the results in other manners and to save the results. To open the context menu right click on one of the results. The menu has the following options:

- Show Frequency view: clicking this option shows both frequency and relative frequency (as a percentage) of the tokens found. The relative frequency is relative to the number of hits.

- Show Frequency view (by frequency): This will display the frequencies, sorted by count.

- Show Alignment view: This option will show you an aligned view of the search results, and there are a number of options you can set. You can change the time scale, hide or show info balloons and set the visible columns (through the context-menu).

- Show hit in transcription: clicking this option shows the transcription in the timeline viewer similar to double clicking an annotation.

- Show Info balloons: by clicking this option you enable ELAN to show you information about a token in an info balloon. This balloon will appear when your mouse cursor is hovering over a token. The information shown in the balloon contains:

  - Transcription file

  - Tier name

  - Tier type

  - Participant

  - Position in tier

- begin time

- end time

- duration

- Context size: this option offers a sub menu that enables you to decrease and increase the context size of the results. Minimum size is 0 and maximum size is 8 tokens.

- Font: click this option to change the font and font size of the results.

- Save hits: when clicking this option, you will be asked to select a directory and enter a file name. The result is a file that contains the following information per token found:

  - Annotation: the annotation token containing the search string.

  - HitPositionInAnnotation: the position of the first character of the search string in the annotation.

  - HitLength: number of characters in the hit

  - HitNumberInAnnotation: if the search string is found more than once in an annotation, this number will give the rank of the hit within the annotation.

  - AnnotationBeginTime: the begin time in ms of an annotation containing the search string.

  - AnnotationEndTime: the end time in ms of an annotation containing the search string.

  - HitPositionInTier: the position of the annotation in a tier.

  - TierName: the name of the tier containing the annotation.

  - TierType: the type of tier containing the annotation.

  - LeftContext: the left context of the annotation.

  - RightContext: the right context of the annotation.

  - TranscriptionName: the path and file name of the transcription in which the annotation is found.

- Save hit statistics: clicking this option lets you save a file that contains hit statistics. The export dialog contains the following options:

  - Separate hit count per hit value: if checked there is a line of statistic for each hit. If not checked, there is line per file.

  - Include file name column.

  - Include file path column.

  - Time format: specify whether the time format should be in milliseconds (ms) or seconds and milliseconds (sec.ms).
  After clicking OK you can enter a file name and click Save to save the statistics file.
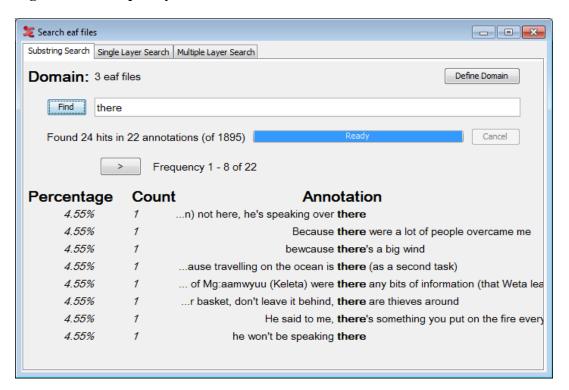
Frequency View

When you are in frequency view or frequency view (by frequency) (Figure 4.21, "Frequency View"), the context menu (right-click) has the following options:

- Show Concordance view: clicking this option will show the annotation results.

- Show hit in transcription: clicking this option shows the transcription in the timeline viewer similar to double clicking an annotation.

- Save frequency info: when clicking this option, you will be asked to select a directory and enter a file name. The result is a file that contains the following information:

  - Annotation

  - Percentage

  - Count

Figure 4.21. Frequency View

Alignment View

The alignment view allows you to view your search results in an aligned time-based view. For detailed information about the Alignment View, see the section called " View search results in Alignment View".

Single Layer Search tab

The Single Layer tab offers a more elaborate search than the Substring Search tab. The first thing that is different from the Substring Search tab is that the Single Layer Search tab has a query history. Clicking the < and > button makes the tab respectively go backward and forward one query. There is also the possibility to save queries, as well as loading previously saved queries.

Figure 4.22. Single Layer Search

Furthermore, the tab offers different modes to restrict the search. The first mode lets you choose the form of the results. There are three options:

- Annotation: the search string is part of or exact match in an annotation.

- N-gram over annotations: each element of the search string (elements are divided by spaces) is part of or exact match in one of several consecutive annotations.

- N-gram within annotation: each element of the search string (elements are divided by spaces) is part of or exact match in one of several consecutive tokens within one annotation.

The following mode offers the straightforward distinction between case sensitive and case insensitive search. The third mode lets the user choose if the element of the first mode should contain the search string (substring match), if the element should exactly match the search string (exact match) or if some regular expression should be used in the match (regular expression).Finally, one can choose to restrict the search to one tier, a tier type or a participant.

