ELAN - Linguistic Annotator

version 4.4.0

This manual was last updated on 2012-07-20

The latest version can be downloaded from: http://tla.mpi.nl/tools/tla-tools/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: Jeroen Geerts

ELAN - Linguistic Annotator: version 4.4.0

version 4.4.0

This manual was last updated on 2012-07-20

The latest version can be downloaded from: http://tla.mpi.nl/tools/tla-tools/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: Jeroen Geerts

Table of Contents

Introduction vii
1. Notation Conventions vii
2. Recent changes to this manual viii
I. Getting started 9
1. Installation 11
2. Selecting a user interface language 12
3. Consulting the manual and website 13
II. USER'S GUIDE 14
4. ELAN documents 15
4.1. Basic Information: Media Files and Annotation Files 15
4.2. How to manage documents 16
4.2.1. Creating a new document 18
4.2.2. Set the author of a document 19
4.2.3. Synchronizing video files 19
4.2.4. Opening an existing document 22
4.2.5. Re-open recently accessed files 23
4.2.6. Switching between multiple files 24
4.2.7. Saving a document 24
4.2.8. Saving a selection as .eaf file 24
4.2.9. Merging transcriptions 26
4.2.10. Saving a template 28
4.2.11. Changing the links to media files 29
4.2.12. Creating automatic backups 30
4.2.13. Printing 32
4.2.14. Configuring the page settings (page setup) 32
4.2.15. Previewing the printed pages 33
4.2.16. Opening a wave file in Praat 38
4.2.17. Exporting a selection to a wave file with Praat 39
4.2.18. Closing a file 39
4.2.19. Exiting ELAN 39
4.3. Import and Export options 40
4.3.1. Import from 40
4.3.2. Export as 52
4.4. Screen display and navigation through a document 77
4.4.1. Basic Information: The ELAN window 77
4.4.2. The Video Viewer 78
4.4.3. The Annotation Density Viewer 80
4.4.4. The Waveform Viewer 81
4.4.5. The Subtitle Viewer 83
4.4.6. The Grid Viewer 84
4.4.7. The Text Viewer 85
4.4.8. The Lexicon Viewer 86
4.4.9. The Metadata Viewer 86
4.4.10. Audio And Video Recognizer 86
4.4.11. The Timeline Viewer and the Interlinear Viewer 87
4.4.12. The Timeseries Viewer 89
4.4.13. The Menu options 94
4.4.14. The Media Player options 95
4.4.15. The color coding 97
4.4.16. Increasing/decreasing the size of the ELAN window 98
4.4.17. Switching Viewers on/off 98
4.4.18. Show / Hide Specific Viewers 98
4.4.19. Increasing/decreasing the size of Viewers 99
4.4.20. Switching tiers on/off 100
4.4.21. Rearranging the order of tiers 101

4.4.22. View tiers by Type/Participant/Annotator 102
4.4.23. Sorting tiers 103
4.4.24. Changing the time zoom 103
4.4.25. Changing the vertical (intensity) zoom 104
4.4.26. Changing the font size 105
4.4.27. Font Browser 106
4.4.28. Editing preferences 107
4.4.29. Importing and exporting preferences 113
4.4.30. Editing shortcuts 113
4.4.31. How to copy current time of the media 114
4.5. How to navigate through a document 114
4.5.1. Accessing points in time 115
4.5.2. Going to the borders of a selection 117
4.5.3. Jumping step by step through a document 117
4.5.4. Viewing a list of annotations within one tier (Grid Viewer) 117
4.5.5. Jumping from annotation to annotation (Interlinear Viewer) 121
4.6. How to play a document 121
4.6.1. Playing a document 121
4.6.2. Playing a selection 122
4.6.3. Playing around a selection 122
4.6.4. Playing in slow motion 123
4.6.5. Changing playback rate and volume via the keyboard 124
4.7. Multiple File Operations 124
4.7.1. How to select multiple files 124
4.7.2. Multiple file import options 125
4.7.3. Multiple file export options 129
4.7.4. Editing multiple files 135
5. Annotations 139
5.1. Basic Information: Annotations, tiers and linguistic types 139
5.2. Data categories 143
5.3. How to define a linguistic type 145
5.3.1. Adding new linguistic types 145
5.3.2. Creating linguistic types 146
5.3.3. Changing linguistic types 148
5.3.4. Deleting linguistic types 149
5.3.5. Importing linguistic types 150
5.3.6. Associating a linguistic type with a Data Category 150
5.4. How to define a tier and its attributes 151
5.4.1. Adding new tiers 151
5.4.2. Importing tiers 153
5.4.3. Creating a tier using the Audio Recognizer 154
5.4.4. Changing tier attributes 155
5.4.5. Deleting tiers 158
5.4.6. Merging tiers 162
5.4.7. Merging tier groups 163
5.4.8. Changing the parent of a tier 163
5.4.9. Add New Participant 165
5.5. Controlled Vocabularies 166
5.5.1. The creation of a CV 166
5.5.2. Associating a CV entry with a Data Category 168
5.5.3. Importing an existing CV 169
5.5.4. Using an External CV 169
5.5.5. Exporting a CV 170
5.5.6. Using CV's 170
5.5.7. Associating a linguistic type with Controlled Vocabularies 170
5.6. Consulting a Lexicon 171
5.6.1. Setting up a Lexicon Service 171
5.6.2. Set the Lexicon Service and entry field information for a Linguistic Type.... 175

5.6.3. Perform a Lexicon Lookup 176
5.7. How to make a selection 177
5.7.1. Making a selection on an independent tier 177
5.7.2. Using the selection controls 179
5.7.3. Creating annotation units without gaps 181
5.7.4. Making and saving a selection on a referring tier 183
5.7.5. Deselecting a selection 184
5.7.6. Changing the boundaries of an existing selection and annotation 184
5.7.7. Shifting annotations 186
5.7.8. Activating and deactivating the Bulldozer mode or Shift mode 187
5.8. How to enter annotations 189
5.8.1. Entering annotations (into the Inline Edit box) 189
5.8.2. Entering annotations (into the Edit Annotation box) 191
5.8.3. Entering annotations from a Controlled Vocabulary 194
5.8.4. Entering annotations before/after other annotations 194
5.8.5. Creating depending annotations for a active annotation 196
5.8.6. Creating annotations on dependent tiers 196
5.8.7. Creating annotations from gaps 196
5.8.8. Creating regular annotations 197
5.8.9. Duplicating annotation values 198
5.8.10. Snap Annotations 198
5.8.11. Automatic labelling and numbering 198
5.8.12. Modifying the content of annotations 200
5.8.13. Moving annotations to another tier 200
5.8.14. Changing the case of Annotations 200
5.8.15. Deleting annotation values 200
5.8.16. Deleting annotations 201
5.8.17. Deleting annotations and annotation values on multiple tiers 201
5.8.18. Split Annotations 202
5.8.19. Merging annotations 203
5.8.20. Entering annotations in different character sets 205
5.8.21. Associating an annotation with a Data Category 210
5.9. How to create annotation units on dependent tiers 211
5.10. How to filter and copy tiers 213
5.10.1. Filtering tiers 213
5.10.2. Copying tiers 214
5.11. How to detect annotation overlaps 215
5.11.1. Creating annotations from overlaps 215
5.11.2. Comparing annotators 220
5.12. Create Annotation By subtraction 220
5.13. Copy and paste annotations 222
5.13.1. Copy annotation 222
5.13.2. Paste annotation 222
5.13.3. Paste annotation here 223
5.13.4. Copy and paste annotation groups 223
5.13.5. Duplicate annotation 224
5.13.6. Synopsis 225
5.14. How to undo an action 225
5.15. Information about annotations 226
5.15.1. Annotations Spreadsheet 226
5.15.2. Annotations Statistics 226
6. Working modes 229
6.1. Annotation mode 229
6.2. Synchronization mode 229
6.3. Transcription mode 229
6.3.1. Selecting the tier types for transcription 230
6.3.2. Selecting tiers for transcription 232
6.3.3. Show or Hide tiers 233

6.3.4. Making tiers non-editable 234
6.3.5. Entering text for a annotation 234
6.3.6. Merging annotations 238
6.3.7. Modifying annotations 238
6.3.8. Shortcut keys used in this mode 238
6.4. Segmentation mode 239
6.4.1. Configuring the behavior of the segmentation keystroke 240
6.4.2. Editing annotations 240
6.4.3. Merging annotations and splitting annotations 241
6.4.4. Step-and-Repeat playback mode 241
6.4.5. Keyboard shortcuts 242
7. Searching 243
7.1. Searching in a single annotation file 243
7.1.1. Advanced searching: an example 247
7.2. Displaying search results 248
7.2.1. The search history 249
7.2.2. Refining a search 250
7.2.3. Saving and loading queries 251
7.2.4. Exporting the search results 252
7.2.5. Search and replace 254
7.3. Jumping from the search results to annotations 255
7.4. Searching through multiple annotation files 256
7.5. Structured search through multiple annotation files 259
7.5.1. Substring Search Tab 261
7.5.2. Single Layer Search tab 264
7.5.3. Multiple Layer Search tab 266
7.6. Search and replace in multiple annotation files 270
8. Help 272
9. Reference guide 273
9.1. The mouse options 273
9.1.1. Resizing Viewers 273
9.1.2. Rearranging the order of tiers 273
9.1.3. Make a tier the active tier 273
9.1.4. Accessing points in time 273
9.1.5. Playing a document 273
9.1.6. Making a selection 273
9.1.7. Adding/modifying an annotation 273
9.2. The shortcut keys 274
9.2.1. File options 274
9.2.2. Playing a document & making selections 274
9.2.3. Working with annotations 274
9.2.4. Copy and paste 275
9.2.5. Working with selections 275
9.2.6. Searching 276
9.2.7. General 276
9.3. Key concepts of ELAN 276
9.3.1. Annotation file (*.eaf) 276
9.3.2. Media file (*.mpg, *.mov, *.wav etc.) 276
9.3.3. Annotation 276
9.3.4. Tier 276
A. REGULAR EXPRESSION SEARCH 278
A.1. Character Classes 281
A.2. Line terminators 282
A.3. Groups and capturing 282
A.4. Comparison to Perl 5 283

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 was 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 language, sign language, and gesture, but it can be used by everybody who works with media corpora, i.e., with video and/or audio data, for purposes of annotation, analysis and documentation.

ELAN supports:

- display a speech and/or video signals, together with their annotations;
- time linking of annotations to media streams;
- linking of annotations to other annotations;
- unlimited number of annotation tiers as defined by the users;
- different character sets;
- export as tab-delimited text files;
- import and export between ELAN and Shoebox;
- search options.

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

Part I explains the installation process.

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

- ELAN documents
- The screen display and the navigation through a document.
- Annotations.
- 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 the relevant chapter before starting to work with your own data.

Part III is the reference guide, i.e., it provides brief information on the following topics: mouse options, menu items, and shortcut keys. In addition, the key concepts of ELAN are defined.

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

1. Notation Conventions

The following notation conventions are used:

- Menu items, icons and screen displays are written in the font sans-serif.
- (Shortcut) keys are written in SMALL CAPS.



Note

Information on troubleshooting is printed like this

2. Recent changes to this manual

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 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.

Part I. Getting started

Table of Contents

1. Installation 11
2. Selecting a user interface language 12
3. Consulting the manual and website 13

Chapter 1. Installation

To install ELAN on your computer, go to http://tla.mpi.nl/tools/tla-tools/elan/ and follow the instructions. On this webpage, you will also find information about software and hardware requirements.

Chapter 2. Selecting a user interface language

Once ELAN has been started, you can change the language of the user interface at any time. To achieve this, 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, Dutch, English, French ,German, Japanese, Portuguese, Spanish and Swedish 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` in 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.

Chapter 3. Consulting the manual and website

This manual is available in three different formats:

- HTML: http://www.mpi.nl/corpus/html/elan/index.html
- PDF: http://www.mpi.nl/corpus/manuals/manual-elan.pdf
- Java Help: via Help -> Help Contents... in the ELAN window. The help files are searchable.

Alongside the manual, the website (http://tla.mpi.nl/tools/tla-tools/elan/) contains a Release History and Forum. You can reach both via Help > Visit the ELAN website >. This menu also lets you go to the Download Page to download a new version or Subscribe to the mailing list.

Part II. USER'S GUIDE

This part of the manual contains the user’s guide. It is organized as follows:

- ELAN documents (see Chapter 4)
- Screen display and navigation through a document (see Section 4.4)
- Annotations (see Chapter 5)
- Search Options (see Chapter 7)

It is recommended that you read the relevant section before starting to work with your own data.

Chapter 4. ELAN documents

4.1. 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):
   - 0 or more video file(s) (`*.mpg`, `*.mov` etc.)
   - and/or one audio file (`*.wav`)

The video file allows you to view the video and listen to the sound. If you want to view the waveform as well, you need to create an additional `*.wav` file through a conversion program that converts the audio data from one `*.mp(e)g` file into a `*.wav` format. In the case of multiple video files, the audio of the first selected video file is played. 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 goes for ELAN.

The following media frameworks are known to work:

- Windows (in order of preference, DirectX being the best solution):
  - JDS (Java Direct Show)

    —Java - Microsoft Media Foundation (.mp4, .m4a,. m4v(win &only), .wmv, .wma, .asf)
  - DirectX/DirectShow (Windows Media Player)
  - QuickTime
  - JMF (Java Media Framework)
- MacOS (in order of preference):
  - QuickTime using the Cocoa Framework (QTKit)
  - QuickTime using QuickTime for Java
- Linux: JMF



Note

- 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. The following formats can be imported:Shoebox/Toolbox files (with the extension `*.txt`)

- CHAT files (`*.cha`)
- Transcriber files (`*.trs`)

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 the annotations, as described in Section 4.2.3.

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 *a.eaf* next to *a.mpg* and *a.wav*.

Imported Shoebox files (`*.txt`) 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 alternatively be saved as ELAN files (`*.eaf`).

All annotation files (`*.eaf`) can be exported as text, Shoebox/Toolbox and CHAT files.

4.2. How to manage documents

To start ELAN, do the following:

1. Double-click on the ELAN icon (on your desktop, or contained within the Java Web Start icon on your desktop).

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

2. Click on the File menu (see Figure 4.1)

   The main options in the File menu are:

   - New (Section 4.2.1)
   - Open (Section 4.2.4)
   - Import (Section 4.3.1.1, Section 4.3.1.2, Section 4.3.1.3, Section 4.3.1.4, Section 4.3.1.5, Section 4.3.1.6 and Section 4.3.1.7)
   - Exit (Section 4.2.19)

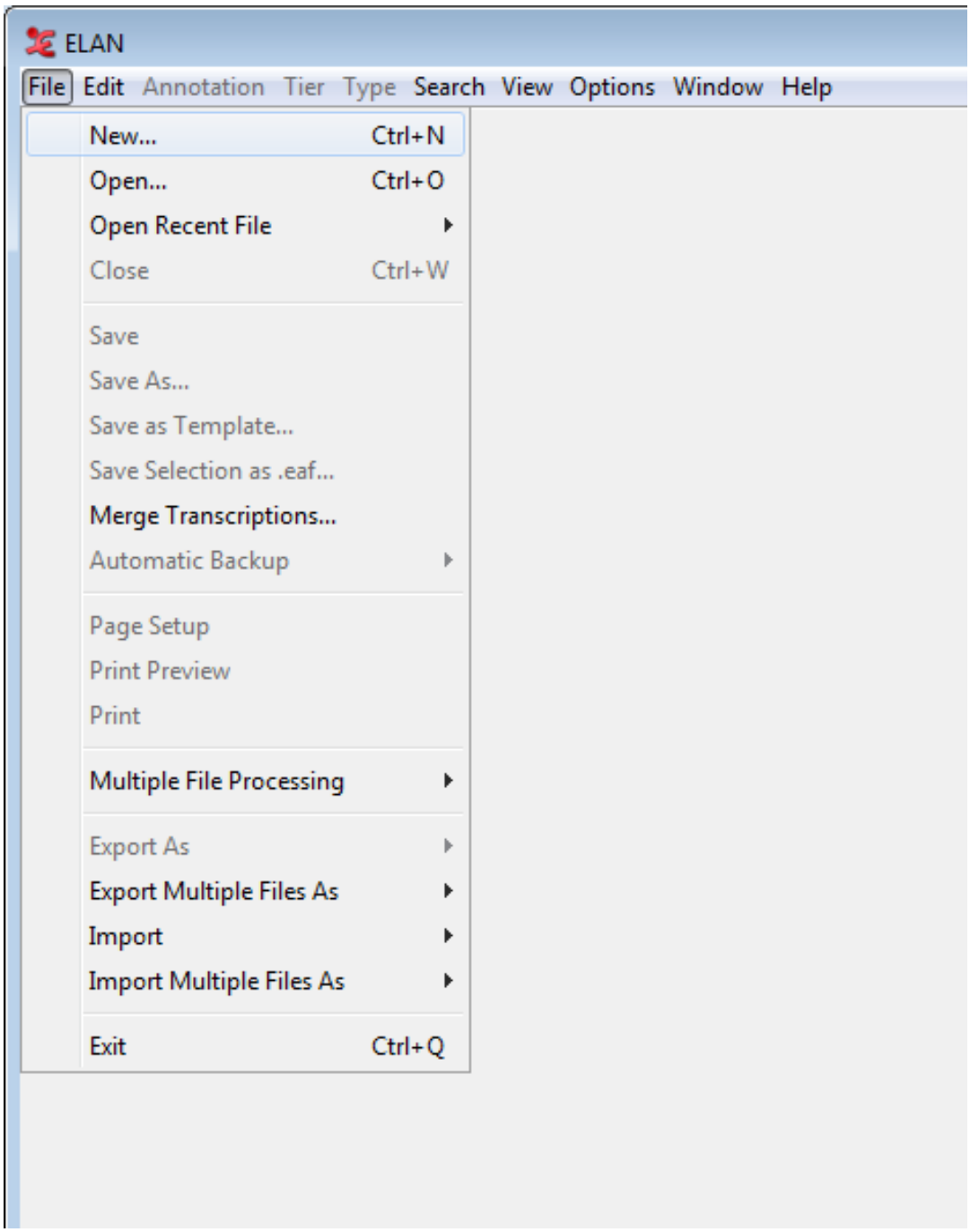


Figure 4.1. File menu

3. Click on:

   - Open... in case there is an ELAN file (`*.eaf`) (Section 4.2.4)
   - New... in case there is a media file (e.g. `*.mpg`, `*.wav`), but no annotation file (`*.eaf`, `*.txt`) (Section 4.2.1).
   - Import > Shoebox File... (Section 4.3.1.3)
   - Import > Toolbox File... (Section 4.3.1.4)
   - Import > FLEx File... (Section 4.3.1.5)

- Import > CHAT File... (Section 4.3.1.2)
- Import > Transcriber File... (Section 4.3.1.1)
- Import > CSV / Tab-delimited Text File... (Section 4.3.1.6)
- Import > Praat TextGrid File... (Section 4.3.1.7)

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

Once you have started ELAN and opened a document, use the File menu to open, create or import a second document.

4.2.1. Creating a new document

In case there is a media file but no annotation file (`*.eaf`, `*.txt`, `*.trs`), click on New in the File menu. The New Transcription dialog window (see Figure 4.2) will be displayed, e.g.:

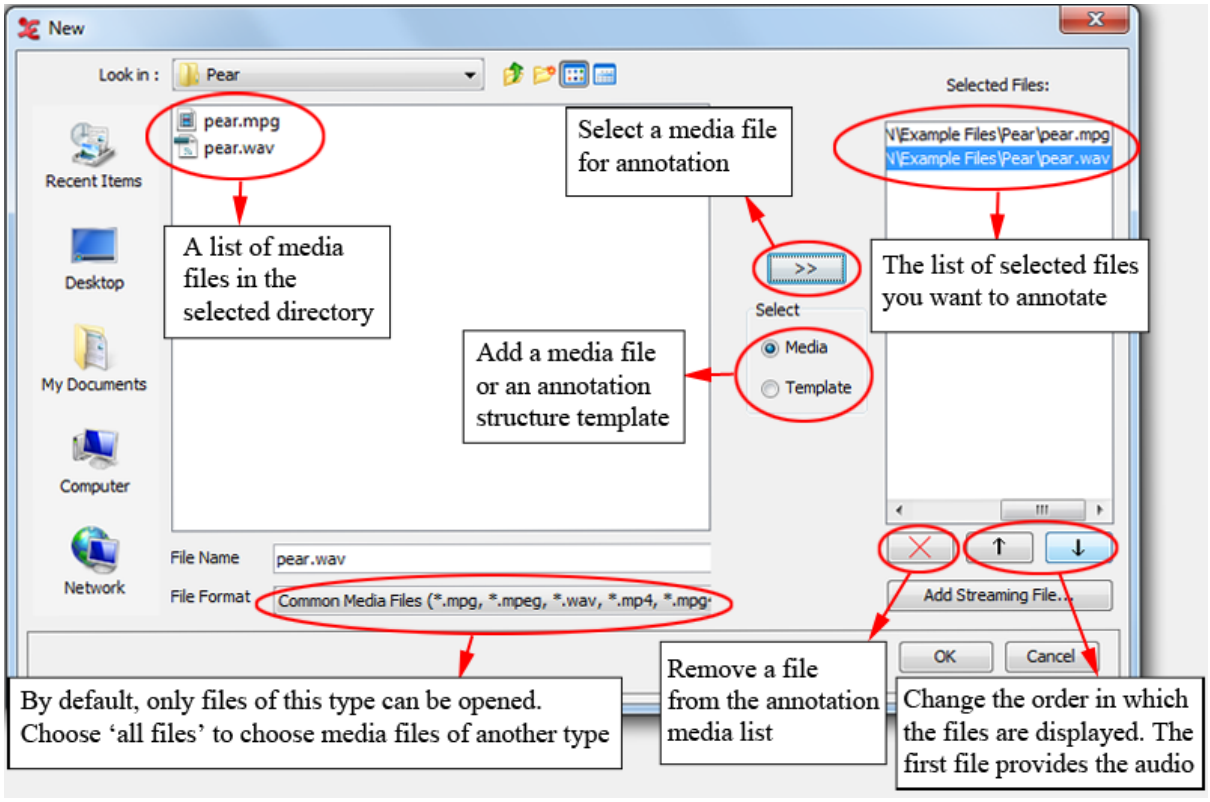


Figure 4.2. The New Transcription dialog window: media

Do the following:

1. Click on the Look in pull down box and browse to the directory that contains the media files
2. If you want to use media files of another type (e.g. QuickTime `*.mov`) then select All Files in the Files of type dropdown menu. If a media type is supported depends on your software configuration.
3. Double-click on the media file (`*.mpg`, `*.mov`, `*.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:

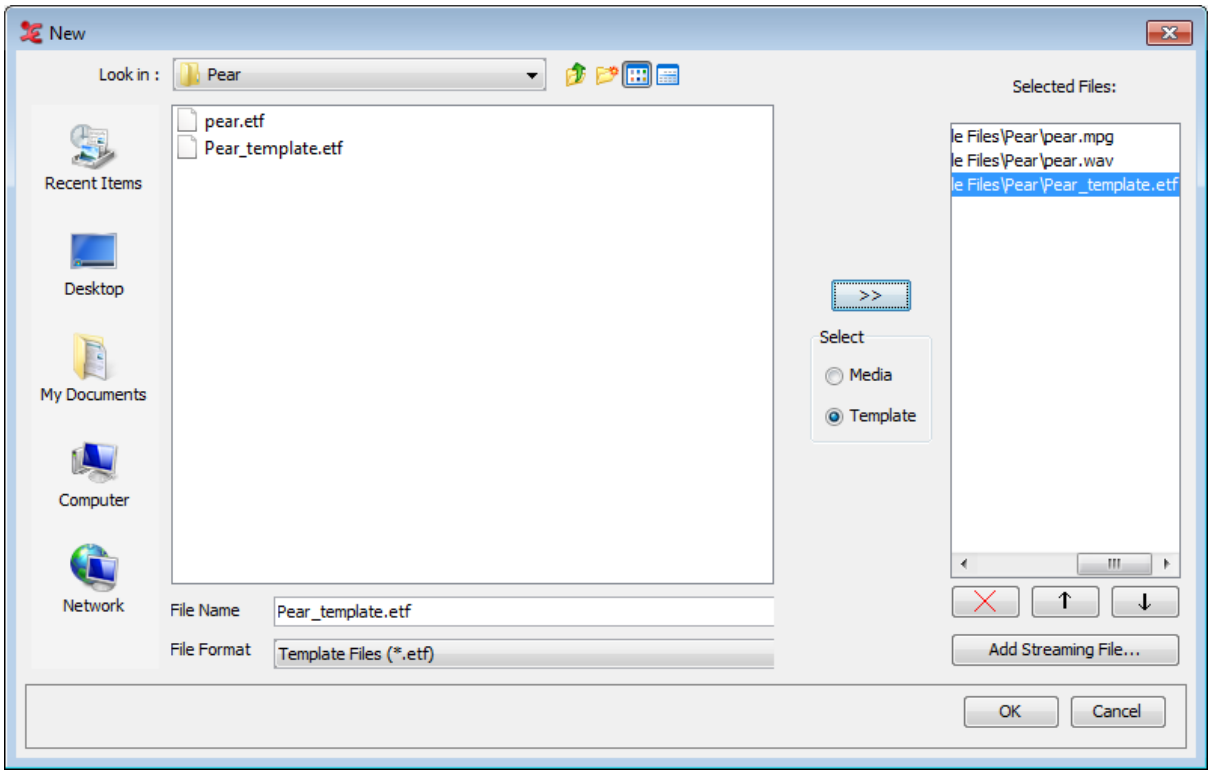


Figure 4.3. The New Transcription dialog window: templates

5. Beside media files on disk you can also add a streaming file of Real Time Streaming Protocol (RTSP). Click on Add Streaming File... and enter the URL of the streaming 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.

4.2.2. Set the author 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.

4.2.3. 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. This can be done as follows:

1. Open the new document with the 2 (or more) video files.

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 expressed as to the 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 Section 4.5.

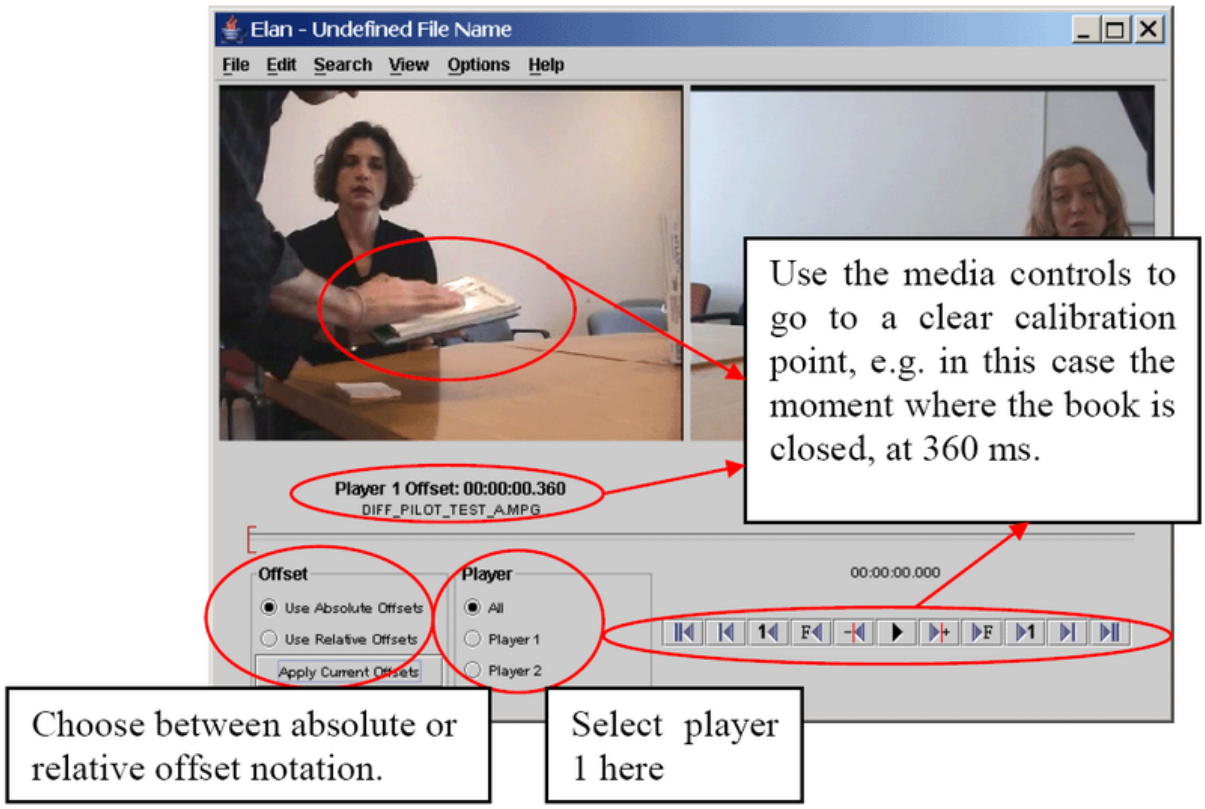


Figure 4.4. Synchronizing video files: Offset of player 1

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



Note

See Section 4.2.11 for changing the order of the videos, i.e. the order of appearing in Player 1, Player 2, etc.



Figure 4.5. 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 Section 5.7.8):

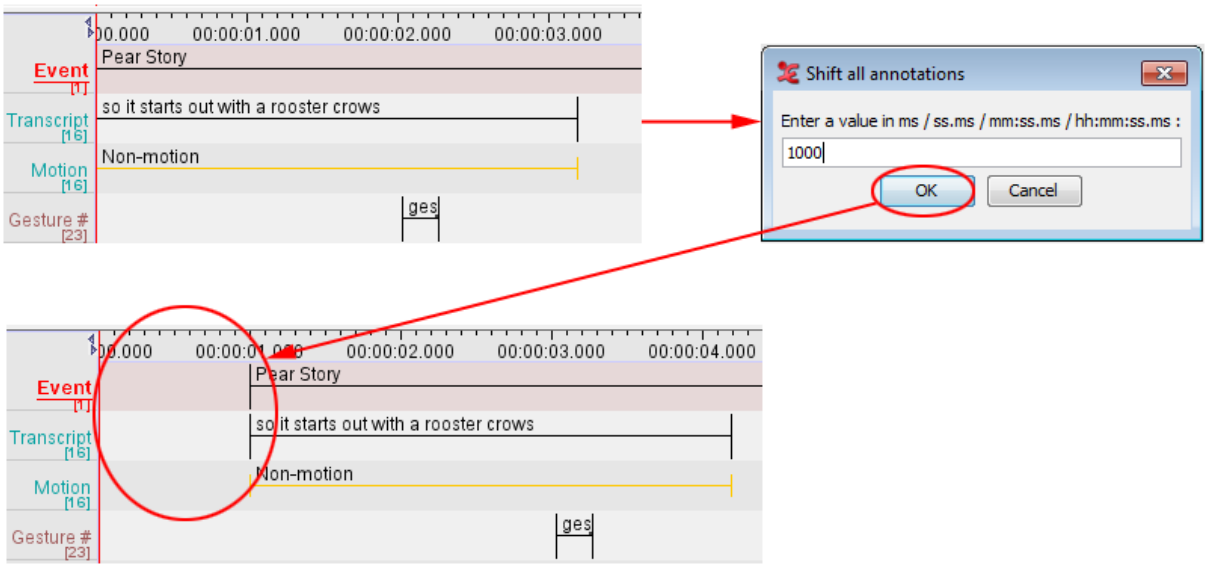


Figure 4.6. Shift all annotations

This process wont delete any annotation. If the annotations are shifted to the left, the maximum shift will be restricted by the leftmost annotation unit.

4.2.4. Opening an existing document

In case there is a media file (`*.mpg`, `*.mov`, `*.wav` etc.) and an ELAN file (`*.eaf`), click Open in the File menu.

The Open dialog window is displayed:

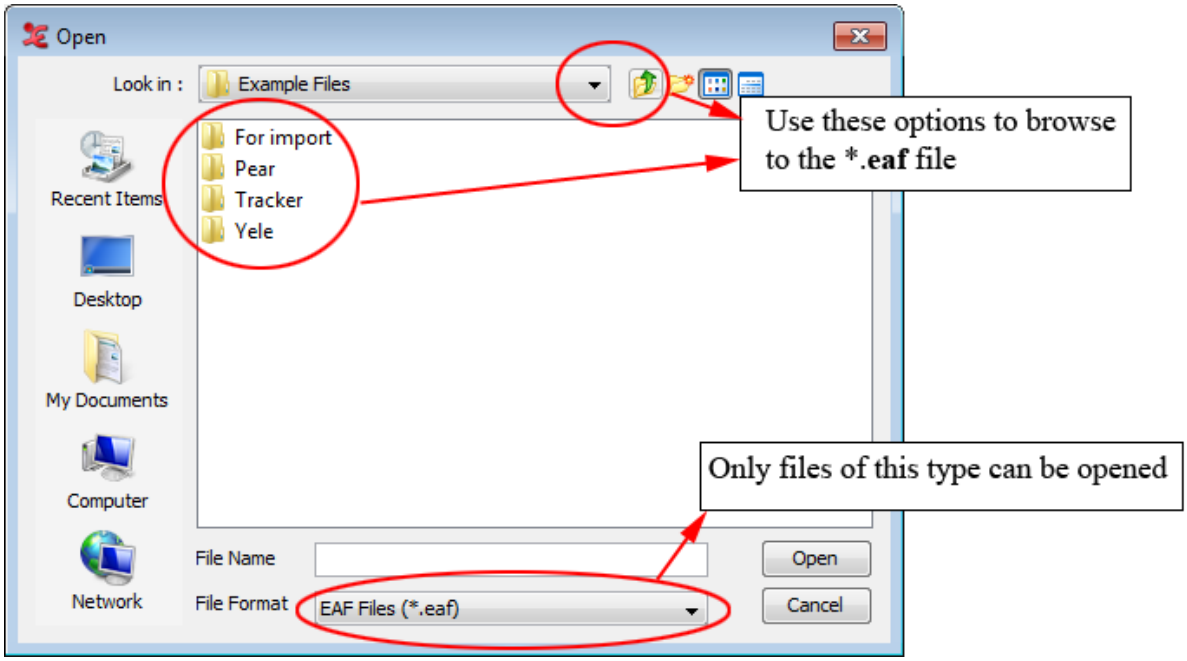


Figure 4.7. Open dialog window

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 appears.

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

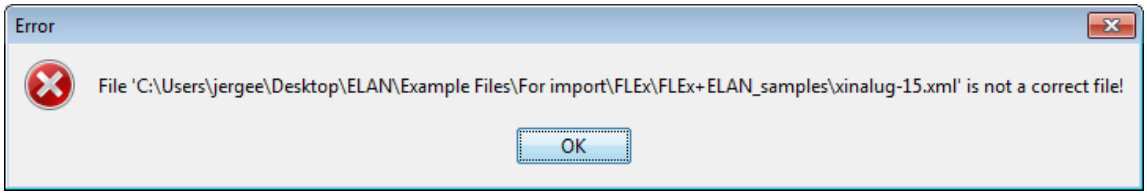


Figure 4.8. Error message: no eaf file



Note

If ELAN cannot find the associated media files (`*.mpg`, `*.mpeg`, `*.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:

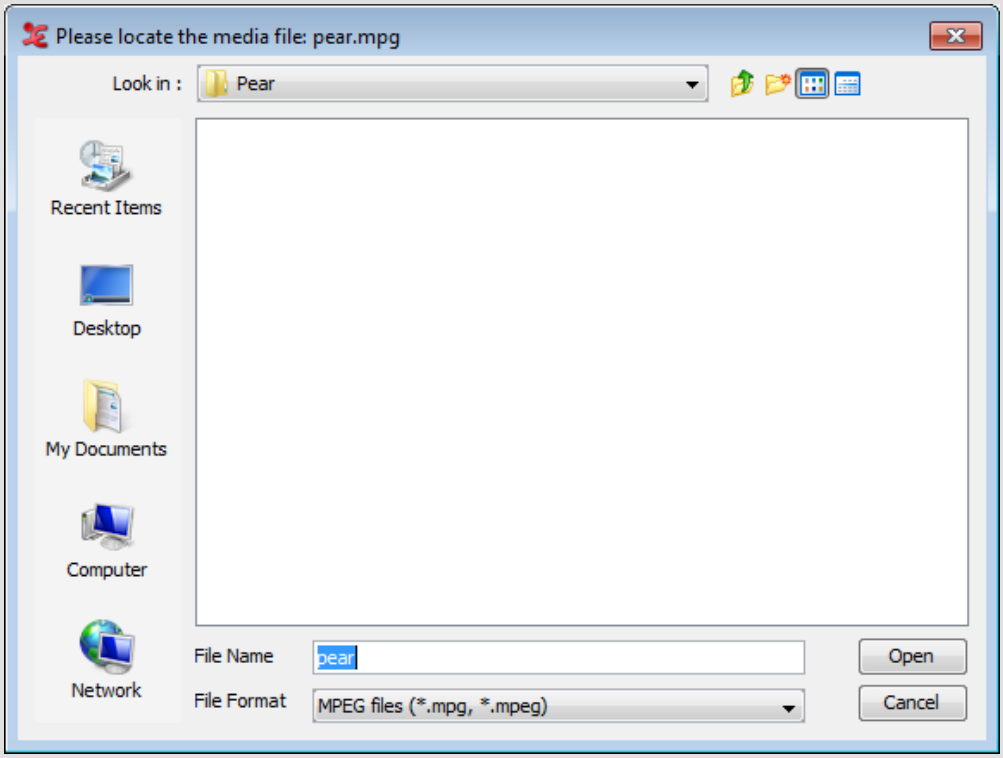


Figure 4.9. Locate media file

4.2.5. Re-open recently accessed files

A convenient way to open a file that you have previously been working with is the File > Open Recent File dropdown menu.

4.2.6. 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:

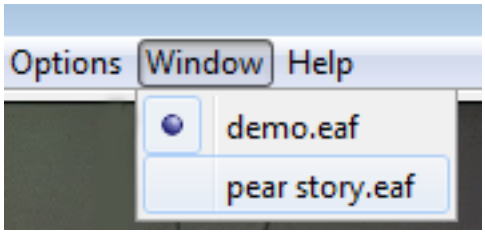


Figure 4.10. Windows dropdown menu

Click on one of the files to activate it.

4.2.7. Saving a document

All documents can be saved as ELAN files (`*.eaf`, "EUDICO Annotation Format"). This includes documents that were created by ELAN itself (see Section 4.2.1) as well as documents that were imported into ELAN from Shoebox/Toolbox, CHAT or Transcriber (see Section 4.3.1.3, Section 4.3.1.2 or Section 4.3.1.1). 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.
2. Or use the shortcut key CTRL+S.



Note

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

4.2.8. 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:

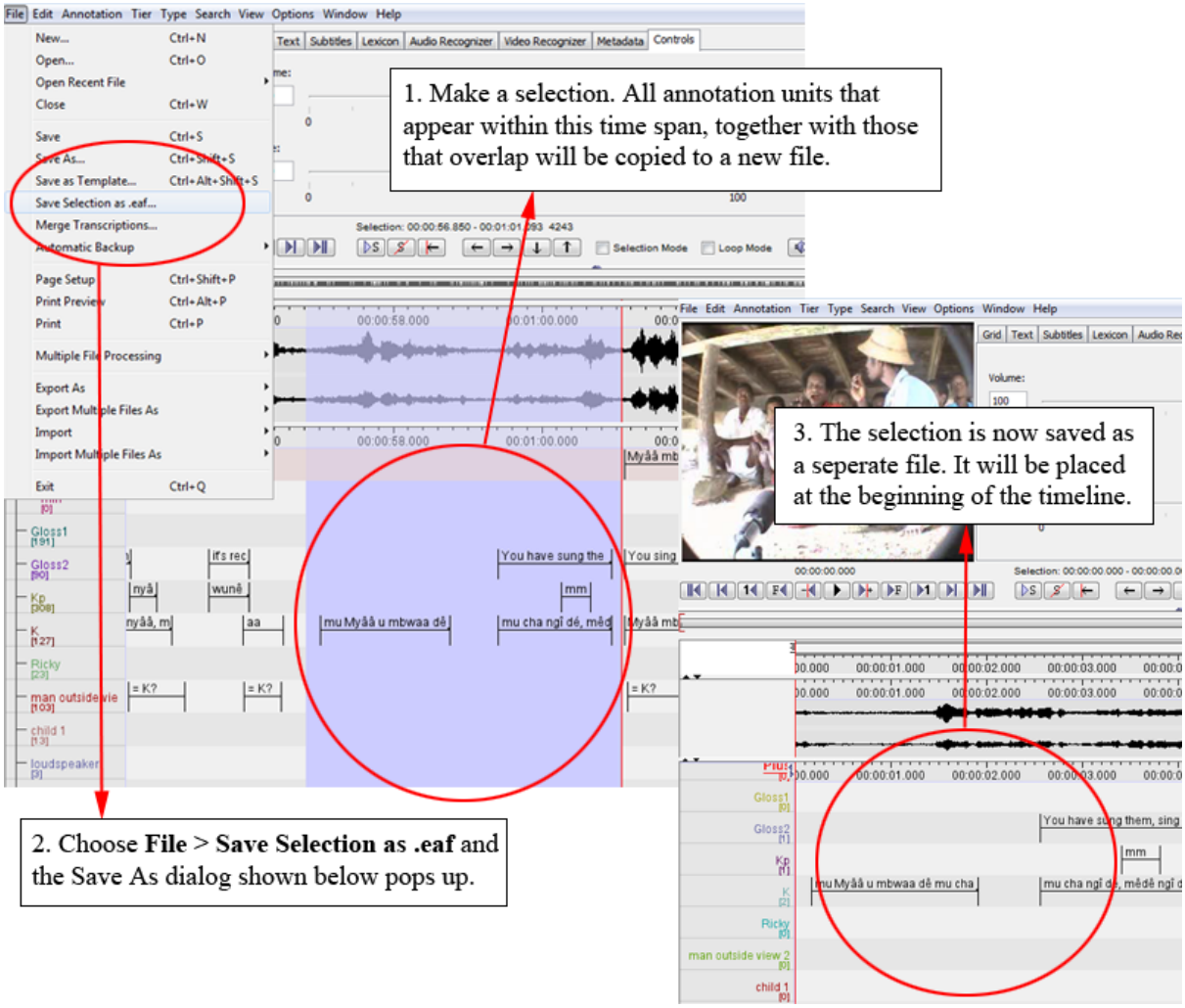


Figure 4.11. Save selection as .eaf file



Note

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

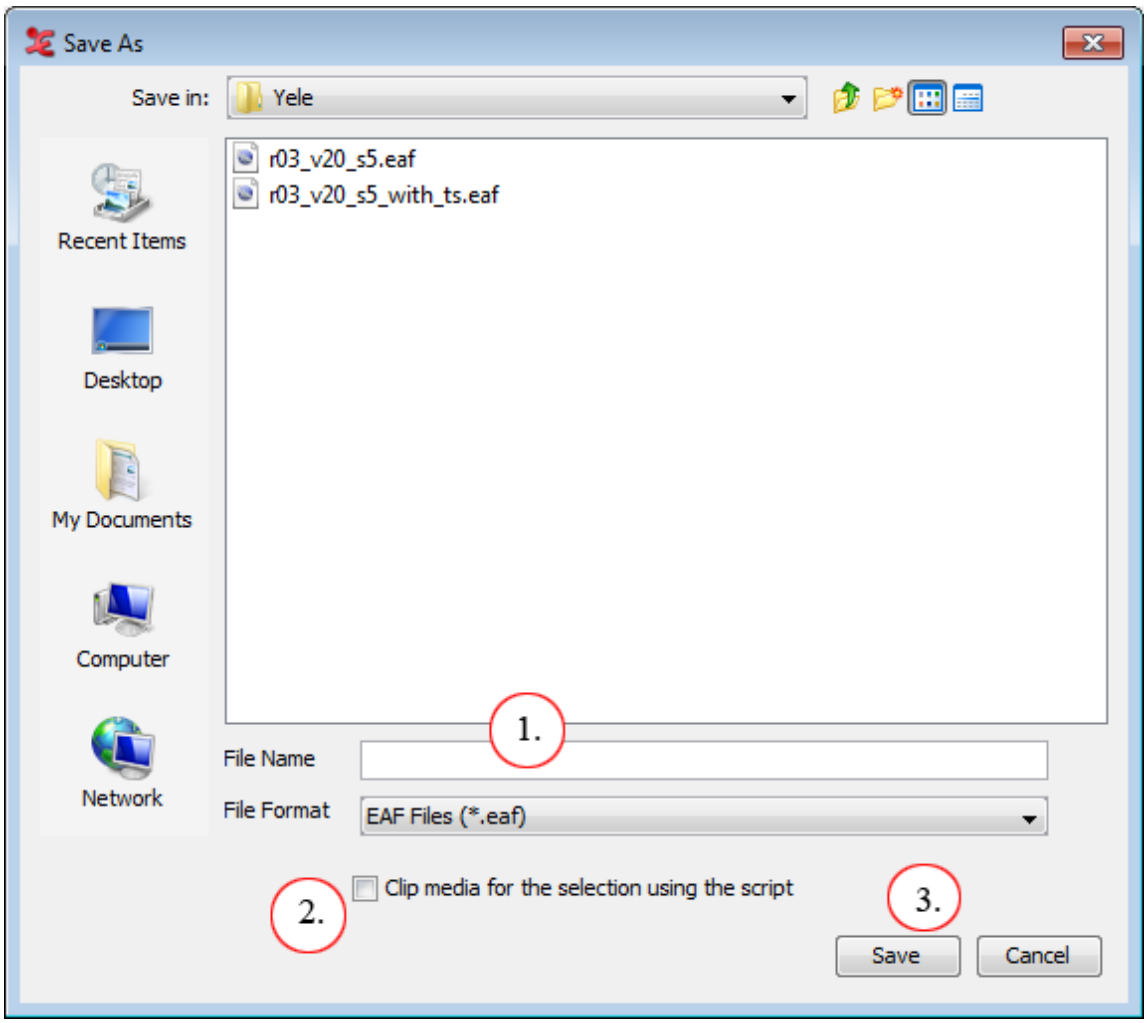


Figure 4.12. Save As Dialog

1. Enter a filename for the new .eaf

2. Check Clip media for the selection using the script also to 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 Section 4.3.2.11 .)

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

4.2.9. 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:

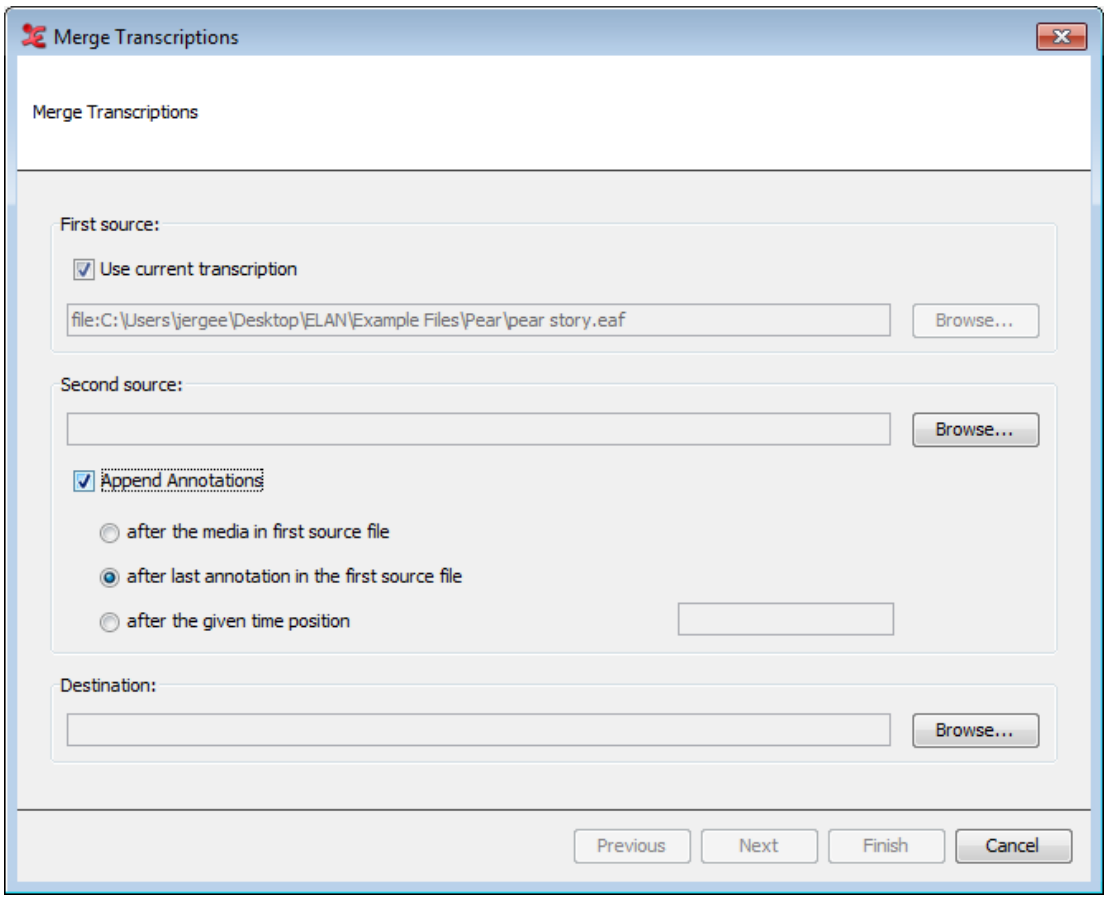


Figure 4.13. Merge transcriptions dialog window

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 below options to append. Otherwise the annotations are append to the start of the first eaf-file.

   - Select after the media in first source file to append annotation after the media duration in the first eaf-file.

   - Select after the last annotation in the first source file to append after end time of the last annotation in the first eaf-file.

   - Select after the given time position to append the annotation from the give time frame.