Wildcards and negation

When you choose an N-gram to be the form of the result, you can use two more options: a wild card and a negation. The wildcard takes the form of a #-sign. For instance, the search string **the # man** with the mode N-gram over annotations would return three annotations per hit: the first annotation contains `the` (or exactly matches that, if the mode exact match is chosen), the second annotation may contain anything due to the use of the wildcard and the third annotation contains or exactly matches `man`. If the mode N-gram within annotation is chosen, each hit contains one annotation. In this annotation there is a N-gram consisting of three tokens where the first token contains or exactly matches `the`, the second may be anything and the third contains or exactly matches `man`.

If you want to find N-grams where a token matches anything but one string, you can use the negation operator NOT(...), where you can fill in the search string not to be matched on the dots. For instance, the search string **the NOT(strange) man** would return 3-grams in same way as describe above, but the hits where the second annotation or token matches *strange* are left out.

Multiple Layer Search tab

The Multiple Layer Search tab houses the most comprehensive search in ELAN. Similar to the Single Layer Search tab a Query History is kept, enabling the user to go back and forward a query by clicking the < and > respectively. It it also possible to either save or load a previously saved query. To do so, click either the Save query or the Load query button. Queries are saved in XML format.

The two modes case sensitive/case insensitive and substring match/exact match/regular expression are also similar to the second tab. The first new element is the Clear-button. Clicking this button will clear all data of a query.

A new option has been included into the menu containing all the different types of matches (i.e. substring match, exact match, regular expression): **variable match**. As the name says, it has to do with using variables, and it can be used every time you want to search for two or more annotations, contained in two or more different tiers, reporting the same text and/or the same time alignment. See the image below for an example:

Figure 4.23. Variable Match

As you can see in the example, the variable 'X' can match any same value of annotations that meet all other constraints. They are in the same time-frame (overlap) and reside in the same file (the base constraint is Must be in the same file) . In this case 'BONE' is found in the tier 'Gloss RH English' and in 'Gloss LH English', the same for the value '(p-) leg dog'.

It is possible to use more than one variable, e.g. X and Y. This is especially useful in those cases where more than two query fields are filled in.

Figure 4.24. Multiple Variable Match

X and Y can either match different values or the same value. If a variable should be unique, i.e. should never match the same value as any of the other variables, it should be preceded by an exclamation mark, e.g. !Y.

The buttons Minimal Duration and Maximal Duration enables you to constrict the minimal and maximal duration of each result. When you click on one of the buttons, a dialog window appears, e.g.:

Figure 4.25. Minimal Duration

Here you can enter the minimal or maximal duration as the total number of milliseconds or in hours:minutes:seconds.milliseconds. A value of 0 milliseconds or 00:00:00.000 yields as undefined. Searching for annotations with a maximum duration being less then the minimum duration is impossible. Hence, entering conflicting values results in an error message saying that the combination is impossible. After entering a correct duration, it will be displayed in the corresponding button.

The buttons Begin After and End Before give a dialog similar to that of the previous two buttons. They give the possibility to restrict the annotations in the result to begin after a certain time and end before a certain time. Entering a Begin After-time that is greater than the End Before-time or vice versa results in an error message saying it is impossible. After entering a correct time, it will be displayed in the corresponding button.

Search string and constraints

Beneath the buttons discussed above, you will find a table consisting of white and green fields. Search strings are entered in the white fields while a green field between two non-empty white fields must contain a constraint. The fields on one row give the search strings and constraints to be matched by annotations on one tier. The result of having two or more rows in the query table is that the search engine may find annotations on two or more tiers as one hit. Furthermore, it is possible to restrict the search to one (type of) tier for each row by choosing the appropriate option in the pull-down menu on the right of each row.

Let us first take a look at search strings and constraints in one row. If you enter two search strings in two white fields separated by a green field, you must fill in that green field i.e. make a constraint. Clicking the arrow on the green field gives a menu offering the following constraints:

- = N annotations: between the annotations containing the two search strings, there must be exactly N annotations.

- > N annotations: between the annotations containing the two search strings, there must be more than N annotations.

- < N annotations: between the annotations containing the two search strings, there must be less than N annotations.

- = X milliseconds: between the annotations containing the two search strings, there must be exactly X milliseconds.

- > X milliseconds: between the annotations containing the two search strings, there must be more than X milliseconds.

- < X milliseconds: between the annotations containing the two search strings, there must be less than X milliseconds.

- No constraints: there are no constraints.

- Clear: clear the current constraint.

When you click on Find and there is an empty constraint between two non-empty search string fields, you will get an error message. You will also get an error message if there is an empty search string field and constraint fields between two non-empty search string fields.

As we saw earlier the search mechanism on this tab has the possibility to construct a query for two or more tiers (up to eight). Besides the constraints on annotations on a tier, one can also apply constraints on annotations on different tiers. This means that if the search engine has found an annotation that matches a search string on one tier, the engine looks if the search string for another tier can be matched on another tier while considering the constraint that is between the two search strings.
The top down hierarchy of the rows in the query table does not reflect the hierarchy of the tiers in your data. That means, for instance, that search strings and constraints in the upper query table row may be matched by a child tier of the tier that matches search strings and constraints in the middle query table row.