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

7. Click on Next. This dialog will appear:

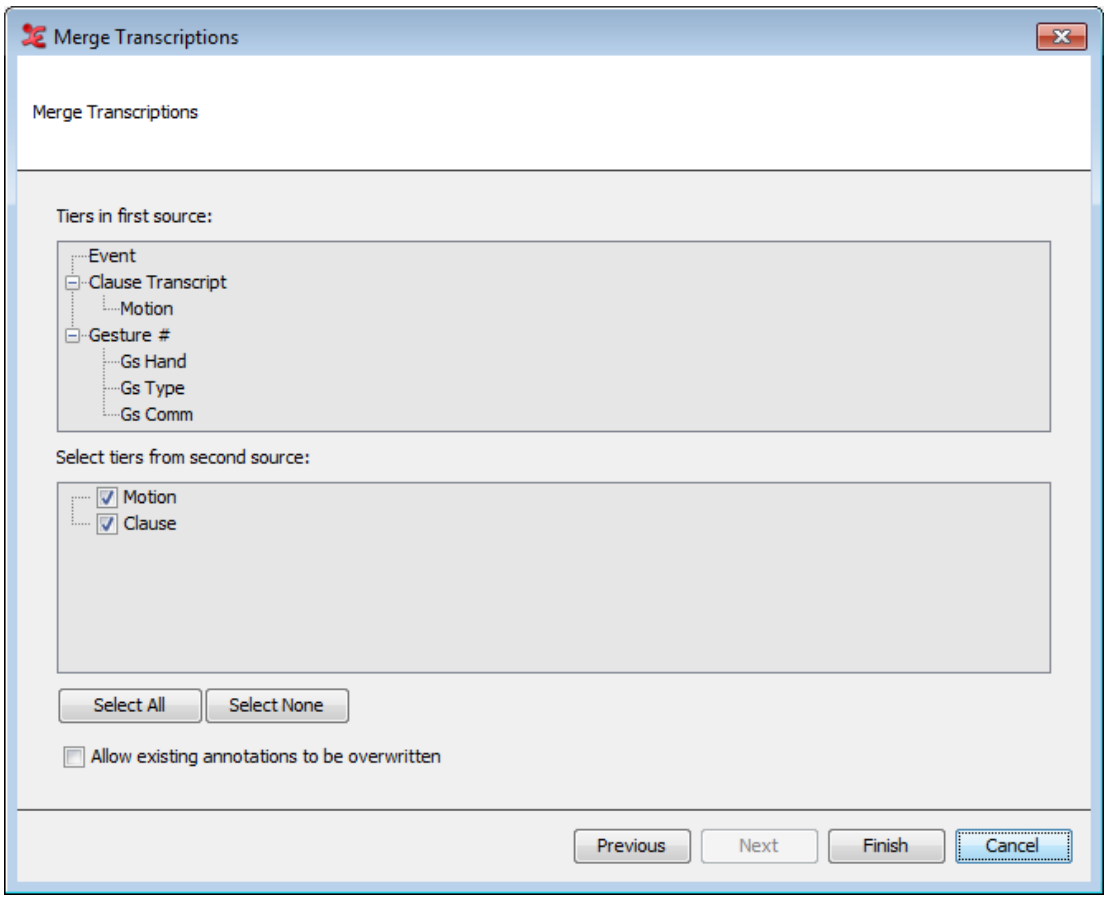


Figure 4.14. Merge transcriptions dialog window

8. Select the tiers of the second source file that you want to merge with the first file.
9. If there are common tiers in both files and you want those of the second file to overwrite those of the first, make sure Allow existing annotations to be overwritten is checked.
10. Click on Finish to start merging.
11. When the merge procedure has been finished you can choose whether to open the result immediately in new ELAN window:

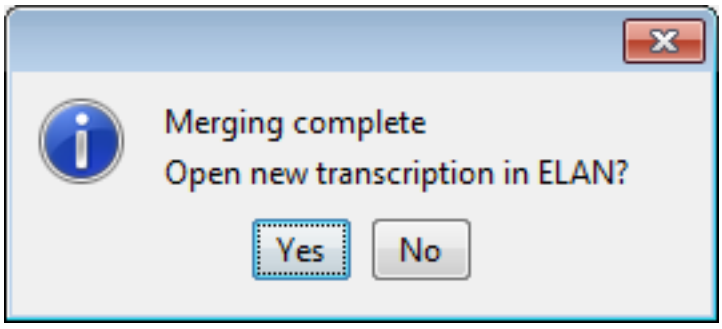


Figure 4.15. Open new transcription

4.2.10. Saving a template

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

- linguistic types (see Section 5.3)
- controlled vocabulary (see Section 5.5.7)
- tiers (see Section 5.4)

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 Section 4.2.1 for a 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.

4.2.11. Changing the links to media files

As from ELAN version 2.4, the possibility exists to explicitly change the links to media files that are linked from an `*.eaf` file. This option can be handy if e.g. you have moved media files to another location after the last time you edited an ELAN file.

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

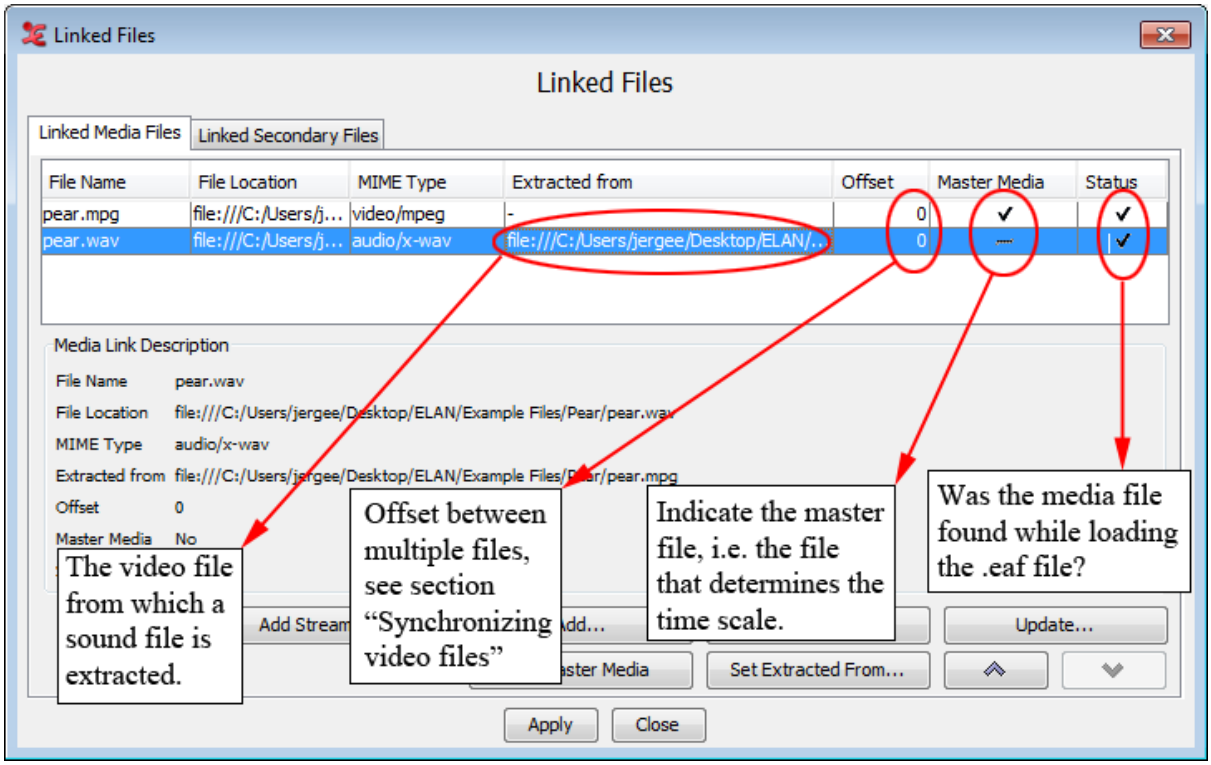


Figure 4.16. Linked files dialog window

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

- Add… : add a link to a new media file to the current `*.eaf` file
- Remove: remove the selected media file
- Update… : specify a new location of the selected file. 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.

The Linked Secondary Files tab shows files that are linked as secondary files. In particular files that contain data that to be displayed by the Timeseries viewer (see Section 4.4.12) is 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. Especially useful if the checkbox Status is not marked. The latter indicates 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.

4.2.12. 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 the time interval after which ELAN should create the backup, e.g., after every 10 minutes.

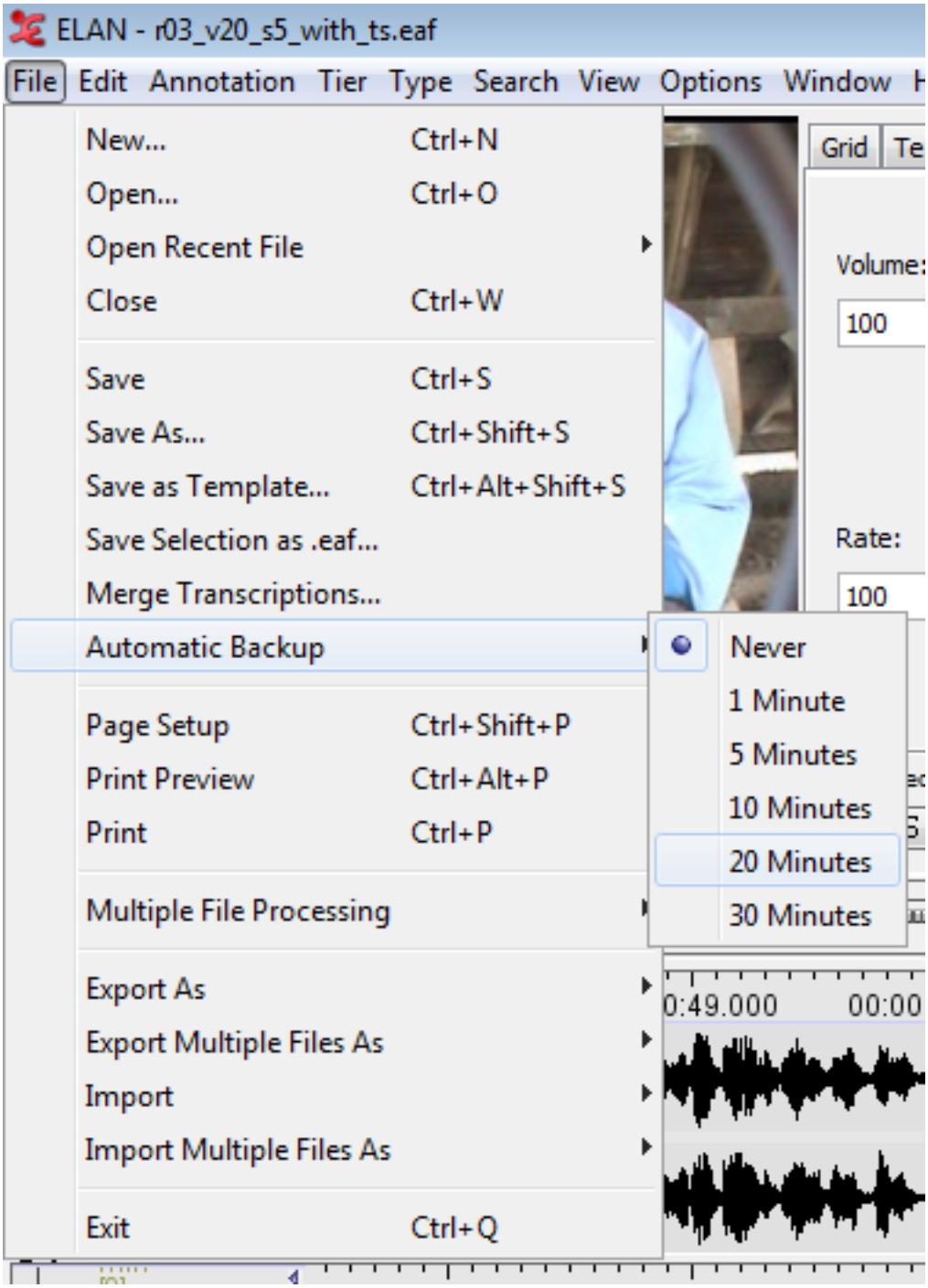


Figure 4.17. Automatic backup

A checkmark appears next to the selected time interval. From now on ELAN will automatically create a backup copy in the same directory as the original file, saving it with the extension `*.eaf.001`. Before opening such a file, rename its extension to .eaf instead of .eaf.001



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.

4.2.13. Printing

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

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

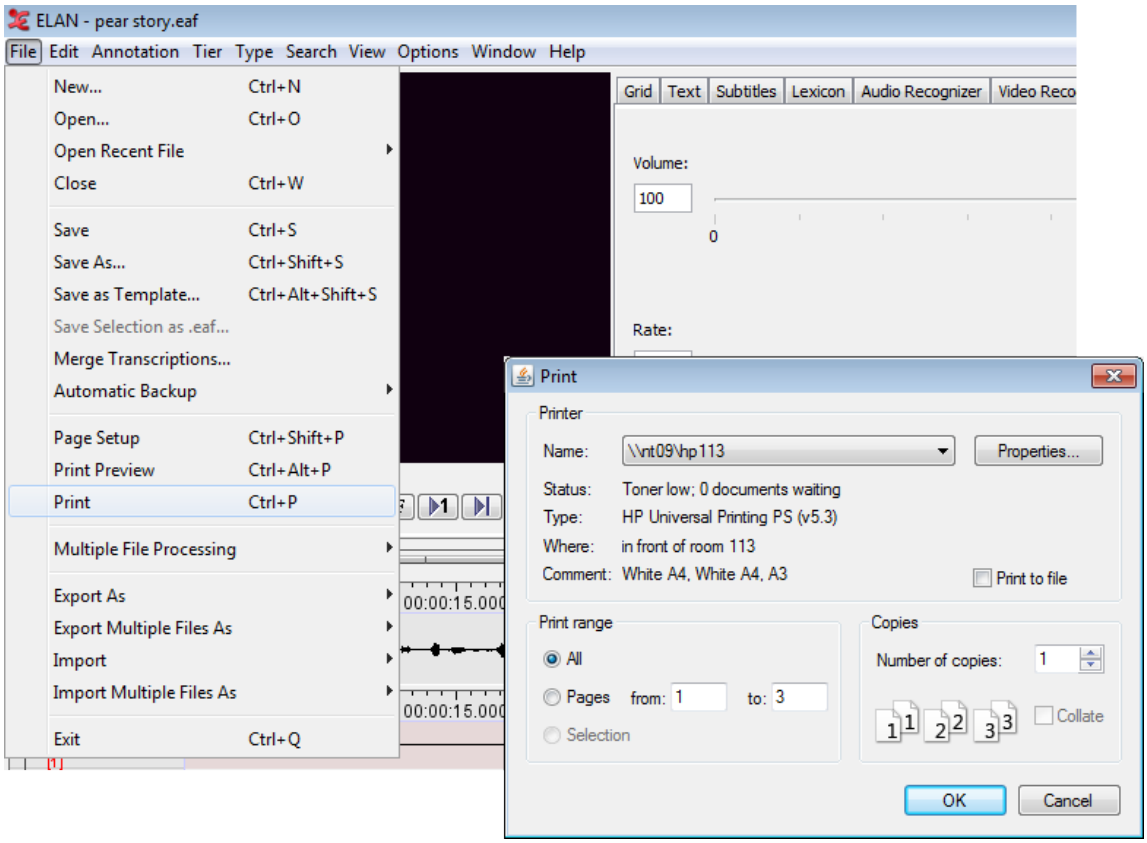


Figure 4.18. Printing

4.2.14. Configuring the page settings (page setup)

Through File > Page setup you can alter the paper size and other settings of the pages to be printed.

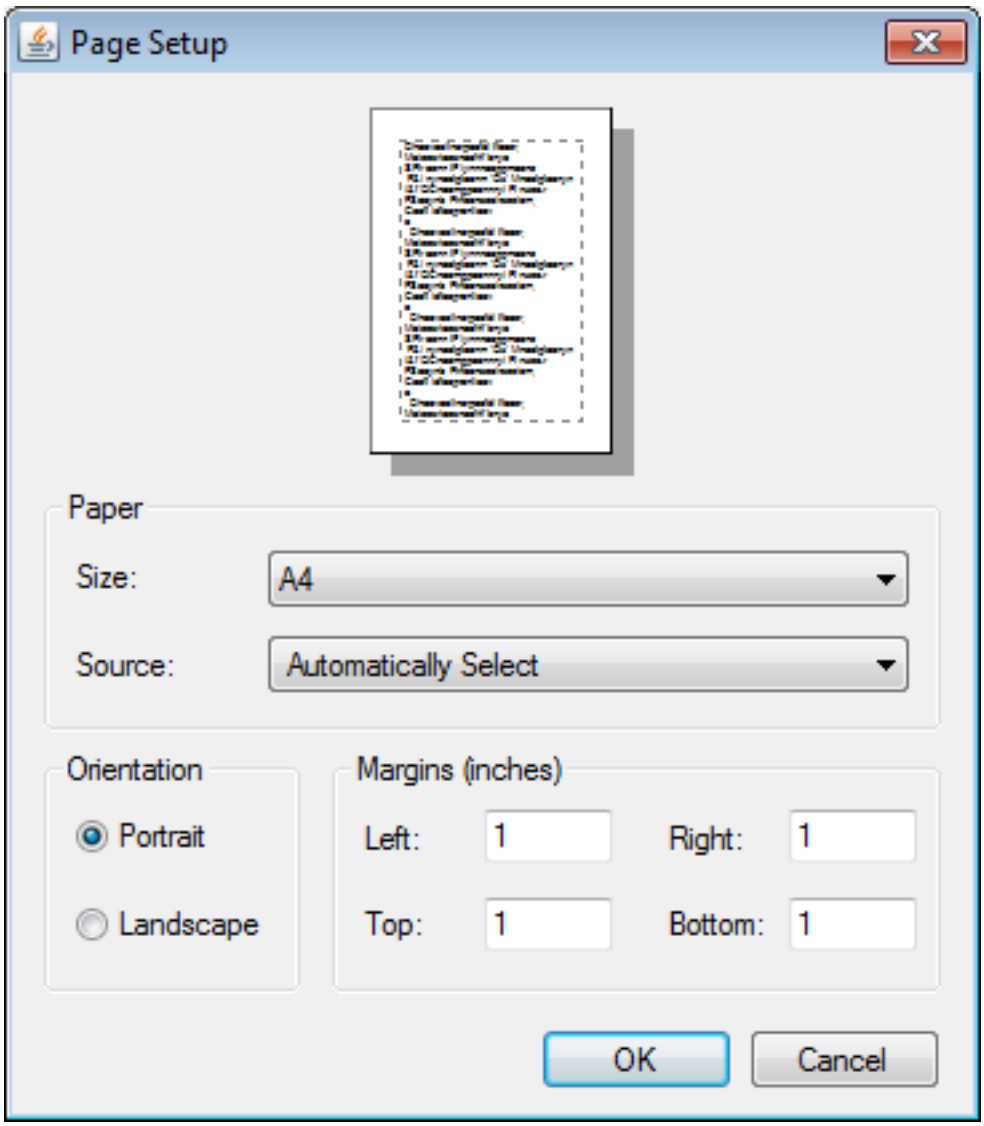


Figure 4.19. Page setup

4.2.15. 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.

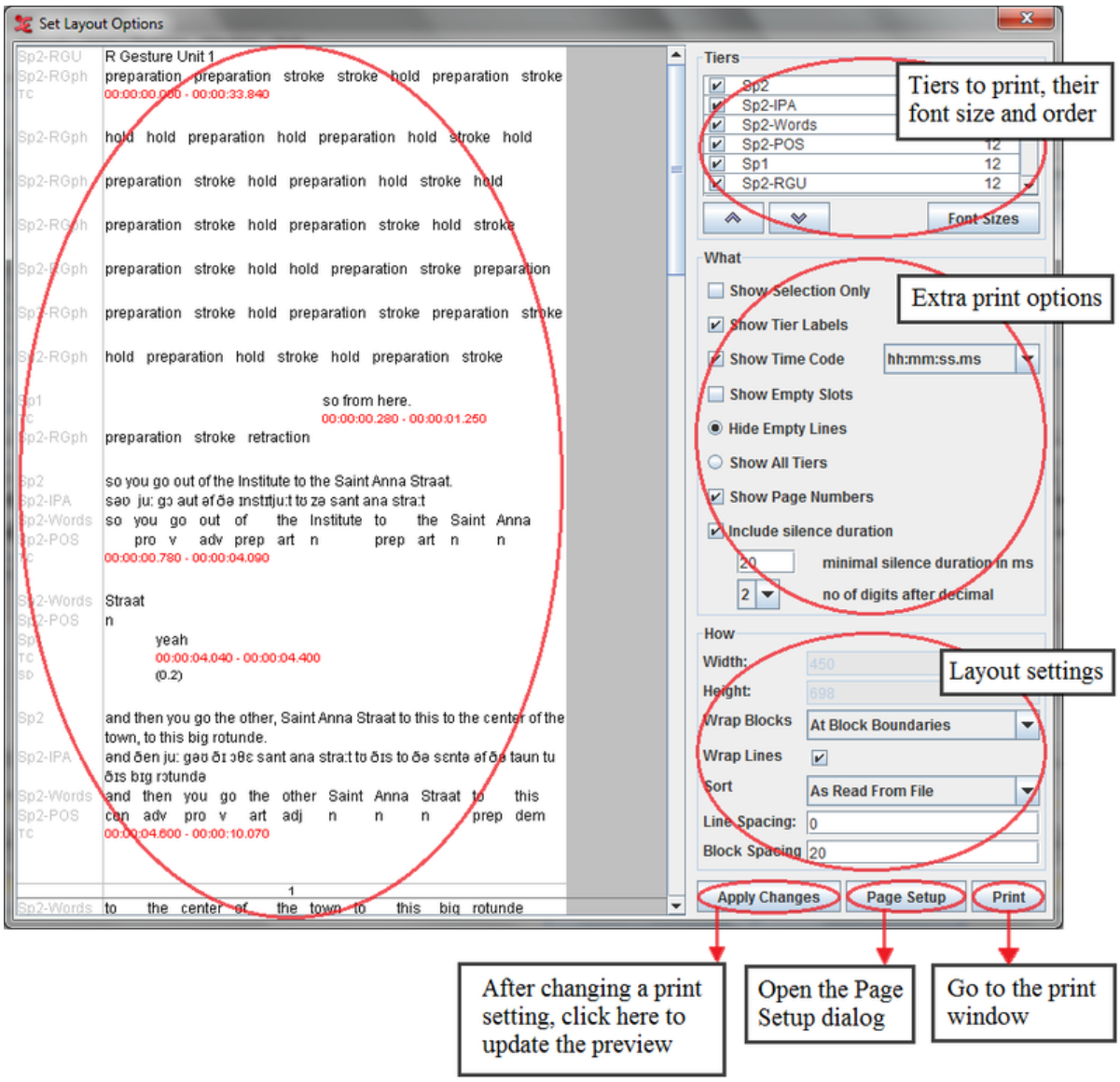


Figure 4.20. Print preview

Tiers settings

- Put a checkmark 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 a advanced selection dialog(see Section 4.3.2.1) .
- The font size of the tiers can be adapted by clicking on the Font Sizes button. A new window will appear:

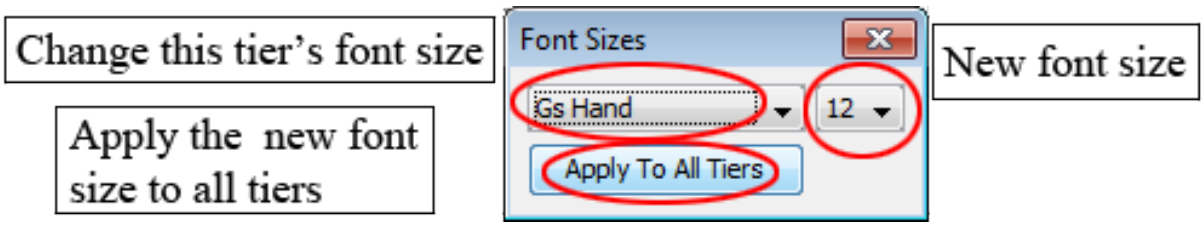


Figure 4.21. Font sizes

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

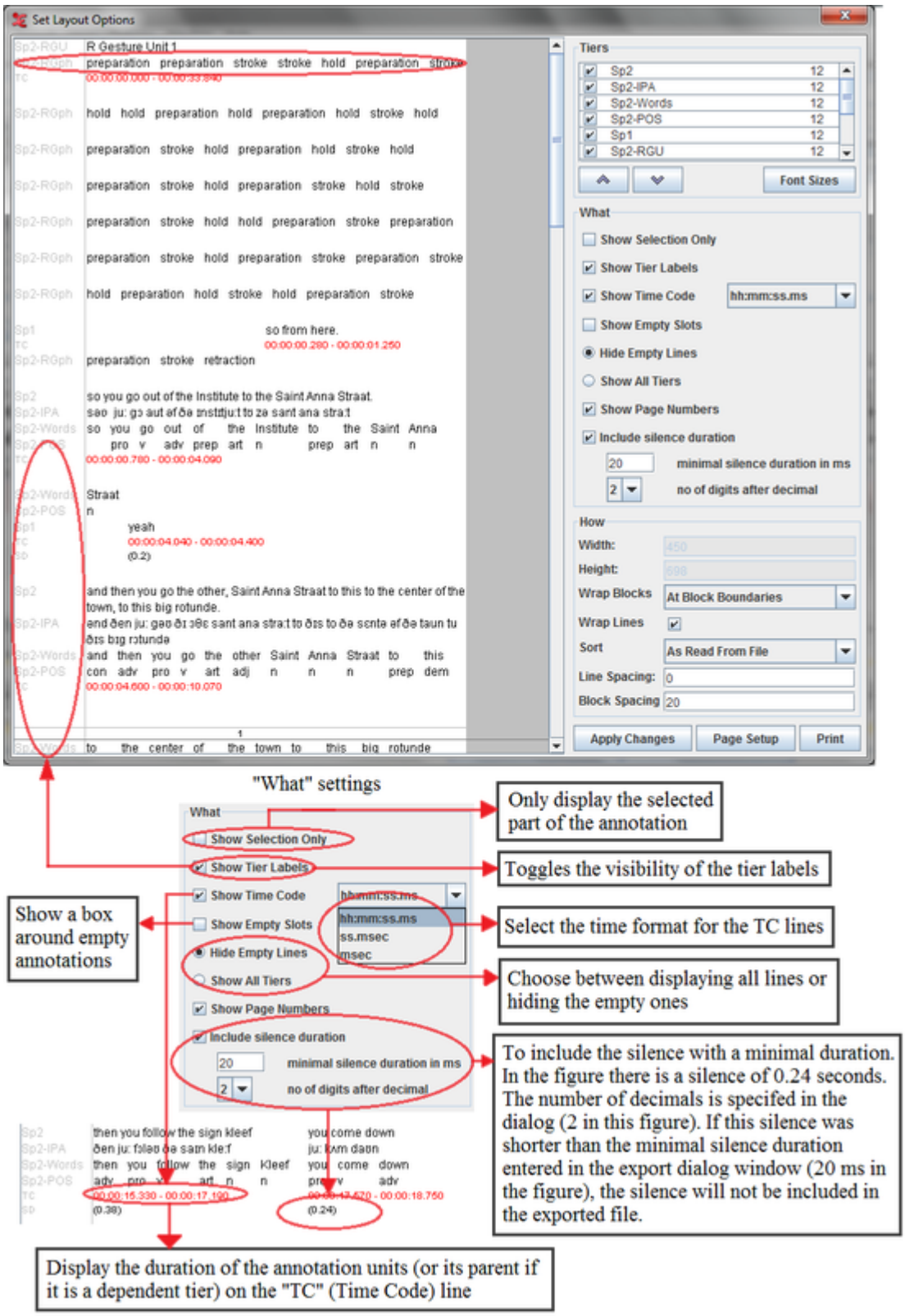
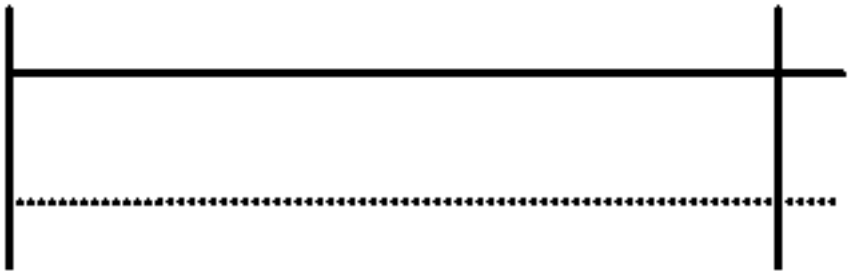


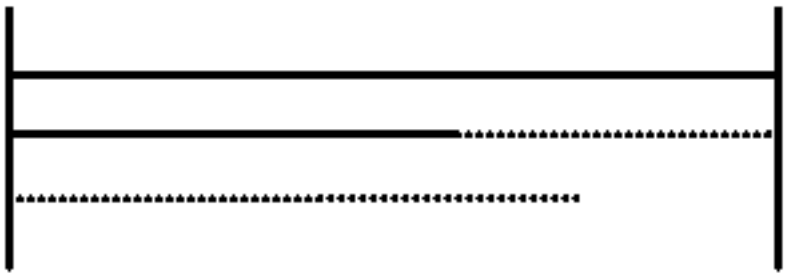
Figure 4.22. Print preview after changes

- 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 Section 4.2.14).
- 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.

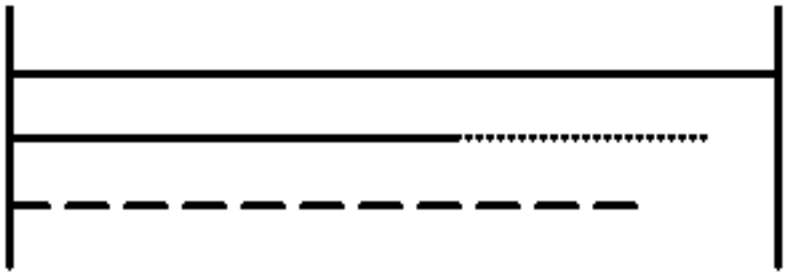
- Wrap Blocks:
  - *No wrapping*: use 1 line for each tier, only usable for files that contain a small amount of annotations.



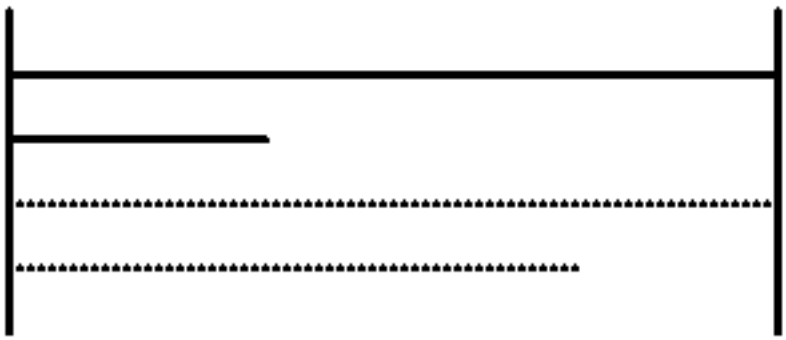
  - *Within block*: wrap blocks, and continue with a new block on the same line if there is space left.



  - 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 Section 4.4.23).
- Line spacing: amount of white space between the lines (default: 0 pixels)

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

4.2.16. Opening a wave file in Praat

ELAN offers the possibility to open wave files (or a part of them) in Praat[1]. To achieve this, follow the steps below:

1. If desired, make a selection first (See Section 5.7)
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 and Sendpraat[2] program yet, you will have to locate them now in the file dialog
5. The Praat program containing the wave file appears

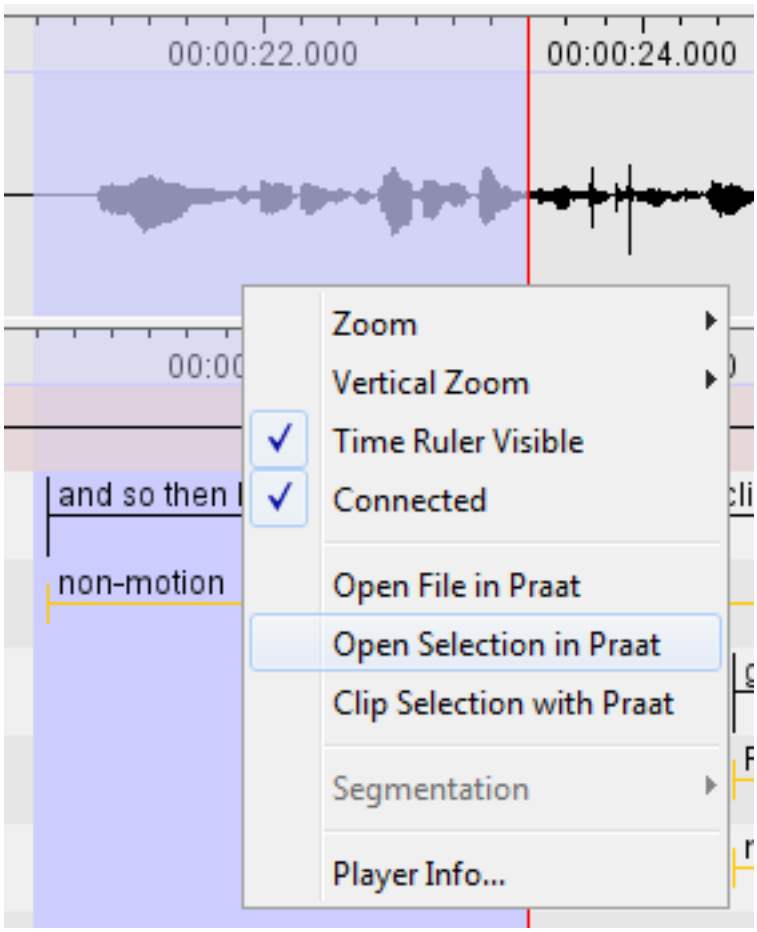


Figure 4.23. Opening a wave pattern in Praat



Note

Make sure you are using a recent version of Praat (higher than 4.0.5), otherwise this feature will not work.

[1] See http://www.fon.hum.uva.nl/praat/

Note for advanced users: compiling SendPraat on Linux

- Make sure a C compiler (like gcc) is installed.
- You will also need the X11 or Xorg development packages (they can be found in a package called xorg-x11-devel or something likewise)
- Replace *if0* with *if1* in *sendpraat.c*, as described on http://www.fon.hum.uva.nl/praat/sendpraat.html
- Link the executable with the X11 libraries when you compile Sendpraat:

```
cc -o sendpraat sendpraat.c -L /usr/X11R6/lib -lX11
```

4.2.17. 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 Section 5.7) 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.

4.2.18. 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.

4.2.19. 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 Section 4.2.5), the Saving transcription dialog window appears, e.g.:

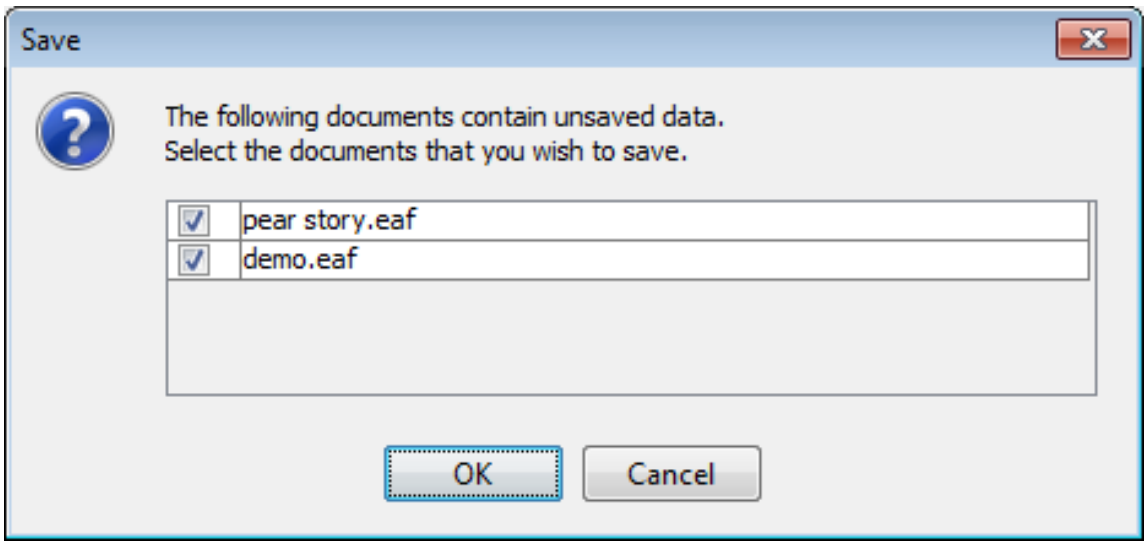


Figure 4.24. Save on exit

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

This behaviour differs from earlier ELAN versions. Therefore you will be warned when exiting. If you don't want this warning again the next time, check *Don't show this message again* and click on OK.

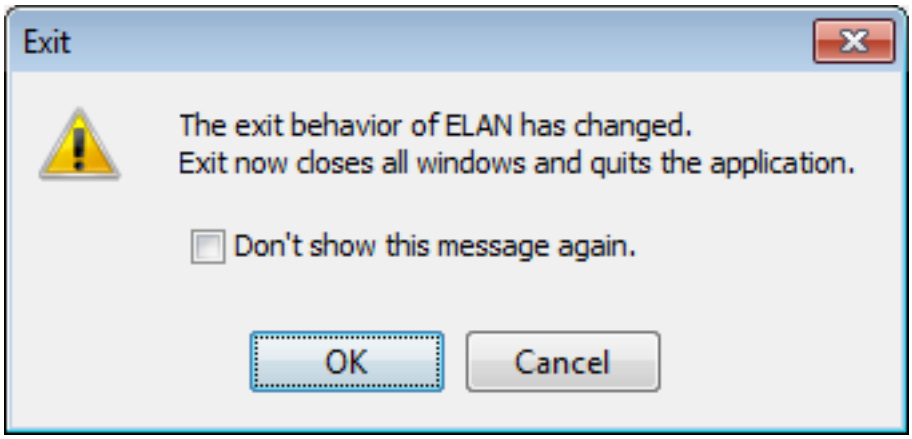


Figure 4.25. Exit behavior

4.3. Import and Export options

4.3.1. Import from

Elan supports importing file from :

- Transcriber file (Section 4.3.1.1)
- CHAT file (Section 4.3.1.2)
- Shoebox file (Section 4.3.1.3)
- Toolbox file (Section 4.3.1.4)
- Fieldworks Language Explorer (FLEx) file (Section 4.3.1.5)
- CSV / Tab-delimited Text Files (Section 4.3.1.6)
- Praat TextGrid file (Section 4.3.1.7)

4.3.1.1. 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 Section 5.1).
- Events are embedded into the annotation text.

4.3.1.2. 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 Linguistic 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

4.3.1.3. Shoebox file[3]

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

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

Optionally you can use the corresponding Shoebox 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 Shoebox database type file, do the following:

1. Open the Shoebox `*.txt |*.sht |*.tbt` file in Shoebox. 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.:

[3] From here on, every appearance of Shoebox can also be read as Toolbox, i.e. the newer version of what was formerly known as Shoebox.

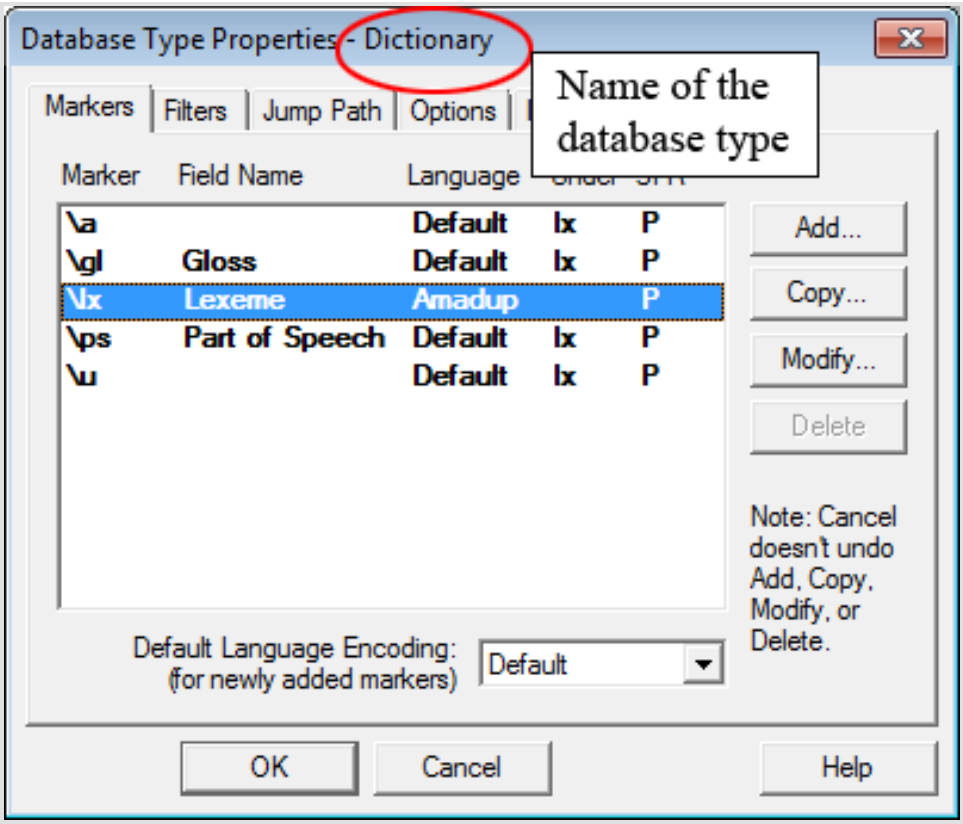


Figure 4.26. Database type properties dialog window

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 Shoebox files with a TYP file

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

1. Click on File > Import > Shoebox File. The Import Shoebox dialog box appears.
2. Specify the name and directory of the two files, e.g.:

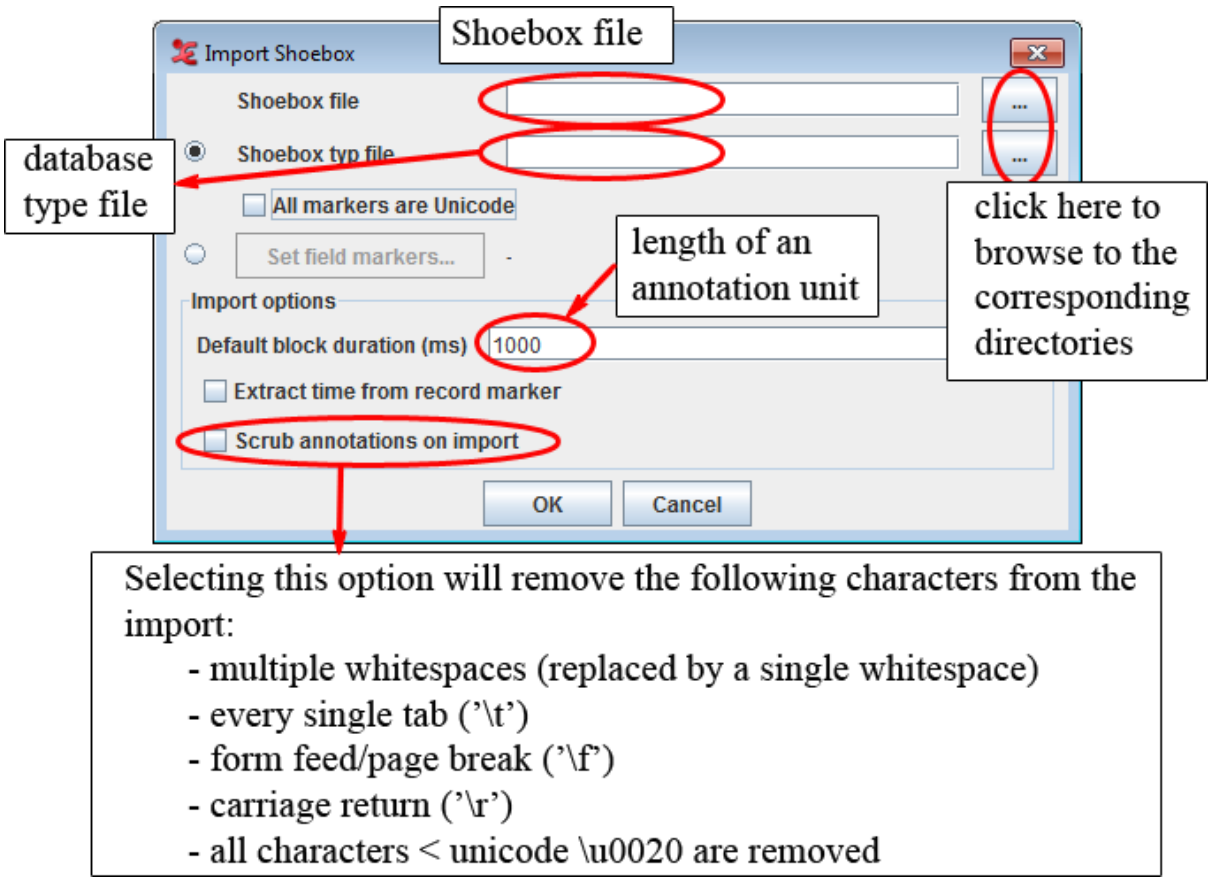


Figure 4.27. Import Shoebox file

3. Like `*.eaf` documents, the Shoebox 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 Section 4.1).

   If the Shoebox 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 Shoebox file appears.

Importing Shoebox files without a TYP file

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

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

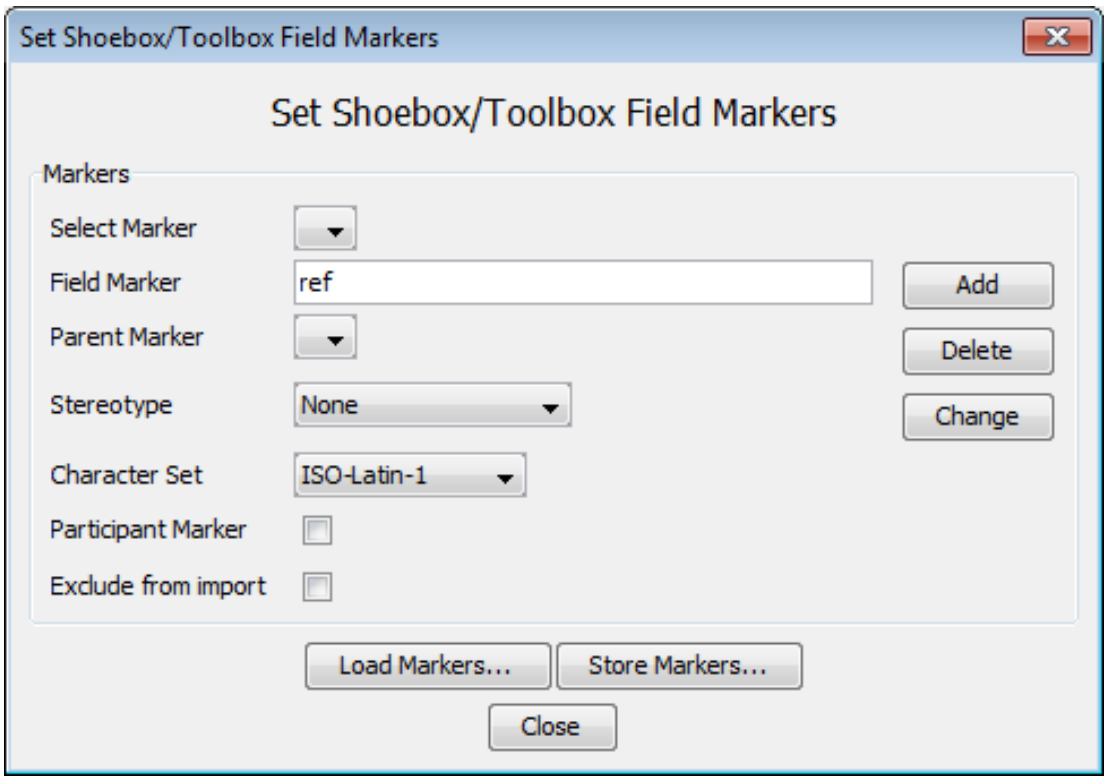


Figure 4.28. Set Shoebox/Toolbox field markers

2. Now fill in a field marker as used in the Shoebox `*.txt|*.sht |*.tbt` file
3. Optionally select a parent marker (see Section 5.1)
4. Optionally select a stereotype (symbolic subdivision or association, see Section 5.1)
5. Choose a character set (Latin-1, SIL IPA or UTF-8) for the tier
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

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 filename, and press save.
- The same way you can open a stored field marker set by clicking on Load Markers…

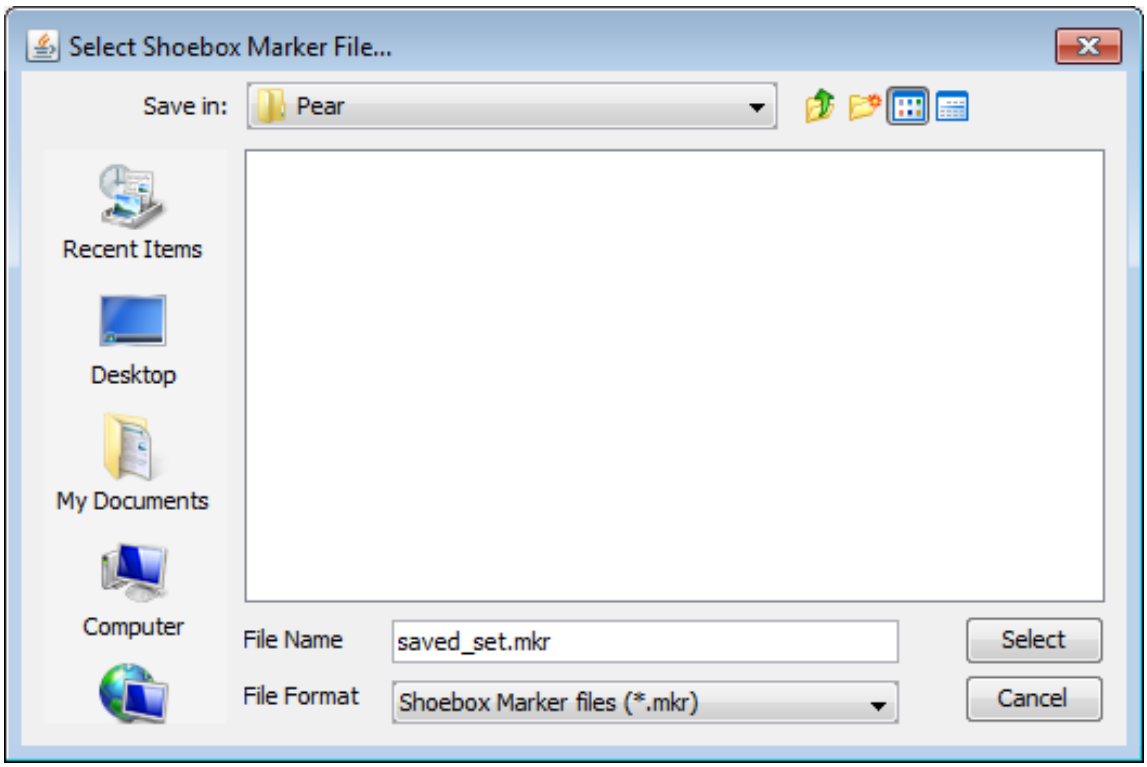


Figure 4.29. Store markers

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 Section 4.2.11. If the imported Shoebox 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 Shoebox files according to the following conventions:

1. The Shoebox 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 Shoebox differ in how they code information about multiple speakers:

   - In ELAN, each speaker is coded on a separate tier.
   - In Shoebox, all speakers are coded using the same field, and their identity is specified in a separate field.

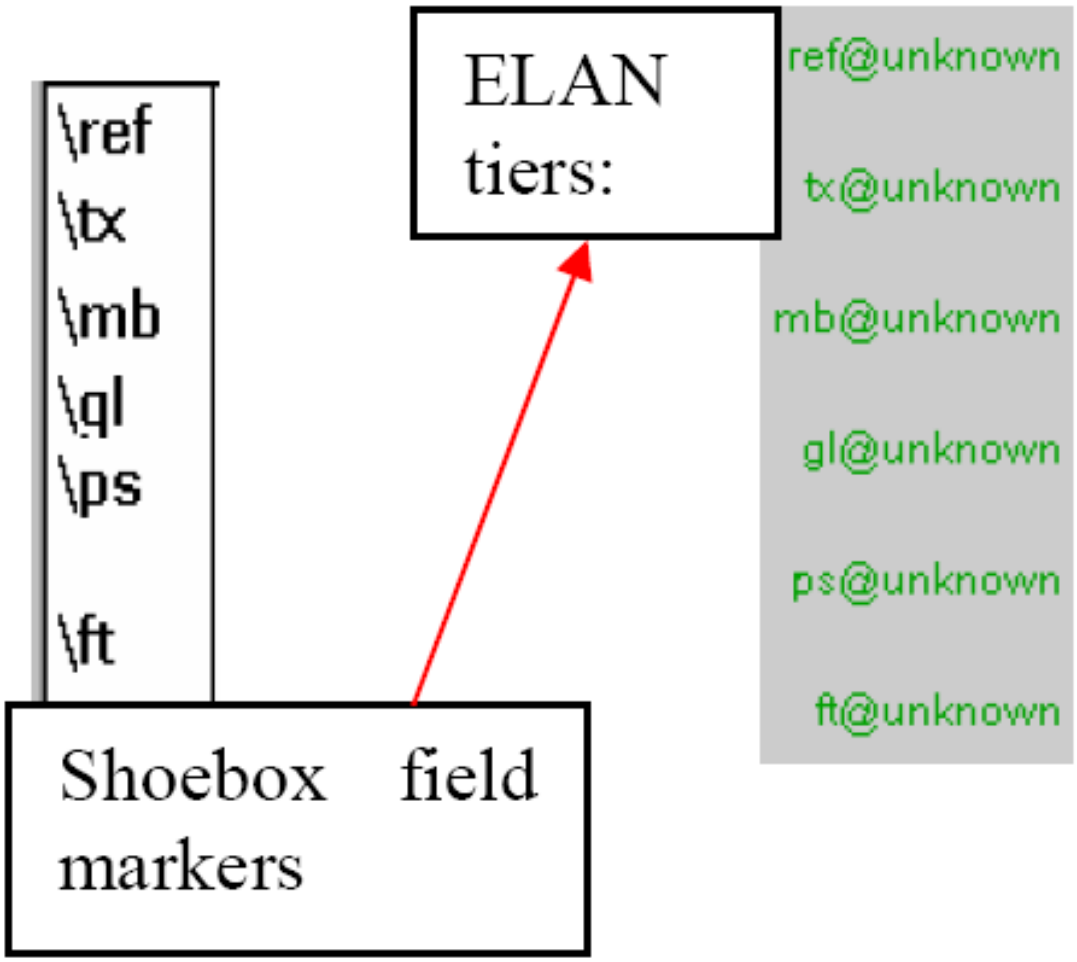


Figure 4.30. Shoebox field markers and ELAN tiers

When importing texts by multiple speakers, ELAN splits each Shoebox 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 Shoebox file, the extension @unknown is added.