Clicking the arrow in the green field between two search strings gives a menu with the following constraints:

- Fully aligned: the begin time and end time of both annotations are the same:

- Overlap: part of both annotations overlap. This includes the other options Fully aligned, Left overlap, Right overlap, Surrounding and Within.

- Left overlap: the begin time and end time of the annotation matching the lower search string lie *before* the begin time and end time of the annotation matching the upper search string:

- Right overlap: the begin time and end time of the annotation matching the lower search string lie *after* the begin time and end time of the annotation matching the upper search string:

- Surrounding: the begin time of the annotation matching the lower search string lies before the begintime of the annotation matching the upper search string and end time of the annotation matching the lower search string lies after the end time of the annotation matching the upper search string:

- Within: the begin time of the annotation matching the lower search string lies after the begintime of the annotation matching the upper search string and end time of the annotation matching the lower search string lies before the end time of the annotation matching the upper search string:

- No overlap: the begin time of the annotation matching a search string lies after the end time of the annotation matching the other search string:

⊢⎯⊣
⊢⎯⊣
or
⊢⎯⊣
⊢⎯⊣

- No annotation: a special case that retrieves annotations matching the upper search string that have no (overlapping) annotation on the lower tier. It is not possible to enter a lower search string; contrary to the No overlap constraint, which still looks for annotations on the lower tier (namely those that don't overlap), this constraint really looks for *no annotation* in the timespan of the upper annotation (*empty slots*). The user interface allows specifying constraints on lower levels and to the left and right of this constraint, but the behavior in that case is undefined!

- begin time - begin time = X milliseconds: the begin time of the annotations matching the upper search string must lie exactly X milliseconds before the begin time of the annotation matching the lower search string.

- begin time - begin time < X milliseconds: the begin time of the annotations matching the upper search string must lie less than X milliseconds before the begin time of the annotation matching the lower search string.

- begin time - begin time > X milliseconds: the begin time of the annotations matching the upper search string must lie more than X milliseconds before the begin time of the annotation matching the lower search string.

- begin time - end time = X milliseconds: the end time of the annotations matching the upper search string must lie exactly X milliseconds before the begin time of the annotation matching the lower search string.

- begin time - end time < X milliseconds: the end time of the annotations matching the upper search string must lie less than X milliseconds before the begin time of the annotation matching the lower search string.

- begin time - end time > X milliseconds: the end time of the annotations matching the upper search string must lie more than X milliseconds before the begin time of the annotation matching the lower search string.

- end time - begin time = X milliseconds: the begin time of the annotations matching the upper search string must lie exactly X milliseconds before the end time of the annotation matching the lower search string.

- end time - begin time < X milliseconds: the begin time of the annotations matching the upper search string must lie less than X milliseconds before the end time of the annotation matching the lower search string.

- end time - begin time > X milliseconds: the begin time of the annotations matching the upper search string must lie more than X milliseconds before the end time of the annotation matching the lower search string.

- end time - end time = X milliseconds: the end time of the annotations matching the upper search string must lie exactly X milliseconds before the end time of the annotation matching the lower search string.

- end time - end time < X milliseconds: the end time of the annotations matching the upper search string must lie less than X milliseconds before the end time of the annotation matching the lower search string.

- end time - end time > X milliseconds: the end time of the annotations matching the upper search string must lie more than X milliseconds before the end time of the annotation matching the lower search string.

• No constraint: there are no constraints.

• Clear: clear the current constraint.

An example of a Multiple Layer Search with constraints is shown below:

Figure 4.26. Multiple Layer query

As you can see the tiers in the result are indicated by #1 and #2, corresponding to the first and second query table row respectively. The annotations in a tier are surrounded by vertical bars indicating their start and end.

It is possible to add or remove columns and/or layers to your search query. To do so, click the respective button:

• Fewer Columns

• More Columns

• Fewer Layers

• More Layers

It is also possible to hide the query once there are search results. This allows you to see more query results within a single window. This can be helpful when using the Alignment View the section called " View search results in Alignment View".

also illustrates what to do if you would like to use both Exact match and Substring match in one query: use the Regular expression. In places where you would like to have an exact match use the ^ and $ signs to match the beginning and end of a string (e.g. ^of$) otherwise just enter a word for the substring match.

The figure also show how to use a wildcard to match anything. Instead of using the **#** as in the Single Layer Search, you can use the regular expression **.+** to indicate any character (the dot) one or more times (the plus). See also Appendix A, *REGULAR EXPRESSION SEARCH* for more on regular expressions. The NOT(...) construction on the other hand can be used in the Multiple Layer Search in the same way as describe in the section called "Single Layer Search tab".

One final but not less important remark concerns the placing of more and less restrictive search strings. Figure 4.26, "Multiple Layer query" shows a very restrictive search string in the upper row: **^n$**. The less restrictive, or should we say non-restrictive, search string .+ is in the middle row. As we saw earlier, the hierarchy of the rows in the query does not reflect the hierarchy in the data. That means that the search string **^n$** could also be placed in the lower row and not affect the outcome of the search. While this is perfectly true, we advise you to place restrictive search strings in the left most field on the upper most row possible and the least restrictive search string in the right most field of the lowest row possible. The reason for this is the order in which the search engine considers the search strings in the query. If it finds a restrictive search string it can filter out all the other possibilities, but if it finds a less restrictive search string it has to consider all the matches of this search string. In the example of Figure 4.26, "Multiple Layer query" it is clear that if **^n$** is in the bottom row, the search engine first considers all annotations matching **.+** which is in fact *all* annotations in the search domain. Because of this, the search takes much more time than if **^n$** was in the upper row.

View search results in Alignment View

From the context-menu (right-click the search results), you van view query results from the Multiple Layer Search in Alignment View:

Figure 4.27. Alignment View

There are a number of options you can set when viewing the query results. Firstly, you can adjust the time scale of the results:

• 1 sec / 2sec / 5sec / 10 sec / 15 sec / 20 sec / User defined / Scale to fit.

   When choosing 'Scale to fit', every query result will be scaled to fit the window, which means the time scale for every result will differ.

   There is also the possibility to hide the alignment time scale altogether. To do so, go to the context-menu (right-click) and uncheck Show alignment timesby clicking on it.

You can set the visible columns to the right of the query results through the context-menu (right-click anywhere in the results). You can show or hide the following columns:

• Tier Type

• Annotator

- Participant

- Begin Time

- End Time

- Duration

The blue bars above every query result graphically show the duration of each annotation and the position of the annotations with respect to each other.

There are also two indicators visible, depending on the length of the query result and the setting of the time scale. These indicators are either red or green.

A green indicator means that the annotation does not fit in the current time scale. In the example above, the bottom annotation '*and then you see um a man in maybe his fifties*' has a duration of 5.060 seconds. The time scale is set to 1 second, so 4.060 seconds are outside the current view.

The red indicator means that the annotation in the query result starts outside of the current time scale. The top annotation '*fifties*' overlaps the bottom annotation, but starts at 9.177 seconds. This causes it not to be visible in the current time scale, which is set to display 1 second. You would need to set the time scale to 10 seconds to see both annotations visualised completely (as the blue bars) and how they overlap.

Chapter 5. Help

There is a help Menu in ELAN which offers the following options:

- Help Contents...

  Select this menu option to start the in-built help of ELAN.

- Check for Updates...

  This menu checks for any new updates of ELAN. If available it gives an overview about the changes made in the new version and leads you to the download page from where you can download the latest version. There is also a automatic check for updates option in ELAN. To set that option, go to Edit -> Preferences -> Edit Preferences... and go to Preferences options (see the section called "Editing preferences") . If check for automatic updates is set true, then elan checks for new updates one in a month and intimates when an update is available.

- Visit the ELAN website

- This menu has a sub menu, which will get to any one of these options in the website : Release History, Download Page, Forum and also allows you to Subscribe to the mailing list.

- About ELAN

Chapter 6. Reference guide

In this reference guide, you find concise descriptions of the mouse options (1), the menu item (2), and the shortcut keys (3). In addition, a brief definition of key concepts is provided (4).

The mouse options

Resizing Viewers

- Click on the up/down arrows in the split-pane to increase/decrease the size of the corresponding Viewer.

- Go with the mouse to the split-pane and move it up/down to increase/decrease the size of the corresponding Viewer.

Rearranging the order of tiers

- Drag a tier label from its place in the Timeline or Interlinear Viewer to any other place in the Timeline Viewer.

Make a tier the active tier

- Double click on a tier label.

- Right click in the Timeline or Interlinear Viewer and choose Active Tier.

Accessing points in time

- Click on any place in the Timeline, Waveform or Annotation Density Viewer to jump to this point in time.

- Click on one of the media controls to go forward/backward in time.

- Click on the time code box and enter a time code to jump to this point in time.

Playing a document

- Click on the Play/Pause icon to start or pause the playback.

- Use the Rate slider (in the Controls tab) to increase/decrease the playback rate.

Making a selection

- Click with the mouse on any point in the Timeline or Waveform Viewer, and drag it to another point to select a time interval. Or click on an annotation in the Grid Viewer and drag the mouse to another annotation.

- Click on an annotation in the Timeline or Interlinear Viewer to select it.

Adding/modifying an annotation

- Double-click on an annotation or selection (at about the height of the tier where you want to enter an annotation). The Inline Edit box appears and you can enter the annotation.

The shortcut keys

The following tables list several of the most common default keyboard shortcut combinations. These combinations can be customized in the Edit Keyboard Shortcuts window (see the section called "Editing shortcuts"). In most (but not all) cases below, the **CTRL** key can be replaced by the **CMD** key on macOS systems.