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

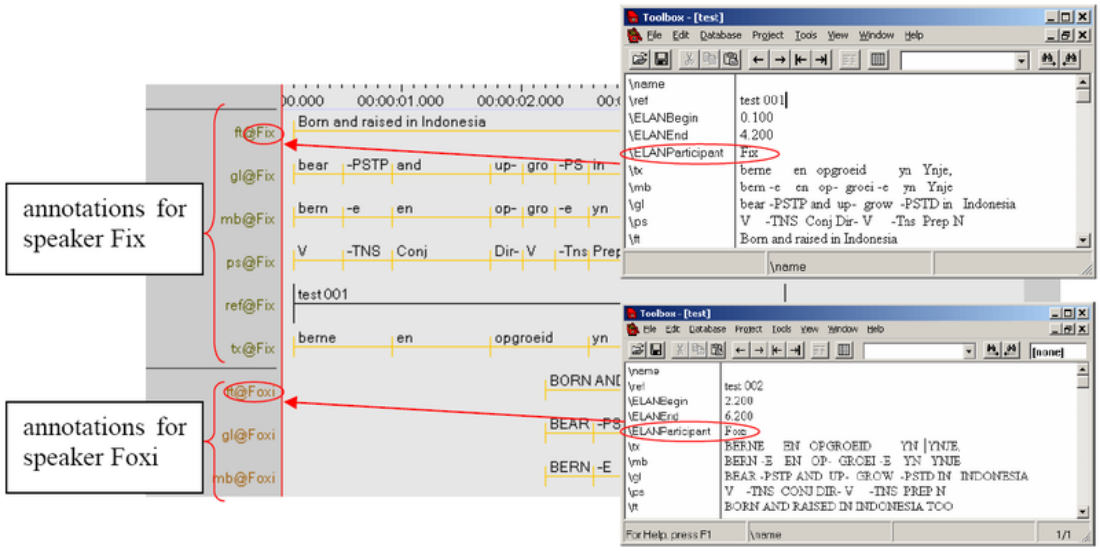


Figure 4.31. Multiple speakers in ELAN

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 Shoebox files without a TYP file above), or if:
- It is coded in a Shoebox 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 Shoebox record, add the field marker \EUDICOp.
  2. For every Shoebox record, enter the relevant speaker-ID into this field.



Note

When the file is exported back to Shoebox (see Section 4.3.2.2), the extension @'Speaker-ID' is automatically dropped from the field marker, and the Shoebox 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 Shoebox database type file, the tiers are brought into a hierarchical relationship and are assigned to linguistic types (see Section 5.1 for details of tier hierarchies and linguistic types). For every tier name a corresponding linguistic type with the same name is created. All of these linguistic types are connected with a stereotype in such a way that it fits with the original Shoebox structure.

   - The Shoebox record marker is assigned to the stereotype None, i.e., it is an independent, time-alignable parent tier.
   - The transcription and parsing fields of Shoebox 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:

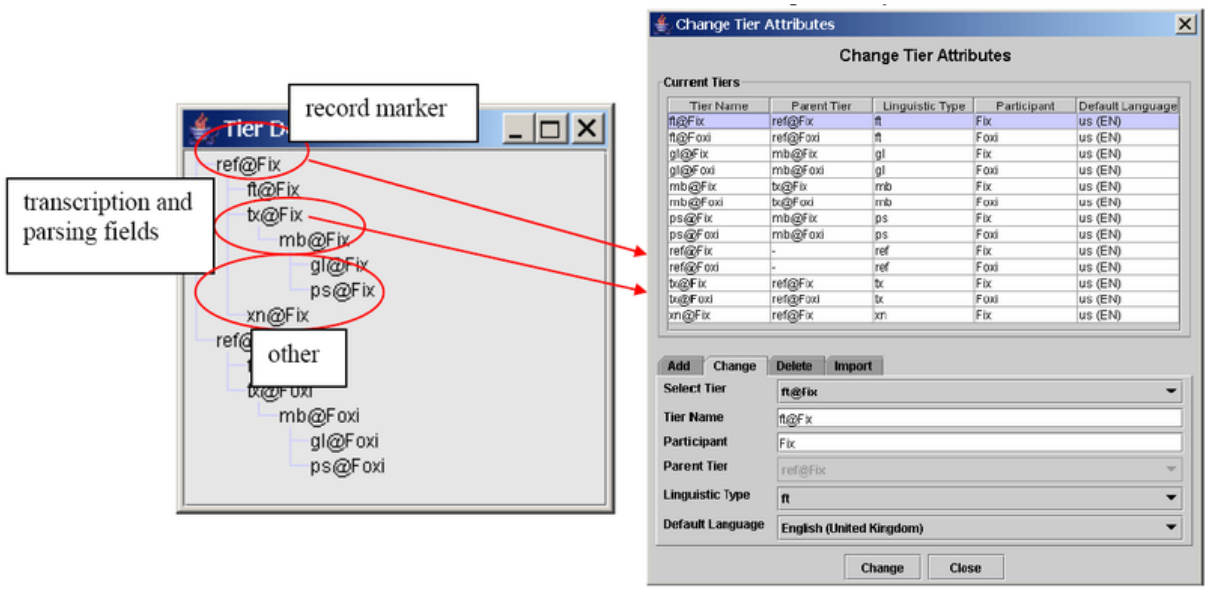


Figure 4.32. Time Subdivision

4. All SIL IPA characters are converted into Unicode characters during import. If you export the file back into Shoebox (see Section 4.3.2.2), the Unicode characters will be converted back into SIL IPA characters.

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

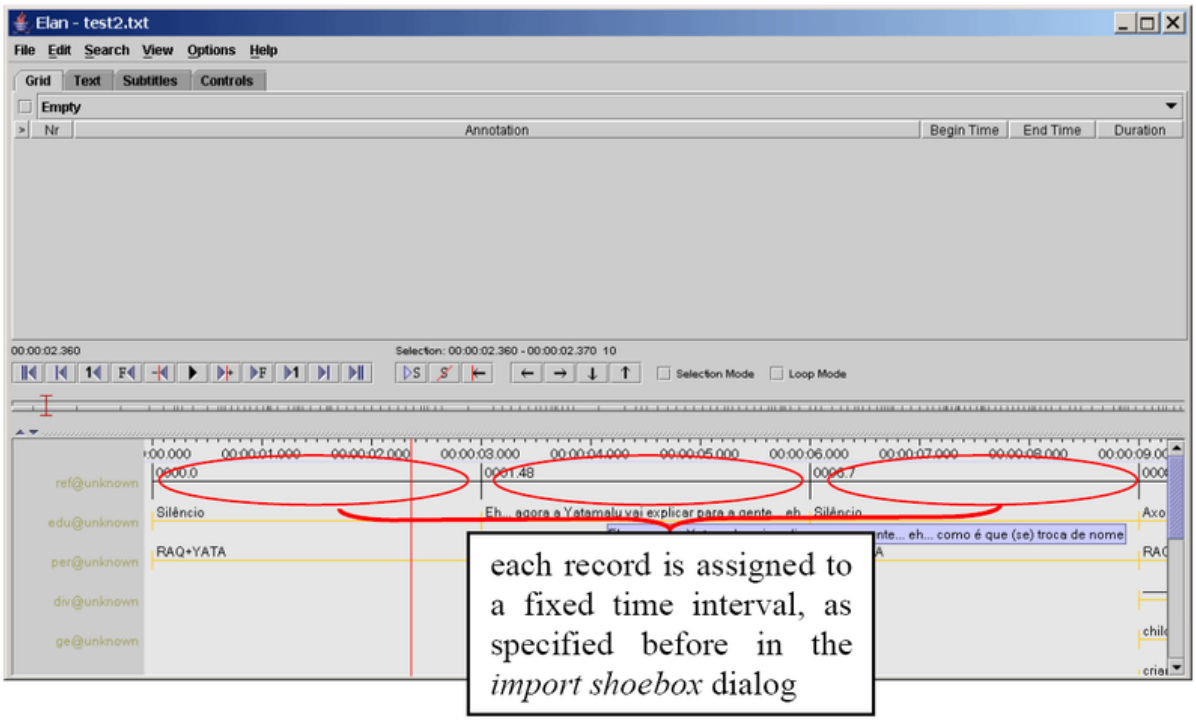


Figure 4.33. Fixed time intervals

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

1. Activate the Bulldozer mode: Click on Options > Propagate Time Changes > Bulldozer Mode (see Section 5.7.8 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 Section 5.7.6 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:

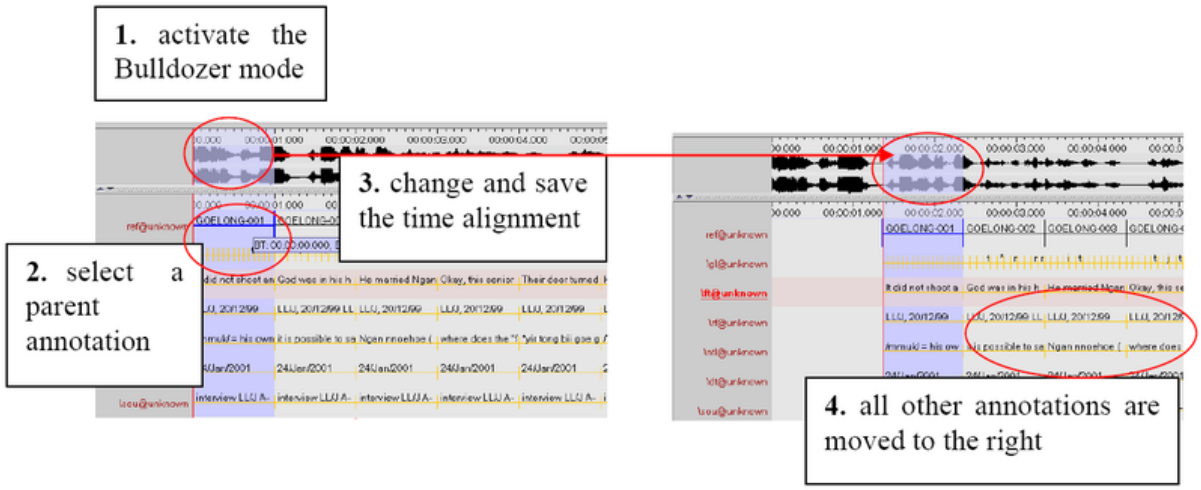


Figure 4.34. Time alignment

After you have done the time-alignment, you can export the file back to Shoebox – in this case, the time code information will be kept (see Section 4.3.2.2). If you then re-import the file back into ELAN, ELAN automatically assigns the Shoebox records to their correct time intervals.

An imported Shoebox file can be saved as an ELAN file (see Section 4.2.5), exported back into Shoebox (see Section 4.3.2.2), or exported as a tab-delimited text (see Section 4.3.2.4).

4.3.1.4. Toolbox file

Importing a document form Toolbox is very much the same as importing a document from Shoebox (see Section 4.3.1.3). The Toolbox import assumes that all markers in the file are Unicode (although it still allows to import files in which all markers are in ISO-Latin if you uncheck All markers are Unicode). This alternative to the Shoebox import attempts to allow more flexibility in terms of tier relations and tries to prevent that words are cut up in case of misalignment. As with the Shoebox 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.

4.3.1.5. Fieldworks Language Explorer (FLEx) file

ELAN can import documents from the SIL Fieldworks Language Explorer (FLEx). To do so click File > Import > FLEx File.... Select the FLEx file and relevant media files by clicking the ...-buttons. Determine whether you want to import the "interlinear-text" and "paragraph" elements and what the smallest time-alignable element should be. Finally enter the duration of the whole file or the initial duration of the smallest time-alignable element you have chosen. Start the import by clicking OK.

4.3.1.6. CSV / Tab-delimited Text Files

A CSV (Comma Separated Values) or Tab-delimited Text 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 4.35. 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.

```
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
```

Figure 4.35. Tab-delimited Text

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 4.36). 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.

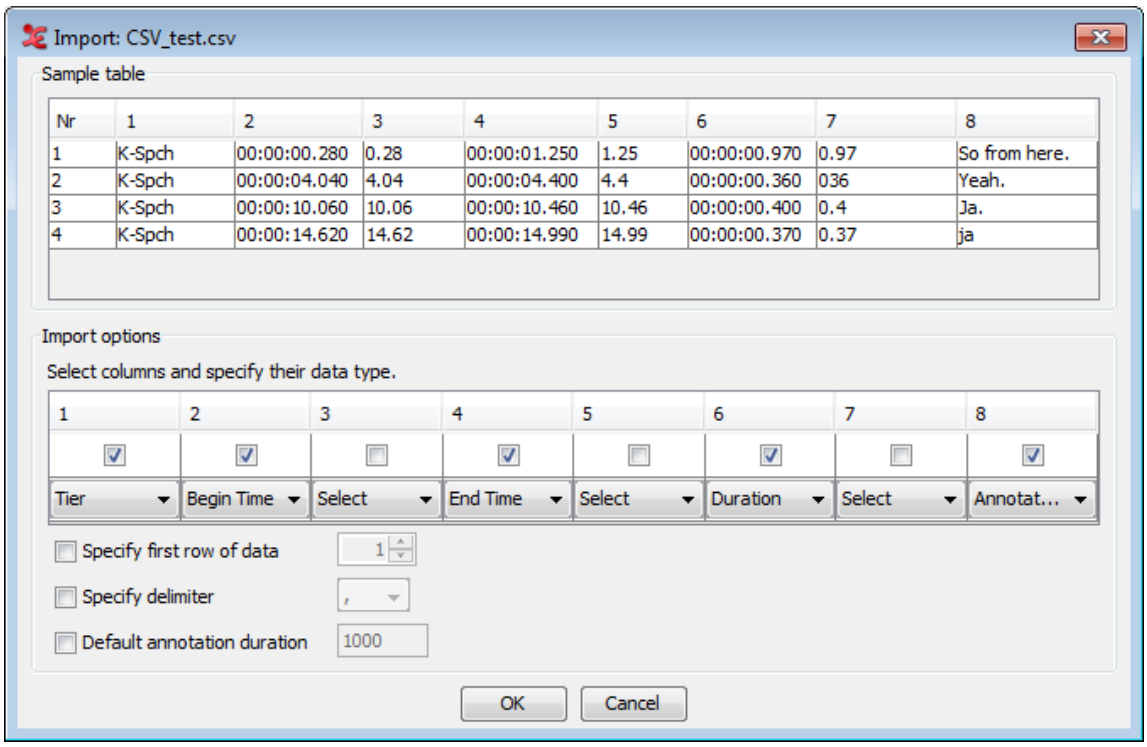


Figure 4.36. Import CSV / Tab-delimited Text

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 and semi-colon.

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.

Finally click OK to import the data. 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:

```
K-Spch          V-Spch
so from here.   so you go out of the Institute to the Saint Anna Straat.
yeah            and then you go the other, Saint Anna Straat to this to the center of the town, to this big rotunde.
ja              and you follow then the sign Kleef
ja              that's the oranje single
```

Figure 4.37. 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:

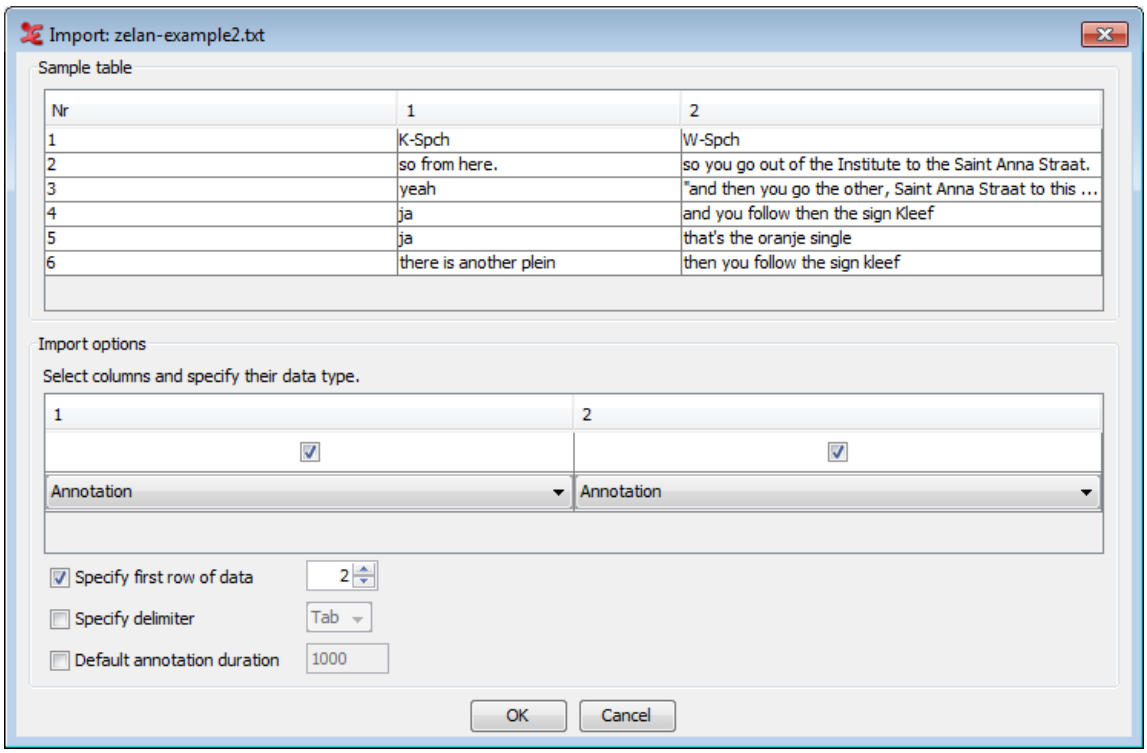


Figure 4.38. Import CSV / Tab-delimited Text

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 part of is specified as 'annotation'. This results in this example in two tiers: K-Spch and W-Spch.

For information on merging a CSV file with an existing *.eaf file, please visit www.lat-mpi.eu/tools/elan/testforum/copy_of_post20091207138021245 6/view []

4.3.1.7. 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. Finally, check Skip empty intervals / annotations if you want to do so.

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.

4.3.1.8. Tiers from recognizers

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 5.15.

4.3.2. Export as

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

- Shoebox file (Section 4.3.2.2)
- Toolbox file (Section 4.3.2.3)
- Tab-delimited text file (Section 4.3.2.4)
- Tiger XML (Section 4.3.2.5)
- CHAT file (Section 4.3.2.6)
- Traditional transcript file (Section 4.3.2.7)
- Praat TextGrid file (Section 4.3.2.8)
- Alphabetical list of words (Section 4.3.2.9)
- Clip of video file (Section 4.3.2.10)
- Media clip using script (Section 4.3.2.11)
- SMIL clip (Section 4.3.2.12)
- QuickTime text (Section 4.3.2.13)
- Subtitle text (Section 4.3.2.14)
- ELAN's document view (Section 4.3.2.15)
- Interlinear text file (Section 4.3.2.16)
- HTML file (Section 4.3.2.17)
- Filmstrip Image (Section 4.3.2.18)
- Tiers for recognizers (Section 4.3.2.19)

4.3.2.1. How to select tiers

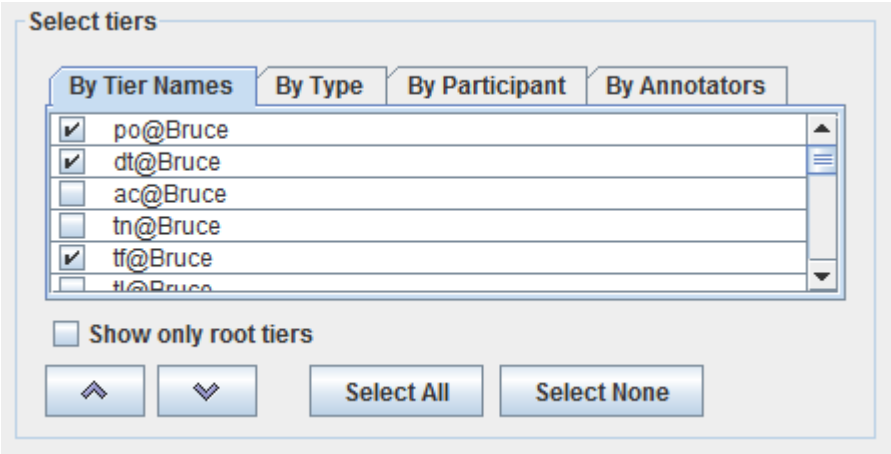


Figure 4.39. Tier Selection panel in most of the dialogs

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 linguistic 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.



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

4.3.2.2. Shoebox file

All Shoebox files that were imported into ELAN (see Section 4.3.1.3) 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 ….

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

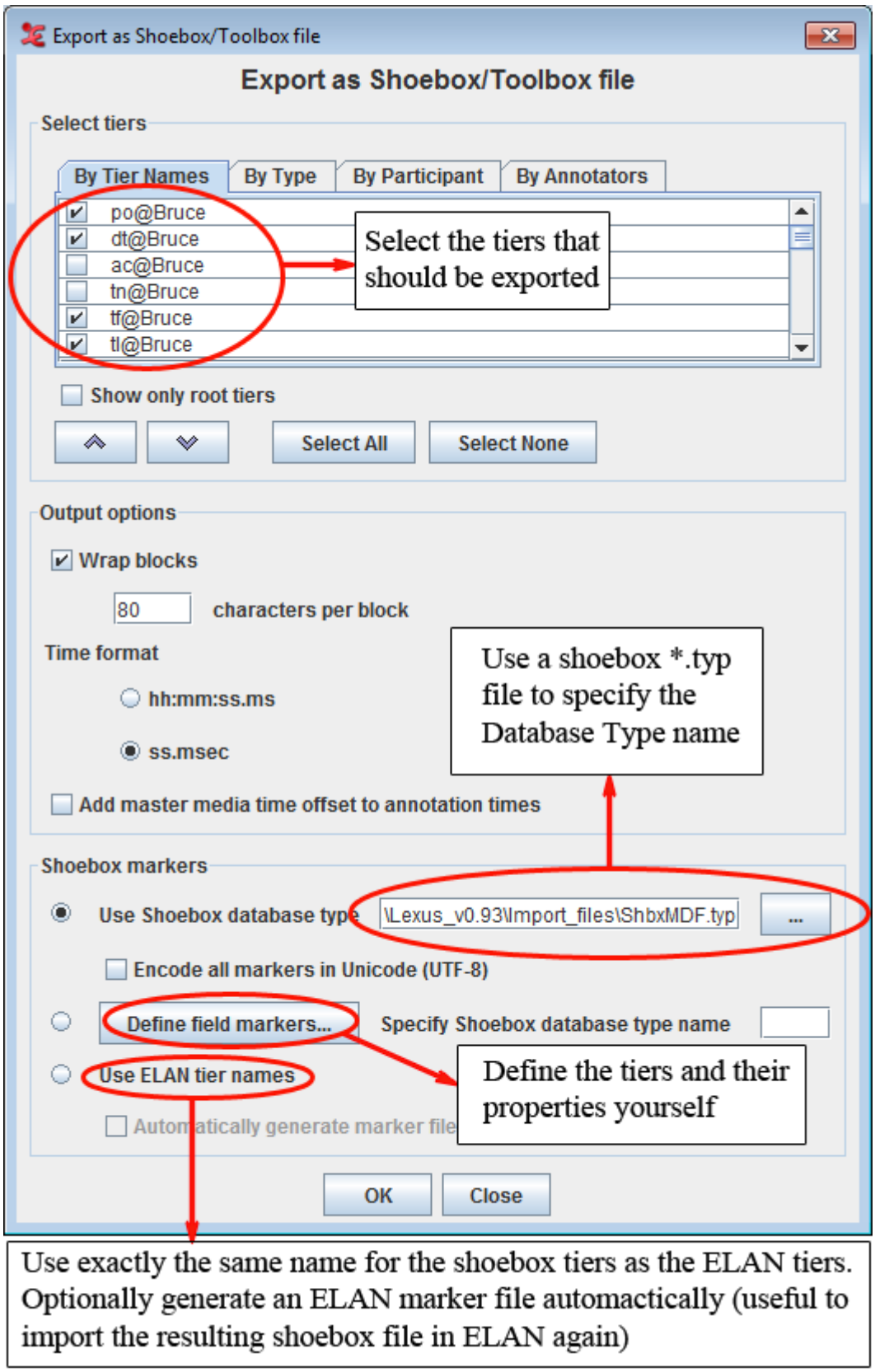


Figure 4.40. Shoebox Export dialog window

- 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 Section 4.2.3).

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

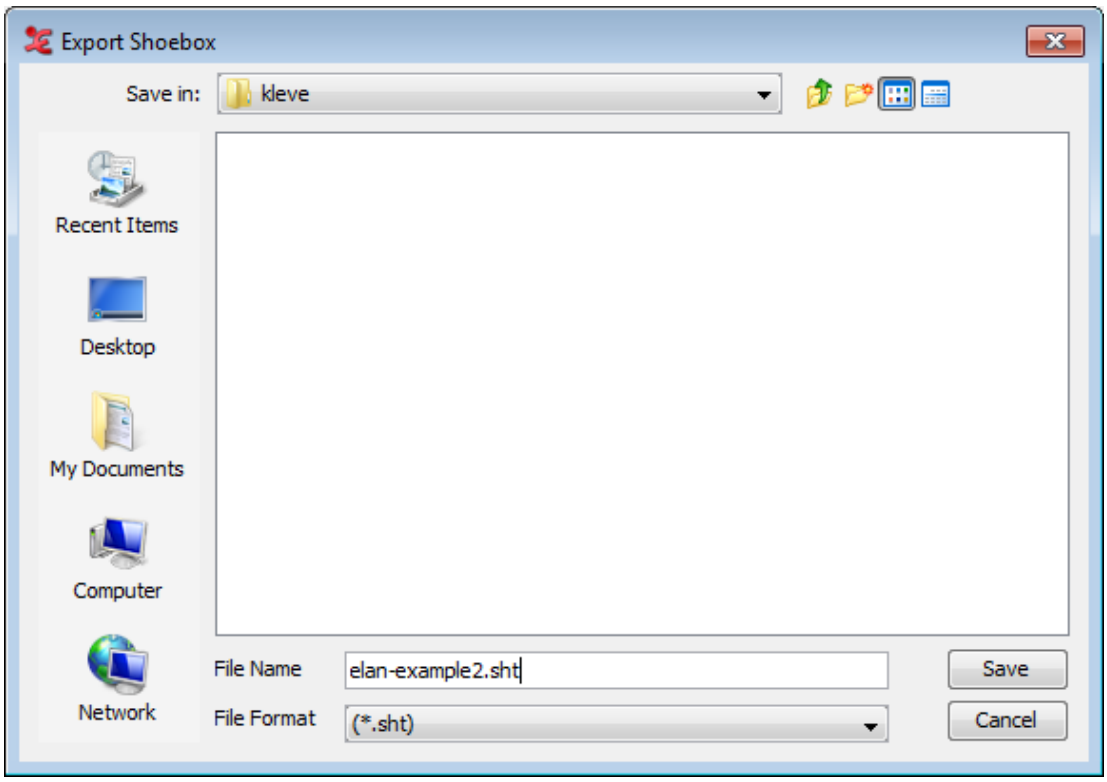


Figure 4.41. Name and directory of exported file

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.:

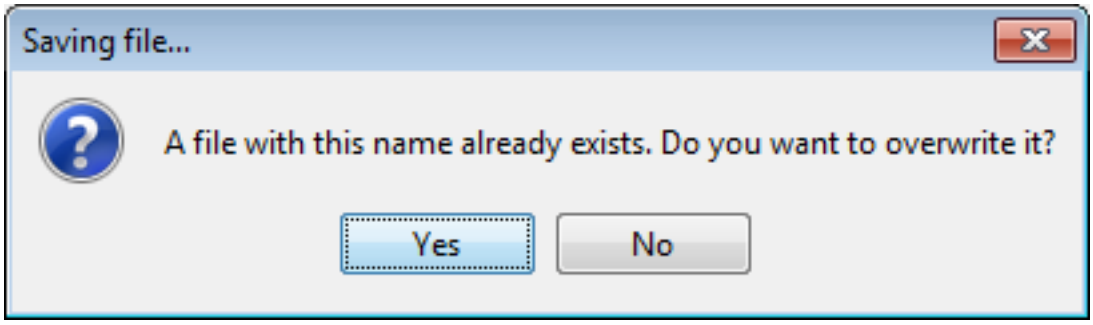


Figure 4.42. File Exists

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 Section 4.3.1.3).

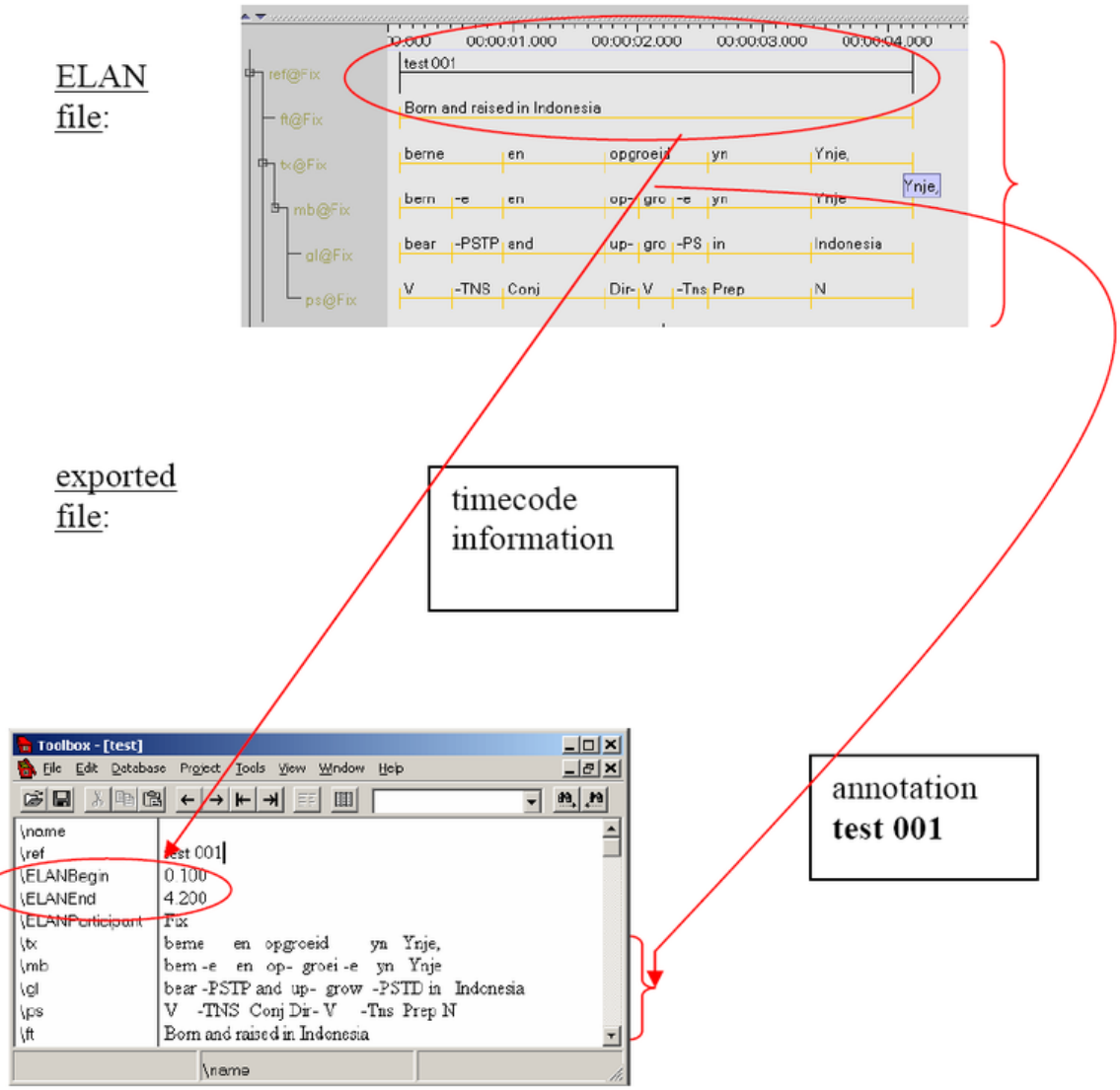


Figure 4.43. ELAN file and exported file

4.3.2.3. Toolbox file(UTF-8)

Similar to exporting a document to Shoebox (see Section 4.3.2.2) 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:

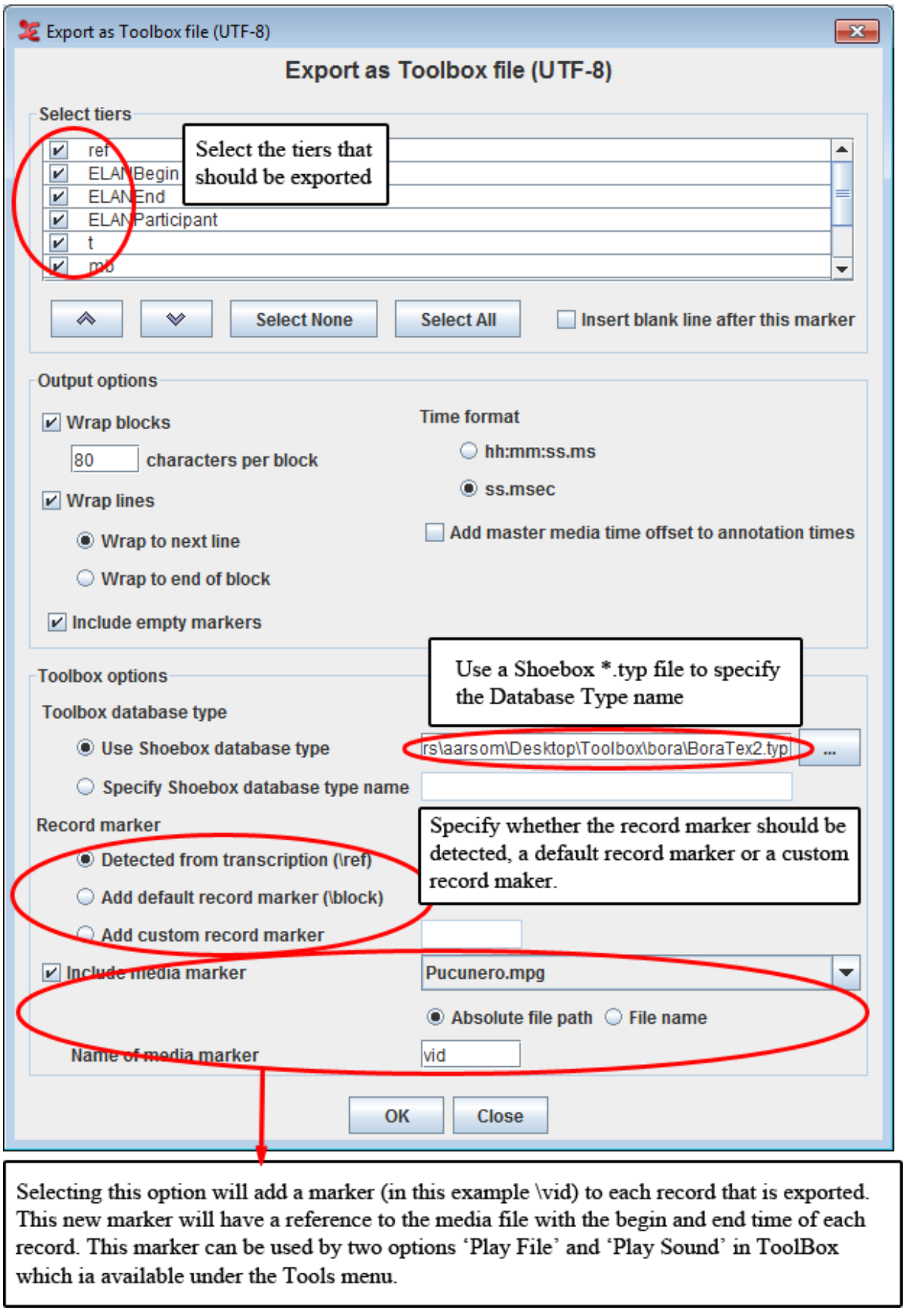


Figure 4.44. Toolbox Export dialog window

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:

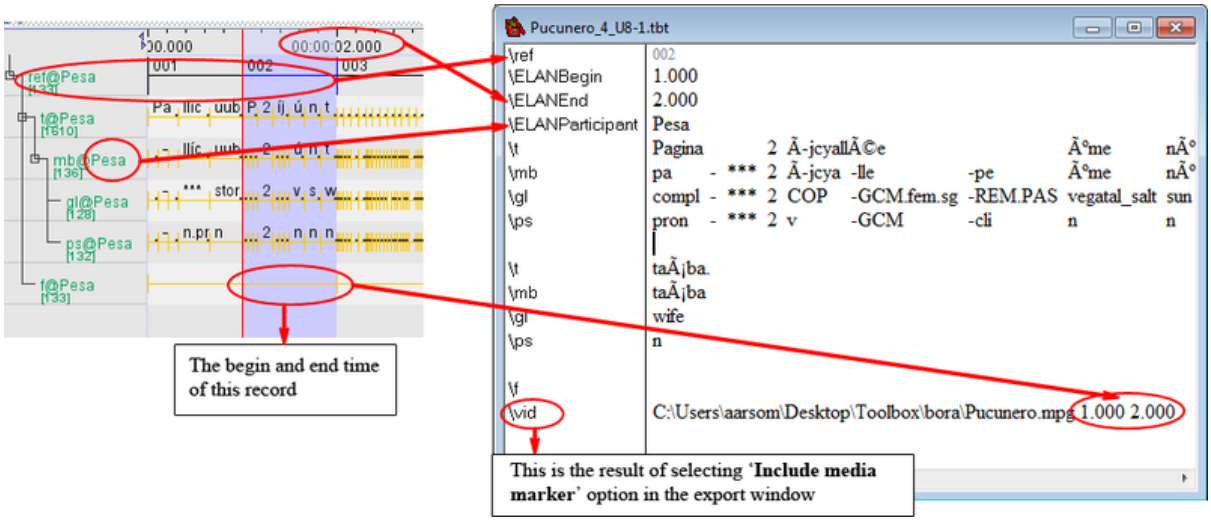


Figure 4.45. ELAN file and exported Toolbox file

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 4.44:

- By first selecting a tier(Section 4.3.2.1) 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 Section 4.2.3).

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 Section 4.3.1.3).

4.3.2.4. 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 Section 4.2.1 and Section 4.2.4) as well as documents that were imported into ELAN from Shoebox (see Section 4.3.1.3) 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.:

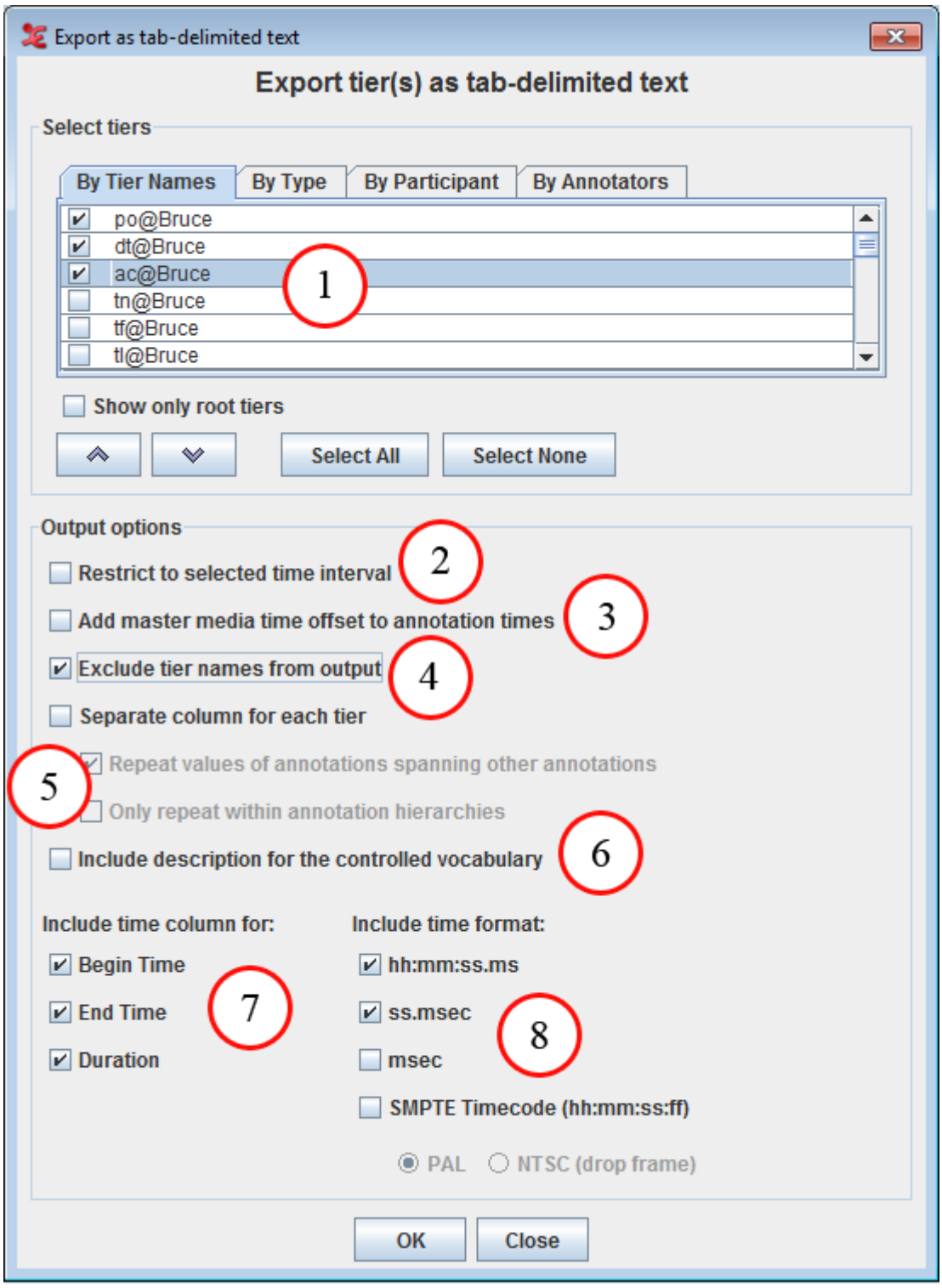


Figure 4.46. Export as tab-delimited text dialog window

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 Section 5.7.1).

      ii. In the Export as tab-delimited text dialog window, click in the box to the left of Restrict to selected time interval. A checkmark 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 checkmark 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.

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 Section 4.2.3).

5. 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 to the same row i.e. the same tab-delimited line.

   - 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.

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

7. 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).

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

9. 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). 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 exported file has the extension `*.txt`.

The exported file contains the following information: participant, begin time of each annotation, end time, total length, content, and tier. It can be opened with any program that can handle tab-delimited texts, e.g., Microsoft Excel.

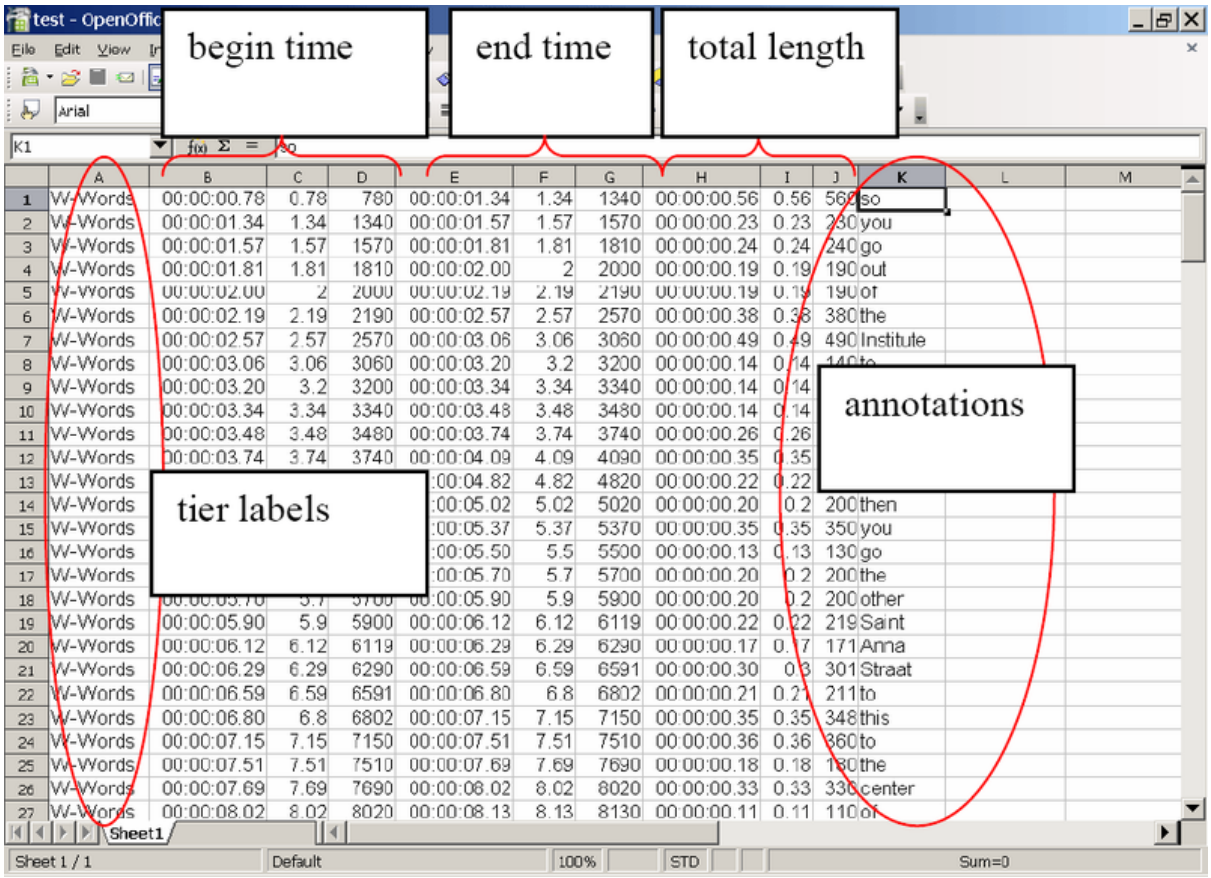


Figure 4.47. Tab-delimited text



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.

4.3.2.5. Tiger XML

If your ELAN annotations contain syntactic elements, it is possible to export these to Synpathy[4] (see http://tla.mpi.nl/tools/tla-tools/older-tools/synpathy/). 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`).

4.3.2.6. CHAT files

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

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).

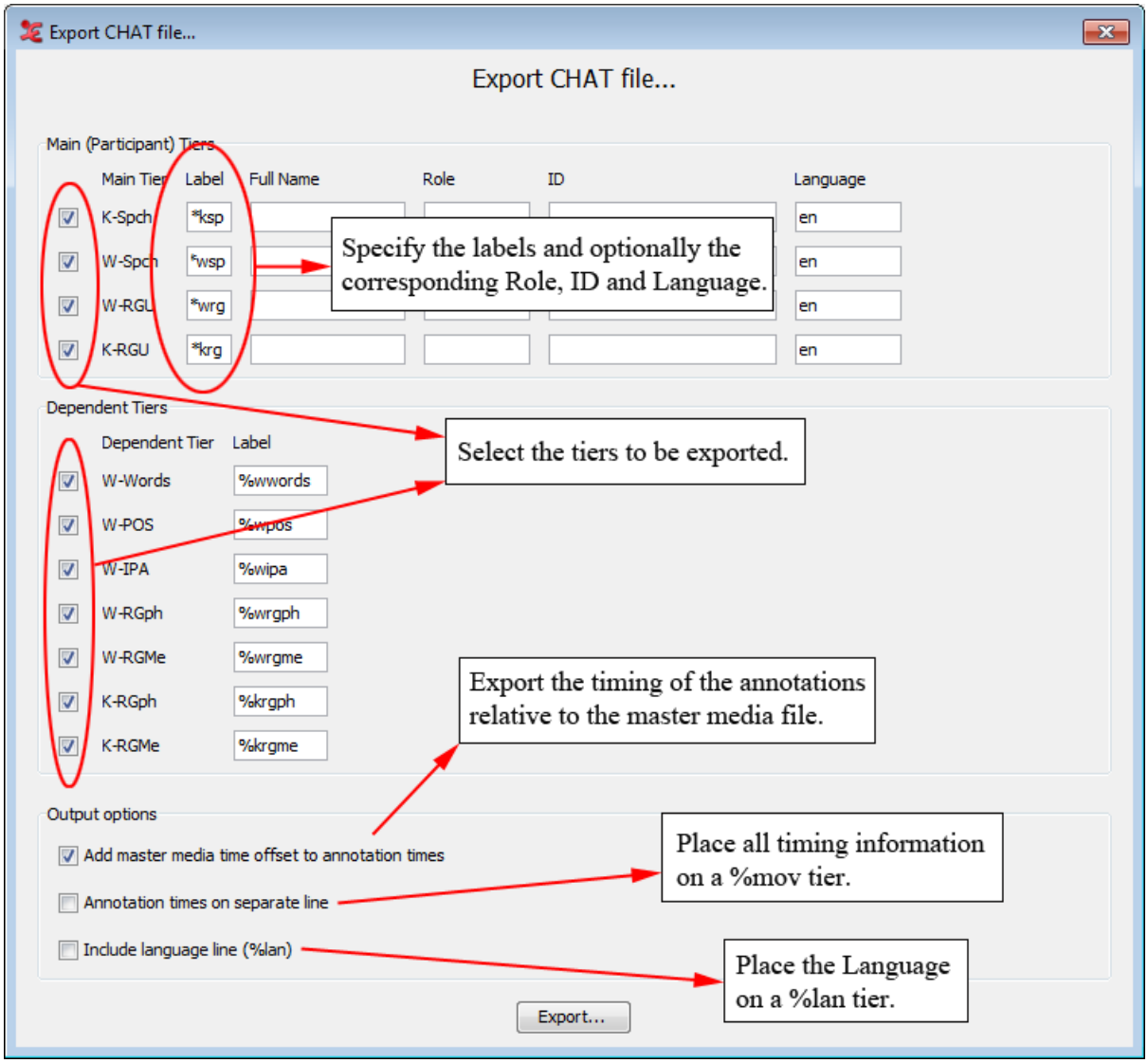


Figure 4.48. Export Chat file

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

4.3.2.7. 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....

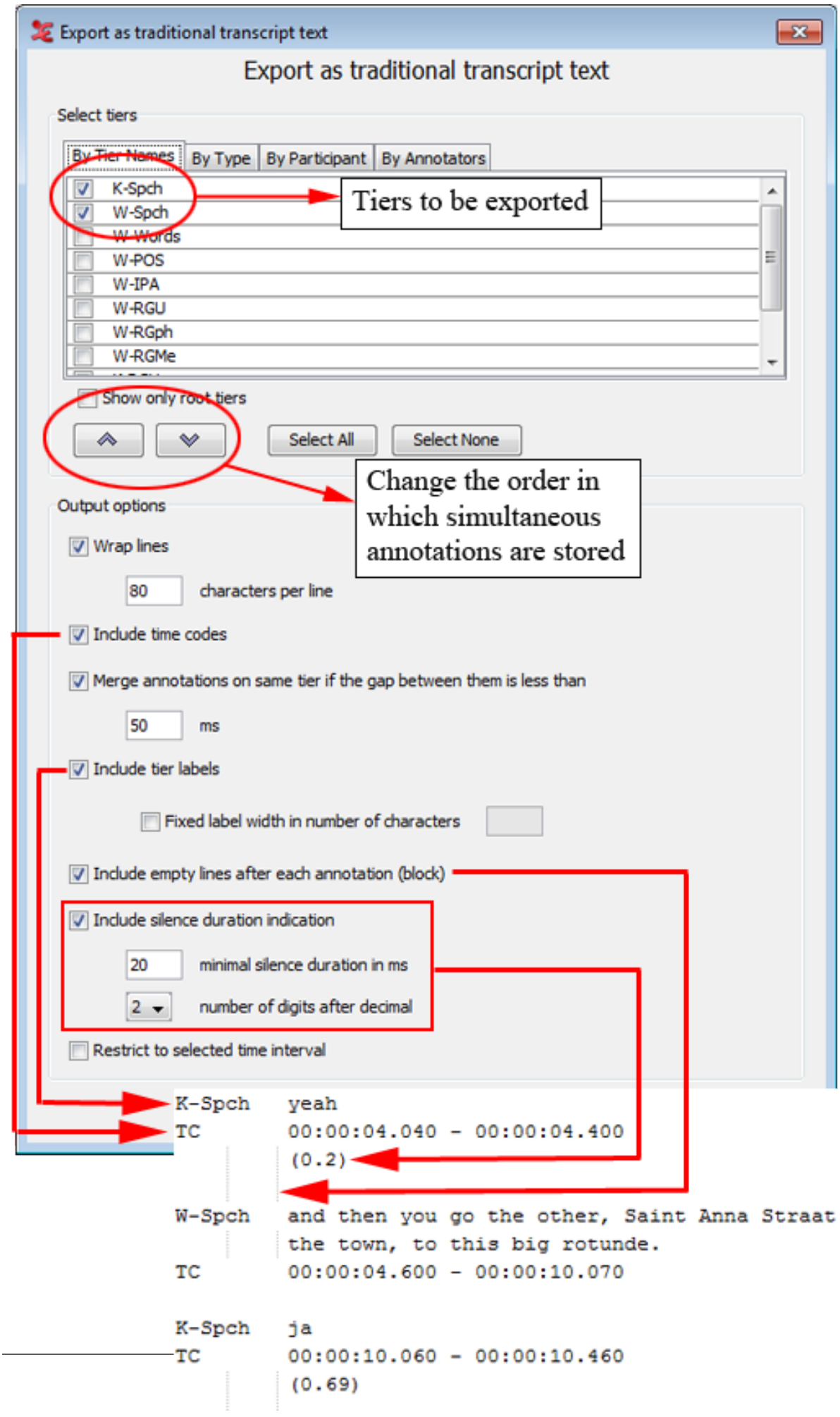


Figure 4.49. Export Transcript Text

As can be seen in one of the options enables you to include silences with a minimal duration. In the figure there is a silence of 0.2 seconds between 'yeah' on the tier K-Spch and 'and the you go the other ...' on the tier W-Spch. The first annotation end at 00:00:04.400 seconds and the second begin 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.

Another option 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.

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.

4.3.2.8. 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(Section 4.3.2.1) and specify whether you want to restrict the output to the selected interval.

After clicking OK, you can enter a filename 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.

4.3.2.9. 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 Section 4.3.2.1) 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. After selecting tiers (or better, deselecting unwanted tiers) you can click OK and choose a filename. Clicking Save will save the word list.

4.3.2.10. Clip of video file

When a command line tool for extracting clips from video files is installed Elan is able to use that tool. At this moment only M2-edit-cl[5] from Mediaware Solutions is supported. If the edit tool is in the user path and a selection is made, there is a menu item to export a video clip of the current selection for each linked video. In that case, follow these steps:

1. Select the part of the video(s) you want to export as (a) clip(s)
2. Choose File > Export As > Media Clip...
3. Enter a filename and press Save

4.3.2.11. Media clip using script

Elan now supports any command line tool to extract clips from the video file. Elan uses a script file named `"clip-media.txt"` which can be found in the Elan folder where Elan is installed. To clip a video file, first the script have to be modified according to specifications of the command line tool used.

For example, the syntax for M2-edit-cl : M2-edit-cl/ in:$begin(fr) / out:$end(fr) $in_file $out_file

M2-edit-cl : the path of the application

[5] See http://www.mediaware.com.au/

in:$begin(fr) : specifies the begin time frame of the clip in frames.

out:$end(fr) : specifies the end time frame of the clip

$in_file : input file

$out_file : output file

Few examples for command line tools are

M2-edit-cl -: windows : M2-edit-cl/ in:$begin(fr) / out:$end(fr) $in_file $out_file

ffmpeX - mac n windows : /Applications/ffmpegX/ffmpeg -sameg -ss $begin(sec.ms) -t -$duration(sec.ms) -i $in_file $out_file

These syntax depends on the command line tool you are using. Look in the script file for more explanation and examples. To clip the video file first make a selection of the video file and select 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 Section 4.4.28.



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 donot 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.

4.3.2.12. SMIL clip

ELAN supports export to SMIL[6]-compliant clips. With a suitable player this enables you to view media files and the associated annotations as a subtitled movie.

4.3.2.12.1. Export SMIL for Real Player

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

For a description of this standard and players see http://www.w3.org/AudioVideo/

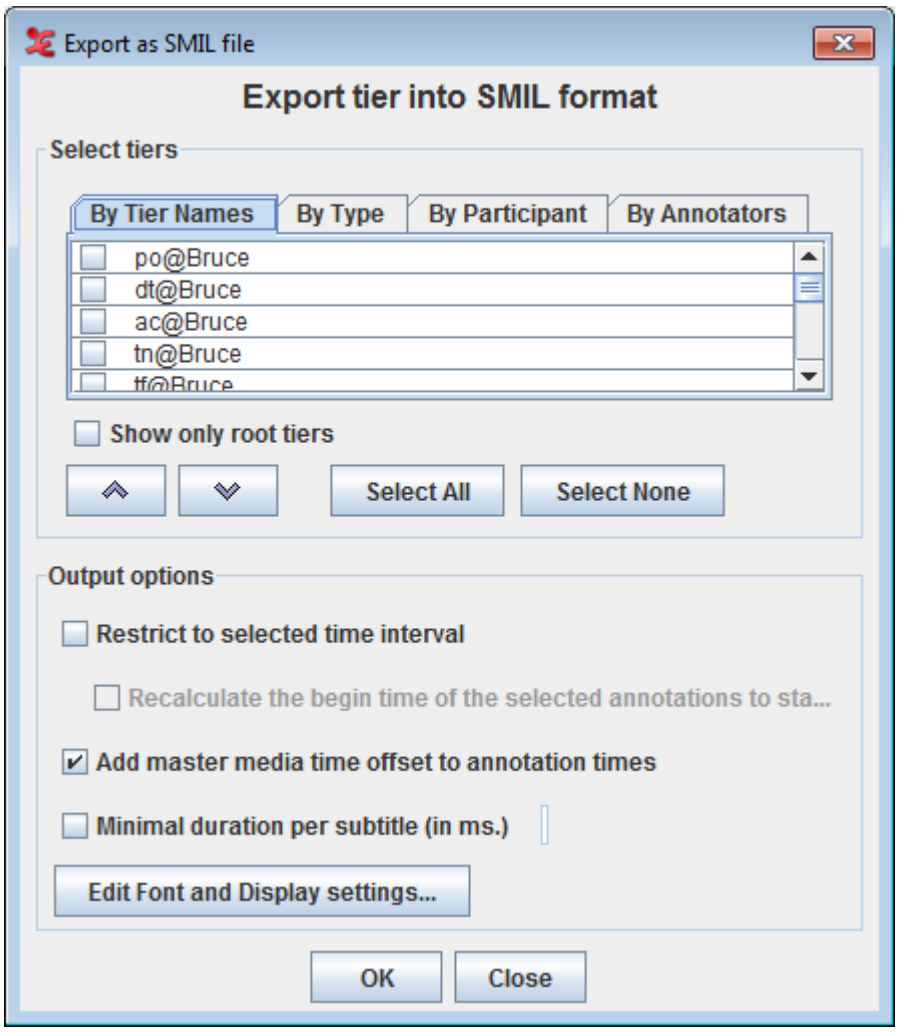


Figure 4.50. Export SMIL Real Player

2. Select the tiers you want to export (see Section 4.3.2.1).

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 Section 4.2.3).

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:

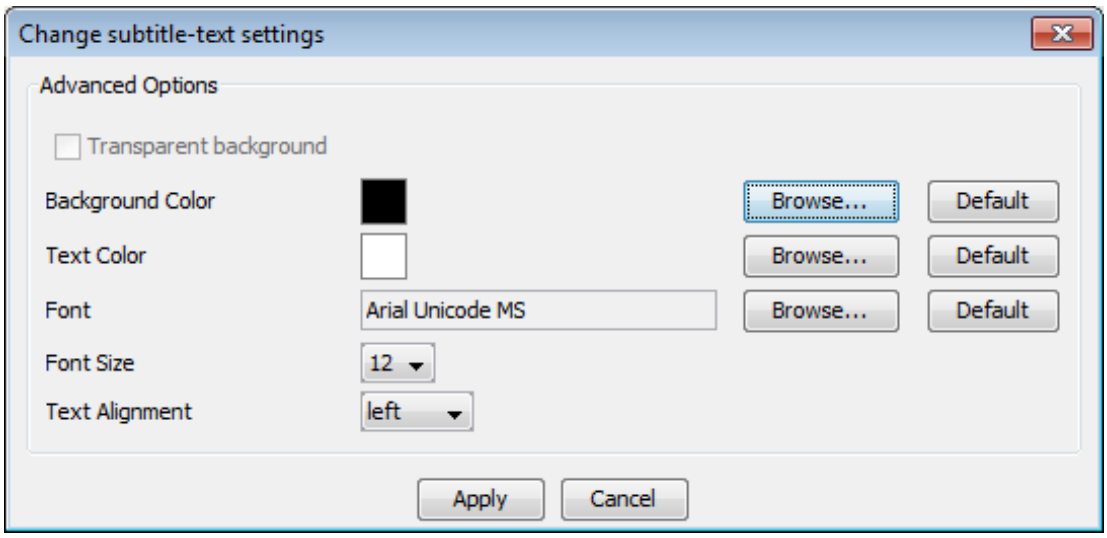


Figure 4.51. Change subtitle text settings

- 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 filename to change the location where the SMIL clip will be saved.

4.3.2.12.2. Export SMIL for Quick Time

Exporting SMIL for Quick time is very much the same as exporting SMIL for real player (see Section 4.3.2.12.1). 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 4.51) which pops up by clicking the Edit Font and Display Settings... button. Finally click on OK to export.

4.3.2.13. 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 Section 4.3.2.1) 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 Section 4.2.3).

- 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 4.51)

Finally click on OK.

4.3.2.14. Subtitle Text

Besides the QuickTime subtitle Text (see Section 4.3.2.13) ELAN can export annotations to there are few other subtitle formats: SubRip (`.srt`), Spruce (`.stl`), Timed Text Marlup Language(ttml) (`.xml`) and LRC (`.lrc`) . Click on File > Export As > Subtitle Text... and select the tiers(see Section 4.3.2.1) 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 form 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).

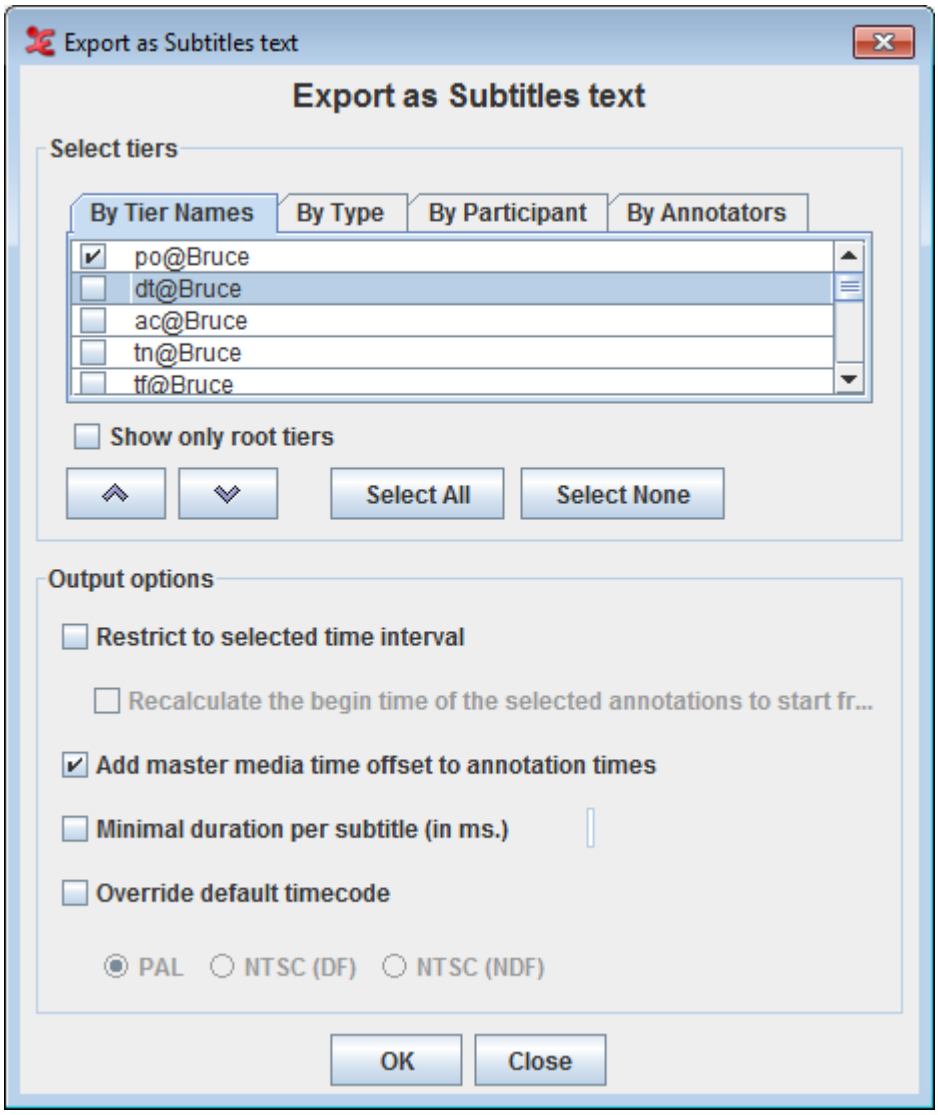


Figure 4.52. Export as Subtitles text

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

4.3.2.15. ELAN's document view

To export ELAN's document view (i.e. to make a screenshot):

1. choose File > Export As > Image from Elan Window...
2. Enter a filename 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.

4.3.2.16. 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 Section 4.2.15. 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.

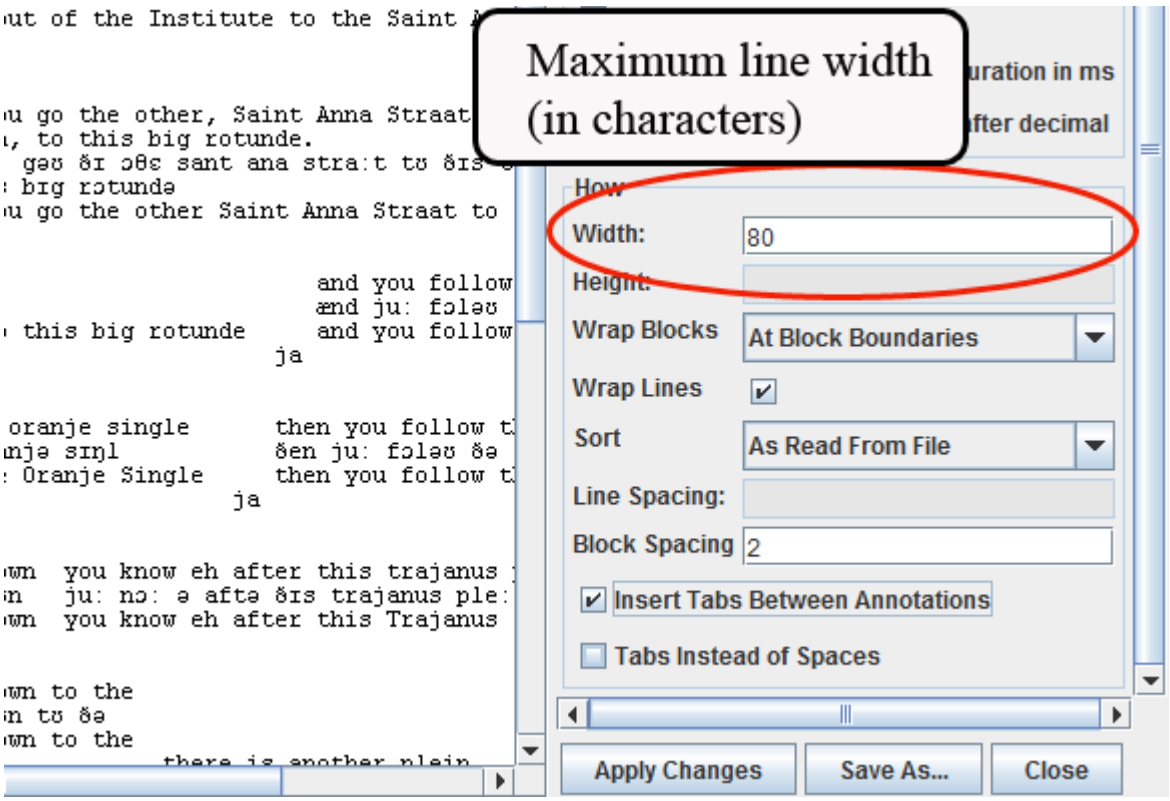


Figure 4.53. Maximum line width

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 whitespace 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 whitespaces. To do that tick Tabs Instead of Spaces box if you prefer to have tabs instead of multiple white spaces.

4.3.2.17. HTML file

Similarly to the export to interlinear text (see Section 4.3.2.16) you can also export annotations to a HTML file, through the File > Export as > HTML... menu.

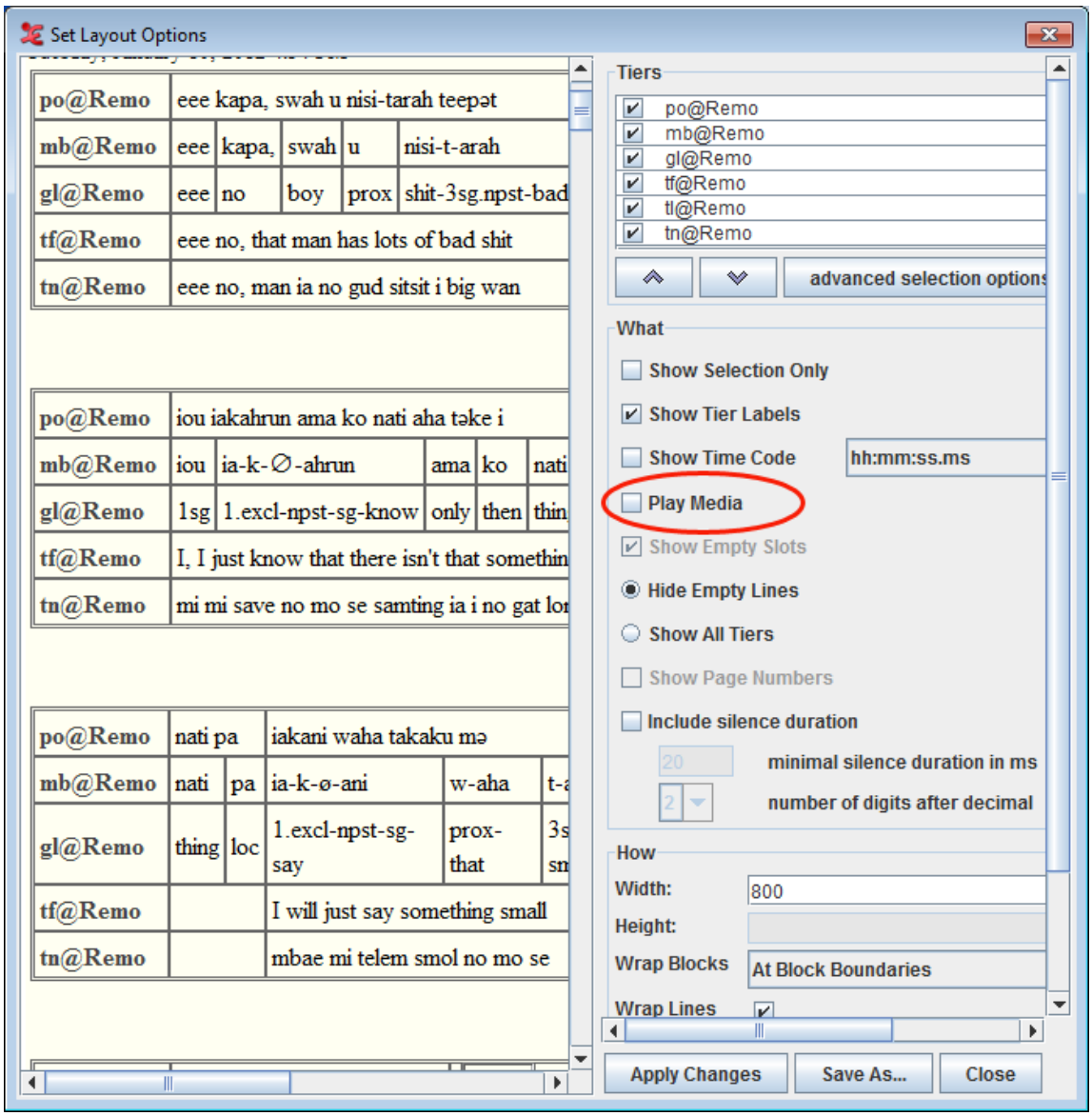


Figure 4.54.

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.

4.3.2.18. Filmstrip Image

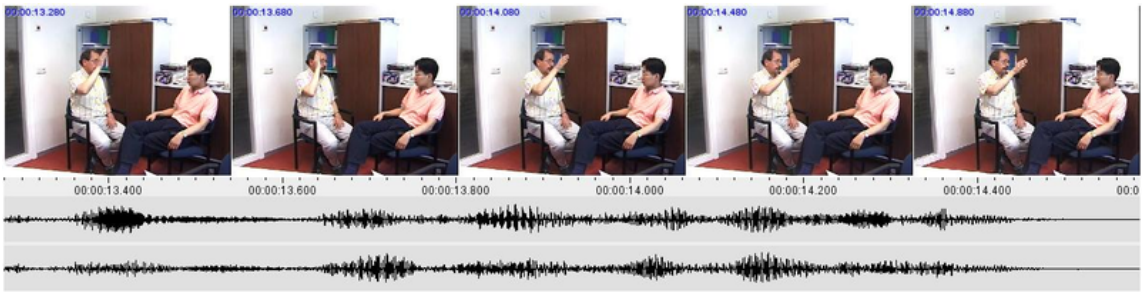


Figure 4.55. An exported filmstrip image

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 4.56) 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 filename and click on Save.

An example or an exported filmstrip image can be seen in Figure 4.55.

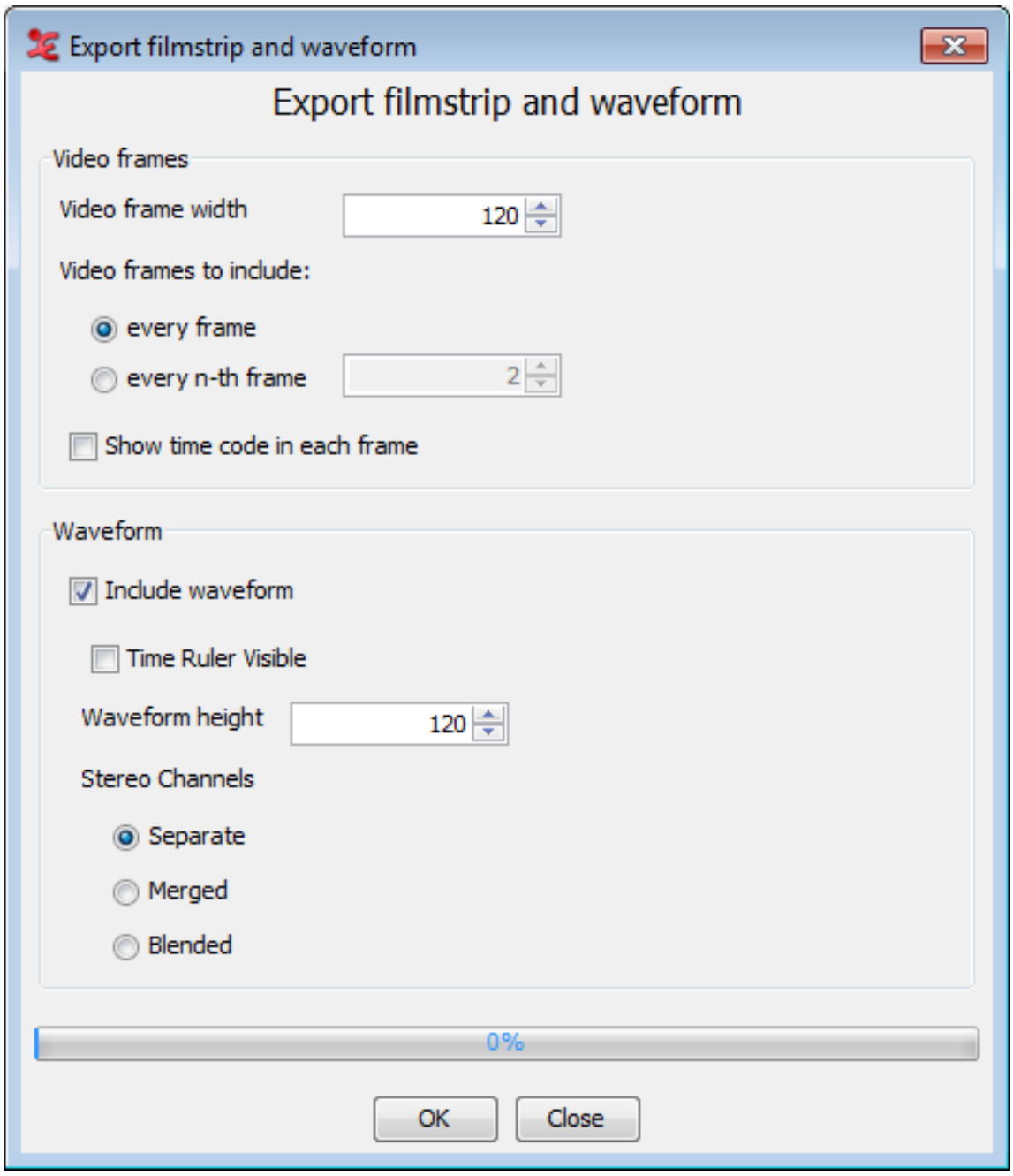


Figure 4.56. Exporting to a filmstrip image

4.3.2.19. Tiers for recognizers

Tiers for the recognizers are exported in the AVATech tier format. For more information on the AVATech tier format see http://tla.mpi.nl/projects_info/avatech/

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

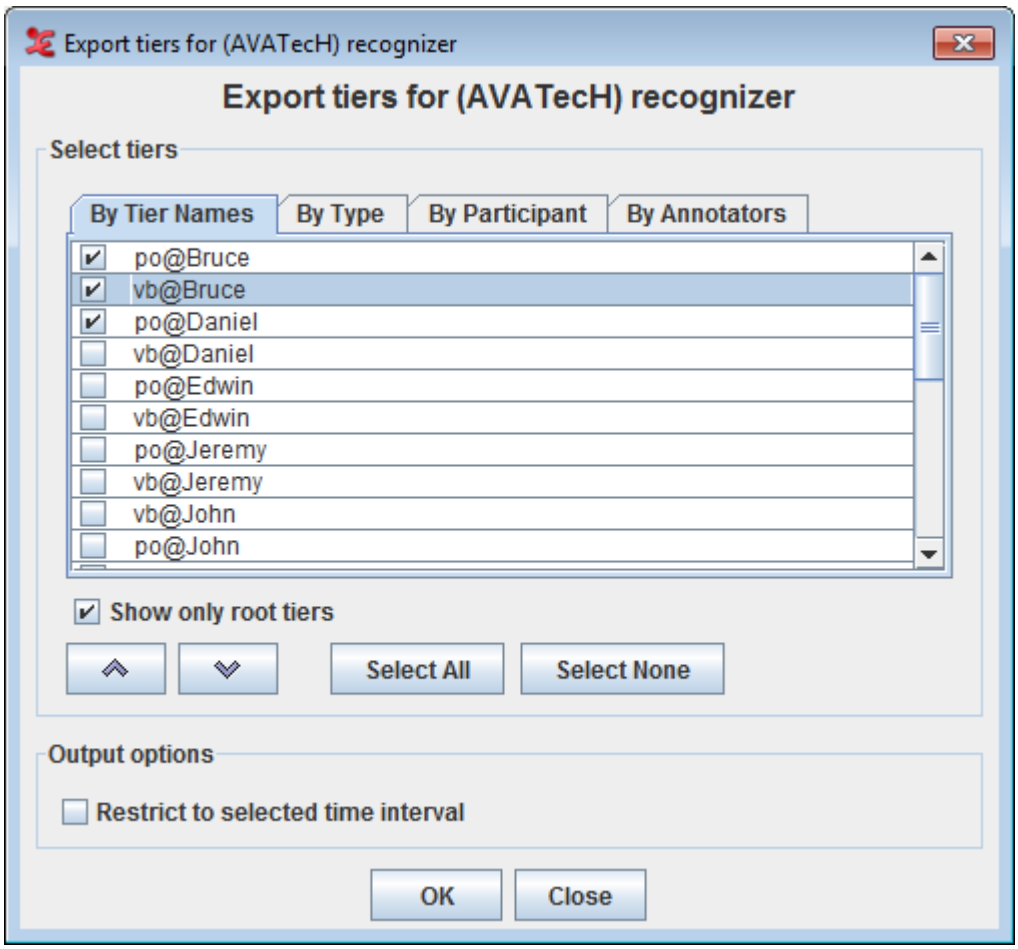


Figure 4.57. Tiers for AVATech recognizers

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. Click OK to export the tiers and give a filename, where the tiers can be exported.

4.4. 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 (Section 4.4.1), and explains the navigation through it (Section 4.5 and Section 4.6).

4.4.1. Basic Information: The ELAN window

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

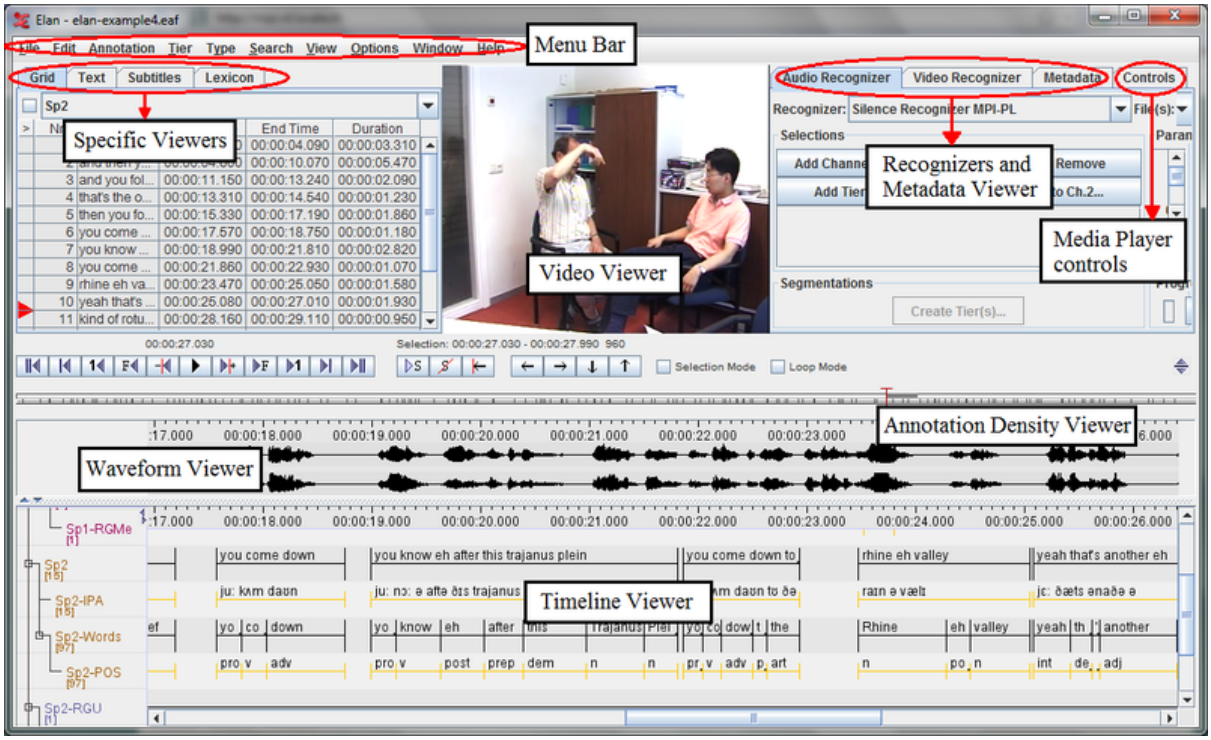


Figure 4.58. The ELAN window

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.

4.4.2. The Video Viewer

The Video Viewer displays up to four video images (of the `*.mpg`, `*.mov` or other 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.

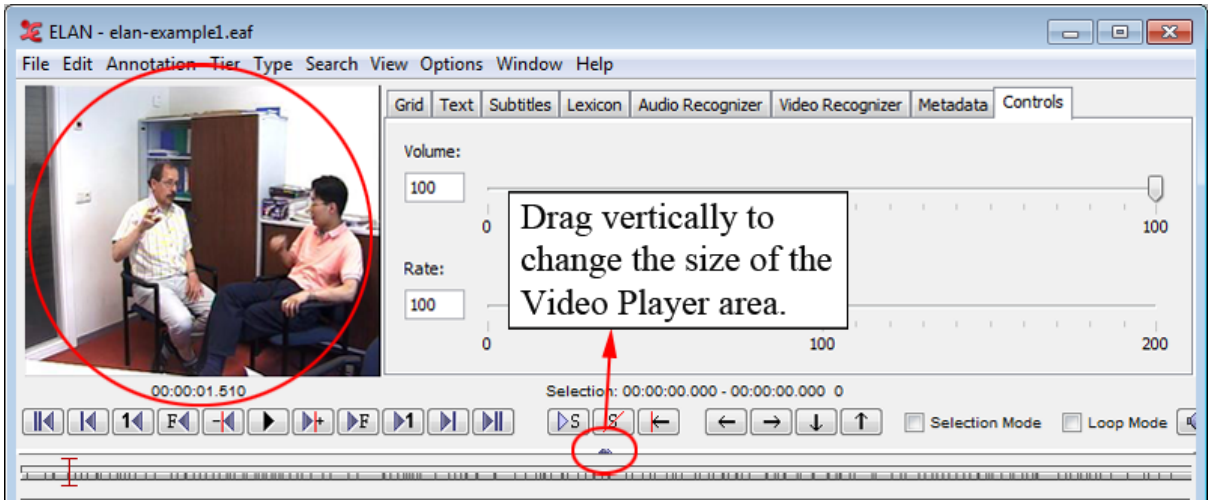


Figure 4.59. The video viewer

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 4.59).



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

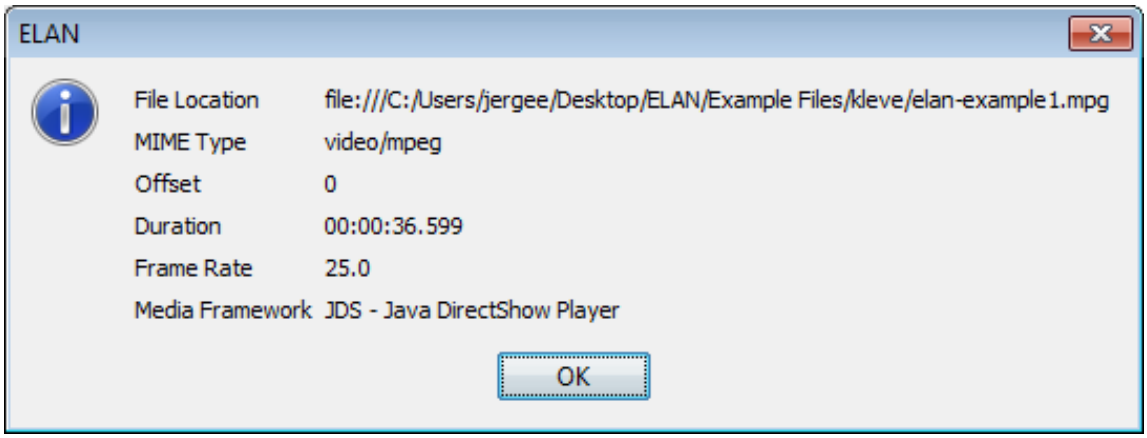


Figure 4.60. Video information dialog window

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…)

[7] For users of a one button mouse on Apple computers: hold the CTRL button and click



Note

Saving a static picture might sometimes cause a freeze of the program on MacOS X. This issue is being investigated.

Clicking on a video 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)

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-ajusted 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 Section 4.4.28 - Media option

4.4.3. 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.

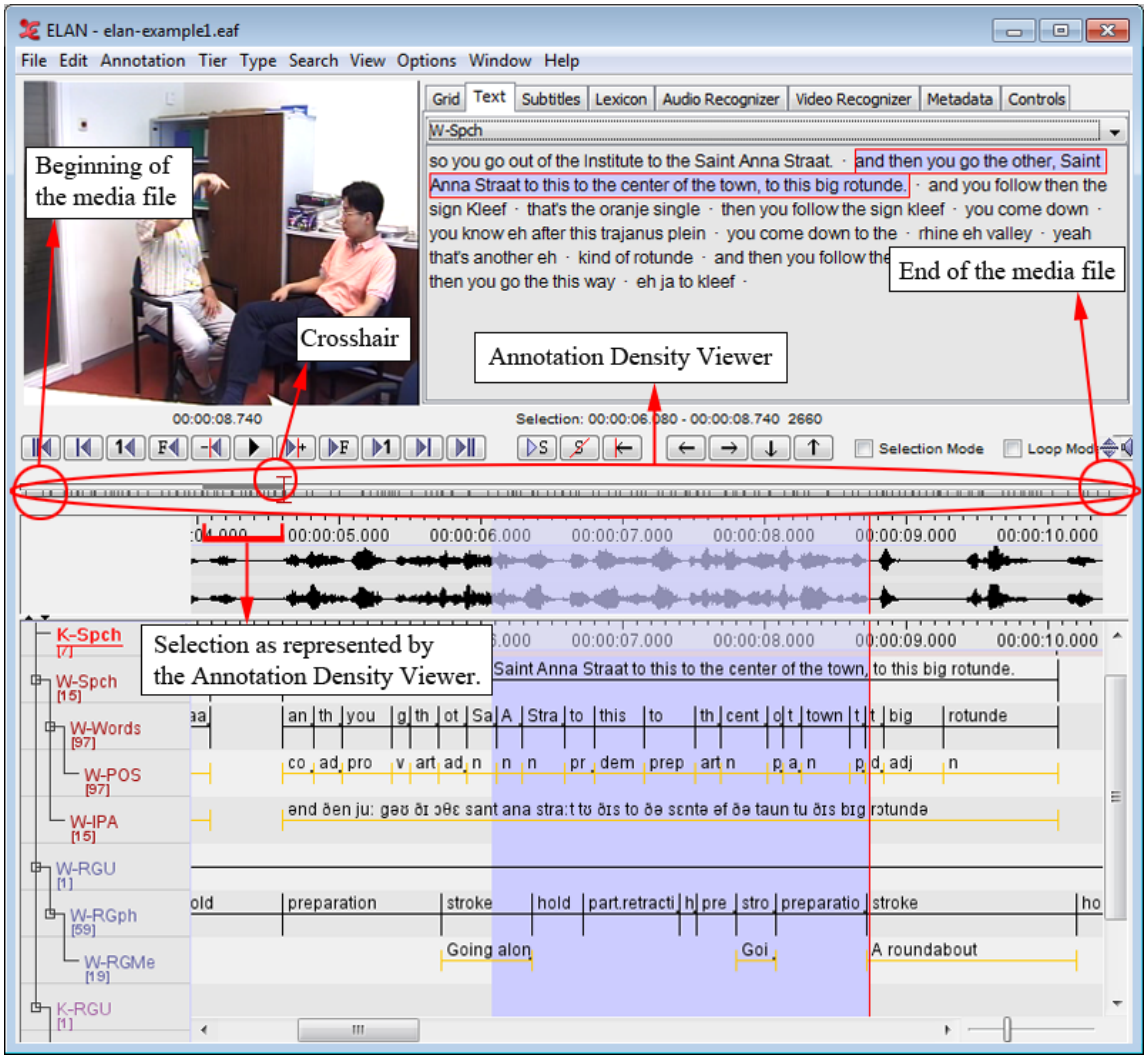


Figure 4.61. Annotation density viewer

4.4.4. The Waveform Viewer

The Waveform Viewer displays the waveform of the audio file[8] (`*.wav`). 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 Section 4.1). 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.

[8] Different resolutions are supported: 8 bits (mono and stereo), 16 bits (mono and stereo) and 24 bits (stereo). Both PCM and A-law encoded wave files can be loaded.

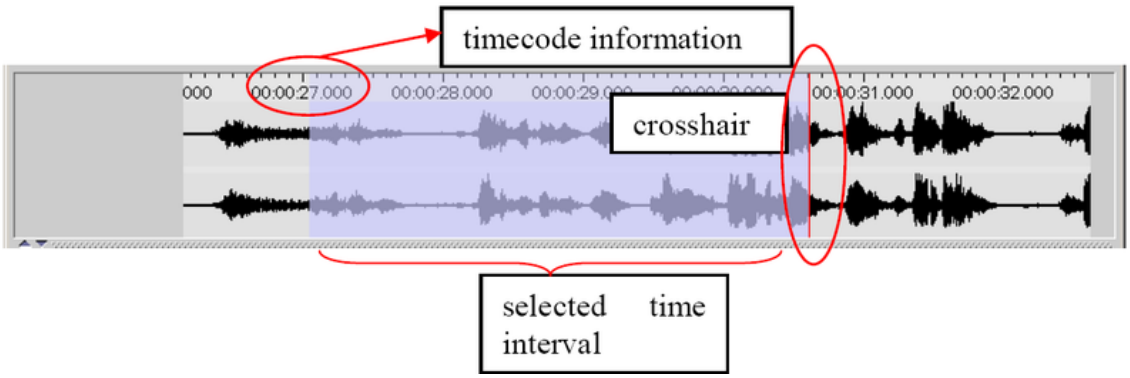


Figure 4.62. The waveform viewer

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 showed, 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:

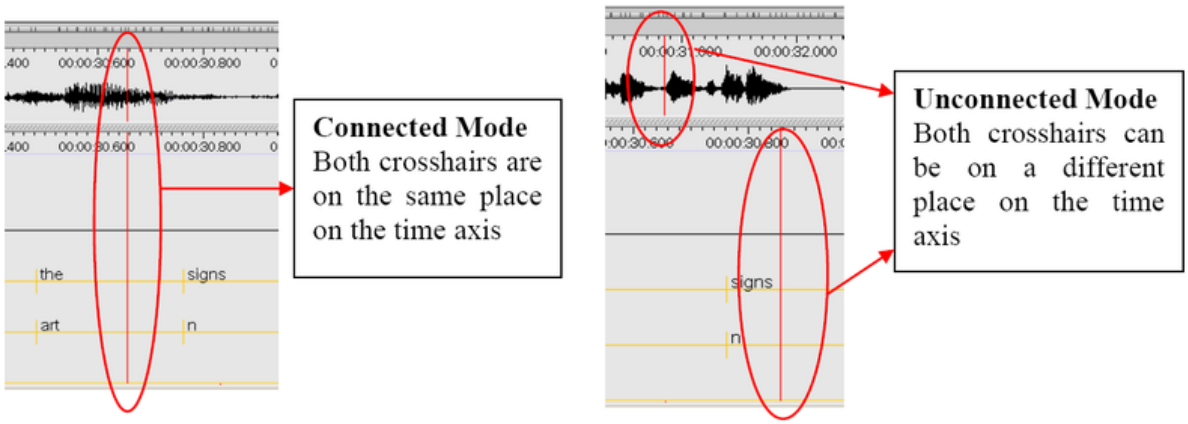


Figure 4.63. Waveformviewer modes

4.4.5. The Subtitle Viewer

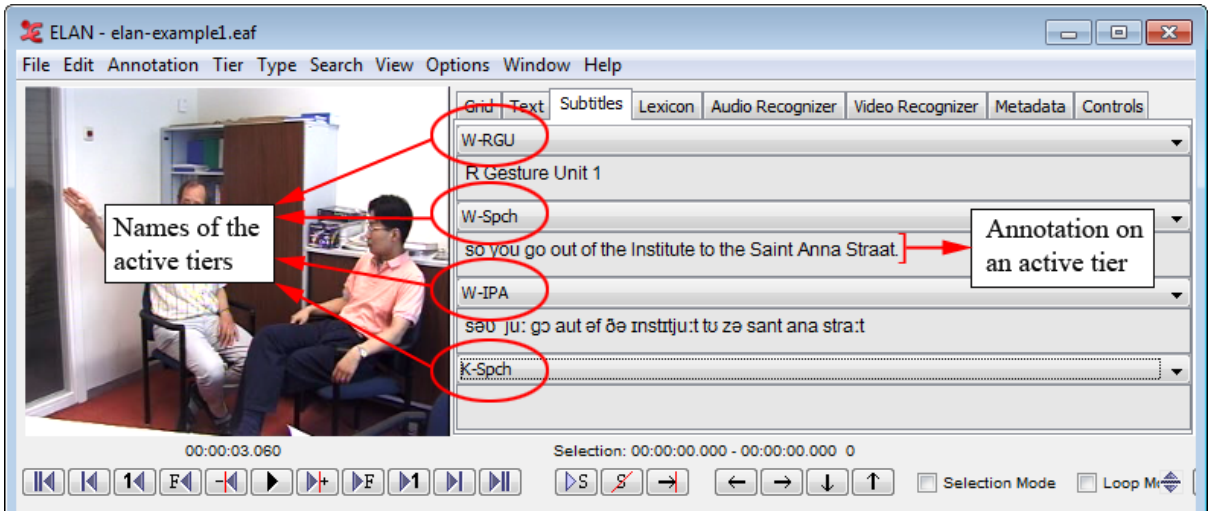


Figure 4.64. The subtitle viewer

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.

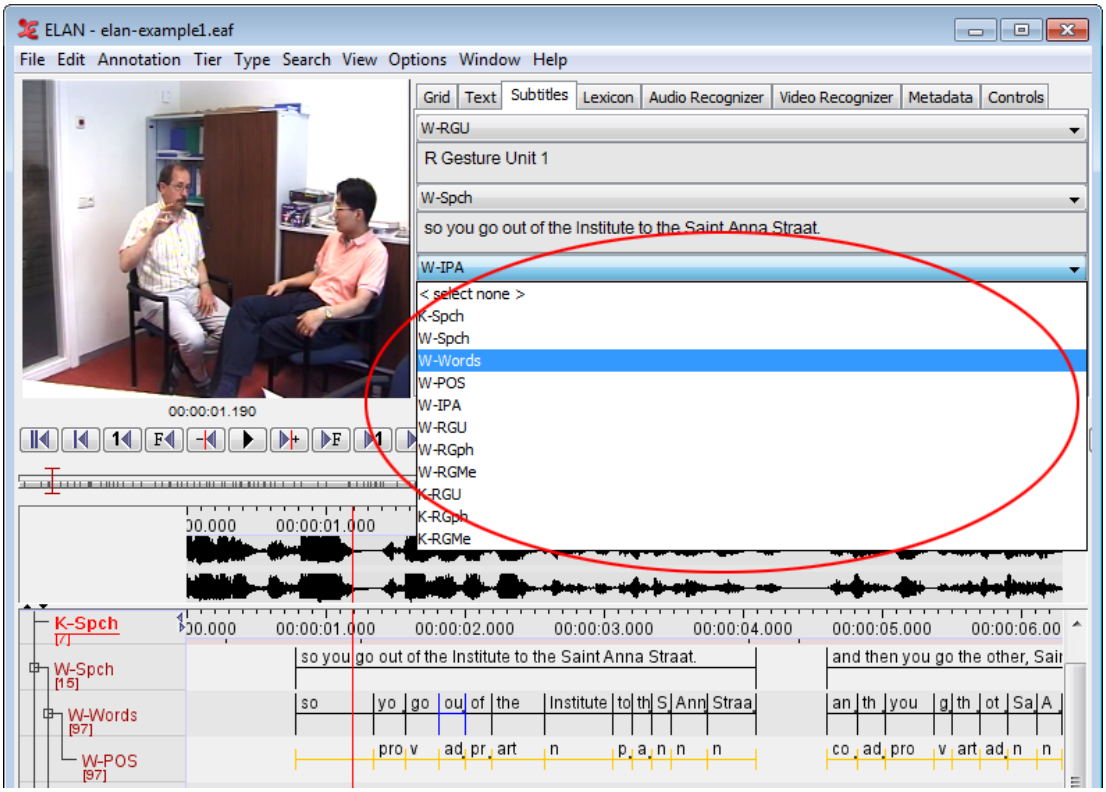


Figure 4.65. Display multiple tiers as a subtitle

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.

4.4.6. 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.

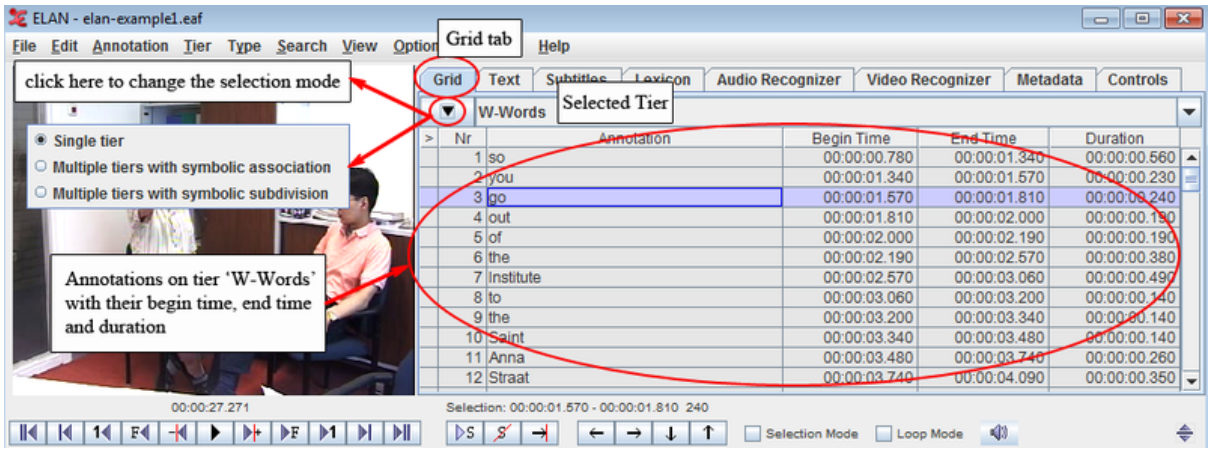


Figure 4.66. The grid viewer

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 Section 5.1). Empty cells of dependent tiers can also be filled in this way.