File options

CTRL+S	Saves the current project
CTRL+SHIFT+S	Save as…
CTRL+SHIFT+ALT+S	Save as template
CTRL+W	Close the current window
CTRL+Q	Exit the application
CTRL+O	Open a document
CTRL+N	Create a new document
CTRL+P	Prints the current document
CTRL+SHIFT+P	Page Setup
CTRL+ALT+P	Print Preview
SHIFT+UP	Activate previous window
SHIFT+DOWN	Activate next window

Playing a document & making selections

CTRL+SPACE	Starts or stops the Media Player
CTRL+B	Go to the beginning of the video/audio fragment
CTRL+PAGE UP	Go to the previous scroll view
SHIFT+LEFT	Go back one second
CTRL+LEFT	Go back one frame
CTRL+SHIFT+LEFT	Go back one "pixel" on the annotation density viewer
CTRL+SHIFT+RIGHT	Go to the next "pixel" on the annotation density viewer
CTRL+RIGHT	Go to the next frame
SHIFT+RIGHT	Go to the next second
CTRL+PAGE DOWN	Go to the next scroll view
CTRL+E	Go to the end of the media fragment
CTRL+ALT+R	Switch to predefined playback rate
CTRL+SHIFT+R	Switch to predefined volume
CTRL+G	Go to time

Working with annotations

CTRL+ALT+N	Opens the Inline Edit box for inserting a new annotation (only available if a time interval or annotation is selected)

CTRL+ALT+M	Opens the Inline Edit box for modifying an annotation (only available if an annotation is selected)
ALT+SHIFT+M	Opens the Local Data Category Selection for associating the annotation with a data category of the ISO Data Category Registry.
CTRL+Delete	Deletes the value of the selected annotation
CTRL+ALT+D	Deletes an annotation (only available if an annotation is selected)
CTRL+SHIFT+N	New annotation before
CTRL+SHIFT+N	New annotation after
CTRL+SHIFT+T	Add new tier type
CTRL+ALT+L	Edit linked file
CTRL+SHIFT+C	Edit controlled vocabulary
ALT+LEFT	Go to previous Annotation
ALT+RIGHT	Go to next Annotation
ALT+UP	Go annotation upward
ALT+DOWN	Go annotation downward
CTRL+T	Add new tier
CTRL+ALT+T	Delete tier
CTRL+UP	Set previous tier active
CTRL+DOWN	Set next tier active
CTRL+ALT+SHIFT	Show empty slots in multitier viewer
CTRL+ENTER	Saves the changes made to the Inline Edit box or the Edit annotation box
ESC	(from within the Inline Edit box or the Edit annotation box): exits the box without saving the changes
ESC	(in all other contexts): deselects a selection
SHIFT+ENTER	Allows you to switch between the Inline Edit box and the Edit annotation box
ALT+click	Allows you to select multiple annotations on multiple tiers.
SHIFT+E	Opens the Inline Edit box for the active annotation

Copy and paste

CTRL+C	Copy Annotation
CTRL+ALT+C	Copy Annotation Group
CTRL+D	Duplicate Annotation
CTRL+V	Paste Annotation
CTRL+ALT+V	Paste Annotation Group
CTRL+SHIFT+V	Paste Annotation Here
CTRL+ALT+SHIFT+V	Paste Annotation Group Here
CTRL+ALT+G	Copy the current time

Working with selections

SHIFT+SPACE	Play selection
CTRL+SHIFT+SPACE	Play around selection
ALT+SHIFT+C	Clear selection
CTRL+SHIFT+Z	Clear selection and cancel selection mode
CTRL+/ or CTRL+SHIFT+K	Move the crosshair to the right or left of the selection
ALT+CTRL+/	Move the crosshair to the center of the selection
CTRL+K	Toggle selection mode
CTRL+L	Toggle loop mode
SHIFT+A	Put the selection in the center of the Timeline Viewer

Searching

CTRL+F	Find
CTRL+SHIFT+F	Search in multiple eaf files

General

CTRL+Z	Undo
CTRL+Y	Redo

Key concepts of ELAN

Annotation file (*.eaf)

An annotation file is the document that contains all the information about tiers (their attributes and dependency relations), annotations and time alignments.

Media file (*.mpg, *.mov, *.wav etc.)

A media file contains the digitized video/audio (e.g. `*.mpg`) or only the audio (`*.wav`) data. It determines the time axis to which the annotation file (`*.eaf`) is linked.

Annotation

An annotation is any type of text (e.g. a transcription, a translation, coding, etc.) that is entered on a tier. It is assigned to a selected time interval of the video/audio file (e.g., to the time interval corresponding to the utterance of a speaker) or to an annotation on another tier (e.g., a translation is assigned to an orthographic transcription).

Tier

A tier is a set of annotations that share the same characteristics, e.g., one tier containing the orthographic transcription, or another tier containing the free translation.

A tier can be 'independent' and 'time-alignable', in which case it is directly linked to a time interval of the media file (e.g., the 'orthographic transcription' tier). Or it can be 'referring', in which case it

is linked to another tier, its so-called parent tier (e.g., the 'orthographic transcription' tier is a parent tier to the 'free translation' tier). The referring tier shares its time alignment with its parent tier. Some referring tiers can be assigned to the time axis, but only to an interval that is contained within the interval of their parent annotation.