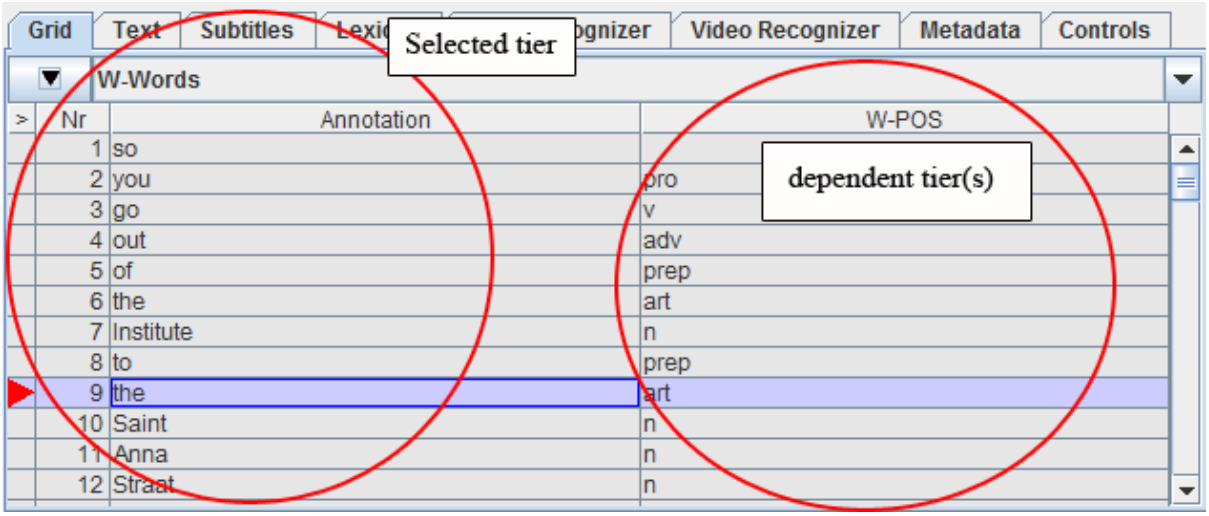


Figure 4.67. Multiple tiers with symbolic association in the grid viewer

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

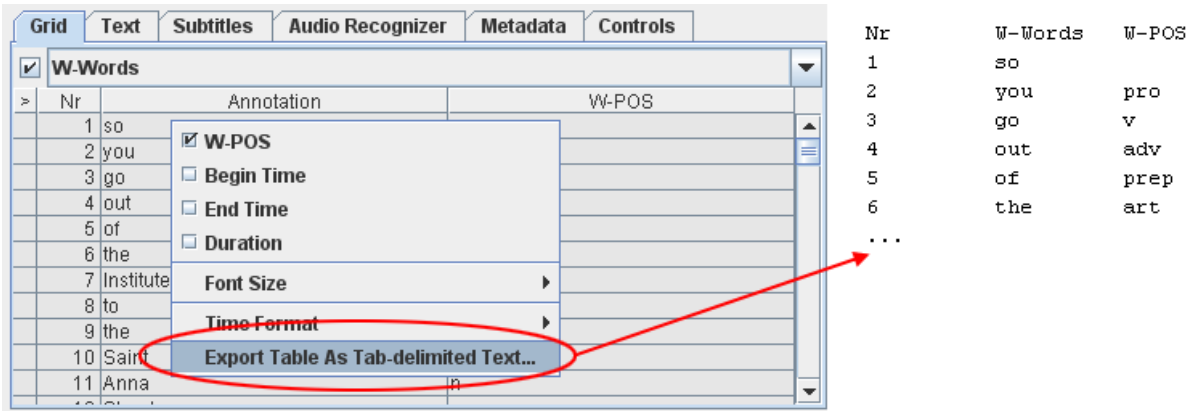


Figure 4.68. Export from the grid viewer

4.4.7. 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 Section 4.4.11)
- 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.

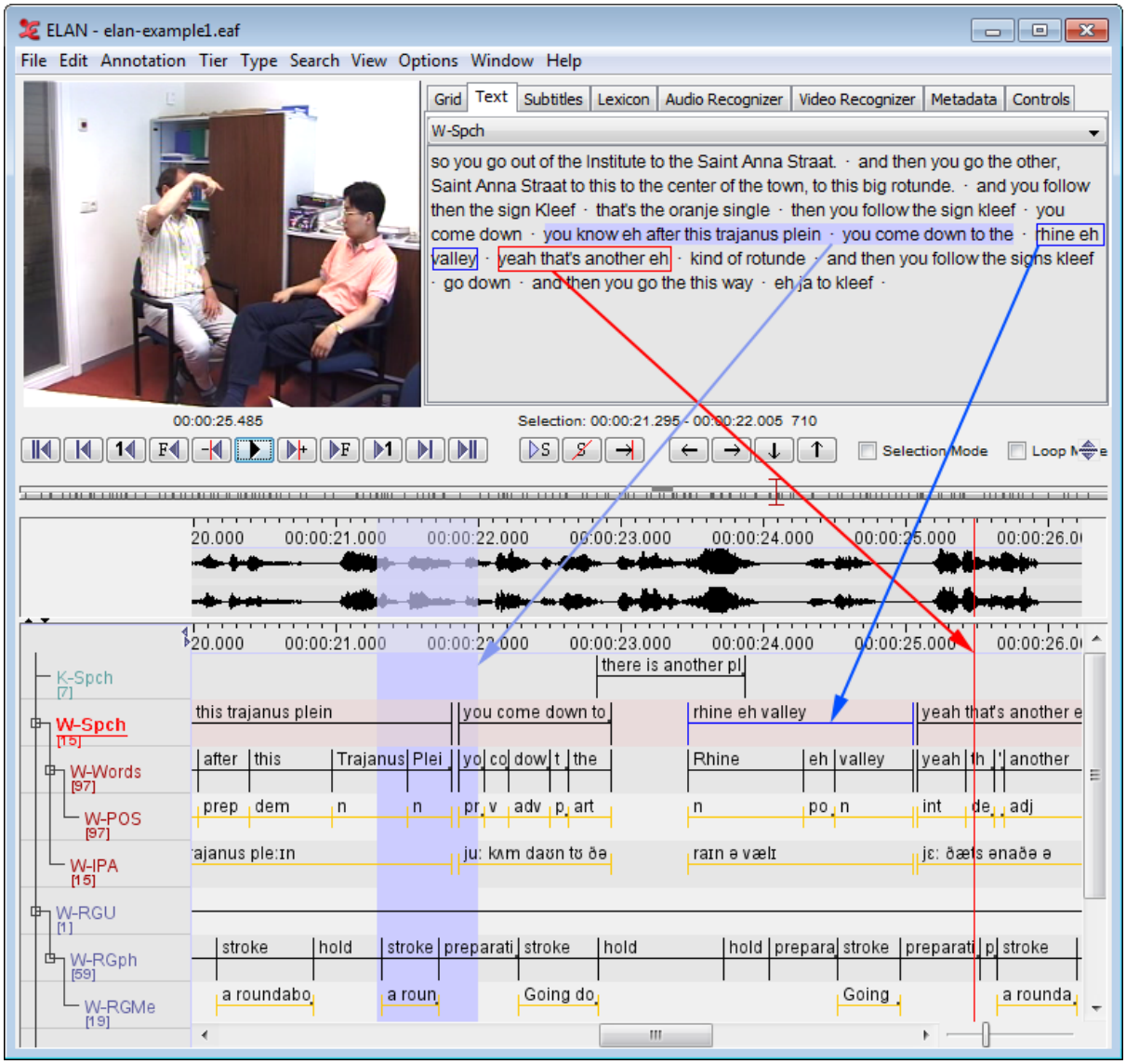


Figure 4.69. The text viewer

4.4.8. The Lexicon Viewer

See Figure 5.34

4.4.9. The Metadata Viewer

On the Metadata tab in the main window the IMDI metadata that belong to the media can be displayed. Click Select Metadata Source..., select an IMDI file and click Select.

The default metadata keys are now displayed (see also Section 4.4.28) in a table. If you want to change which keys are displayed, click Configure... and (de)select the keys.

The metadata can also be displayed in a tree structure instead of a table. Right click the table and select Tree View. Right click and select Table View to return to the table view.

4.4.10. Audio And Video Recognizer

The Audio Recognizer and Video Recognizer tabs provide 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. There is one audio recognizer included in the ELAN distribution, the Silence Recognizer (seeSection 5.4.3). More information on how to interact with the recognizers will be included in this manual at a later stage. The AVATecH project page (http://tla.mpi.nl/projects_info/avatech/) can be consulted for more information (including an interface specification).

4.4.11. 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.

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.

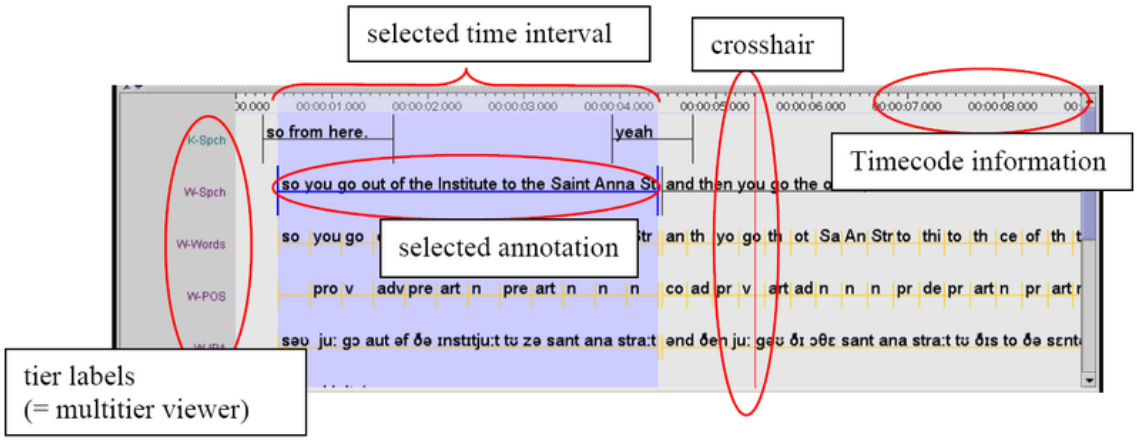


Figure 4.70. The timeline viewer

In the Timeline Viewer you can (a) select and modify time intervals (see Section 5.7) and (b) enter annotations (see Section 5.8).

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.

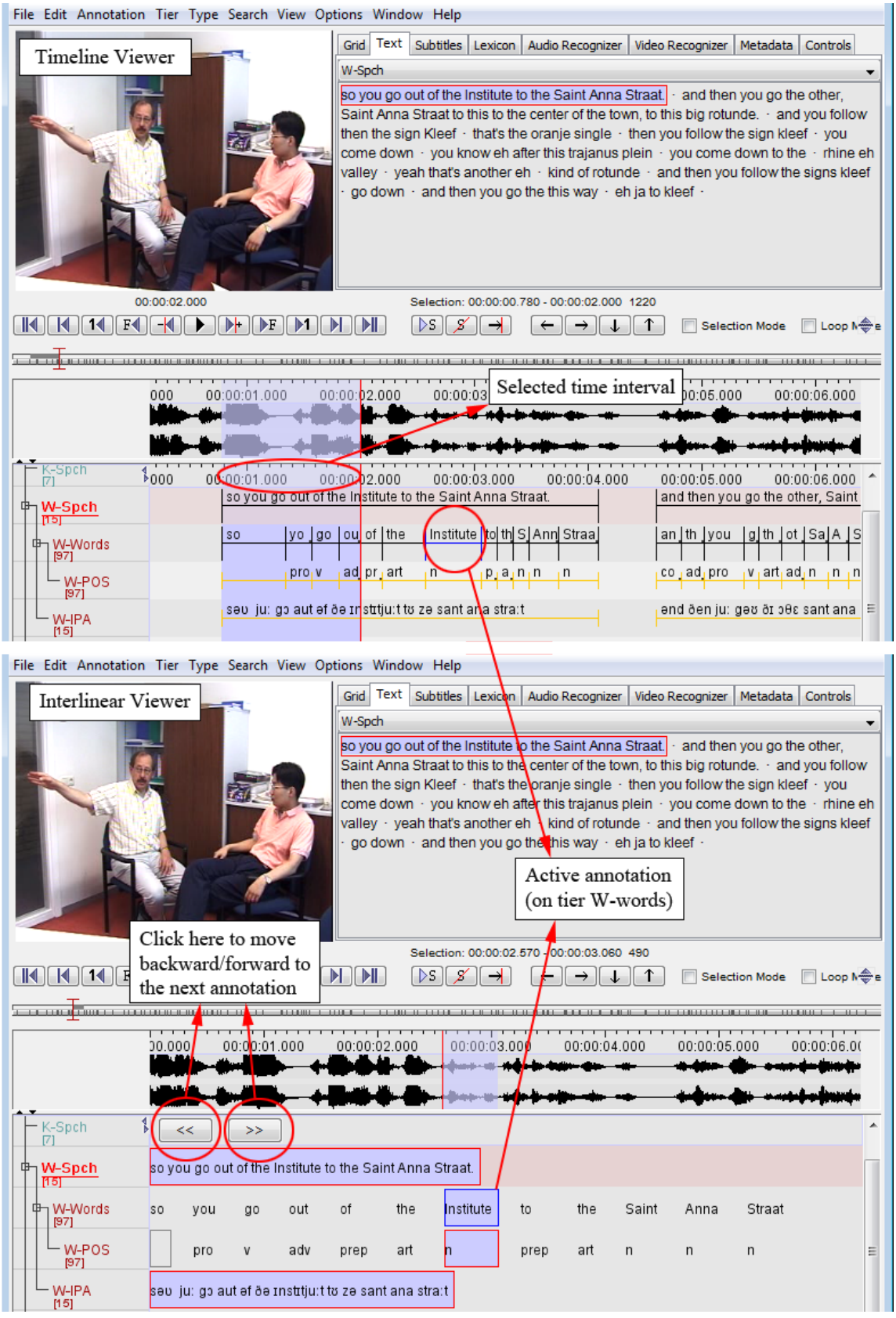


Figure 4.71. The ELAN window with timeline/interlinear viewer

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 Section 5.1). 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.

The Tier Name Panel

The Tier Name Panel gives an overview of the different existent tiers. Both tier names (see also Section 5.1 and Section 5.4) and tier types are defined by the user (see also Section 5.3). 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).

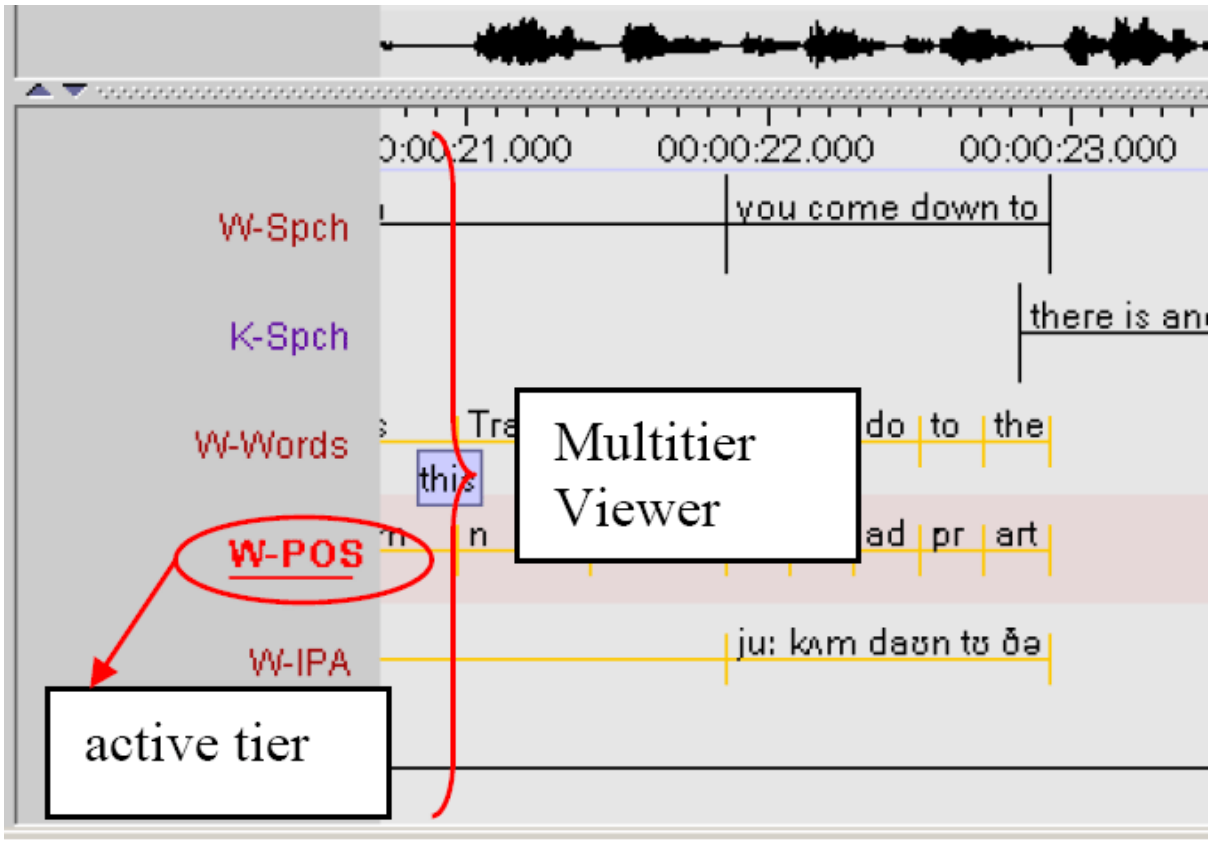


Figure 4.72. The tier name panel

To make a tier the active tier, choose one of the following actions:

- Double click on a tier label.
- Right click in the Timeline or Interlinear Viewer and choose Active Tier.
- Select the active tier with the keyboard shortcut CTRL + ARROW UP/DOWN

To select the tiers to display (and their order) see Section 4.4.20 and Section 4.4.23.

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.

4.4.12. 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 “trackpanels” and each trackpanel can display multiple “tracks”. Trackpanels and tracks can be added and removed via a popup menu. Each trackpanel 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[9] 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”.

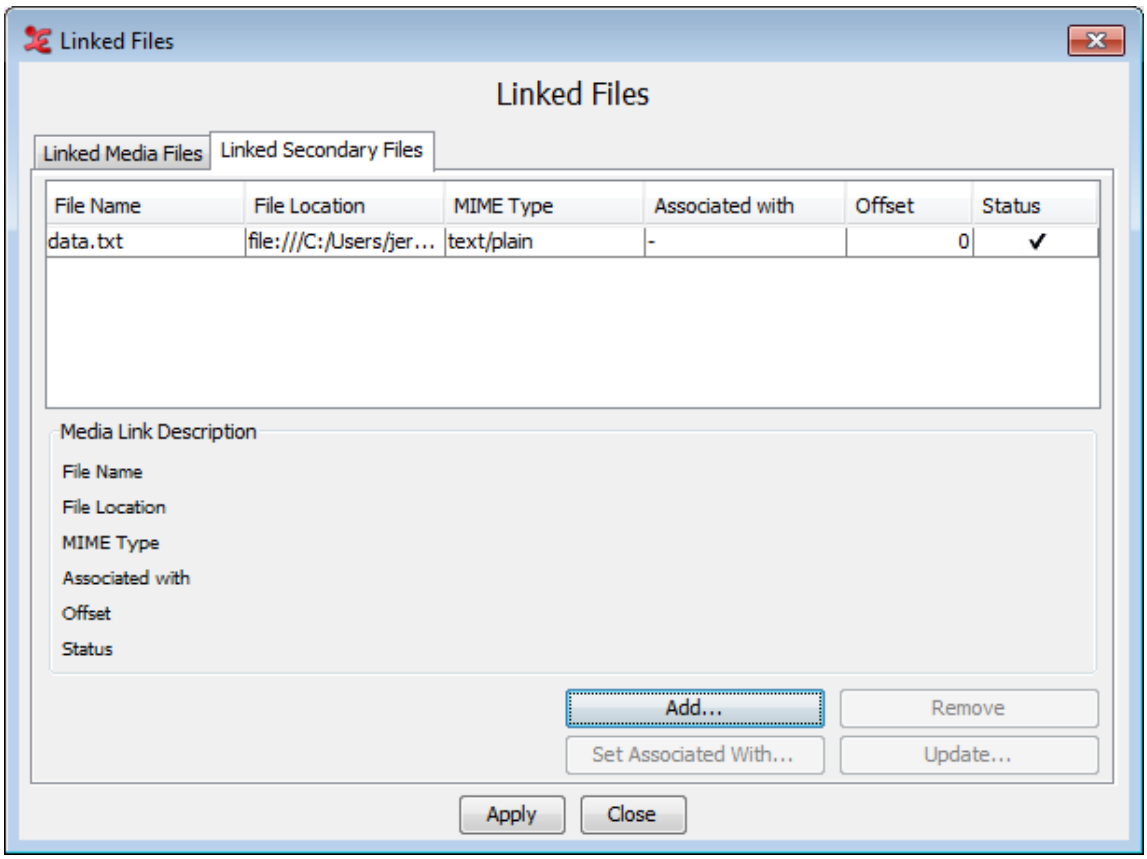


Figure 4.73. Linking timeseries data files

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....

[9] Currently supported file formats are a proprietary .log file produced by MPI CyberGlove software, a special kind of plain text (.txt) file, containing a time-value pair on each line, Praat .PitchTier and .IntensityTier files and CSV/Tab delimited text files. 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).

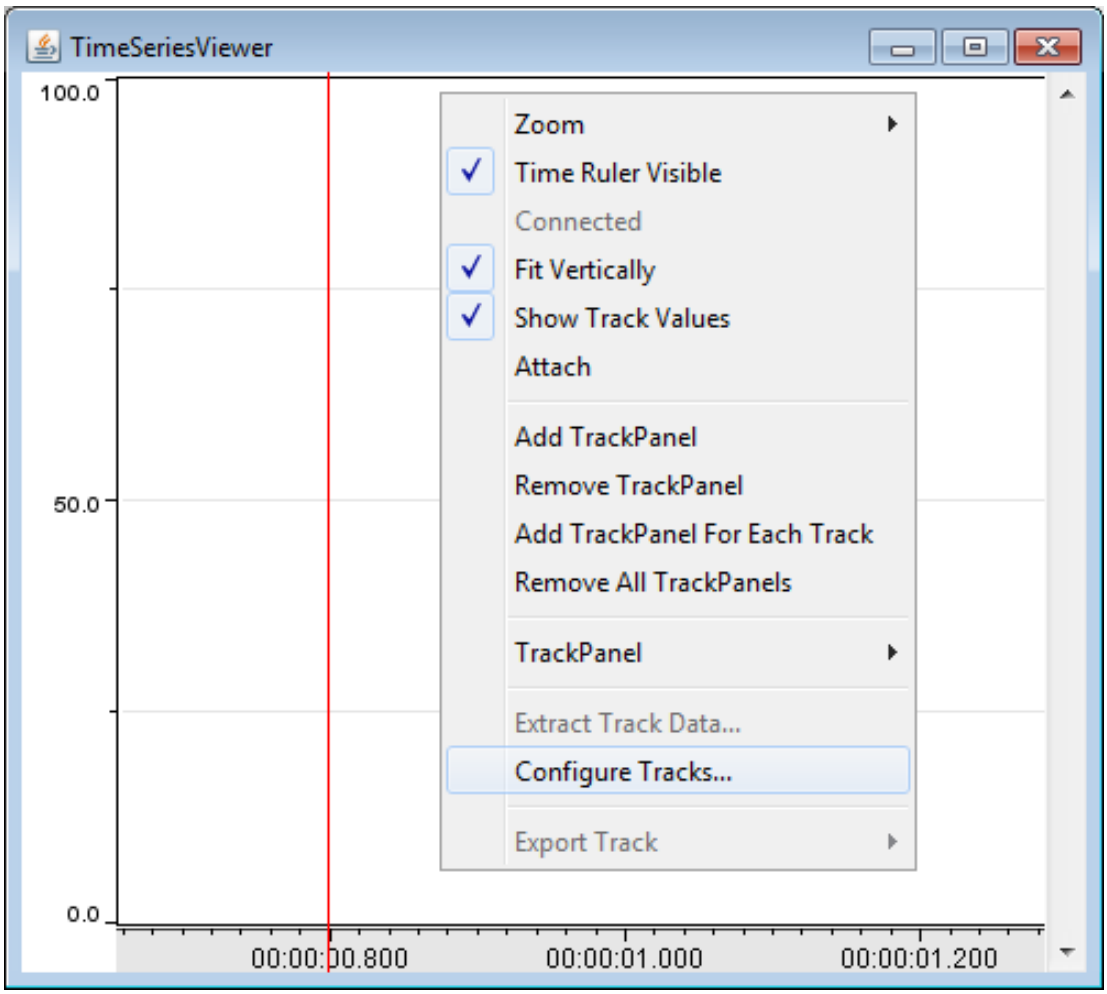


Figure 4.74. Timeseries Viewer popup menu

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.

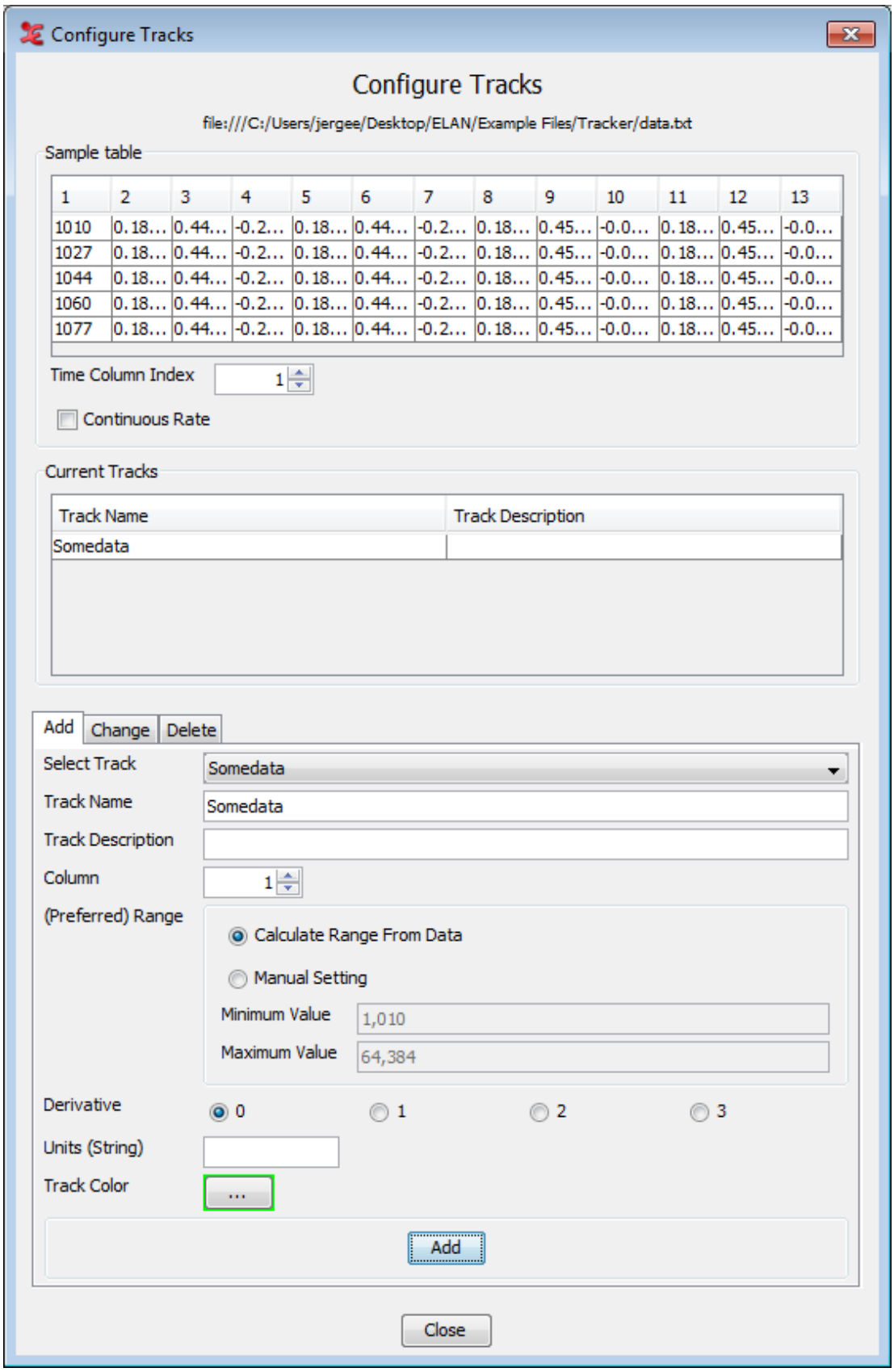


Figure 4.75. Timeseries Viewer: Configure tracks

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 trackpanel. Right click the new trackpanel and select TrackPanel > Add Track. A list of not yet displayed tracks is displayed. Click one to add it to the trackpanel.

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 trackpanel(s) vertically to the Timeseries Viewer window.
- Attach: attaches of detaches the Timeseries Viewer to the main window.

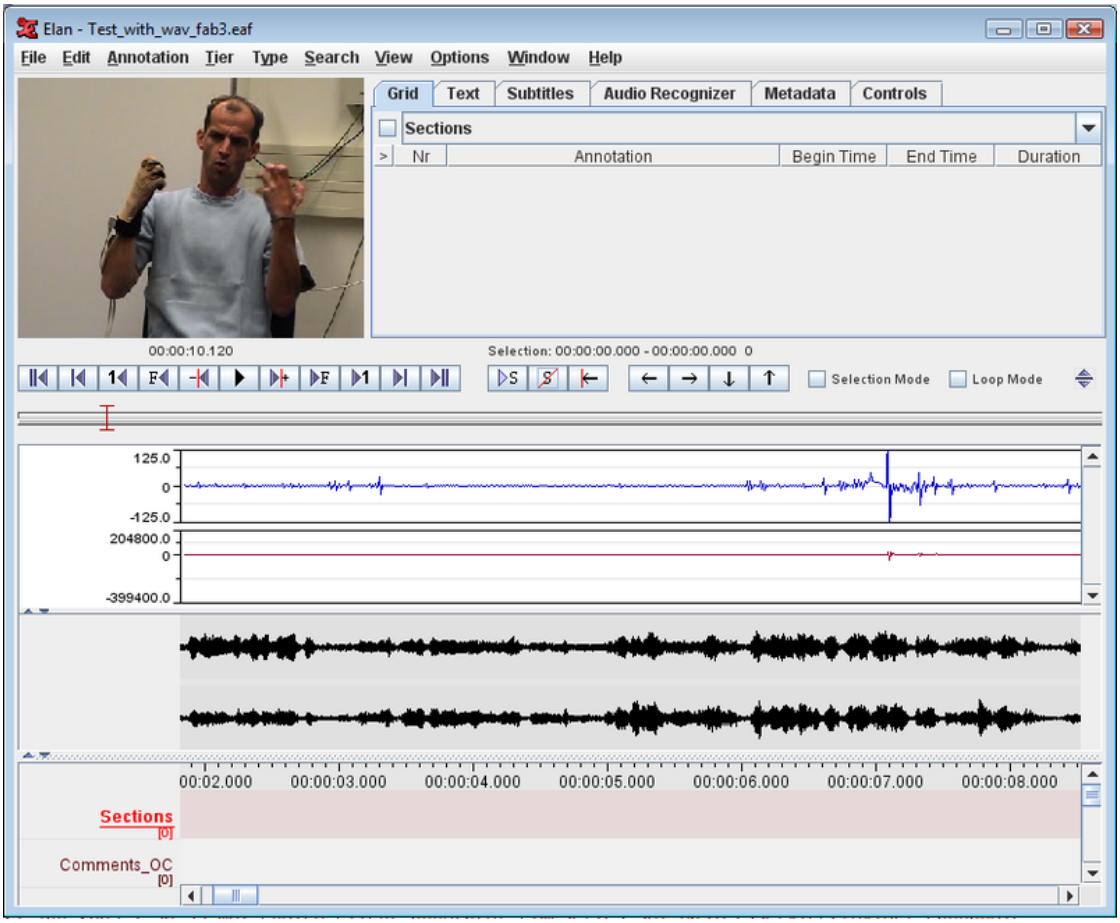


Figure 4.76. Timeseries viewer in the main window

- Add TrackPanel: create a trackpanel
- Remove TrackPanel: remove current trackpanel.
- Add TrackPanel For Each Track: create a trackpanel for each of the existing tracks.
- Remove All TrackPanels: remove all trackpanels 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 trackpanel.
- TrackPanel > Add All Tracks: add all tracks to the current trackpanel.
- TrackPanel > Remove All Tracks: remove all tracks from the current trackpanel.
- Extract Track Data: Extract data from a track and add it to a tier.

4.4.13. 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 Section 4.4) and to configure automatic backups.
- Edit: use this menu to define, modify and delete annotations, tiers and linguistic types (see Chapter 5).

- Search: use this menu to search for text (see Section 5.9).
- View: use this menu to get an overview of the tier dependencies (see Section 5.1) and the shortcut keys.
- Options: use this menu to (de)activate the Bulldozer mode (see Section 5.7.8), to choose between annotation mode and synchronization mode and to select a language and video standard.
- Help: use this menu to read information about ELAN.

4.4.14. 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:

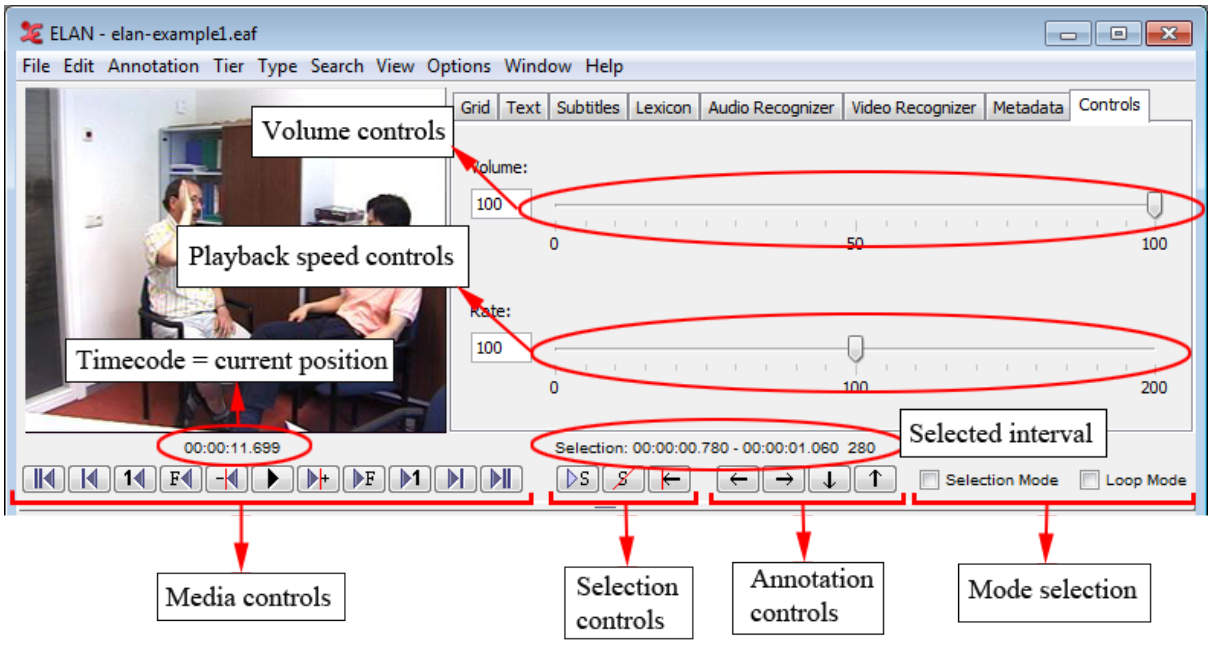


Figure 4.77. The media player options

Table 4.1. 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
	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,	CTRL+SHIFT+LEFT

Icon	Meaning	Shortcut
	depends on the zoom factor of the timeline viewer, default value 10 ms)	
	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 4.2. 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 4.3. 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 4.4. Selection Mode

Icon	Meaning	Shortcut
	While playing, select an interval automatically	CTRL+K

Icon	Meaning	Shortcut
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).

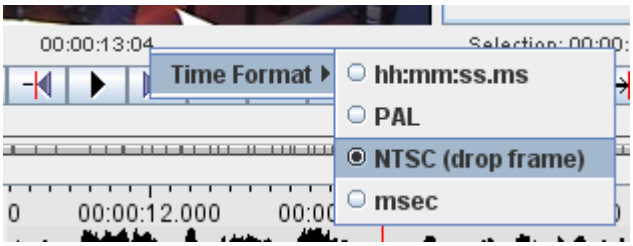


Figure 4.78. Timecode formats

- The sliders available when the Controls tab is selected allow you to control the playback rate and the volume.

4.4.15. 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 4.79. Color coding

How to change the display of the ELAN window

The ELAN window setup as described and illustrated in Section 4.4.1 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 (Section 4.4.16);
- switching Viewers on/off (Section 4.4.17);
- increasing/decreasing the size of Viewers (Section 4.4.19);
- switching tiers on/off (Section 4.4.20);
- rearranging the order of tiers (Section 4.4.21 and Section 4.4.23);
- displaying a tier in the any of the tab panes viewers (Section 4.4.5 and further);
- changing the time resolution (Section 4.4.24);
- changing the font size (Section 4.4.25);
- de-attach/re-attach the video window (Section 4.4.2).

4.4.16. 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.

4.4.17. 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 4.91) 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 Section 4.4.18.



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.

4.4.18. 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 4.91 .

4.4.19. 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.

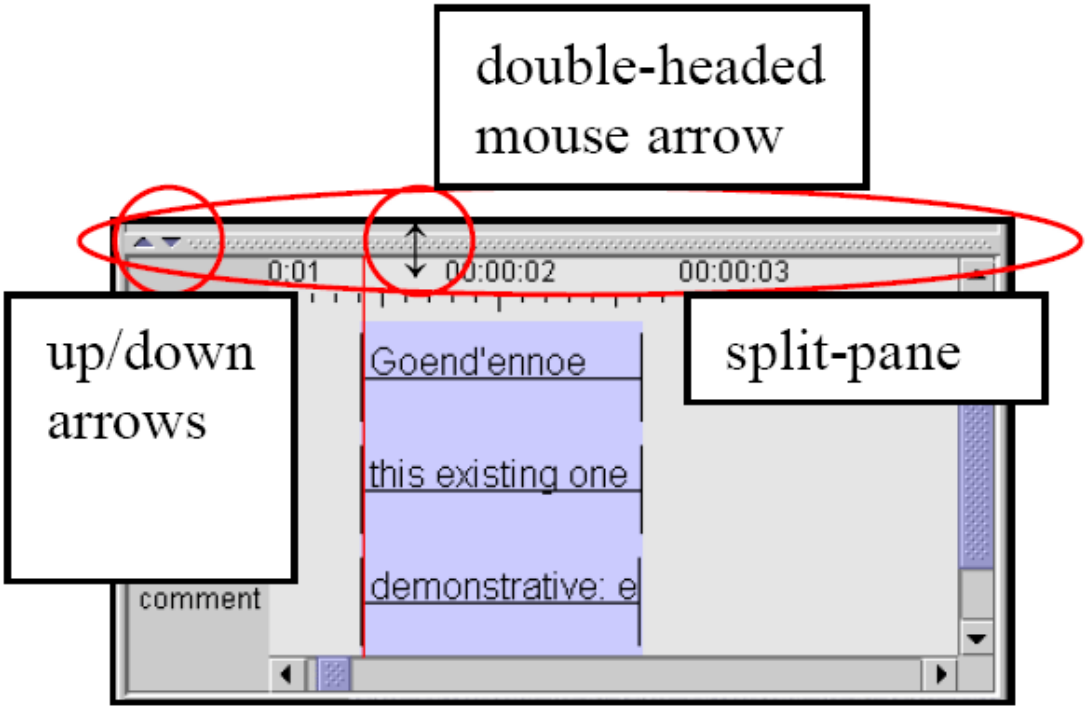


Figure 4.80. Changing size of the timeline viewer

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 4.81. Changing the size of the tier label panel

4.4.20. 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 4.84) in the popup menu.

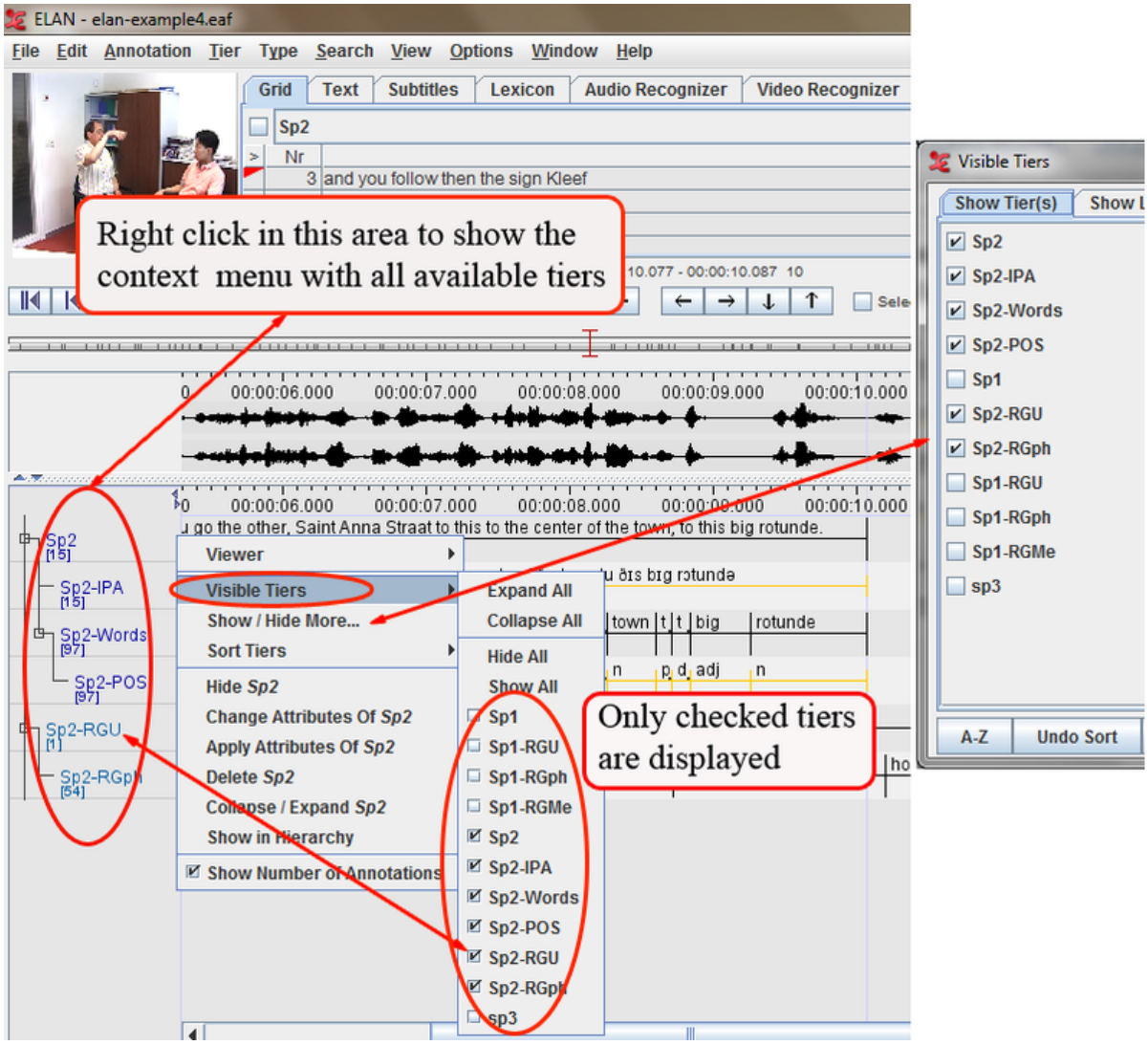


Figure 4.82. Switching tiers on/off

If you switch a tier on, it will be put on the place where you clicked.

- Show / Hide More... : (see Figure 4.84)
- Sort Tiers : (see Section 4.4.23)

4.4.21. 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.

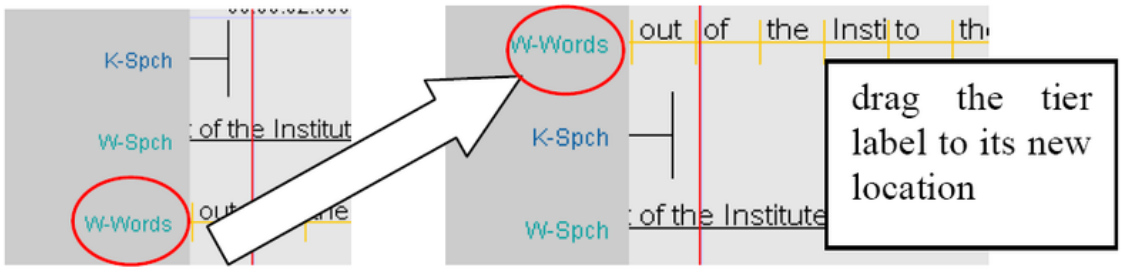


Figure 4.83. Rearranging the order of tiers

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.

4.4.22. View tiers by Type/Participant/Annotator

The tiers can also be viewed by its linguistic 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:

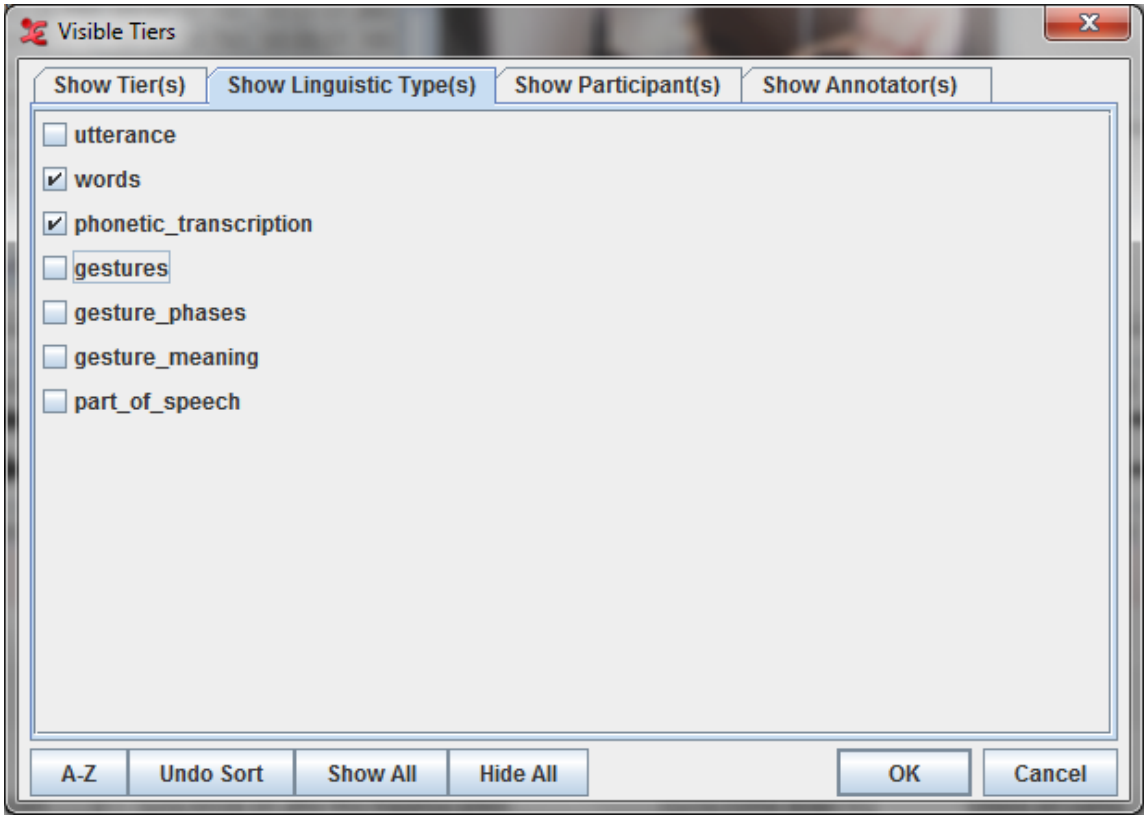


Figure 4.84. Visible Tiers

- Show Tier(s)

  Displays a list of all tiers in the transcription and the selected tiers are the visible tiers

- Show Linguistic Type(s)

  This shows a list of all the linguistic 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.

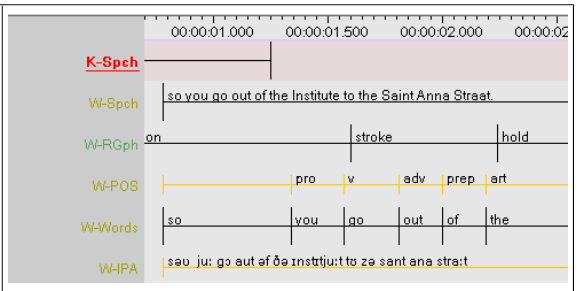
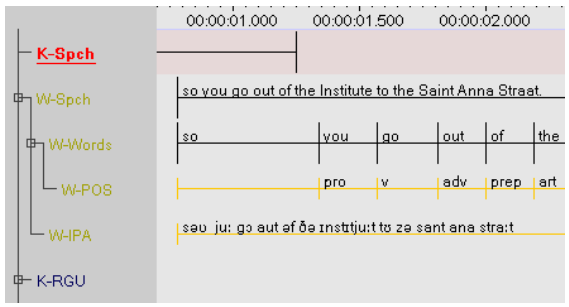
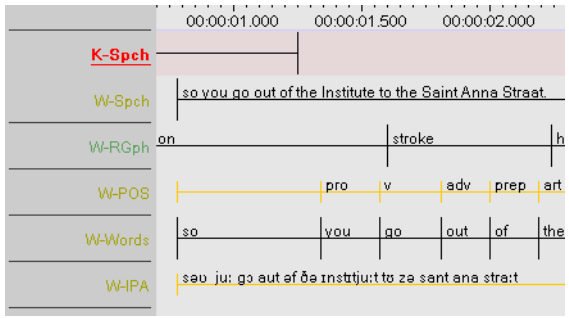
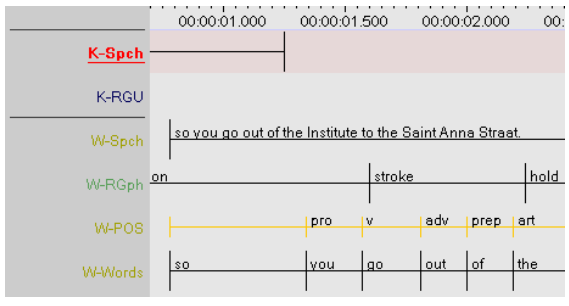
- Undo Sort

  This is to undo the sort and restore the order back.