It is possible to build nested hierarchies, e.g., the 'orthographic transcription' tier is the parent tier to a 'word' tier, and the 'word' tier is the parent tier to a 'morpheme break' tier.

Tiers are assigned to tier types, which specify certain constraints. The following constraints exist: None (independent, time-alignable tiers), Time Subdivision (the annotation on the referring tier can be subdivided and linked to the time axis), Symbolic Subdivision (the annotation on the referring tier can be subdivided, but not linked to the time axis), Symbolic Association (one annotation on the referring tier corresponds to exactly one annotation on the parent tier).

Appendix A. REGULAR EXPRESSION SEARCH

Brief Background

A regular expression consists of a character string where some characters are given special meaning with regard to pattern matching. Regular expressions have been in use from the early days of computing, and provide a powerful and efficient way to parse, interpret and search and replace text within an application.

Supported Syntax

📋 Note

Source: https://docs.oracle.com/en/java/javase/14/docs/api/java.base/java/util/regex/Pattern.html

Table A.1. Characters

x	The character x
`\\`	The backslash character
`\0n`	The character with octal value 0n (0 <= n <= 7)
`\0nn`	The character with octal value 0nn (0 <= n <= 7)
`\0mnn`	The character with octal value 0mnn (0 <= m <= 3, 0 <= n <= 7)
`\xhh`	The character with hexadecimal value 0xhh
`\uhhhh`	The character with hexadecimal value 0xhhhh
`\t`	The tab character ('`\u0009`')
`\n`	The newline (line feed) character ('`\u000A`')
`\r`	The carriage-return character ('`\u000D`')
`\f`	The form-feed character ('`\u000C`')
`\a`	The alert (bell) character ('`\u0007`')
`\e`	The escape character ('`\u001B`')
`\cx`	The control character corresponding to x

Table A.2. Character classes

`[abc]`	a, b, or c (simple class)
`[^abc]`	Any character except a, b, or c (negation)
`[a-zA-Z]`	a through z or A through Z, inclusive (range)
`[a-d[m-p]]`	a through d, or m through p: `[a-dm-p]` (union)
`[a-z&&[def]]`	d, e, or f (intersection)
`[a-z&&[^bc]]`	a through z, except for b and c: `[ad-z]` (subtraction)
`[a-z&&[^m-p]]`	a through z, and not m through p: `[a-lq-z]`(subtraction)

Table A.3. Predefined character classes

.	Any character (may or may not match line terminators [http://docs.oracle.com/javase/7/docs/api/java/util/regex/Pattern.html#lt])
\d	A digit: `[0-9]`
\D	A non-digit: `[^0-9]`
\s	A whitespace character: `[ \t\n\x0B\f\r]`
\S	A non-whitespace character: `[^\s]`
\w	A word character: `[a-zA-Z_0-9]`
\W	A non-word character: `[^\w]`

Table A.4. POSIX character classes (US-ASCII only)

\p{Lower}	A lower-case alphabetic character: `[a-z]`
\p{Upper}	An upper-case alphabetic character:`[A-Z]`
\p{ASCII}	All ASCII:`[\x00-\x7F]`
\p{Alpha}	An alphabetic character:`[\p{Lower}\p{Upper}]`
\p{Digit}	A decimal digit: `[0-9]`
\p{Alnum}	An alphanumeric character:`[\p{Alpha}\p{Digit}]`
\p{Punct}	Punctuation: One of `!"#$%&'()*+,-./:;<=>?@[\]^_`{|}~`
\p{Graph}	A visible character: `[\p{Alnum}\p{Punct}]`
\p{Print}	A printable character: `[\p{Graph}]`
\p{Blank}	A space or a tab: `[ \t]`
\p{Cntrl}	A control character: `[\x00-\x1F\x7F]`
\p{XDigit}	A hexadecimal digit: `[0-9a-fA-F]`
\p{Space}	A whitespace character: `[ \t\n\x0B\f\r]`

Table A.5. Classes for Unicode blocks and categories

\p{InGreek}	A character in the Greek block (simple block [http://docs.oracle.com/javase/7/docs/api/java/util/regex/Pattern.html#ubc])
\p{Lu}	An uppercase letter (simple category [http://docs.oracle.com/javase/7/docs/api/java/util/regex/Pattern.html#ubc])
\p{Sc}	A currency symbol
\P{InGreek}	Any character except one in the Greek block (negation)
[\p{L}&&[^\p{Lu}]]	Any letter except an uppercase letter (subtraction)

Table A.6. Boundary matchers

^	The beginning of a line
$	The end of a line

`\b`	A word boundary
`\B`	A non-word boundary
`\A`	The beginning of the input
`\G`	The end of the previous match
`\Z`	The end of the input but for the final terminator [http://docs.oracle.com/javase/7/docs/ api/java/util/regex/Pattern.html#lt], if any
`\z`	The end of the input

Table A.7. Greedy quantifiers

`X?`	X, once or not at all
`X*`	X, zero or more times
`X+`	X, one or more times
`X{n}`	X, exactly n times
`X{n,}`	X, at least n times
`X{n,m}`	X, at least n but not more than m times

Table A.8. Reluctant quantifiers

`X??`	X, once or not at all
`X*?`	X, zero or more times
`X+?`	X, one or more times
`X{n}?`	X, exactly n times
`X{n,}?`	X, at least n times
`X{n,m}?`	X, at least n but not more than m times

Table A.9. Possessive quantifiers

`X?+`	X, once or not at all
`X*+`	X, zero or more times
`X++`	X, one or more times
`X{n}+`	X, exactly n times
`X{n,}+`	X, at least n times
`X{n,m}+`	X, at least n but not more than m times

Table A.10. Logical operators

`XY`	X followed by Y
`X|Y`	Either X or Y
`(X)`	X, as a capturing group [http://docs.oracle.com/ javase/7/docs/api/java/util/regex/ Pattern.html#cg]

Table A.11. Back references

`\n`	Whatever the nth capturing group [http:// docs.oracle.com/javase/7/docs/api/java/util/ regex/Pattern.html#cg] matched

Table A.12. Quotation

\	Nothing, but quotes the following character
\Q	Nothing, but quotes all characters until \E
\E	Nothing, but ends quoting started by \Q

Table A.13. Special constructs (non-capturing)

(?:X)	X, as a non-capturing group
(?idmsux-idmsux)	Nothing, but turns match flags on - off
(?idmsux-idmsux:X)	X, as a non-capturing group [http://docs.oracle.com/javase/7/docs/api/java/util/regex/Pattern.html#cg] with the given flags on - off
(?=X)	X, via zero-width positive lookahead
(?!X)	X, via zero-width negative lookahead
(?<=X)	X, via zero-width positive lookbehind
(?<!X)	X, via zero-width negative lookbehind
(?>X)	X, as an independent, non-capturing group

Backslashes, escapes, and quoting

The backslash character ('\') serves to introduce escaped constructs, as defined in the table above, as well as to quote characters that otherwise would be interpreted as unescaped constructs. Thus the expression \\ matches a single backslash and \{ matches a left brace.

It is an error to use a backslash prior to any alphabetic character that does not denote an escaped construct; these are reserved for future extensions to the regular-expression language. A backslash may be used prior to a non-alphabetic character regardless of whether that character is part of an unescaped construct.

Backslashes within string literals in Java source code are interpreted as required by the Java Language Specification [http://java.sun.com/docs/books/jls/second_edition/html/j.title.doc.html] as either Unicode escapes [http://java.sun.com/docs/books/jls/second_edition/html/lexical.doc.html#100850] or other character escapes [http://java.sun.com/docs/books/jls/second_edition/html/lexical.doc.html#101089]. It is therefore necessary to double backslashes in string literals that represent regular expressions to protect them from interpretation by the Java byte code compiler. The string literal "\b", for example, matches a single backspace character when interpreted as a regular expression, while "\\b" matches a word boundary. The string literal "\(hello\)" is illegal and leads to a compile-time error; in order to match the string (hello) the string literal "\\(hello\\)" must be used.

Character Classes

Character classes may appear within other character classes, and may be composed by the union operator (implicit) and the intersection operator (&&). The union operator denotes a class that contains every character that is in at least one of its operand classes. The intersection operator denotes a class that contains every character that is in both of its operand classes.

The precedence of character-class operators is as follows, from highest to lowest:

1	Literal escape	\x
2	Grouping	[...]

3	Range	`a-z`
4	Union	`[a-e][i-u]`
5	Intersection	`[a-z&&[aeiou]]`

Note that the set of meta characters that is in effect inside a character class is different from the set that is outside a character class. For instance, the regular expression . loses its special meaning inside a character class, while the expression - becomes a range forming metacharacter.

Line terminators

A *line terminator* is a one- or two-character sequence that marks the end of a line of the input character sequence. The following are recognized as line terminators:

- A newline (line feed) character ('\n'),

- A carriage-return character followed immediately by a newline character ("\r\n"),

- A standalone carriage-return character ('\r'),

- A next-line character ('\u0085'),

- A line-separator character ('\u2028'), or

- A paragraph-separator character ('\u2029').

If UNIX_LINES [http://docs.oracle.com/javase/7/docs/api/java/util/regex/ Pattern.html#UNIX_LINES] mode is activated, then the only line terminators recognized are newline characters.

The regular expression . matches any character except a line terminator unless the DOTALL [http:// docs.oracle.com/javase/7/docs/api/java/util/regex/Pattern.html#DOTALL] flag is specified.

By default, the regular expressions ^ and $ ignore line terminators and only match at the beginning and the end, respectively, of the entire input sequence. If MULTILINE [http://docs.oracle.com/javase/7/ docs/api/java/util/regex/Pattern.html#MULTILINE] mode is activated then ^ matches at the beginning of input and after any line terminator except at the end of input. When in MULTILINE [http:// docs.oracle.com/javase/7/docs/api/java/util/regex/Pattern.html#MULTILINE] mode $ matches just before a line terminator or the end of the input sequence.