4.4.23. 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 4.5. Sort tier options

Unsorted: no specific order.	
Sort by Hierarchy: display a tree with the hierarchical structure of the tiers	
Sort by Linguistic Type: group tiers by their linguistic 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.

4.4.24. 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. Click with the right mouse button on either the Waveform Viewer or the Timeline Viewer.

   b. Go to Zoom. The following menu appears:

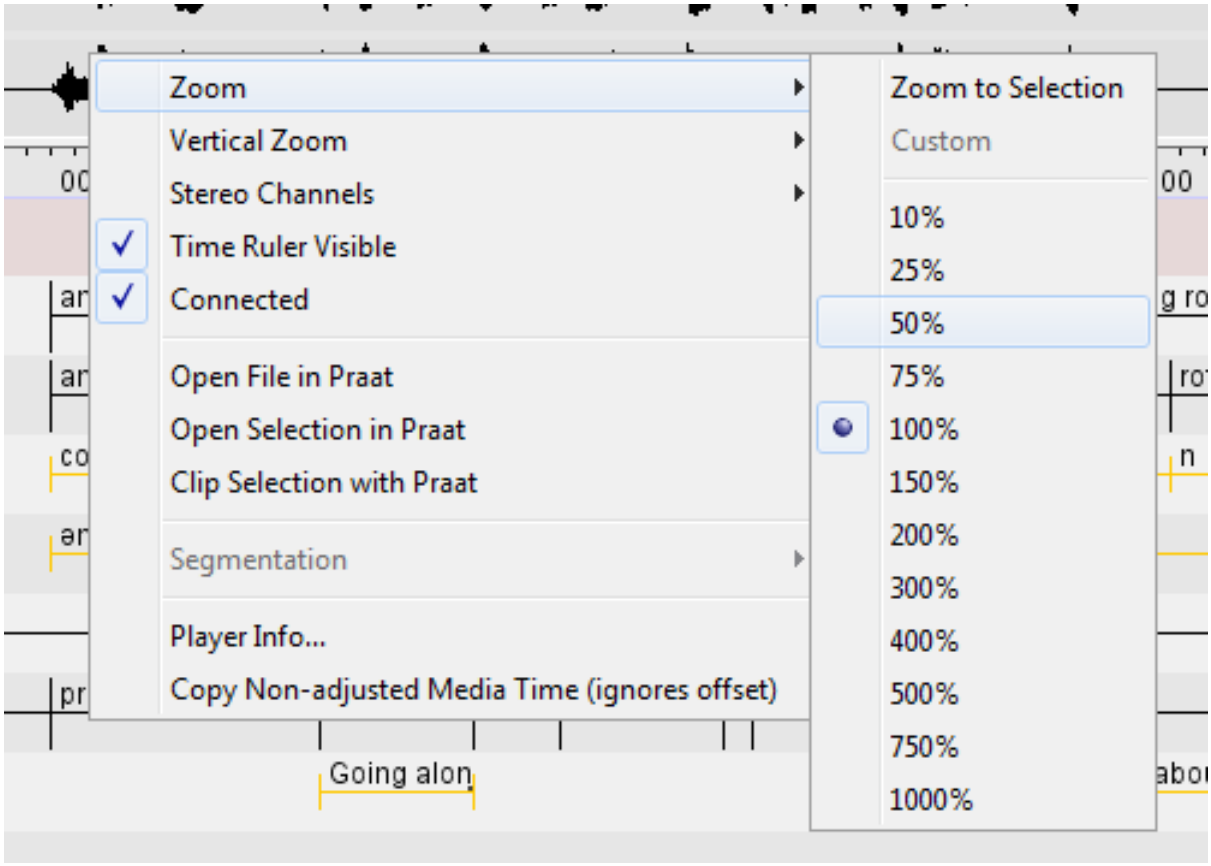


Figure 4.85. Changing time zoom

   c. Click on a zoom rate to select it. A checkmark 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.

There is another zoom option called Zoom to Selection (see Figure 4.85). To use it, first make a selection (see Section 9.1.6). 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.

4.4.25. 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:

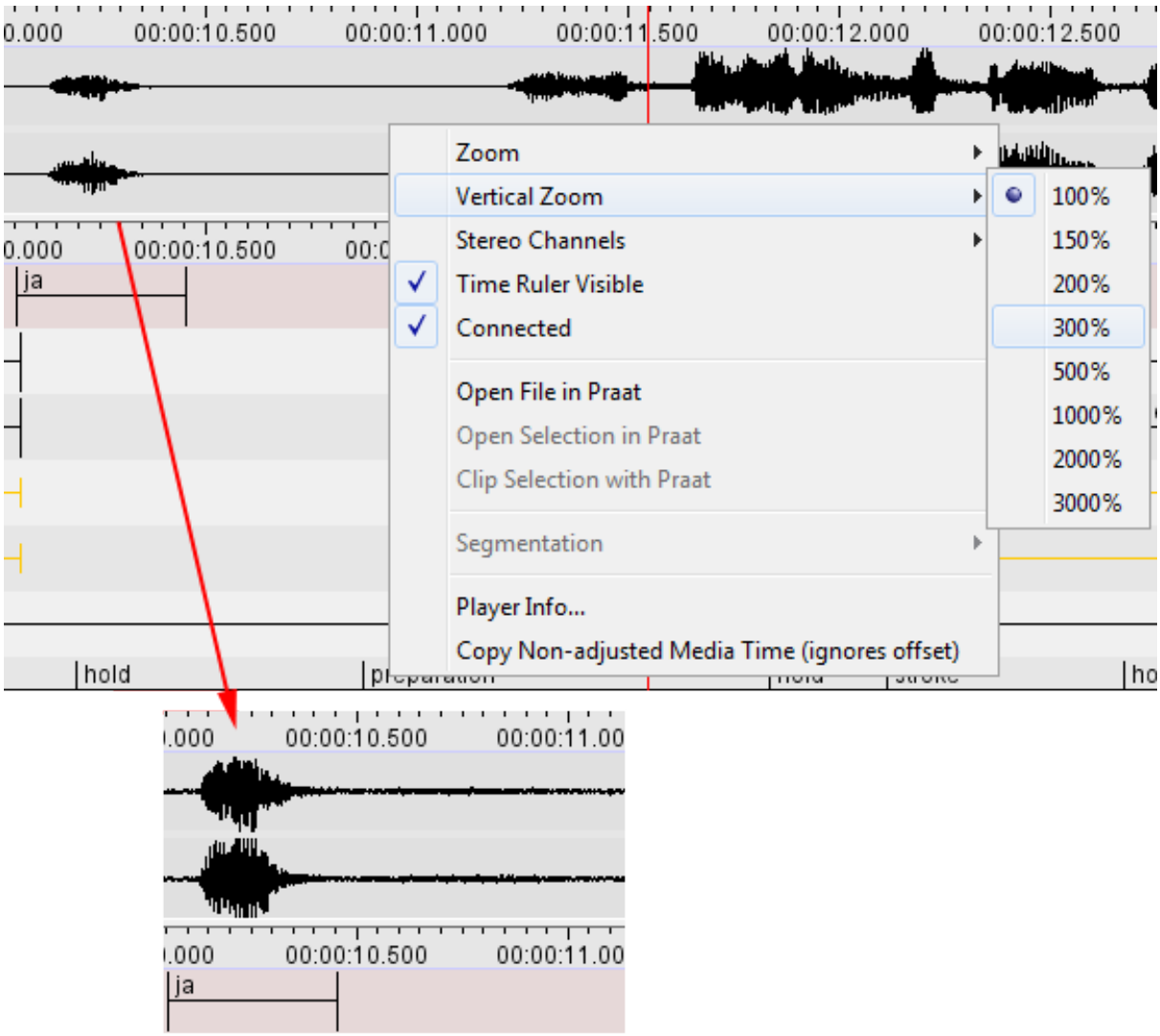


Figure 4.86. Changing vertical (intensity) zoom

Please note that this vertical zoom does not change the audio characteristics in any way.

4.4.26. 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:

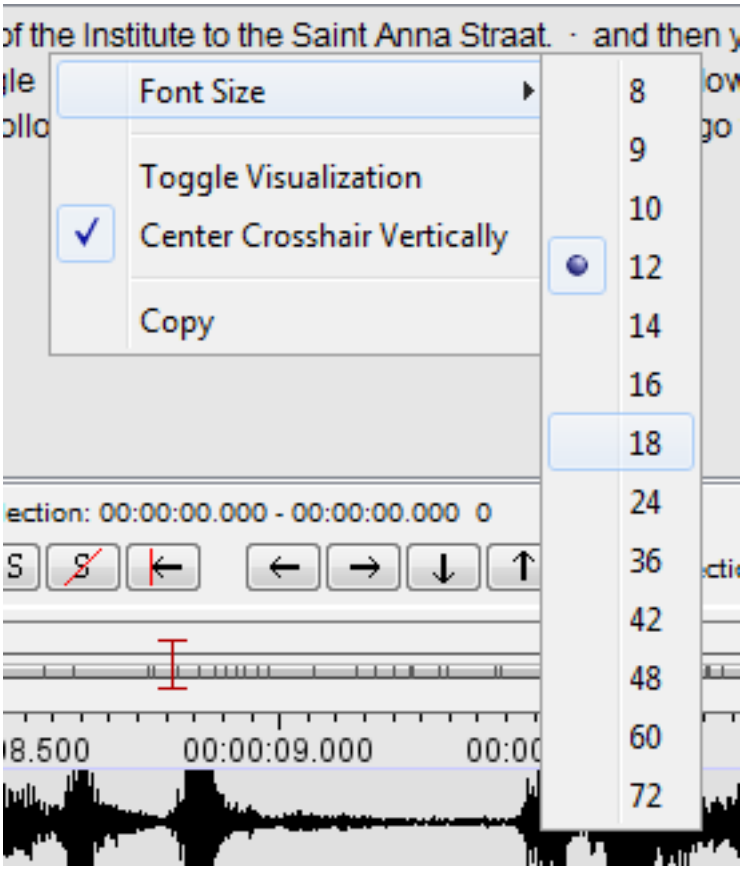


Figure 4.87. Changing the font size

3. Click on a font size to select it. A checkmark appears next to the selected font size.

4.4.27. 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 4.88). 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.

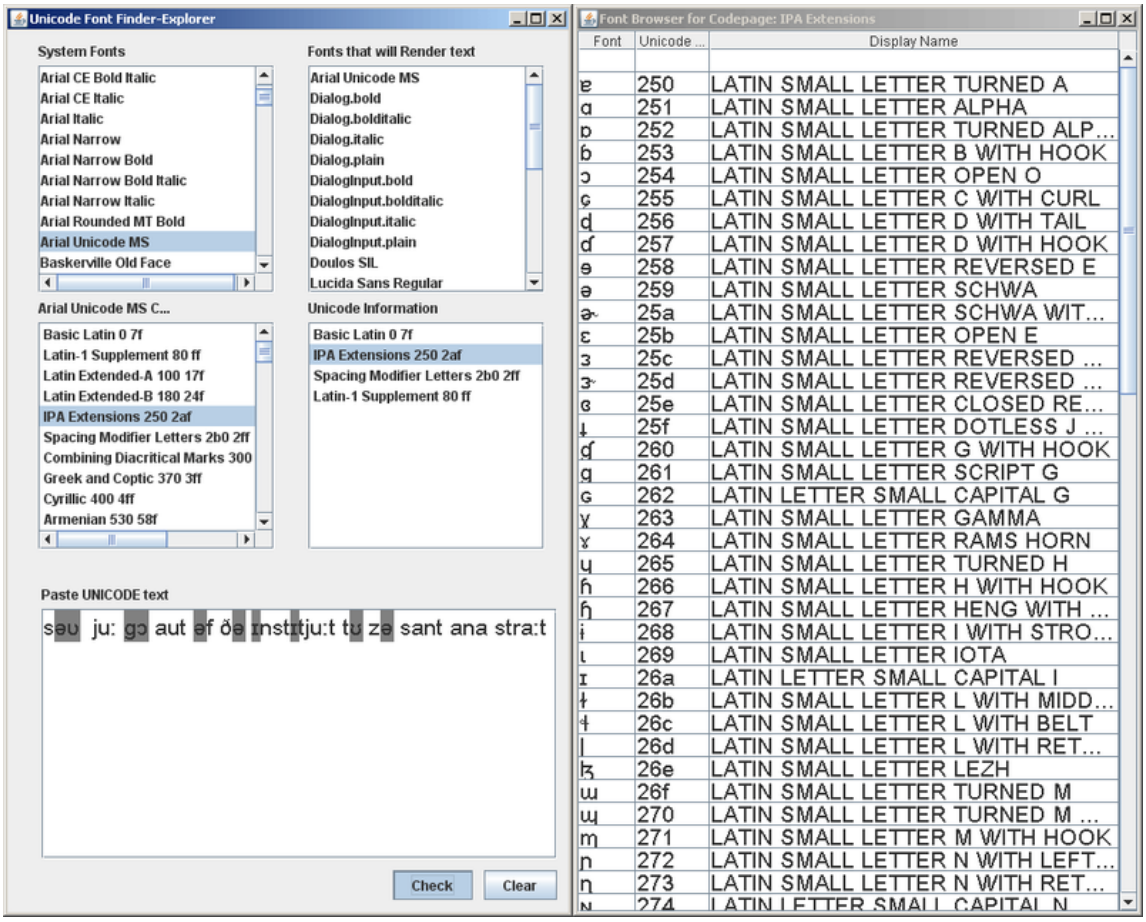


Figure 4.88. Font Finder-Explorer

4.4.28. Editing preferences

Some 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:

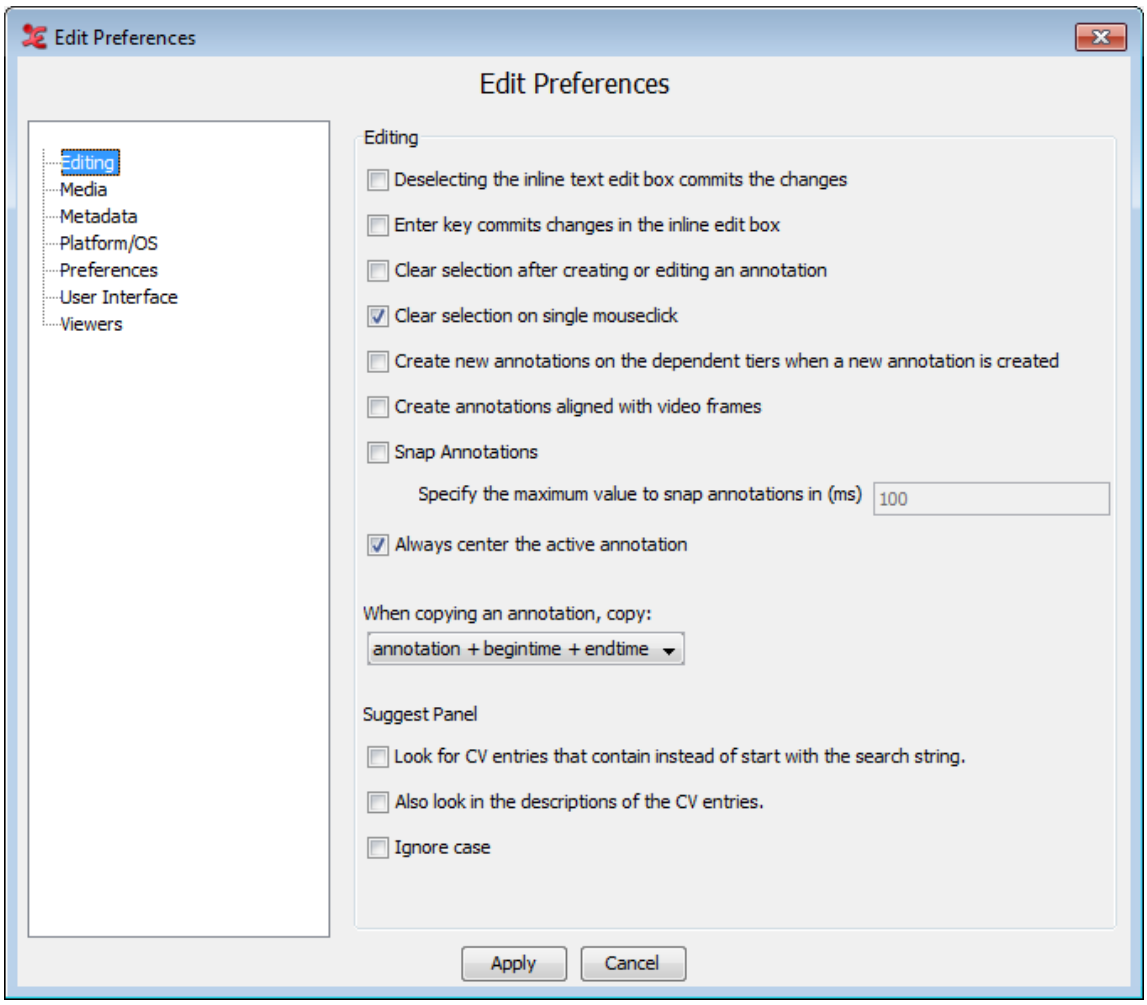


Figure 4.89. Edit preferences

- 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 mouseclick

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.

- 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, casing will be ignored when entering a search string in the CV suggest panel.

- Media

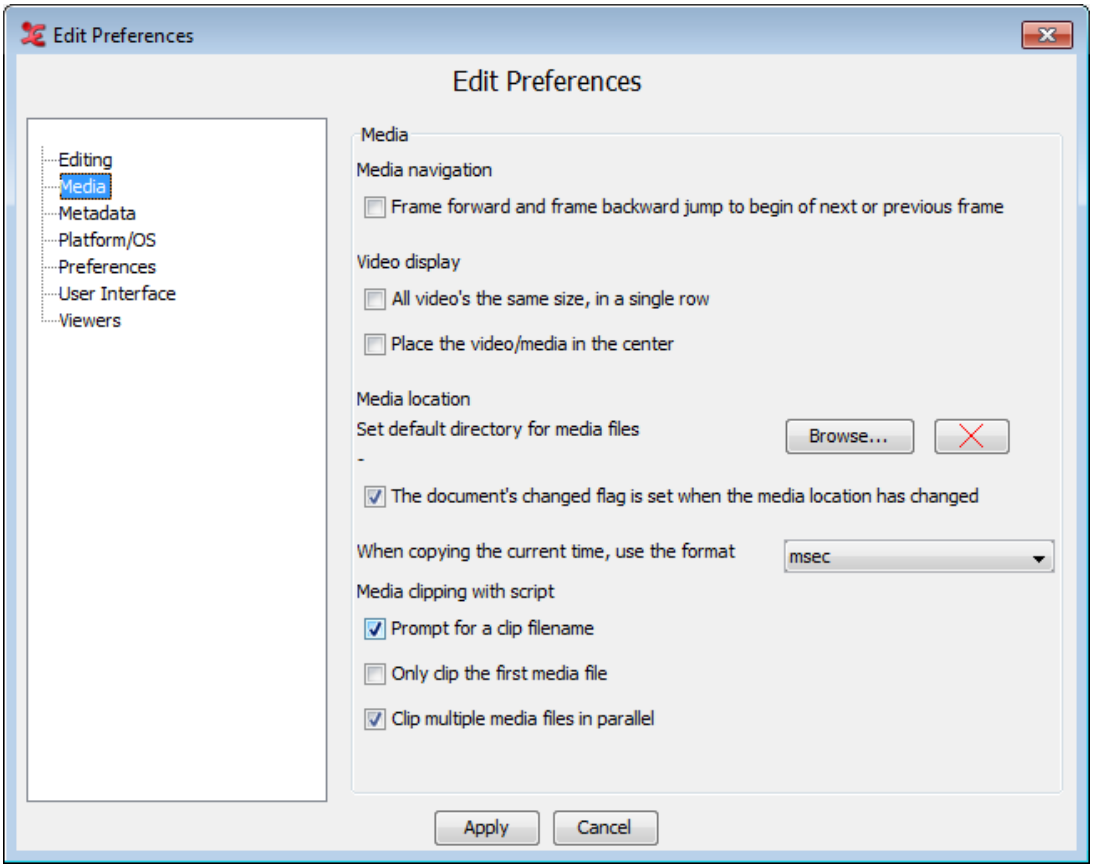


Figure 4.90. Edit preferences

– 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 Section 4.4.14) 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.

– 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.

– 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.
- When copying the current time, use the format

  Click on the list box to select the format of the time, when the current time of the media is copied.
- Prompt for a clip filename

  When this is checked (default), ELAN will prompt for a filename 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 Section 4.3.2.11
- Only clip the first media file

  This option is unchecked by default. If errors occur clipping multiple videofiles and a wavefile at once, check this and only the first (or master) mediafile 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.

- Metadata
  - Check the metadata keys you want to display in the Metadata tab in the main window by default. (See also Section 4.4.9.)
- Platform/OS

  Mac OS X:
  - Use screen menu bar: if checked ELAN will use the screen menu bar in Mac OS
  - Use Mac Look and Feel: if checked ELAN will use the Mac OS Look and Feel. Otherwise a platform independent (i.e. Java) look and feel is used. Note that if the option Use screen menu bar is checked, ELAN will use the Mac Look and Feel, even if you have Use Mac Look and Feelunchecked.
  - 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 Cocoa QT or Quicktime for Java
  - Use detached media window: if checked ELAN will show the media in a separate window by default.

  Windows:
  - Media Framework: Select either

    —Java -DirectShow Framework

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

    —Native Windows Media Player (WMP)

    —QuickTime for Java

    —Java Media Framework
- Preferences

- Click on Browse... to browse to and select the directory where ELAN should look for preference files by default.

- Selecting this optionAutomatic check for updates will enable the automatic checking for new updates feature, which will constantly check for updates once a month and intimates you if any updates were available.

- User Interface

  - Number of recent items: select the number of recent edited item ELAN should remember.

  - 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: Select Show annotation count to show the number of annotations per tier in the menus and tier list in the viewers.

- Viewers

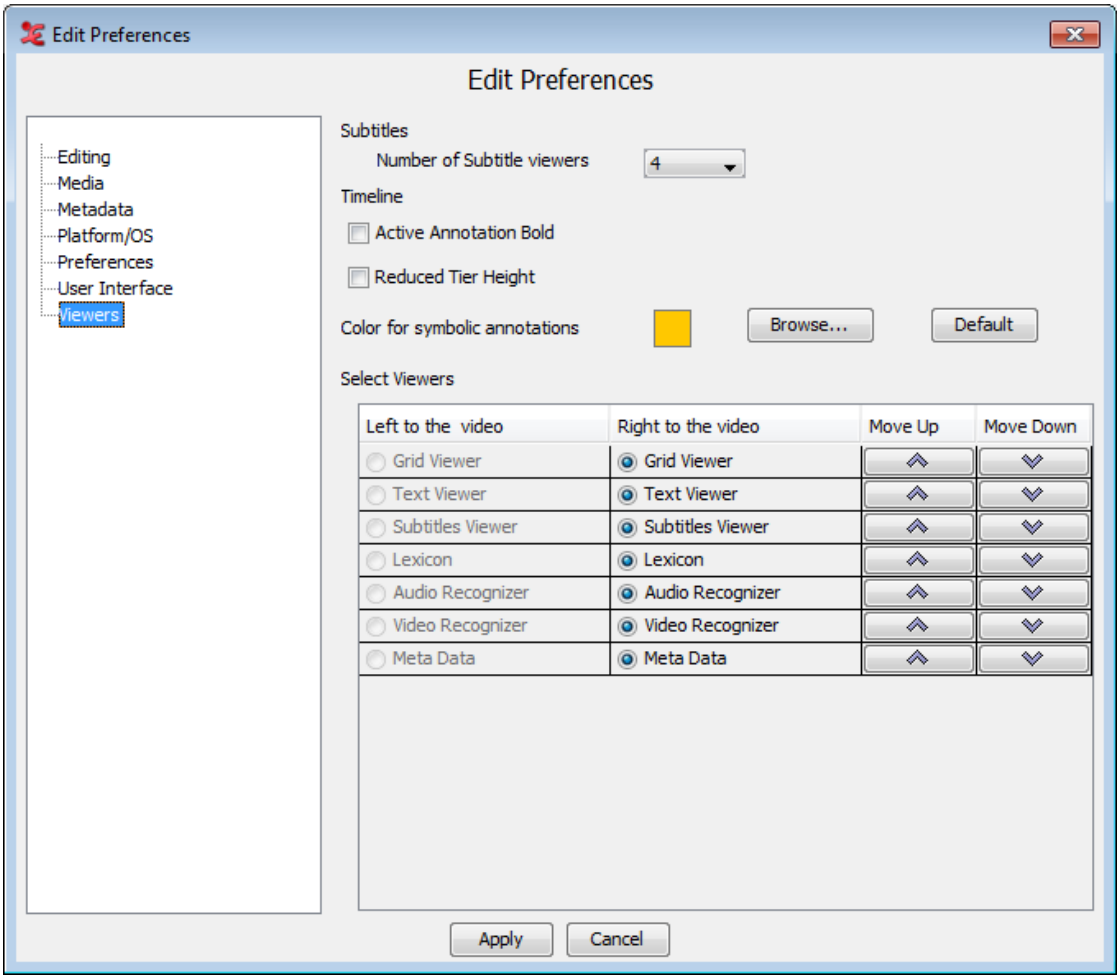


Figure 4.91. Edit preferences

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

  - 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.

– 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.

– 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.

4.4.29. 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 filename and click on Save. To import preferences click Edit > Import Preferences..., look up the preference file and click on Select.

4.4.30. Editing shortcuts

ELAN has shortcuts for many of it's functions. The default shortcuts, which are mentioned throughout this manual (for an overview see Section 9.2), can be changed via Edit > Preferences > Edit Shortcuts.... This dialog window appears.

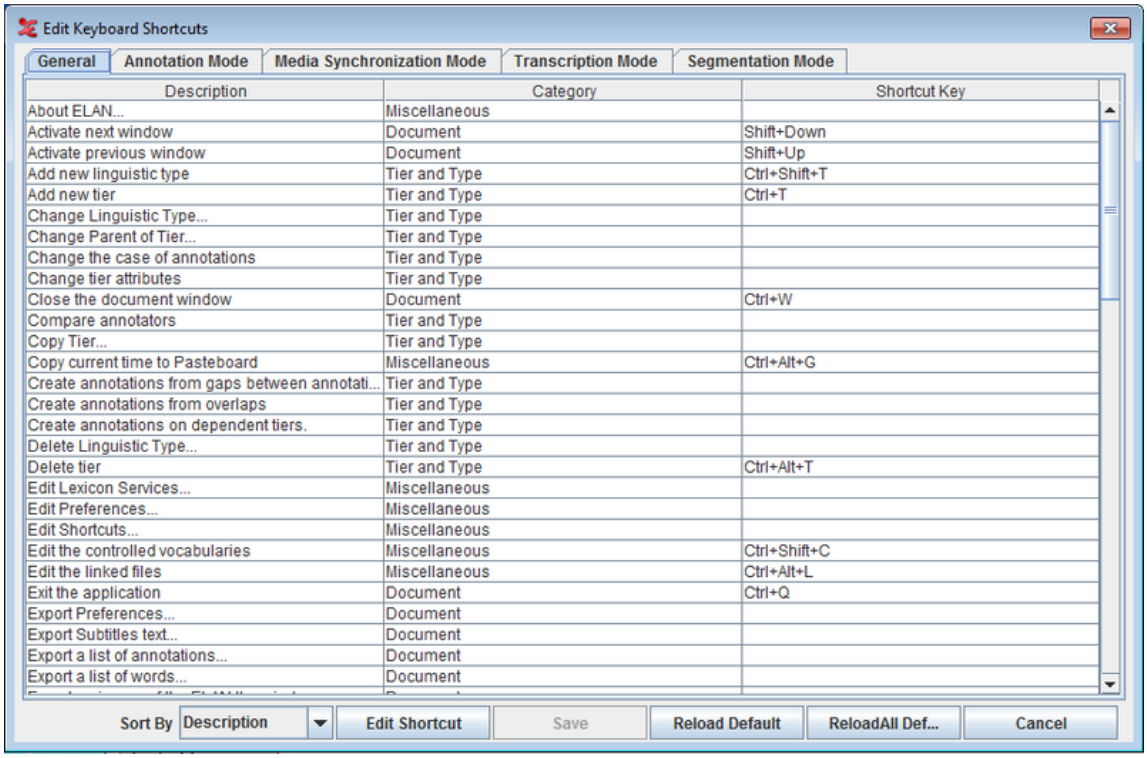


Figure 4.92. Edit KeyBoard Shortcuts

- 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

4.4.31. 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

4.5. How to navigate through a document

ELAN supports the following options for navigating through a document:

- accessing points in time (Section 4.5.1);
- going to a selection (Section 4.5.2);
- jumping step by step through a document (Section 4.5.3);
- navigating using the Timeline Viewer (Section 4.4.14)
- navigating using the Grid Viewer (Section 4.5.4);
- navigating using the Interlinear Viewer (Section 4.5.5).

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).

4.5.1. 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.

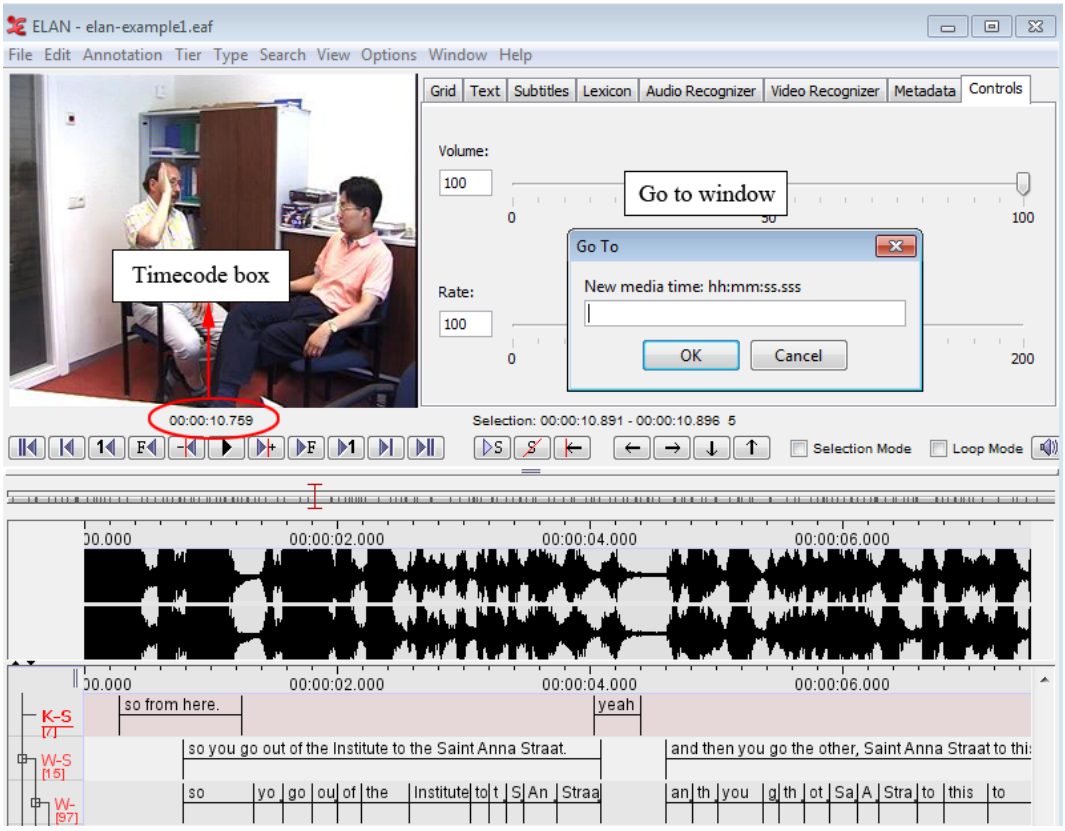


Figure 4.93. Click the timecode box

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

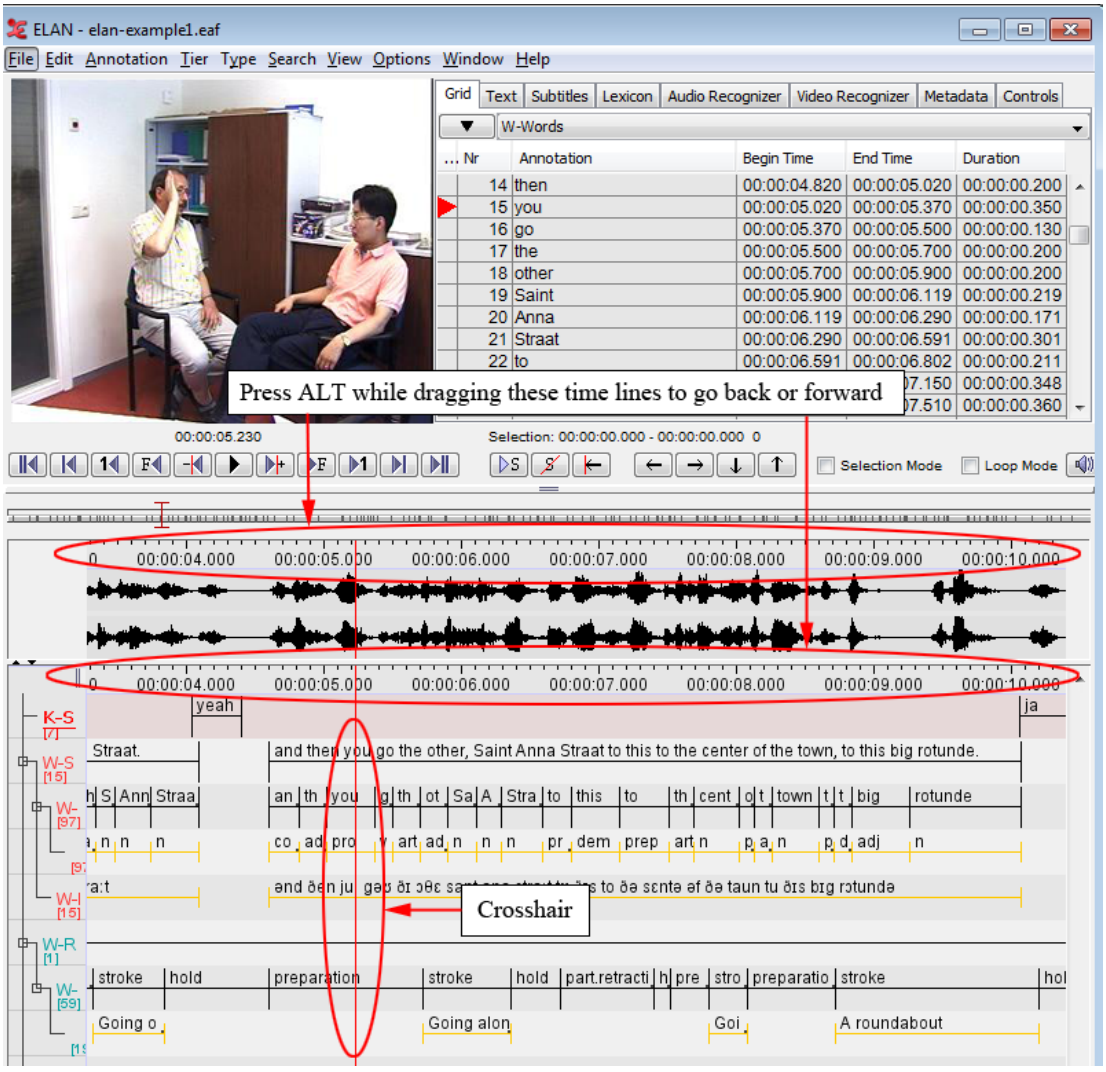


Figure 4.94. Use the crosshair

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 Section 4.4.14

4.5.2. Going to the borders of a selection

If you have made a selection (see Section 5.7), 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.

4.5.3. 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 Section 4.4.14.

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:

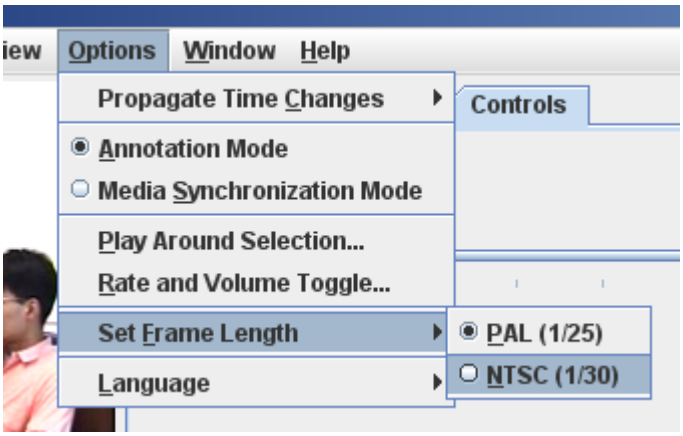


Figure 4.95. Set Frame Length menu

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.

4.5.4. 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:

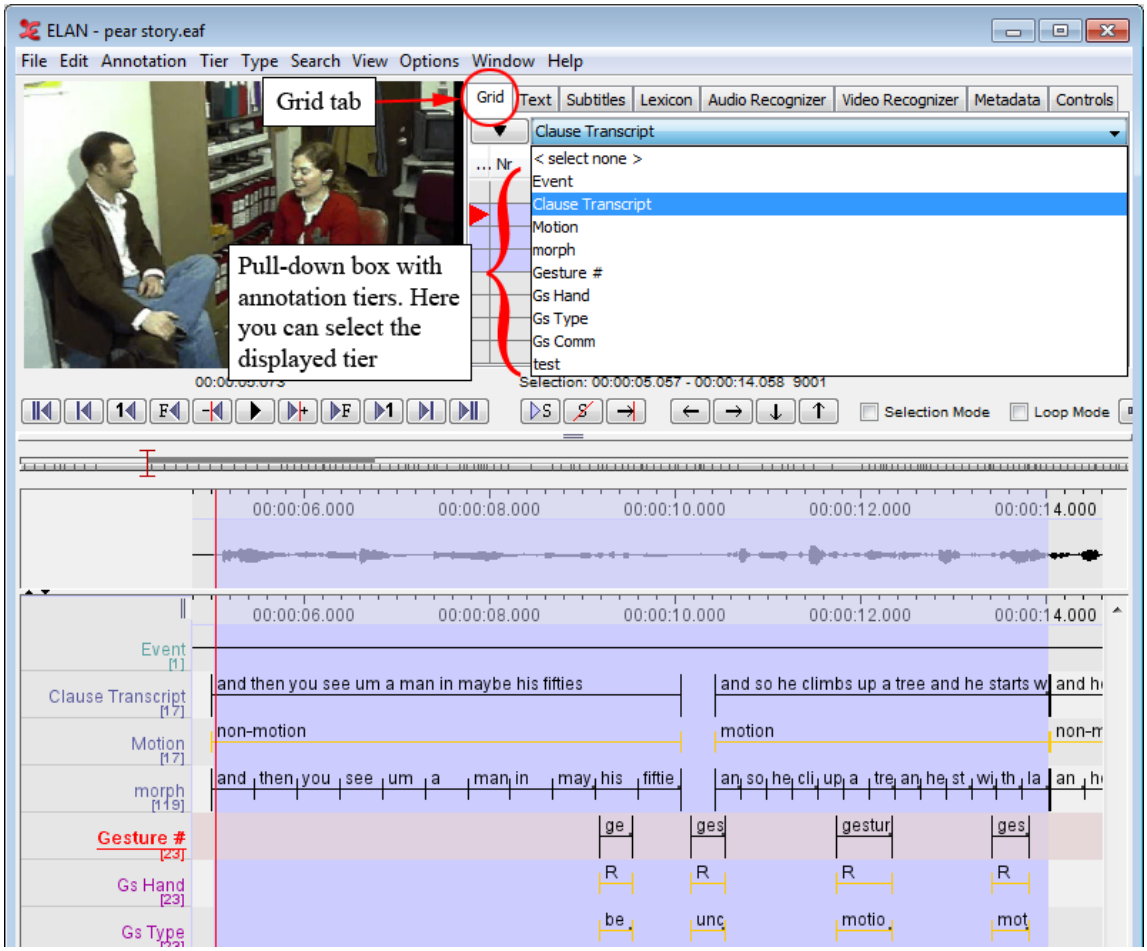


Figure 4.96. The grid viewer

The structure of the Grid View window is as follows:

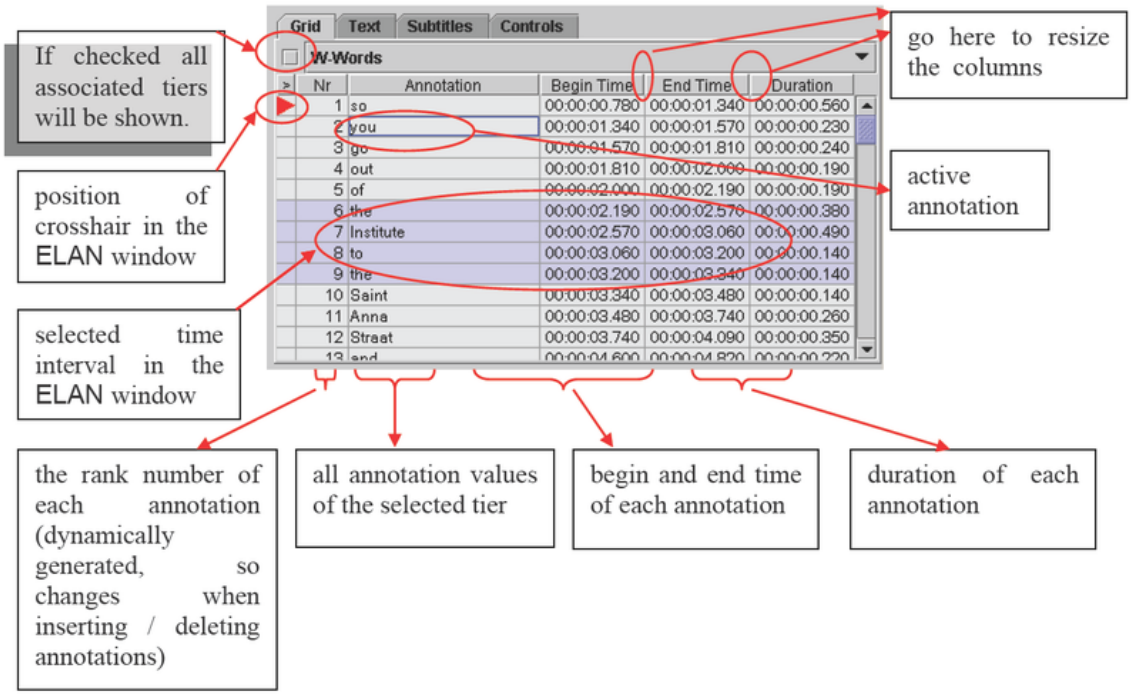


Figure 4.97. Structure of the grid viewer

You can change the Grid Viewers interface at any moment by right clicking in the Grid Viewer. A context menu will appear:

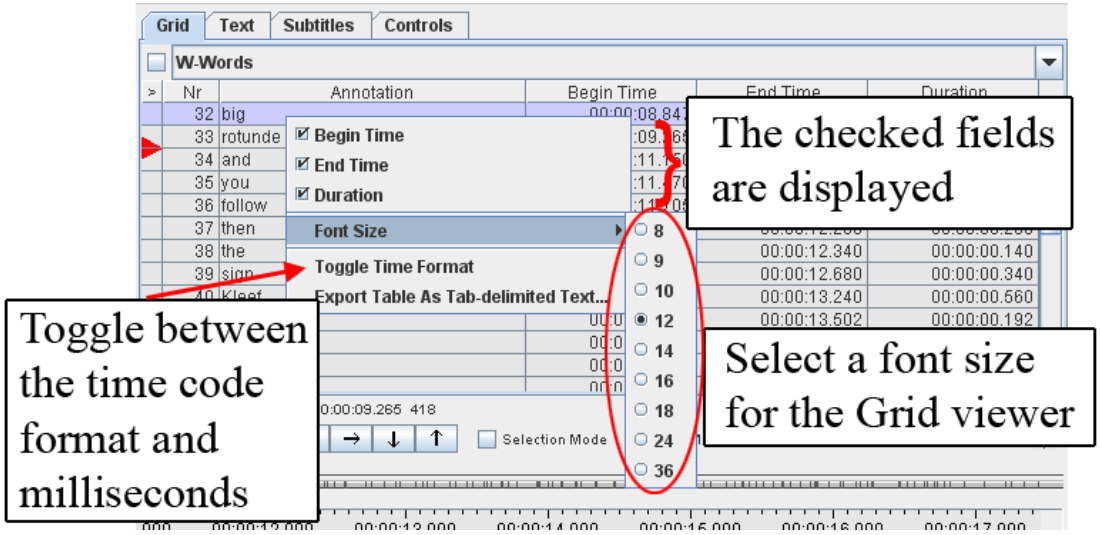


Figure 4.98. Grid viewer with context menu

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.

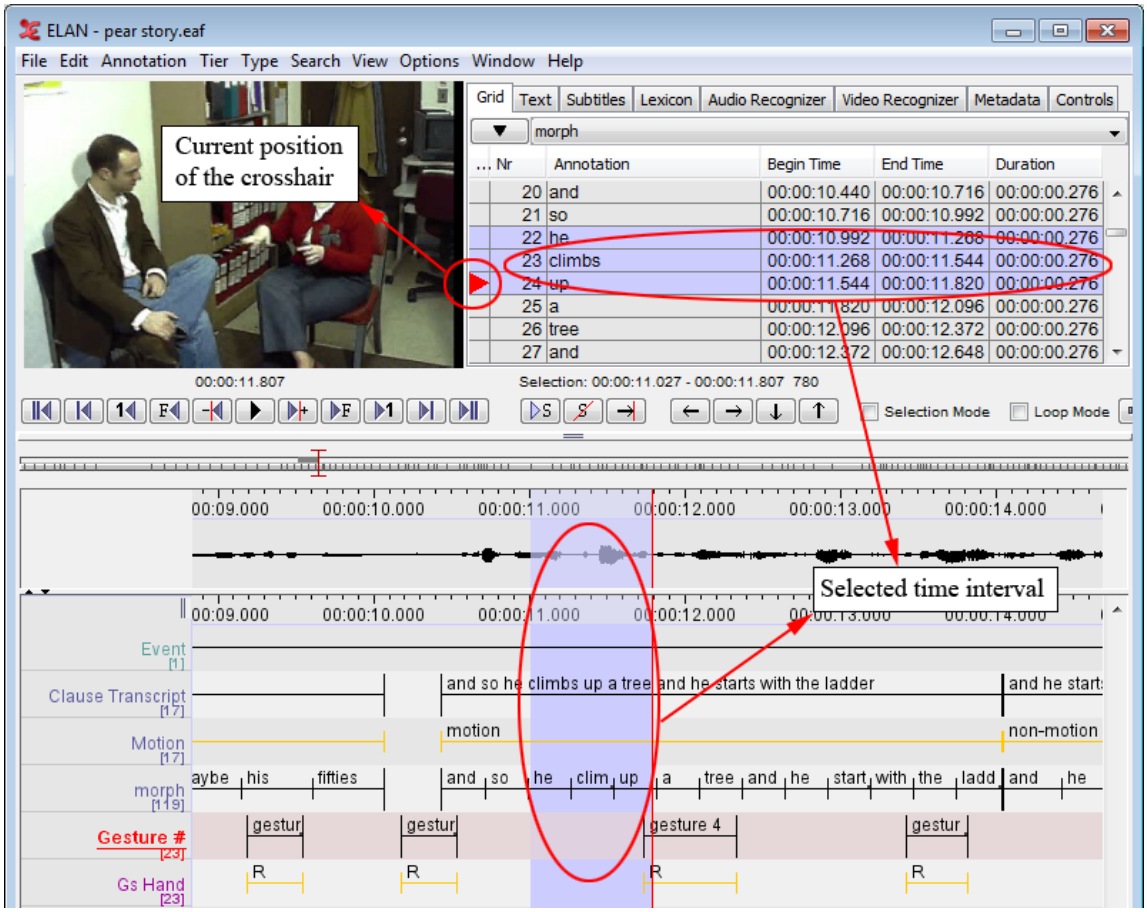


Figure 4.99. Grid viewer with selected annotations

4.5.5. 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.

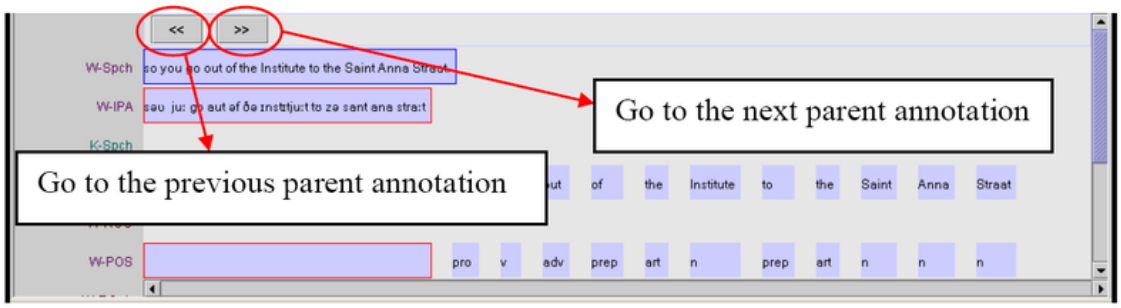


Figure 4.100. Interlinear viewer

4.6. How to play a document

ELAN supports the following options for playing a document:

- playing a document (Section 4.6.1);
- playing a selection (Section 4.6.2);
- playing around the crosshair/selection (Section 4.6.3);
- playing in slow motion (Section 4.6.4).

4.6.1. 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.

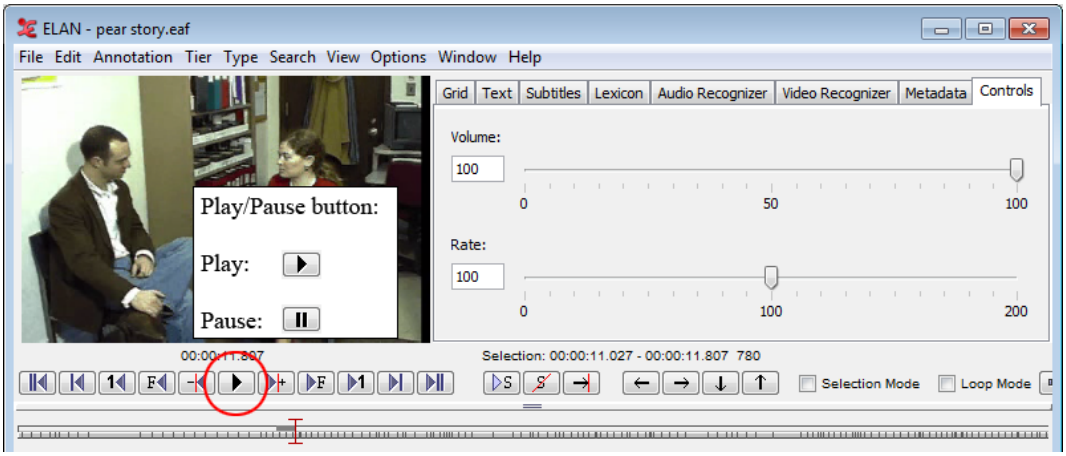


Figure 4.101. Play/pause button

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.

4.6.2. Playing a selection

It is possible to play only a selected part of the document. Do the following:

1. Make a selection (see Section 5.7), or click on an annotation.

2. To play the selection, do one of the following:

   a. Click on the Play Selection  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 Section 4.5.3), 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).

4.6.3. 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:

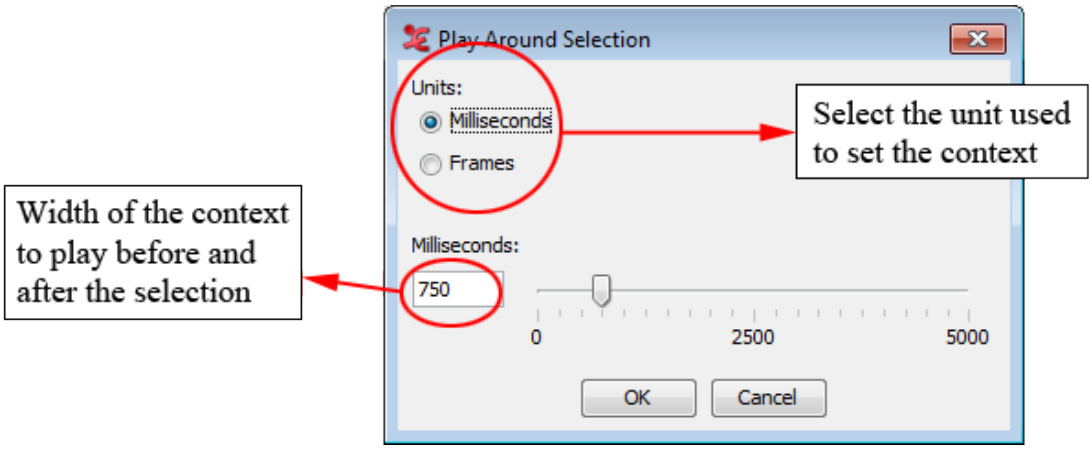


Figure 4.102. Play around selection

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.

4.6.4. 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%.

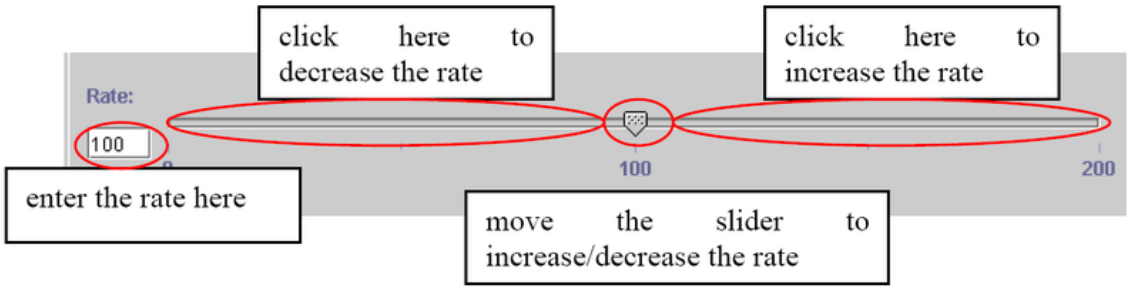


Figure 4.103. Playback rate

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.

4.6.5. 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

4.7. Multiple File Operations

4.7.1. 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:

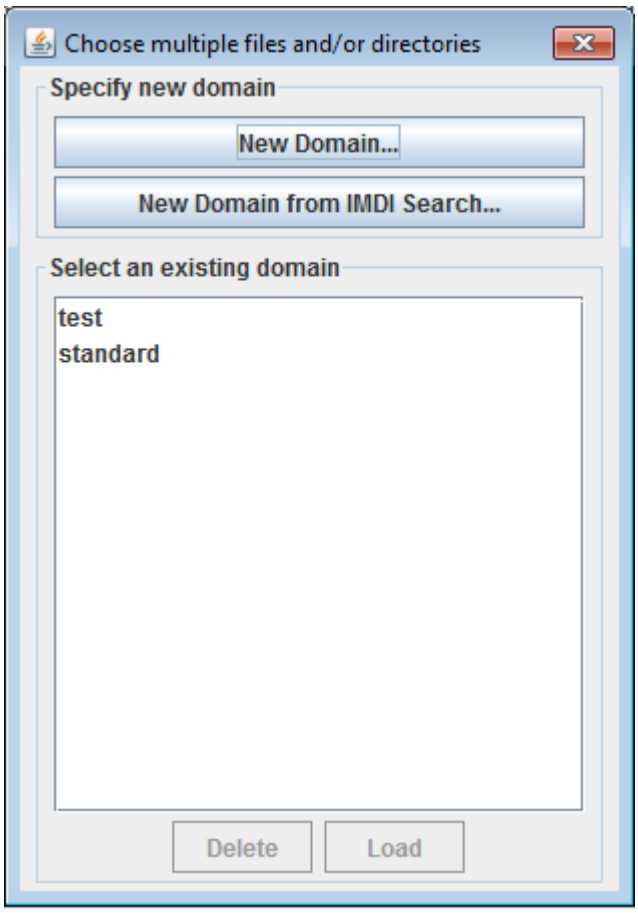


Figure 4.104. Multiple file selection dialog

- 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.

4.7.2. 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 Section 4.7.3.1)
- Praat TextGrid... (see Section 4.7.3.2)

4.7.2.1. 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.

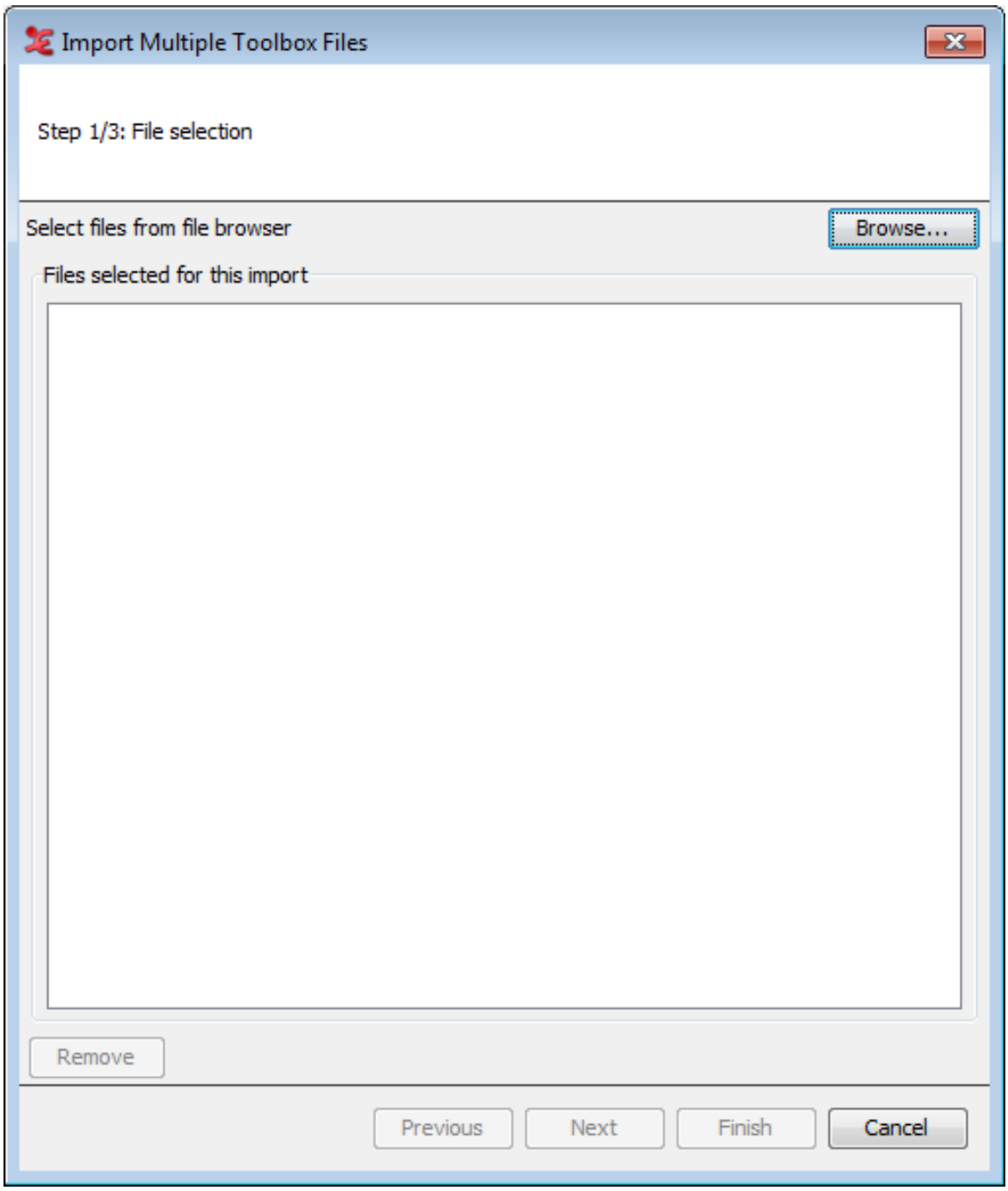


Figure 4.105. File selection

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 Section 4.3.1.3 for working with *.typ files and field markers.

3. Lastly, you will have to configure the Save as settings.

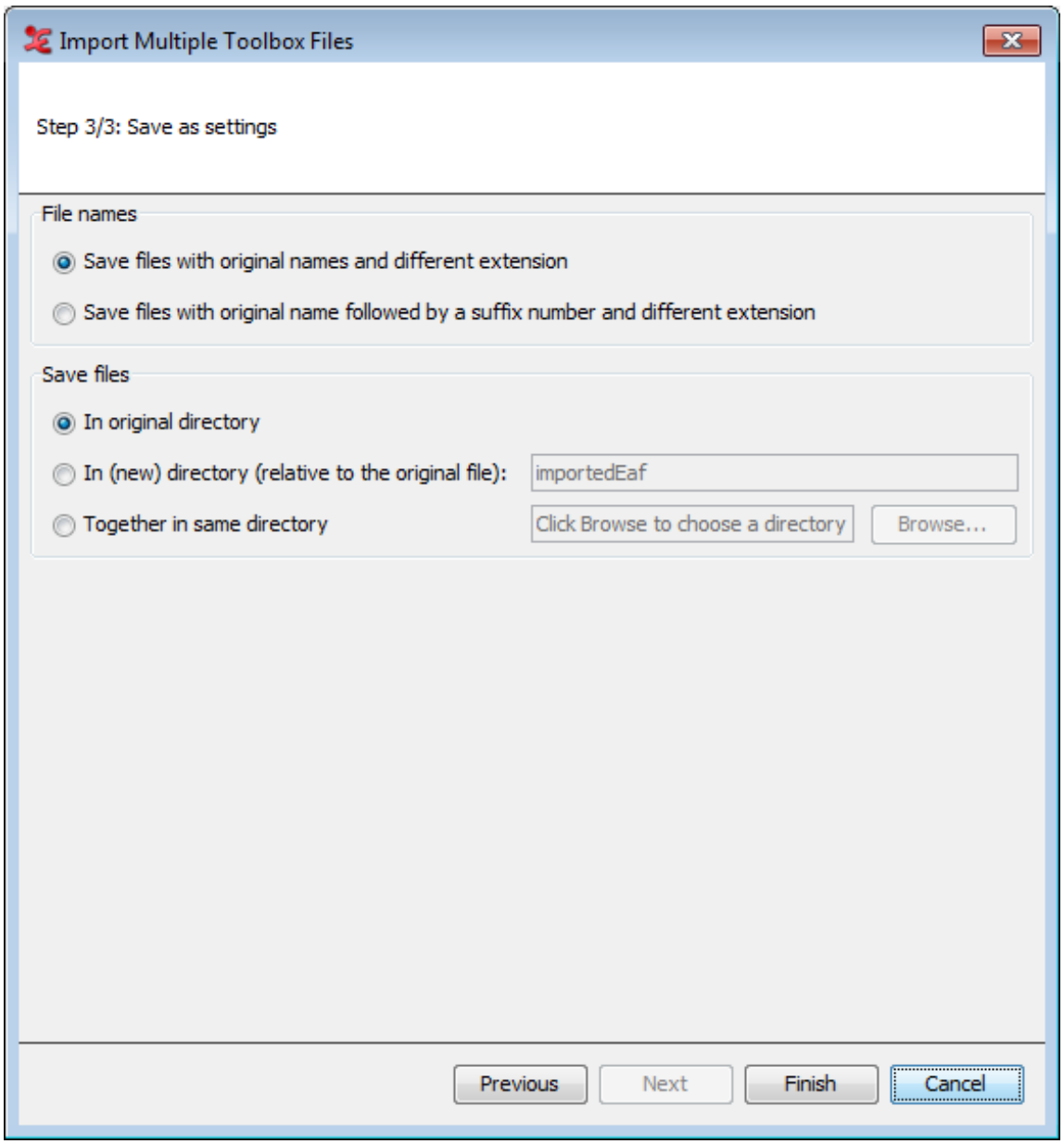


Figure 4.106. Save as Settings

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.

4.7.2.2. 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:

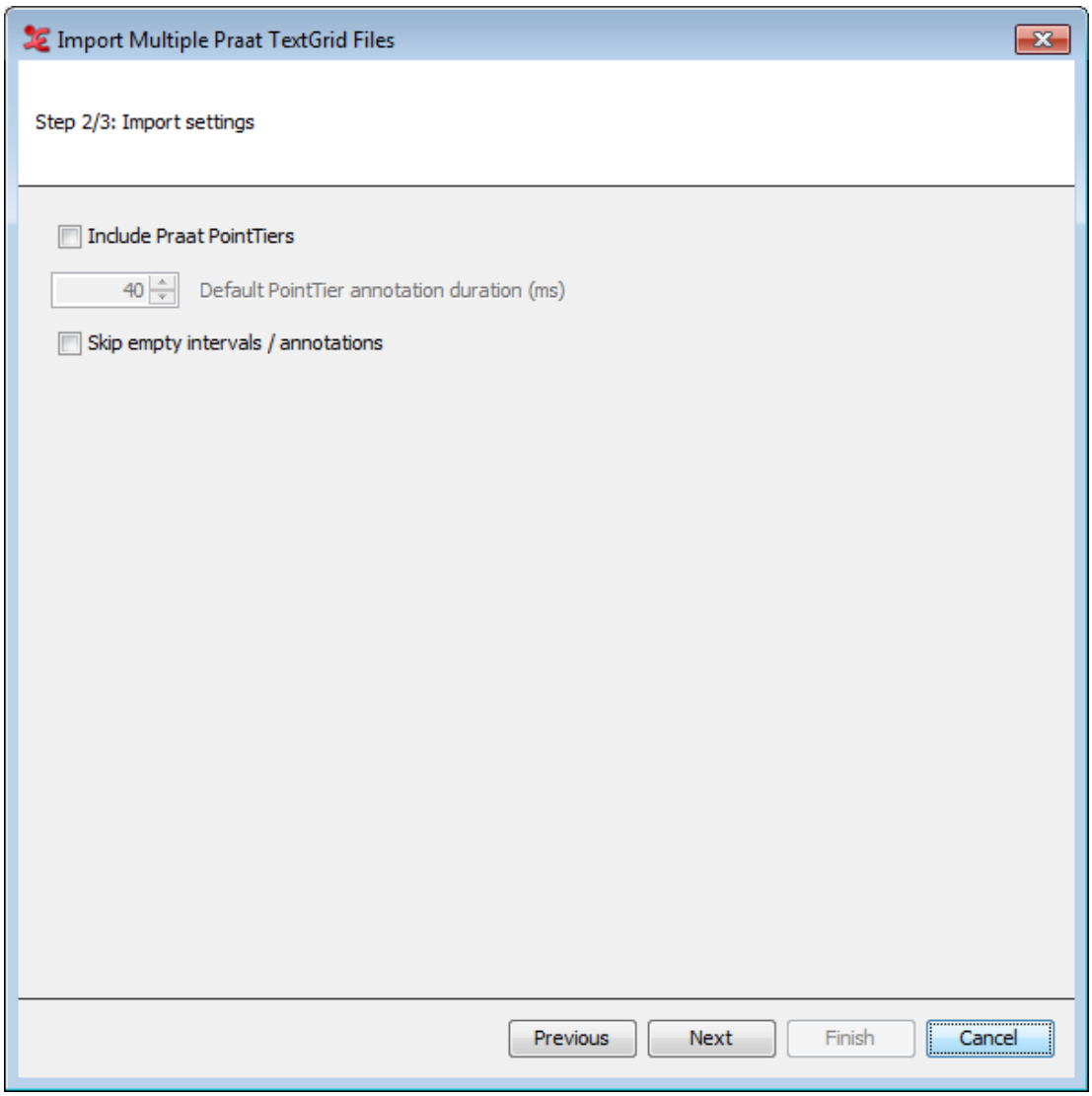


Figure 4.107. Import Settings

In this dialog, you can choose to include Praat PointTiers and if empty annotations or intervals should be skipped or not.

3. Lastly, you will be asked how the files should be saved and in what location.

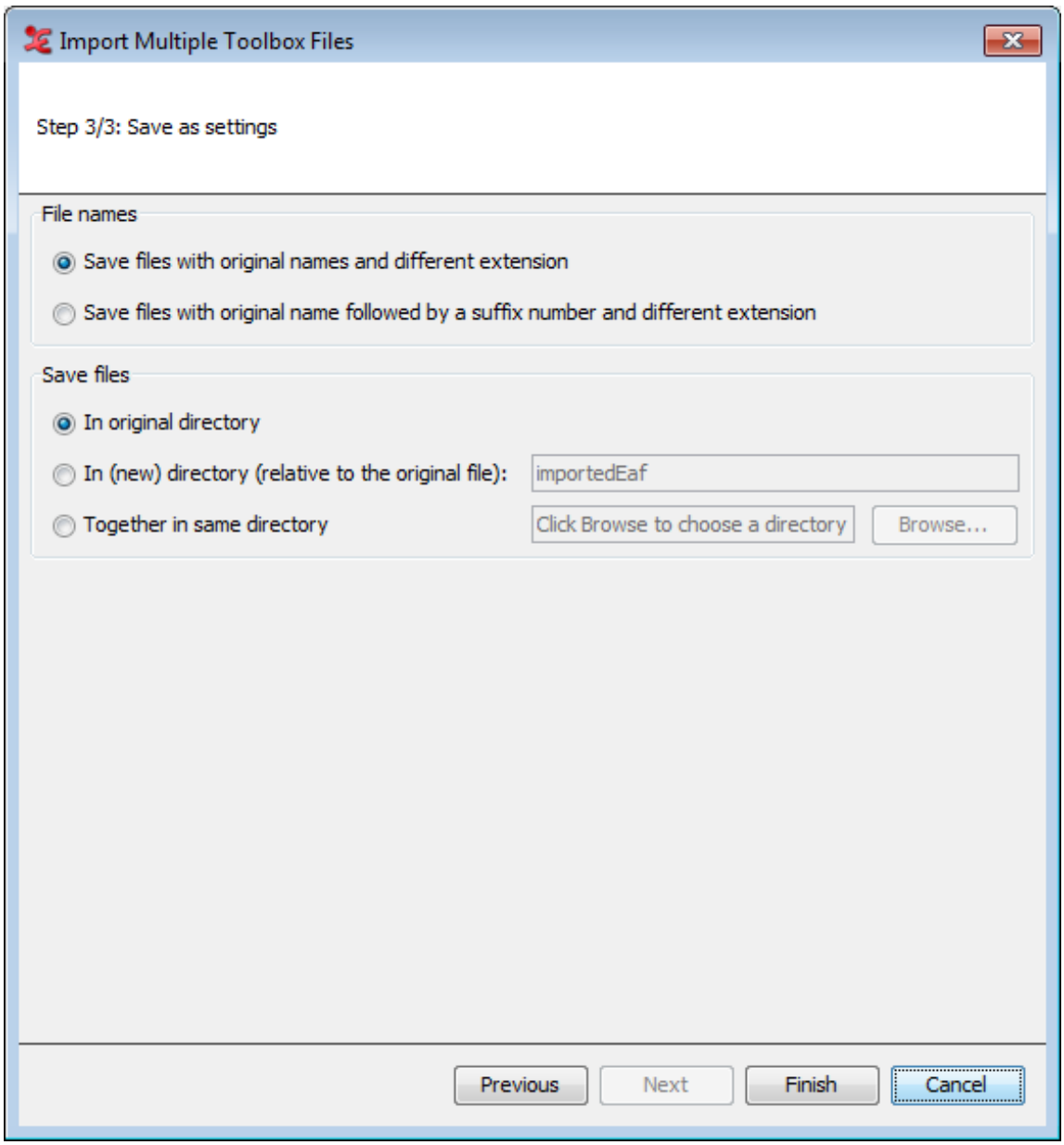


Figure 4.108. Save as Settings

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.

4.7.3. 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 Section 4.7.3.1)
- Praat TextGrid... (see Section 4.7.3.2)
- Tab-delimited Text... (see Section 4.7.3.3)
- List of Annotations... (see Section 4.7.3.4)
- List of Words... (see Section 4.7.3.5)

- Selected Tiers as EAF... (see Section 4.7.3.6)
- Annotation Overlaps Information...(see Section 4.7.3.7)

4.7.3.1. 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** (see Figure 5.79)

   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 : Section 4.7.1

   b. Next select the tiers which are to be used for the export process.

   c. Insert blank line after this marker (see Section 4.3.2.3)

   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**

   In this step you can define output settings and the Toolbox options. The option are more clearly defined in Section 4.3.2.3

3. **Step 3/3: Save as Settings**

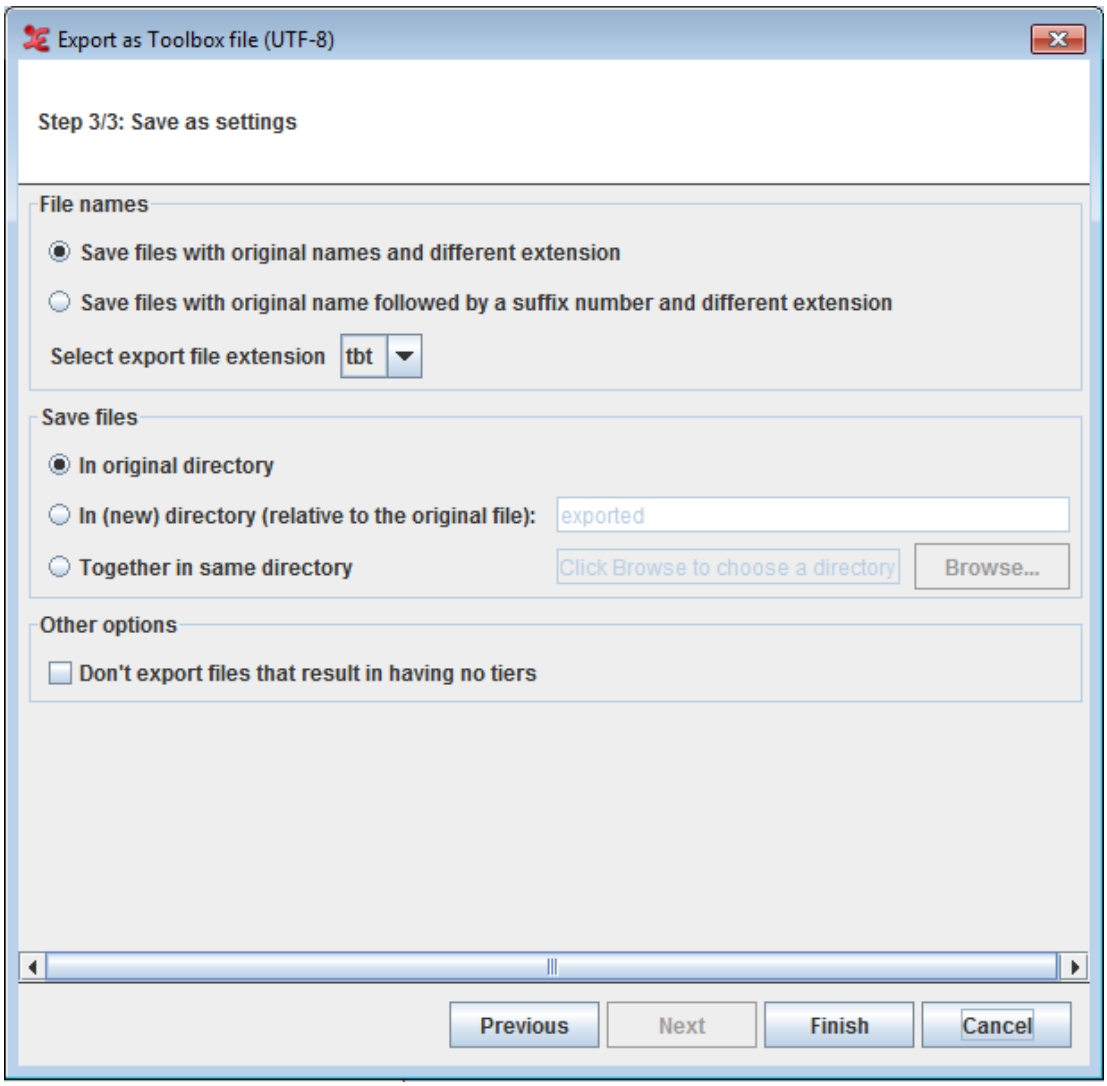


Figure 4.109. Save as Settings

4.7.3.2. Praat TextGrid

To export multiple files as praat textgrid, click on File > Export Multiple Files As... > Praat TextGrid .... This process involves 2 steps. See Section 4.3.2.8 for more details.

1. **Step 1/2: File and Tier Selection** (see 1)

2. **Step 2/2: Export Settings** (see xref3

4.7.3.3. 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 Section 4.7.1 . Select or create a domain and click on ok to continue.

2. Then in the next dialog that appears, select tiers and options as you would do when exporting a single Tab-delimited Text file (see Section 4.3.2.4). A column containing the file names is added to the export file.



Note

The following options cannot be checked:

- Restrict to selected time interval: the selected time interval does not necessarily correspond to an interval in the files selected for export.
- Add master media time offset to annotation times: the master media (time offset) of the file currently opened in ELAN does not necessarily correspond to the media (time offset) of the files selected for export.

4.7.3.4. 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 Section 4.7.1 . Select or create a domain and click on ok to continue.
2. Then in the next dialog that appears, select the tiers (see Section 4.3.2.1) 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.

4.7.3.5. 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 Section 4.7.1 . 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 Section 4.3.2.9).

4.7.3.6. 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 Section 4.7.1) 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.

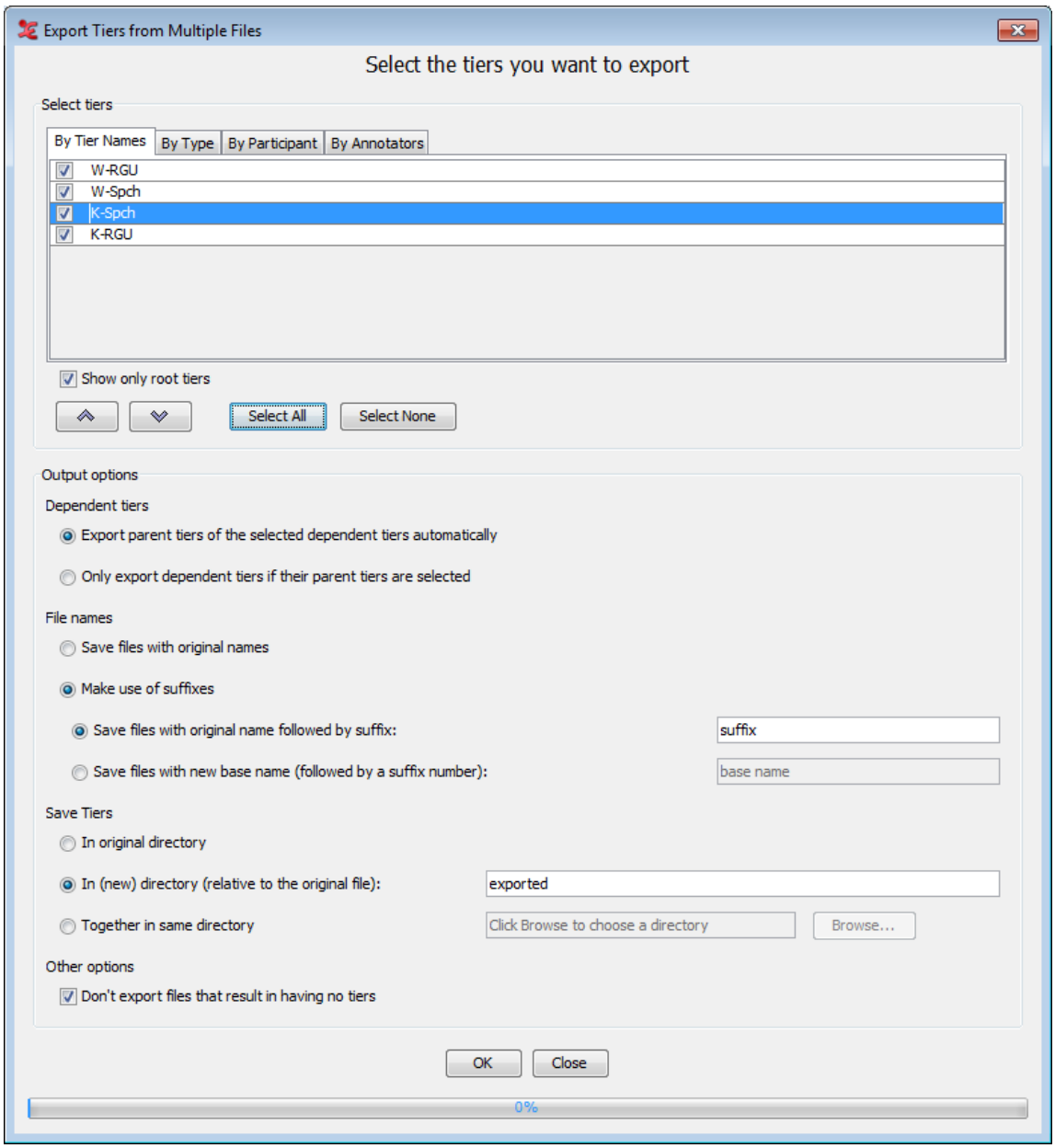


Figure 4.110. Exporting by selecting tiers

To export, do the following:

1. Select tiers (see Section 4.3.2.1) 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.

4.7.3.7. 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.

4.7.4. Editing 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).

4.7.4.1. Create multiple transcriptions

When you choose this option Create Transcription Files for Multiple Media Files..., you see the following dialog.

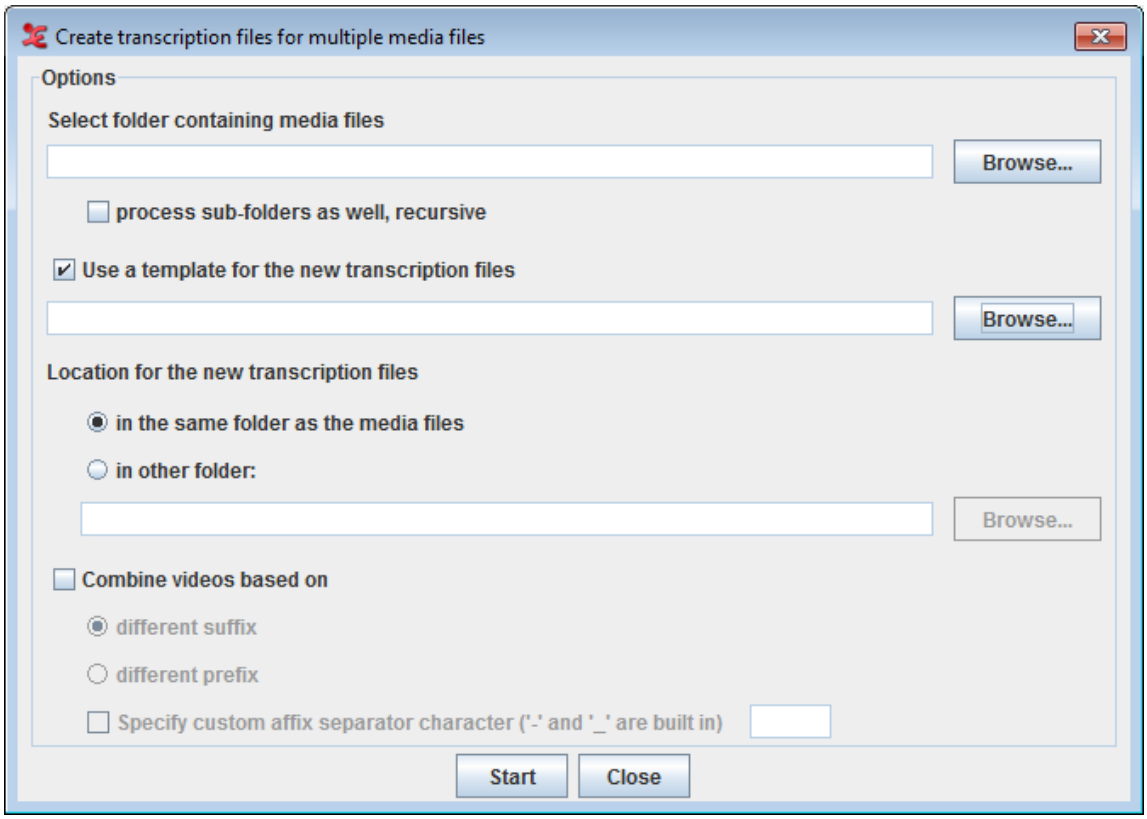


Figure 4.111. Create transcription files for multiple media files Dialog

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 preffix value and has the same file name.

  To specify a separator in the filename to identify the suffix or preffix, check this option Specify custom affix separator character ('-' and '_' are built in).

Click on Start to create the transcriptions based on the options set.

4.7.4.2. 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 Section 4.7.1). After loading a domain, the data is shown in the table. In this table you can edit tiers on the Tiers tab and linguistic types on the Linguistic 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 linguistic type of a tier, select one from the drop down menu. You can add a tier by clicking Add tier and 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.

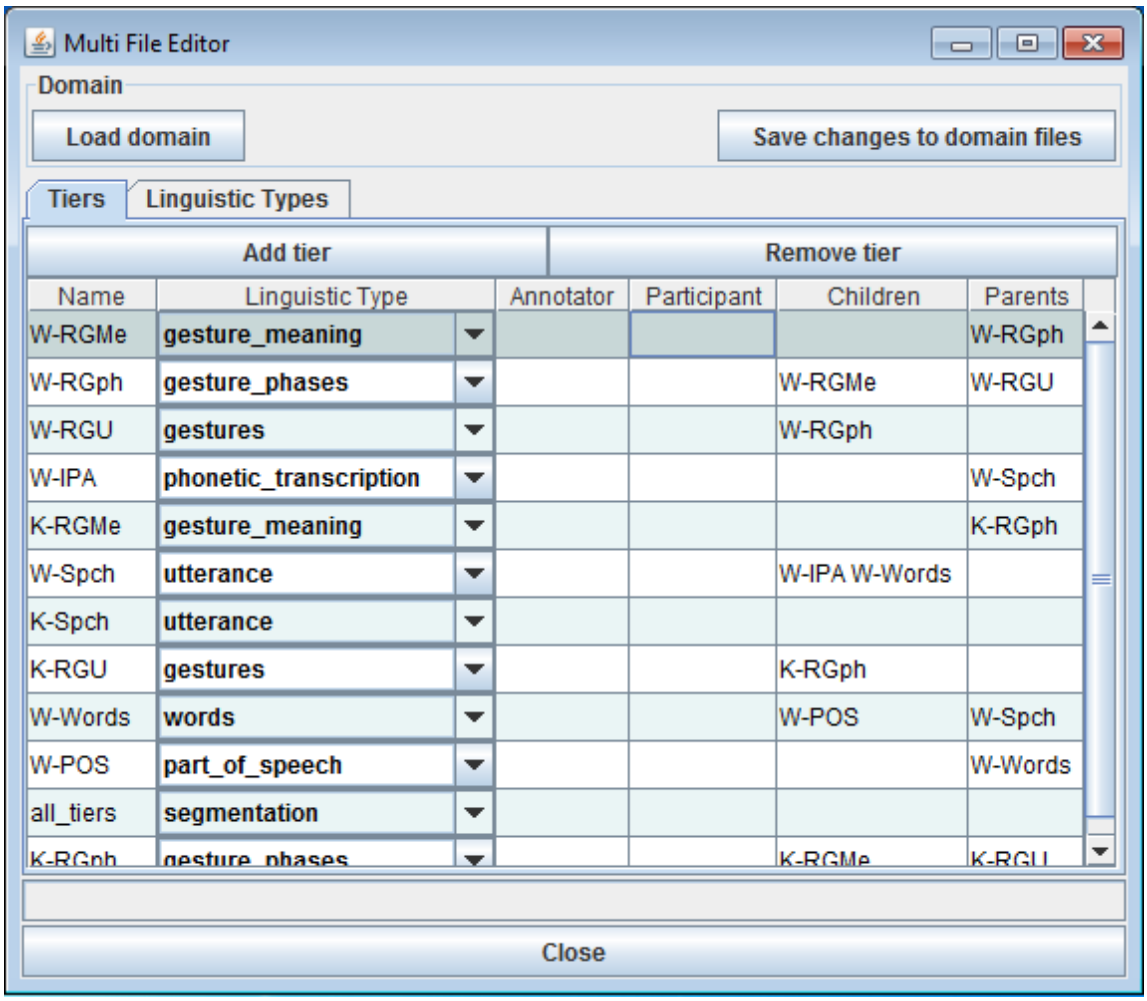


Figure 4.112. Multiple File Editor

On the Linguistic Types tab, the name of a linguistic 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.

4.7.4.3. 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 whitespace characters which are often overlooked by the user but are still saved in the file. To select, create or delete a domain see Section 4.7.1.

In the next dialog, you can specify what characters to delete, new line characters, tab characters and/or whitespace 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.

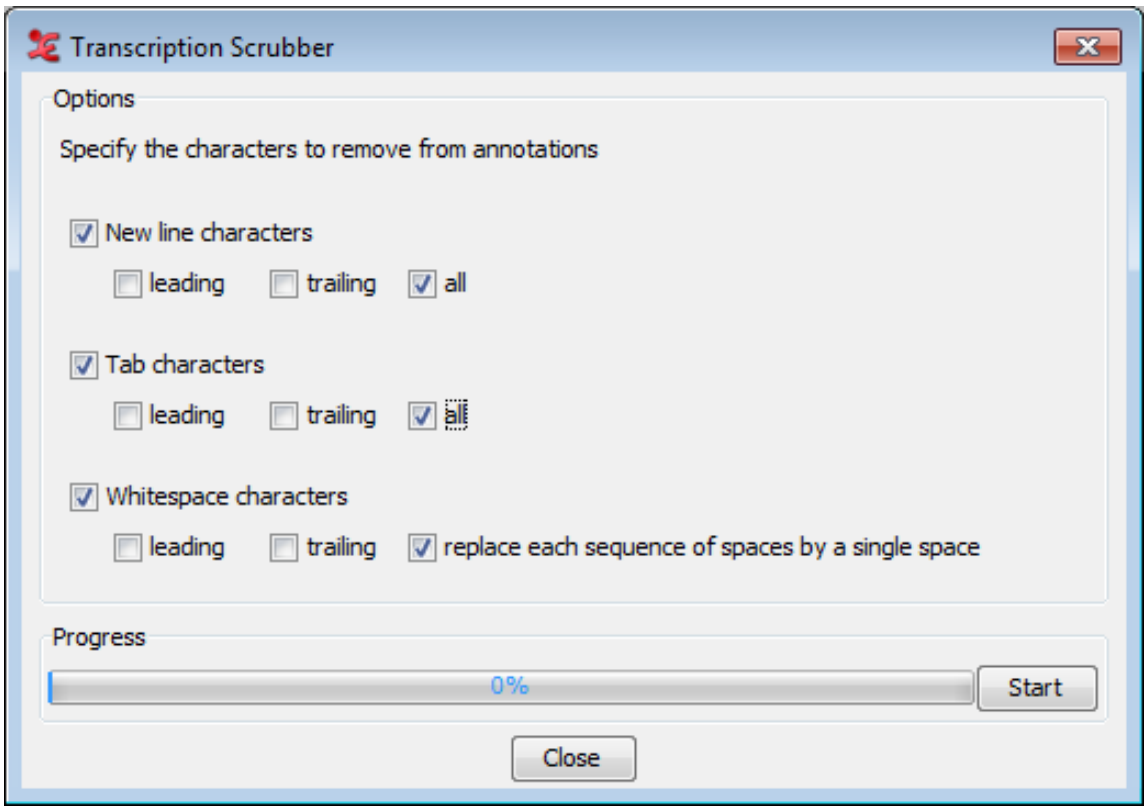


Figure 4.113. Transcription Scrubber

4.7.4.4. 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 Section 5.11.1), 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 Section 4.7.1 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.

4.7.4.5. 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 Section 5.12), 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 Section 4.7.1 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.

4.7.4.6. Annotation statistics

This function Statistics for Multiple Files is similar to annotation statistics for the current file (see Section 5.15.2). 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.

Chapter 5. Annotations

You can use the ELAN program for annotating your data. This annotation process involves three steps: defining linguistic types and tiers (see Section 5.3.1 and Section 5.4), selecting time intervals (see Section 5.7), and entering annotations (see Section 5.8).

5.1. Basic Information: Annotations, tiers and linguistic types

The following illustration shows an example of an annotation document:

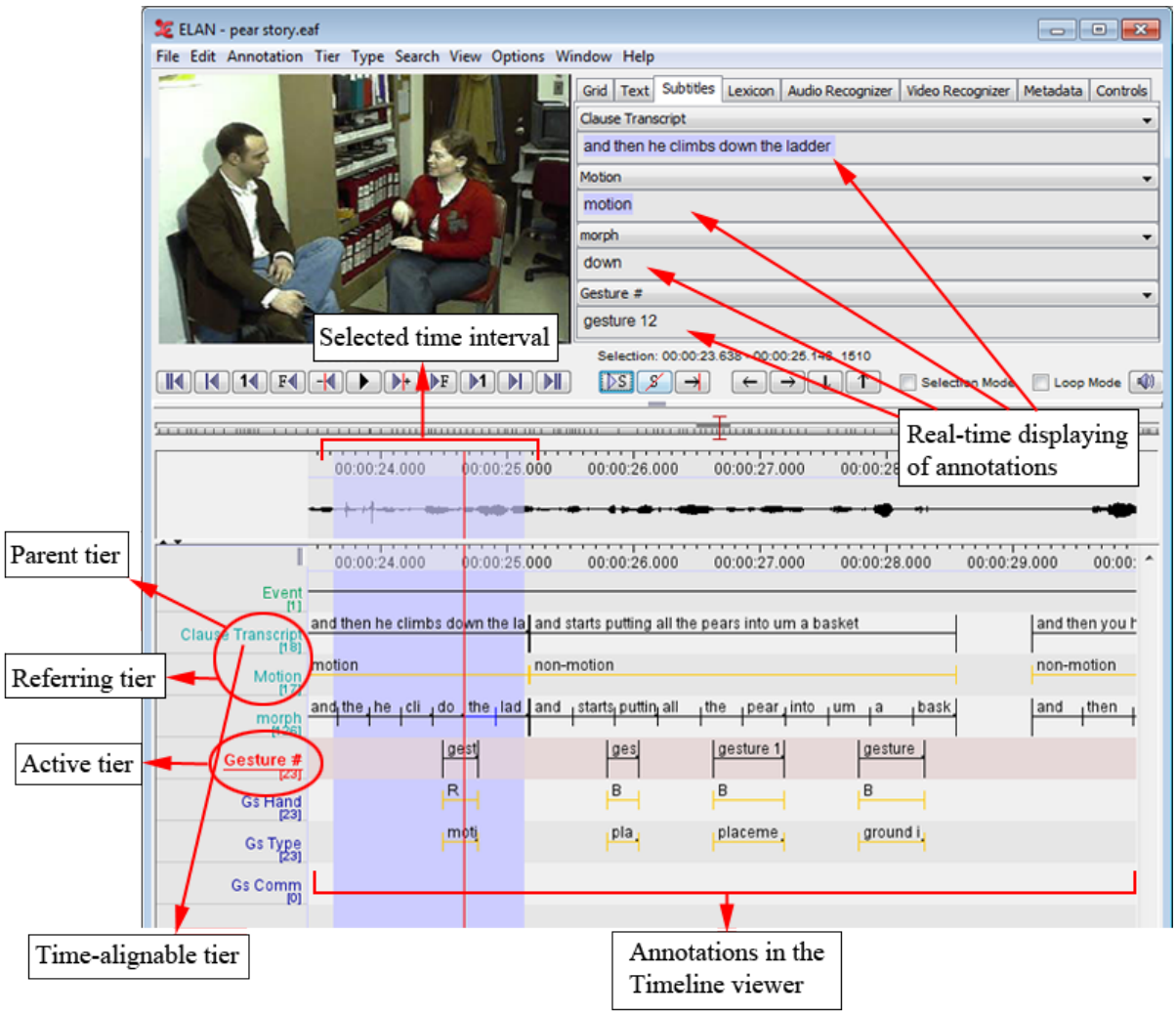


Figure 5.1. Annotation basics

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 (Section 4.4.11), 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 Section 5.8).

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 5.1. 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	-

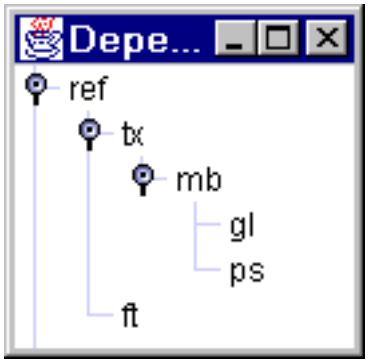


Figure 5.2. Tier dependencies

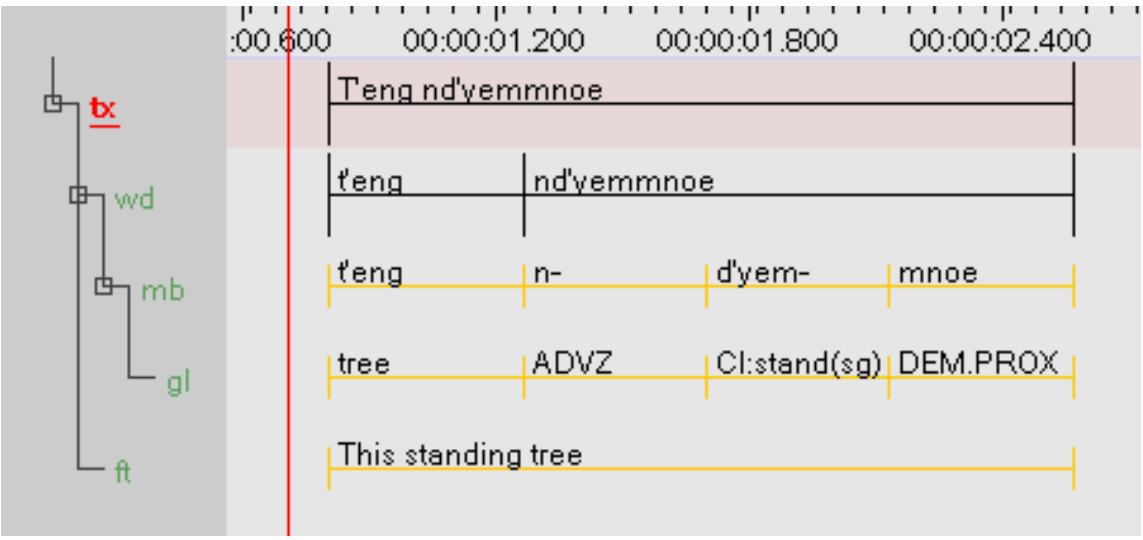


Figure 5.3. Tier dependencies in the timeline viewer



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 linguistic type (see also Section 5.3). A linguistic type denotes the linguistic data that is contained in the referring tier. Examples of names for linguistic types are *utterances*, *words*, *orthography*, *phonetic transcription*, *PoS* (part of speech), but any name can be used. Each linguistic 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 four stereotypes are currently available:

Table 5.2. Linguistic 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.
• 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. 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. 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[a]:	All annotations fall within the borders of the parent tier. However, there can be gaps between the child annotations. 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. E.g., one sentence on a parent tier has exactly one free translation. Or one word has exactly one gloss.

A similar stereotype exists in Media Tagger, so it is especially useful for the import of such files.

The following example illustrates the four different stereotypes (see also Figure 5.3):

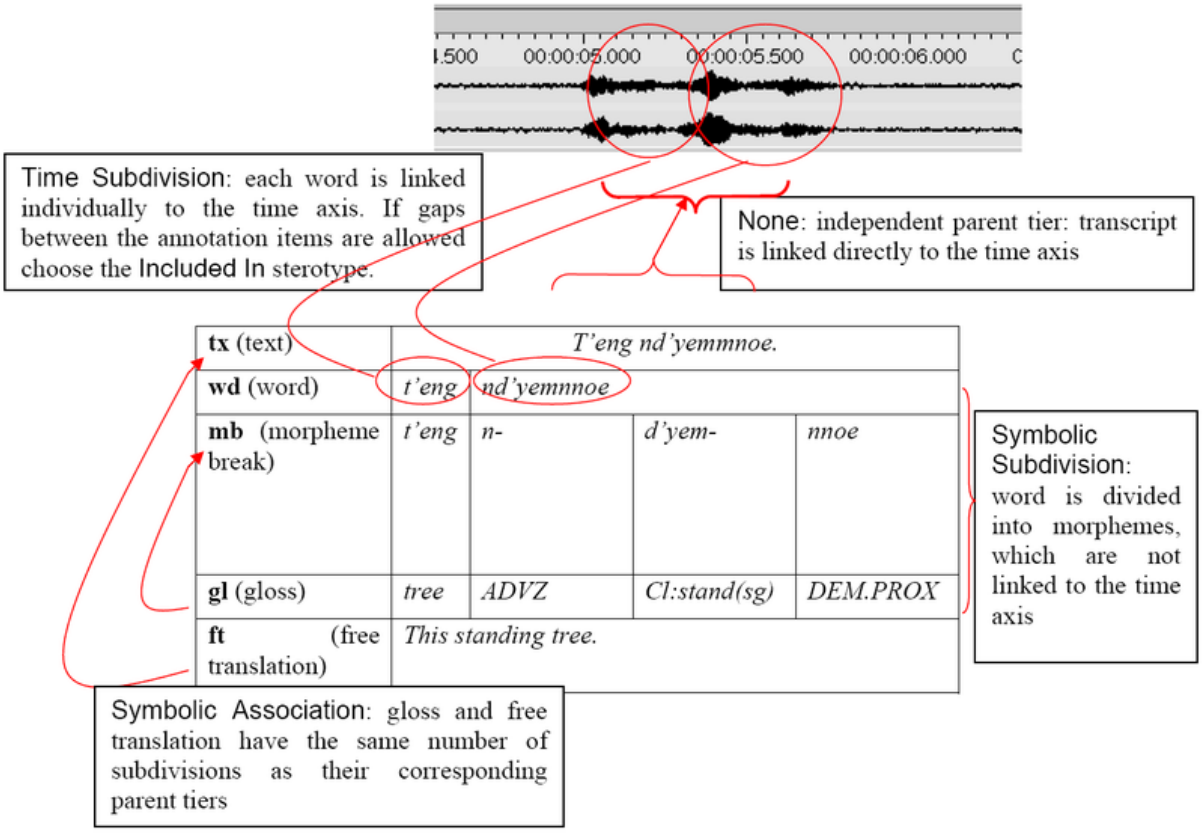


Figure 5.4. Examples of the four stereotypes

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 Section 5.4.8) or to copy a tier (Section 5.10.2) and to alter the copy.

5.2. Data categories

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 Section 5.5) and templates (see Section 4.2.10) 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 (linguistic type (Section 5.3.6), controlled vocabulary entry (Section 5.5.2) or annotation (Section 5.8.21)), the following or a similar window is displayed.