Groups and capturing

Capturing groups are numbered by counting their opening parentheses from left to right. In the expression ((A)(B(C))), for example, there are four such groups:

1	`((A)(B(C)))`
2	`(A)`
3	`(B(C))`
4	`(C)`

Group zero always stands for the entire expression.

Capturing groups are so named because, during a match, each subsequence of the input sequence that matches such a group is saved. The captured subsequence may be used later in the expression, via a back reference, and may also be retrieved from the matcher once the match operation is complete.

The captured input associated with a group is always the subsequence that the group most recently matched. If a group is evaluated a second time because of quantification then its previously-captured

value, if any, will be retained if the second evaluation fails. Matching the string "aba" against the expression (a(b)?)+, for example, leaves group two set to "b". All captured input is discarded at the beginning of each match.

Groups beginning with (? are pure, *non-capturing* groups that do not capture text and do not count towards the group total.

Unicode support

This class follows *Unicode Technical Report #18: Unicode Regular Expression Guidelines [http:// www.unicode.org/unicode/reports/tr18/]*, implementing its second level of support though with a slightly different concrete syntax.

Unicode escape sequences such as \u2014 in Java source code are processed as described in paragraph 3.3 [https://docs.oracle.com/javase/specs/jls/se14/html/jls-3.html#jls-3.3] of the Java Language Specification. Such escape sequences are also implemented directly by the regular-expression parser so that Unicode escapes can be used in expressions that are read from files or from the keyboard. Thus the strings "\u2014" and "\\u2014", while not equal, compile into the same pattern, which matches the character with hexadecimal value 0x2014.

Unicode blocks and categories are written with the \p and \P constructs as in Perl. \p{prop} matches if the input has the property prop, while \P{prop} does not match if the input has that property. Blocks are specified with the prefix In, as in InMongolian. Categories may be specified with the optional prefix Is: Both \p{L} and \p{IsL} denote the category of Unicode letters. Blocks and categories can be used both inside and outside of a character class.

The supported blocks and categories are those of *The Unicode Standard, Version 3.0 [http://www.unicode.org/unicode/standard/standard.html]*. The block names are those defined in Chapter 14 and in the file Blocks-3.txt [http://www.unicode.org/Public/3.0-Update/ Blocks-3.txt] of the Unicode Character Database [http://www.unicode.org/Public/3.0-Update/ UnicodeCharacterDatabase-3.0.0.html] except that the spaces are removed; "Basic Latin", for example, becomes "BasicLatin". The category names are those defined in table 4-5 of the Standard (p. 88), both normative and informative.

Comparison to Perl 5

Perl constructs not supported by this class:

The conditional constructs (?{X}) and (?(condition)X|Y),

The embedded code constructs (?{code}) and (??{code}),

The embedded comment syntax (?#comment), and

The preprocessing operations \l \u, \L, and \U.

Constructs supported by this class but not by Perl:

Possessive quantifiers, which greedily match as much as they can and do not back off, even when doing so would allow the overall match to succeed.

Character-class union and intersection as described above [http://docs.oracle.com/javase/7/docs/api/ java/util/regex/Pattern.html#cc].

Notable differences from Perl:

In Perl, \1 through \9 are always interpreted as back references; a backslash-escaped number greater than 9 is treated as a back reference if at least that many sub-expressions exist, otherwise it is interpreted, if possible, as an octal escape. In this class octal escapes must always begin with a zero. In this class, \1 through \9 are always interpreted as back references, and a larger number is accepted

as a back reference if at least that many sub-expressions exist at that point in the regular expression, otherwise the parser will drop digits until the number is smaller or equal to the existing number of groups or it is one digit.

Perl uses the g flag to request a match that resumes where the last match left off. This functionality is provided implicitly by the Matcher [http://docs.oracle.com/javase/7/docs/api/java/util/regex/Matcher.html] class: Repeated invocations of the find [http://docs.oracle.com/javase/7/docs/api/java/util/regex/Matcher.html#find%28%29] method will resume where the last match left off, unless the matcher is reset.

In Perl, embedded flags at the top level of an expression affect the whole expression. In this class, embedded flags always take effect at the point at which they appear, whether they are at the top level or within a group; in the latter case, flags are restored at the end of the group just as in Perl.

Perl is forgiving about malformed matching constructs, as in the expression *a, as well as dangling brackets, as in the expression abc], and treats them as literals. This class also accepts dangling brackets but is strict about dangling metacharacters like +, ? and *, and will throw a PatternSyntaxException [http://docs.oracle.com/javase/7/docs/api/java/util/regex/PatternSyntaxException.html] if it encounters them.

For a more precise description of the behavior of regular expression constructs, please see *Mastering Regular Expressions*, 2nd Edition, Jeffrey E. F. Friedl, O'Reilly and Associates, 2002 [http://www.oreilly.com/catalog/regex2/].