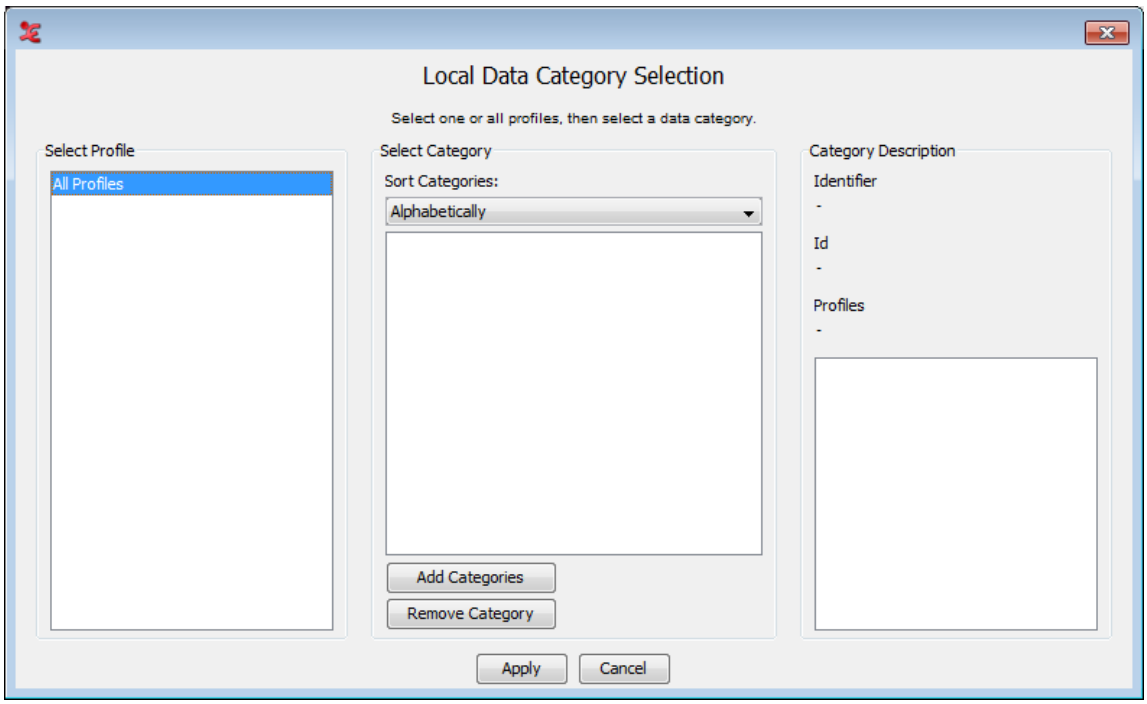


Figure 5.5. Local Data Category Selection

The left panel shows the profiles for which there are data categories in stored on your local system. In Figure 5.5 the only option is All Profiles. Selecting this option will show all data categories of all profile in the left panel. Since there are no profiles in the left panel, the middle panel does not display any data category. To add categories, click on Add Categories. The following window appears:

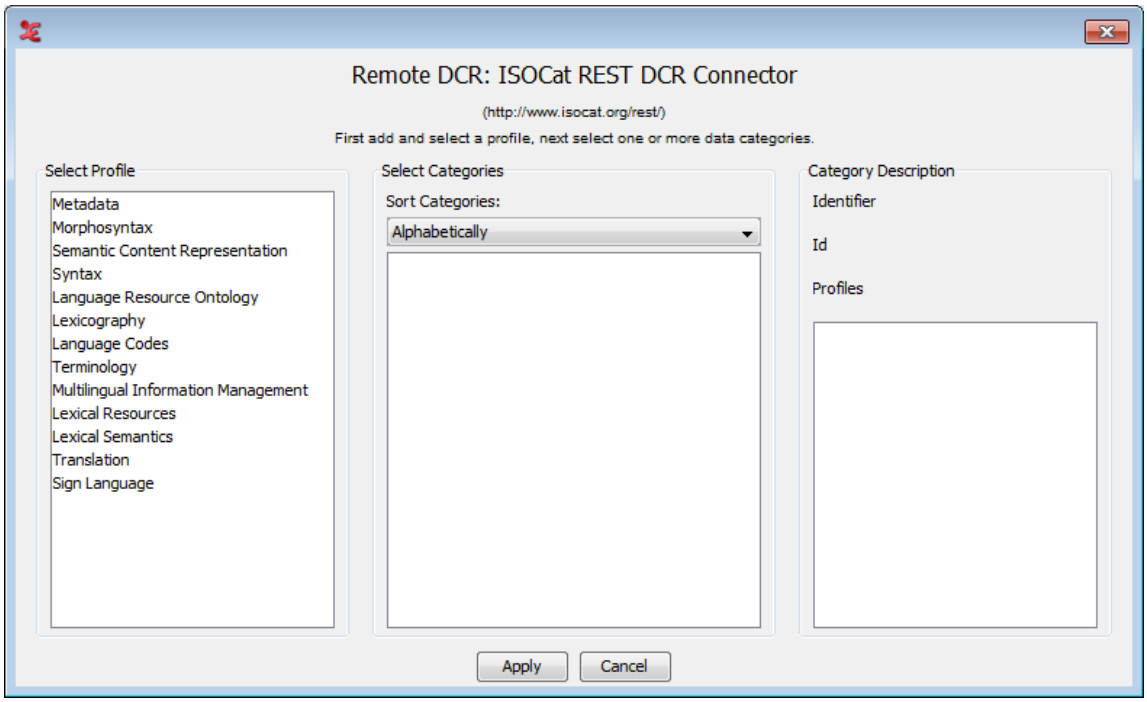


Figure 5.6. Remote DCR

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 in Figure 5.6. 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 5.7 shows the result on the Local Data Category Selection:

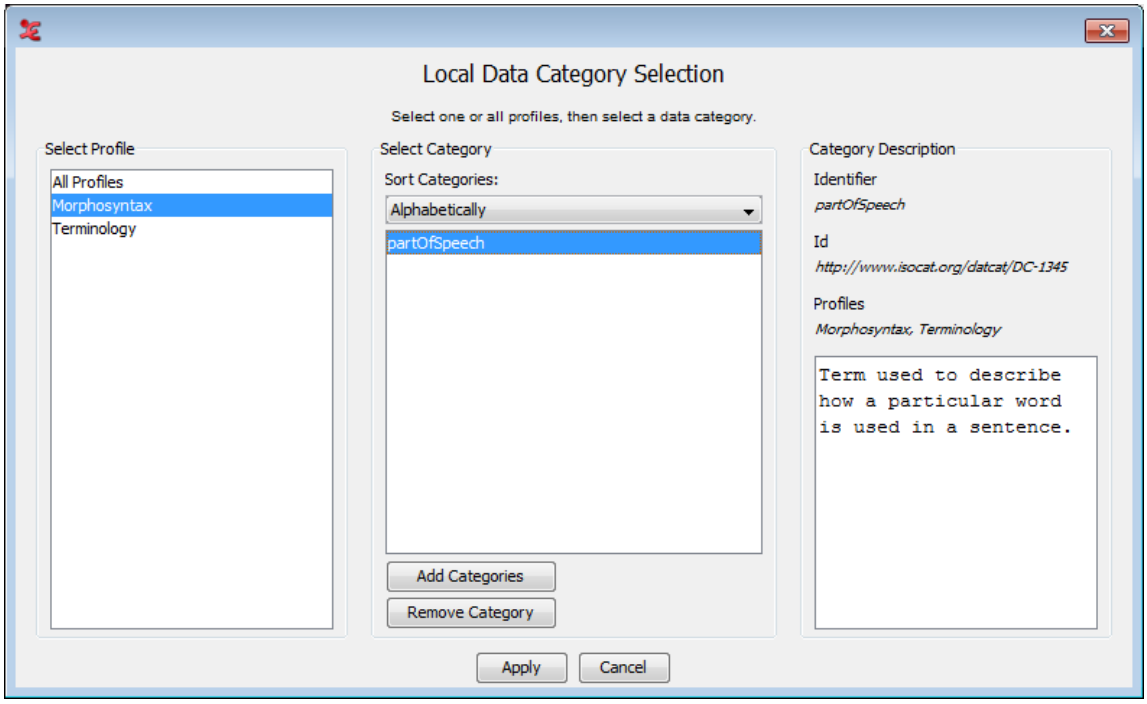


Figure 5.7. Local Data Category Selection: Result

In the same way as described above more data categories, also from other profiles, can be selected and stored on your local system.

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 linguistic type, depending on the point from which you entered the Local Data Category Selection.

5.3. How to define a linguistic type

ELAN supports the following options:

- adding a new linguistic type (Section 5.3.1 and Section 5.3.2);
- changing the attributes of a linguistic type (Section 5.3.3);
- deleting a linguistic type (Section 5.3.4);
- importing a linguistic type (Section 5.3.5).

5.3.1. Adding new linguistic types

Do the following to add a new linguistic type:

1. Click on Edit menu.

2. Click on Add New linguistic type… The Add type dialog window appears.

3. Define the linguistic type (see Section 5.3.2).

4. Click Add to save the linguistic type. Otherwise click Close to exit the window without saving.

5.3.2. Creating linguistic types

Every tier is assigned to a linguistic type. The linguistic 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 Section 5.4.1), and for all tiers belonging to one linguistic 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 linguistic type, do the following:

1. Click on the Type menu.

2. Either go to Add New linguistic type…. In the latter case, click on the linguistic type that you want to change when the Add dialog window appears:

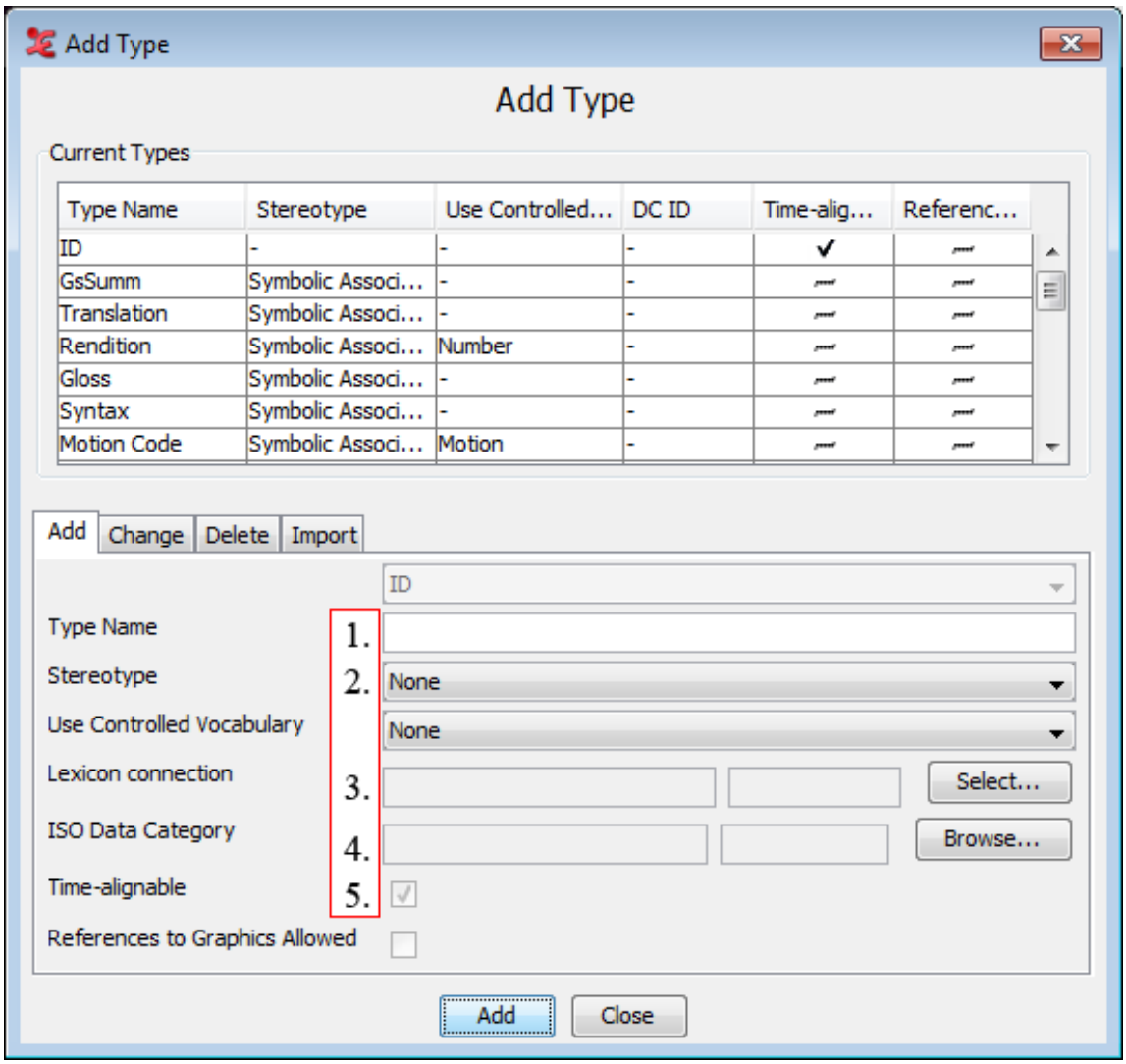


Figure 5.8. Add a linguistic type

1. *Enter a label for the new type*
2. *Select the stereotypical constraint:*
   - *None*
   - *Time Subdivision*
   - *Included In*
   - *Subsymbolic Subdivision*
   - *Symbolic Association*
3. *Associate the new type with a Lexicon Service and Lexicon Entry Field via this option*
4. *Associate the new type with an ISO Data Category via this option*
5. *Ignore this option*

Enter the following information:

1. Go to Type name. Enter/change the name for the type.
2. Go to Stereotype. Select the stereotypical constraint relevant to its tiers (see Section 5.1).

3. Optionally select a controlled vocabulary (see Section 5.5.7).

After you have selected a stereotypical constraint, ELAN automatically enters the correct value next to the box Time-alignable.

Click Add to save the settings; otherwise click Close to exit the window without saving them.

5.3.3. Changing linguistic types

Do the following to change the attributes of an already existing linguistic type:

1. Click on Edit > Change linguistic type… The Change type dialog window appears.

2. The labels of all available linguistic types are displayed in the pull down menu, e.g.:

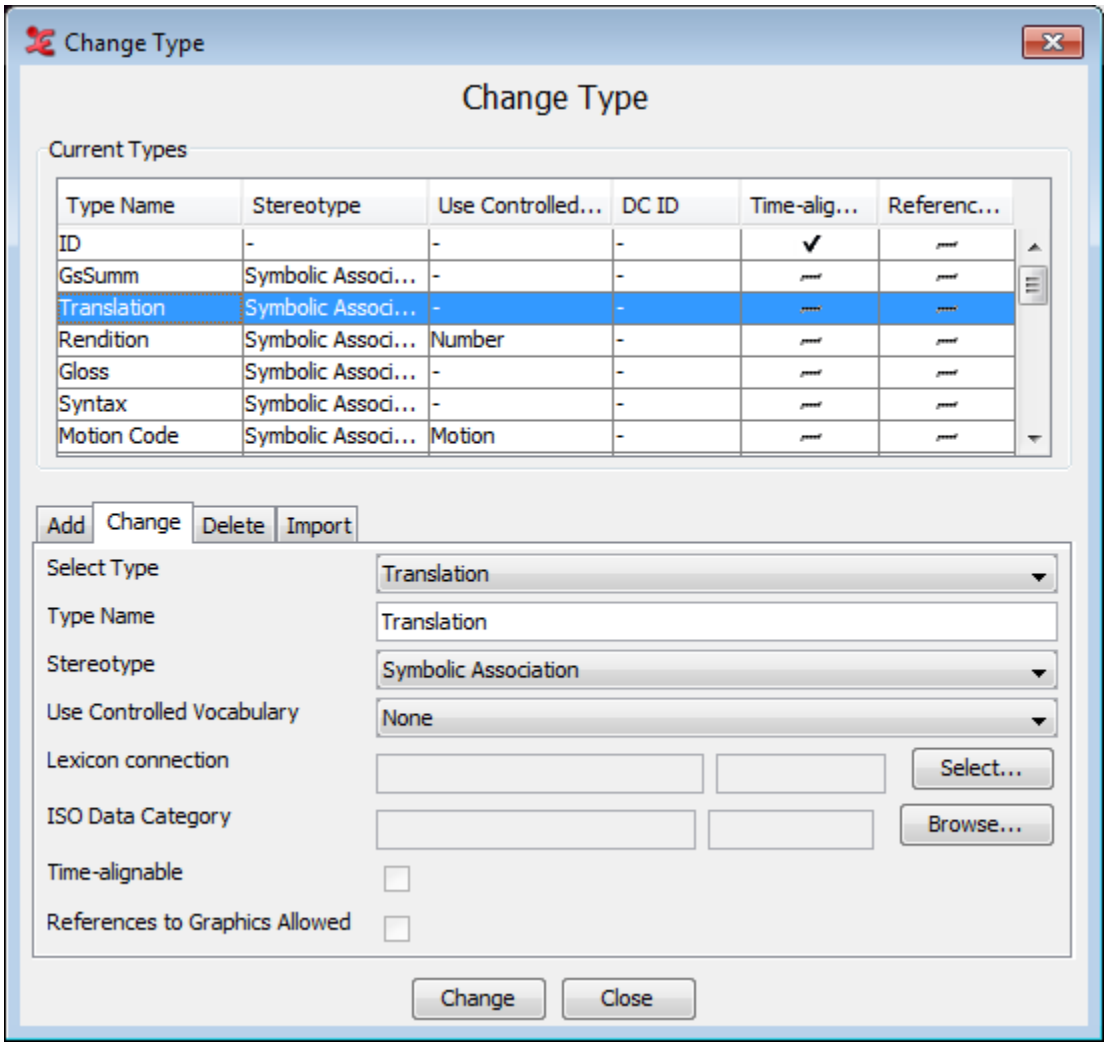


Figure 5.9. Change a linguistic type

3. Click on the linguistic type whose attributes you want to change.

4. Change the settings.

5. Click Change to save the changes. Otherwise click Close to exit the window without saving the changes.

5.3.4. Deleting linguistic types

To delete a linguistic type, do the following:

Click on Edit menu.

1. Go to Delete linguistic type...
2. A dialog window appears. The names of all available linguistic types are displayed in the pull down menu, e.g.:

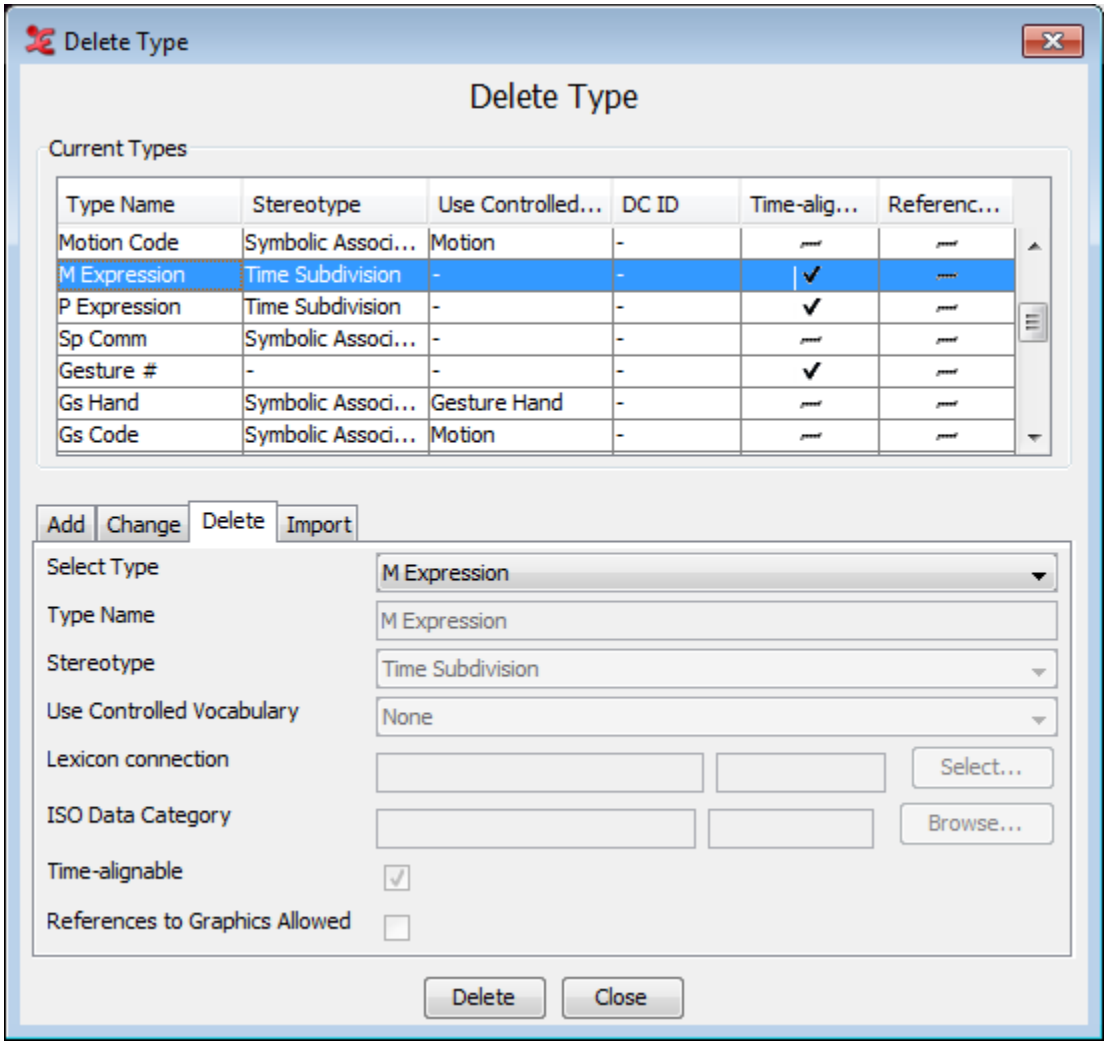


Figure 5.10. Delete a linguistic type

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 linguistic type if it is not used by any of the tiers. If it is used, the following error message appears:

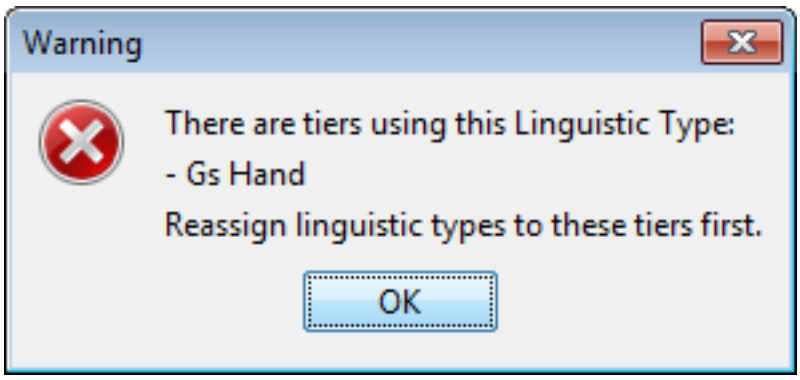


Figure 5.11. Linguistic type is used

5.3.5. Importing linguistic types

It is possible to import linguistic 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 linguistic types from the selected file.

5.3.6. Associating a linguistic type with a Data Category

In ELAN it is possible to associate a linguistic 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 Linguistic 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 5.12), first select a profile in the left panel and then select a data category in the middle panel. Finally click on Apply to associate the selected data category to the linguistic type.

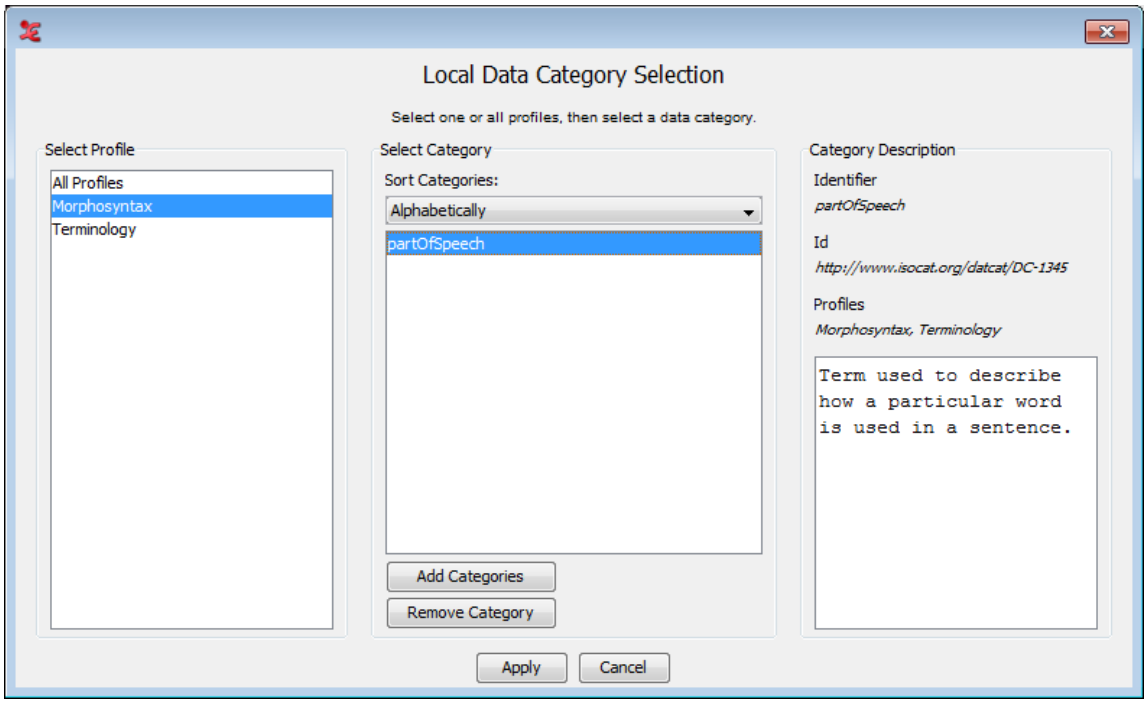


Figure 5.12. Local Data Category Selection

More information about the ISO DCR and how to use it can be found in Section 5.2.

5.4. How to define a tier and its attributes

ELAN offers the following options:

- adding a new tier (Section 5.4.1);
- changing the tier attributes (Section 5.4.4);
- deleting a tier (Section 5.4.5).

5.4.1. 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:

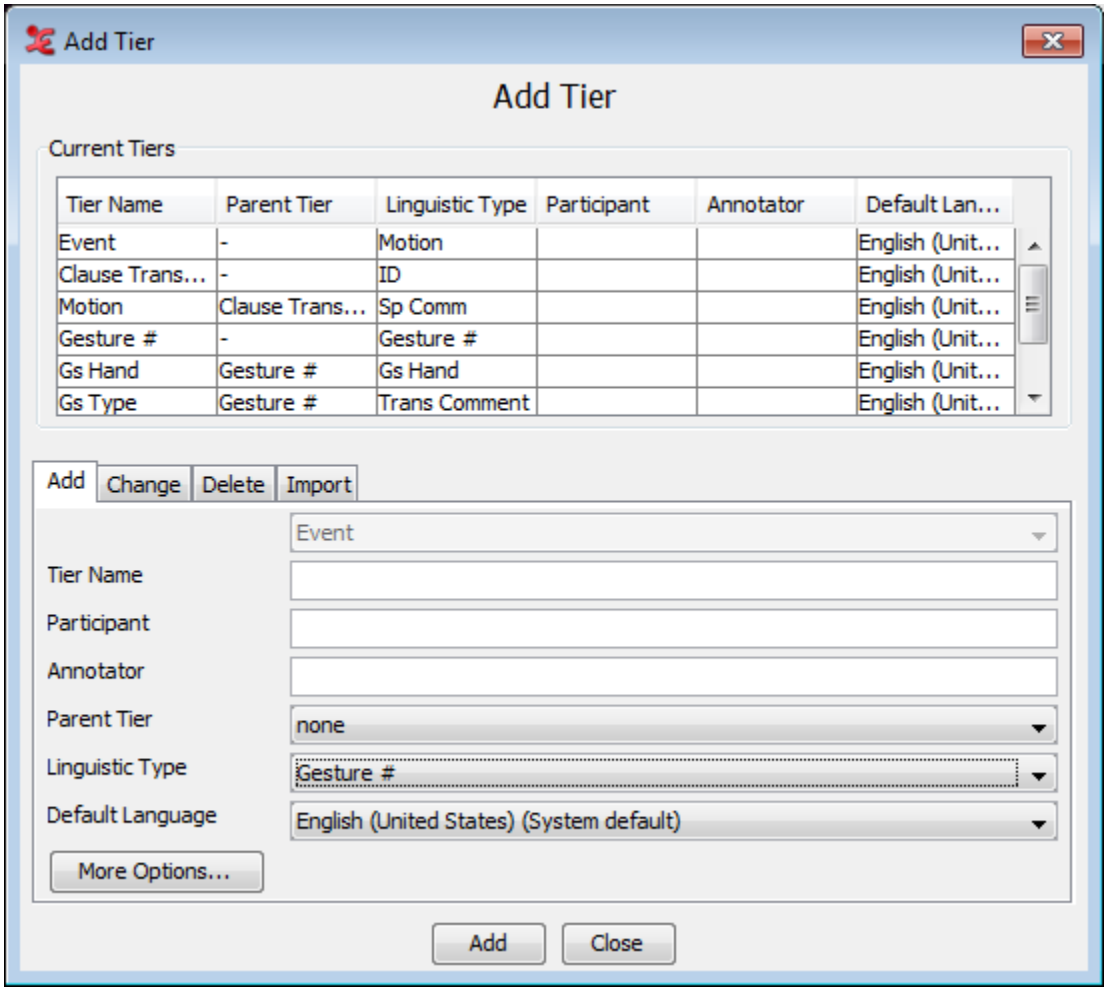


Figure 5.13. Tier attributes

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 is being transcribed.
3. 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.
4. Go to Linguistic type. Select a linguistic type from the predefined list in the pull-down menu (see Section 5.3).



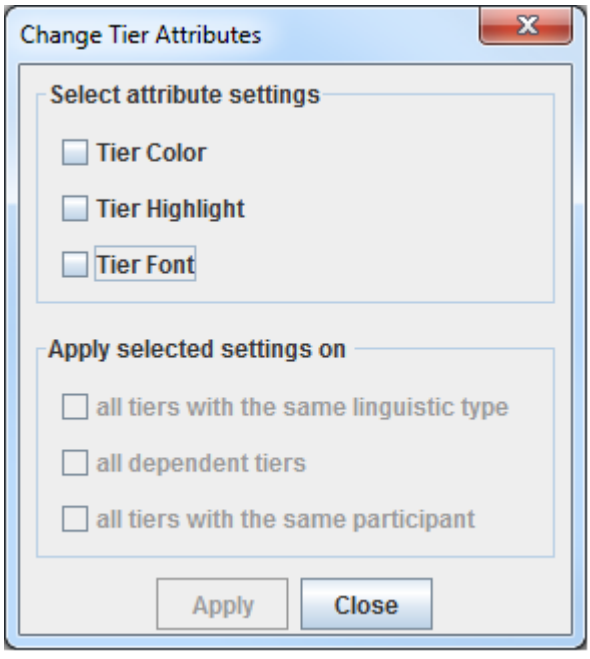
Note

The list of possible linguistic 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 linguistic types to choose from are of the stereotype "none" (see Section 5.3).

5. Go to Default Language. Select the default character set from the pull-down menu or select None. If None is selected, no language will be saved in the tier setting. This means that the system's default language will always be used for this tier.

   Whenever you enter or change annotations on that tier, the text entry box is automatically preconfigured for the default character set.

6. 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:



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.

5.4.2. Importing tiers

You can reuse the tiers (together with their linguistic types) you created before in other eaf or etf (templates, see Section 4.2.10) 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.

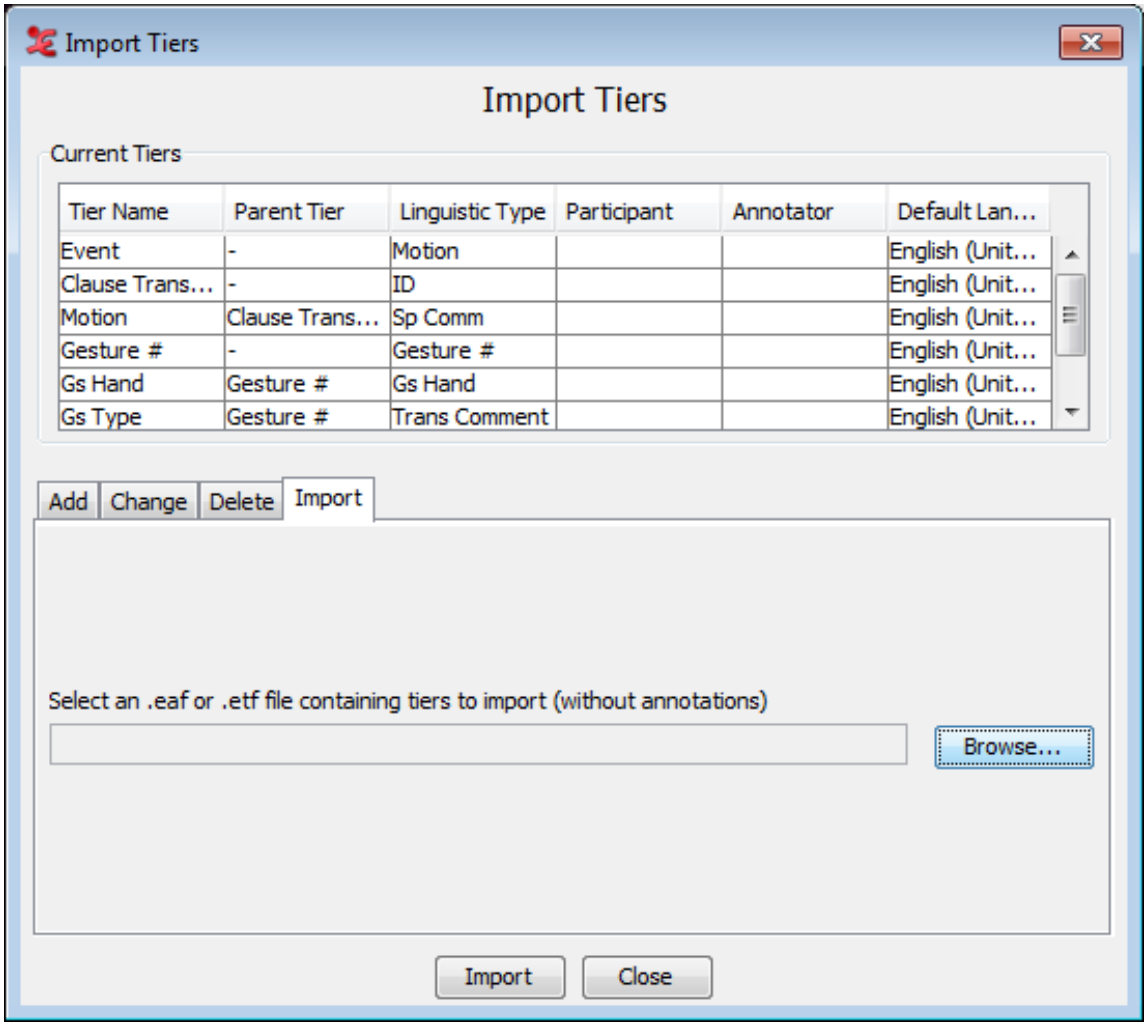


Figure 5.14. Import tiers



Note

If you import a linguistic type that already exists, a postfix like –cp1 or –cp2 will be added to the imported version.

5.4.3. Creating a tier using the Audio Recognizer

As from version 3.6 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 Audio Recognizer in the main ELAN window. Choose the appropriate recognizer from the pull down menu at the top of the tab, in this case Silence Recognizer (see also Figure 5.15).

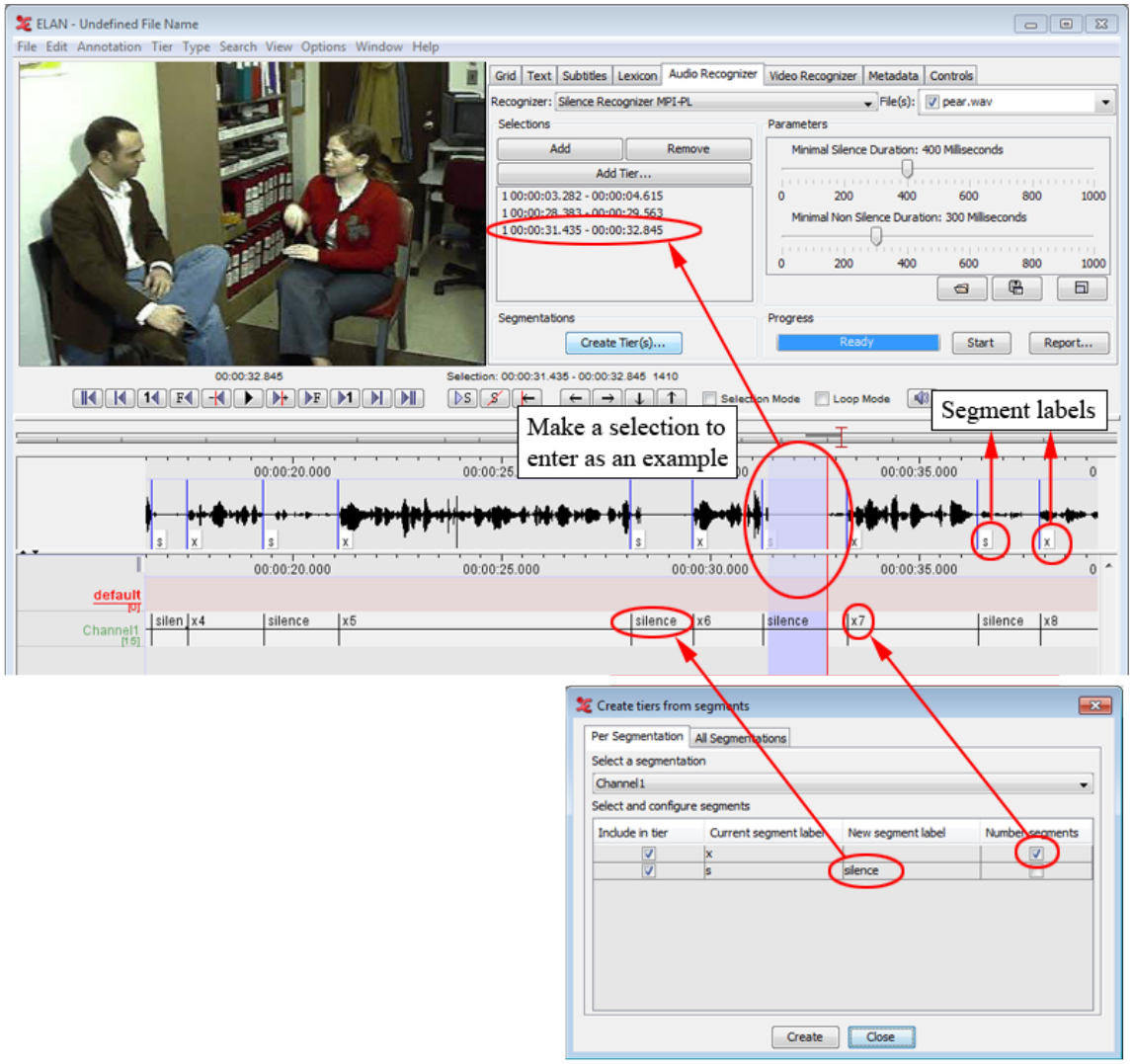


Figure 5.15. Audio Recognizer

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 Section 9.1.6). Then click on Add Channel1 (or Add Channel2 if there is a second channel and that is the channel in focus) in the Audio Recognizer tab. 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 Remove 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.

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 Audio Recognizer 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.

The controls we have seen so far are common for all recognizers. There are also controls that are specific for a recognizer. These controls can be found in the right half of the Audio Recognizer 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, please read the relevant section on the Source Release Notes page (http://tla.mpi.nl/tools/tla-tools/elan/src-release-notes.html).

5.4.4. Changing tier attributes

Do one of the following to change the attributes of an already existing tier:

1. Use the mouse.

    a. In the Timeline Viewer, right-click at about the height of the tier that you want to change. A pull-down menu appears.

    b. In the pull-down menu, click on Change tier attributes. The Change tier attributes dialog window for that tier appears.

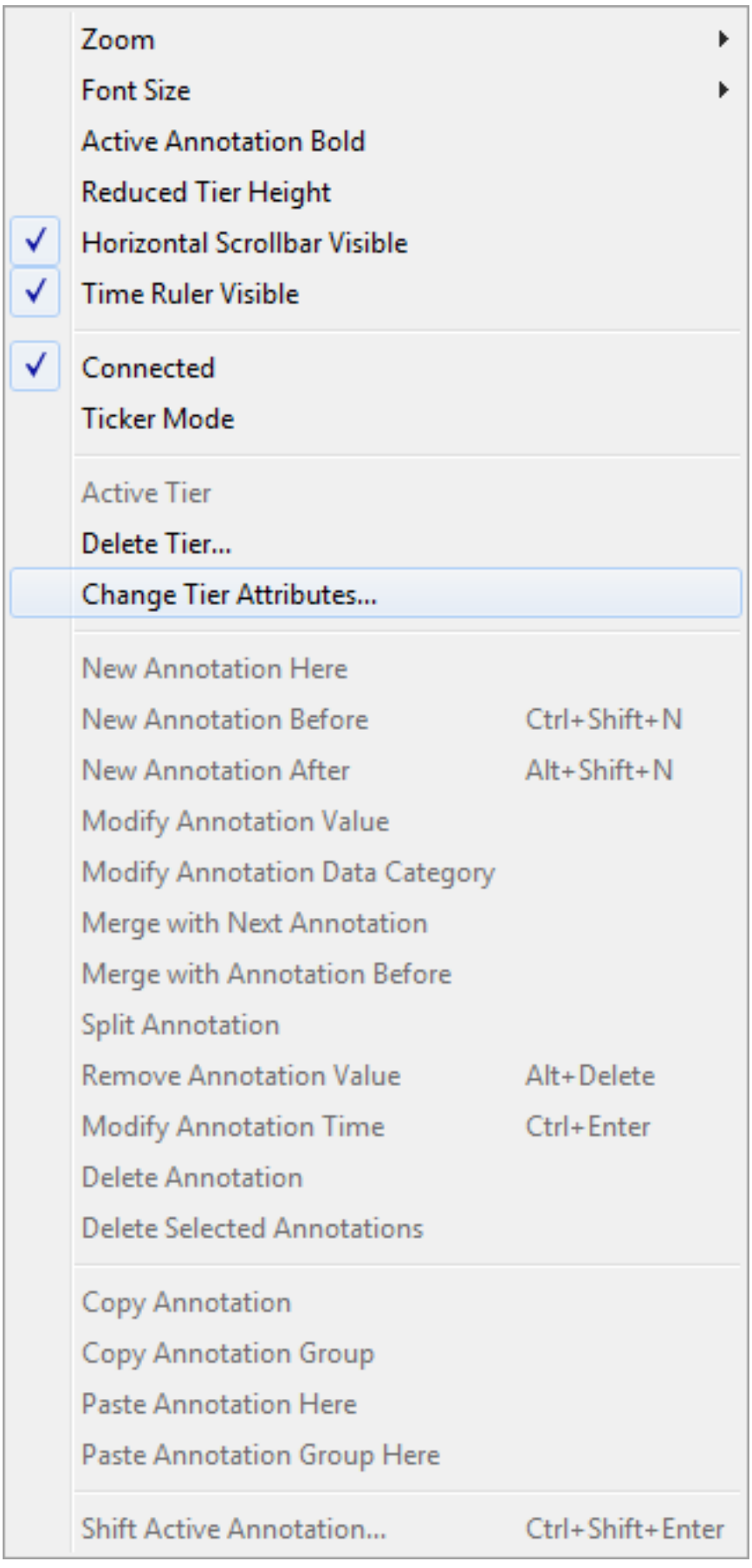


Figure 5.16. Timeline context menu

c. Change the tier attributes in the dialog that now appears and confirm by clicking on Change.

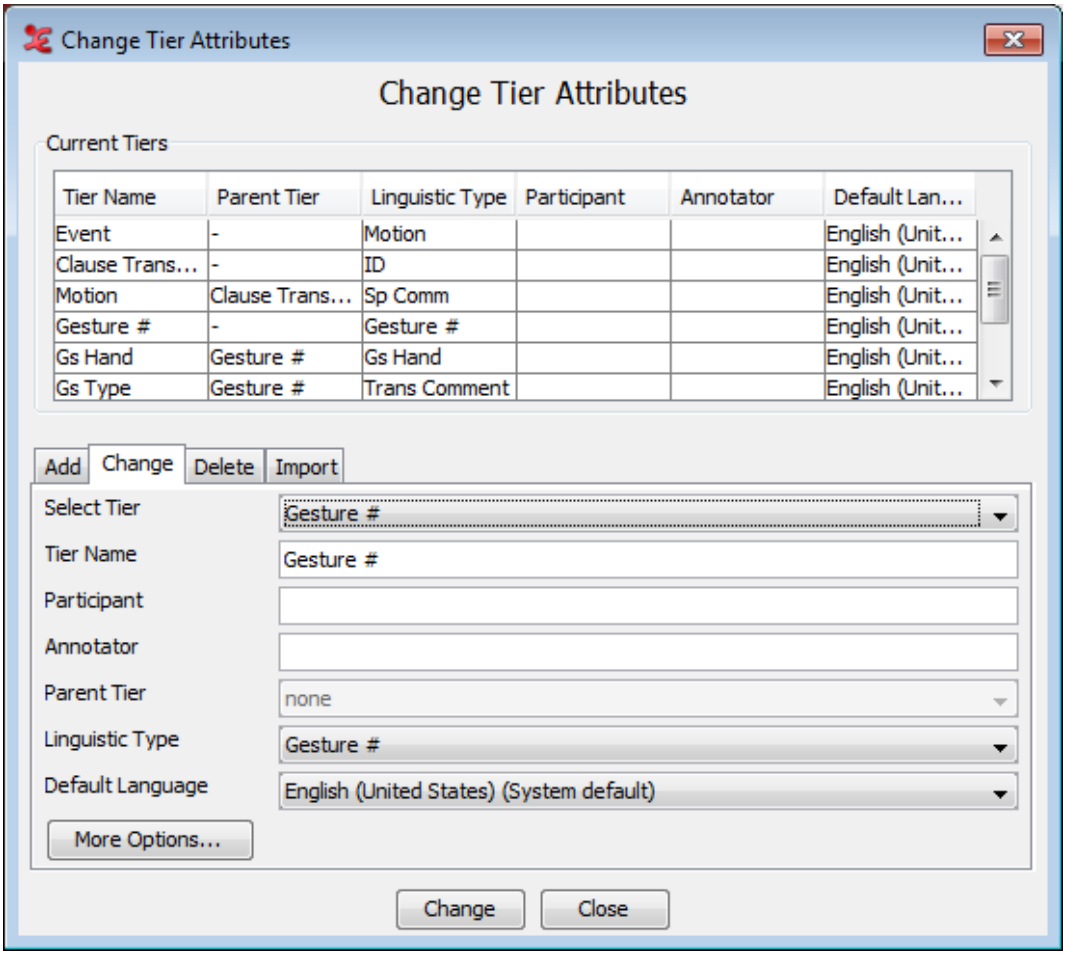


Figure 5.17. Change tier attributes



Note

You cant change the Parent Tier nor the Linguistic type in this dialog because of possible data loss this alterations would cause. To change these settings in a safe way (i.e. creating a copy of the tier and changing this copy) see Section 5.4.4 and Section 5.4.8).

2. Use the Tier menu.

   a. Click on Tier menu.

   b. Go to Change tier attributes…

   c. 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.

   d. After making the changes, click on Change to save them. Otherwise click Cancel to exit the window without saving.

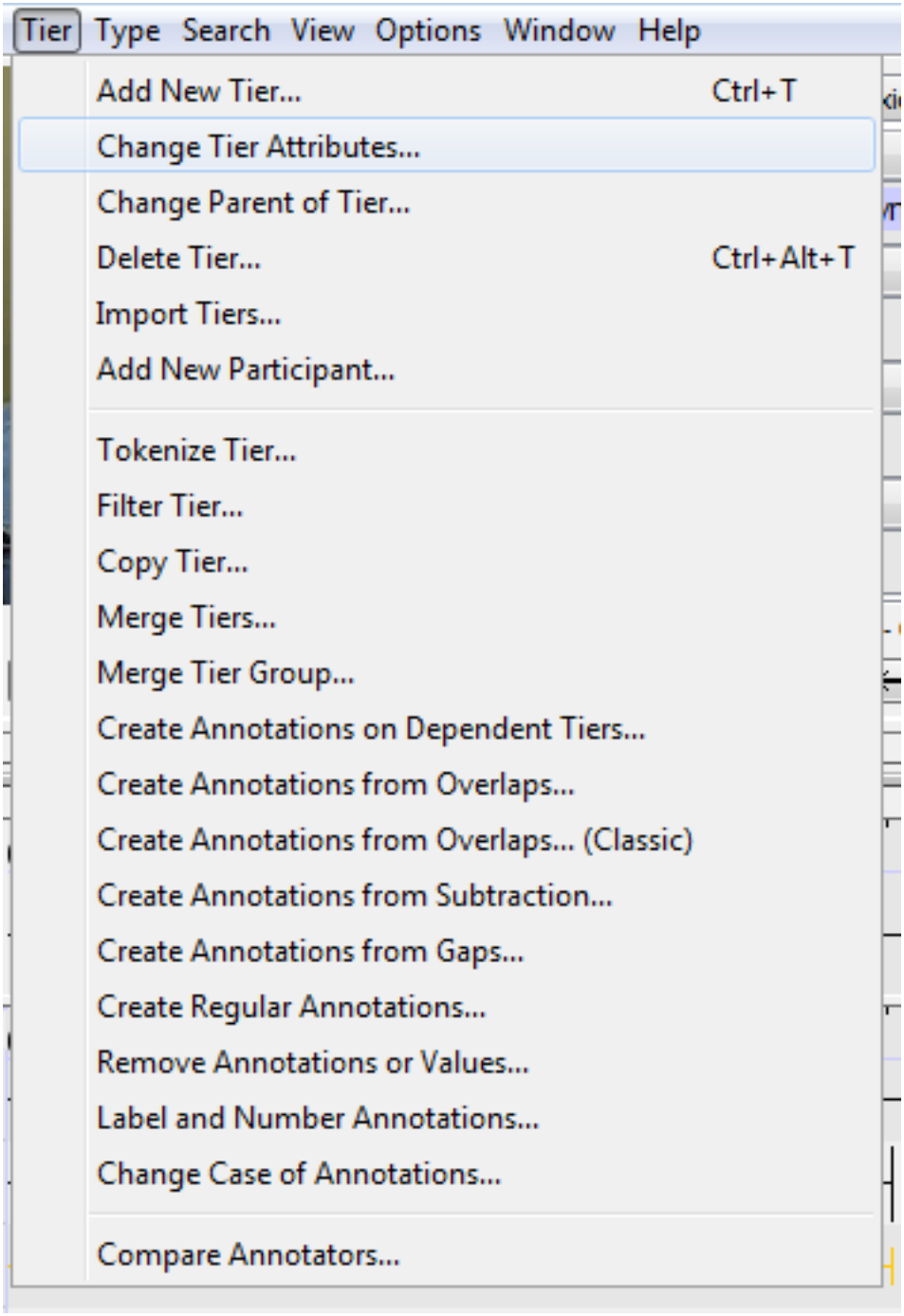


Figure 5.18. Tier menu

Besides the options in the dialog window Change tier attributes there are three more options that can be reached by clicking on the More Options... button. In the new dialog window you can change tier name color, tier highlight color (painted as background color) and tier font by clicking Browse..., selecting a color or font and clicking OK. To apply the new color and font click Apply. A window for changing tier attributes settings for multiple tiers will then appear, please see Section 5.4.1 for more details.

5.4.5. Deleting tiers

Do one of the following to delete a tier:

1. Use the mouse.

a. In the Timeline Viewer, right-click at about the height of the tier that you want to delete. A pull-down menu appears.

b. In the pull-down menu, click on Delete tier.

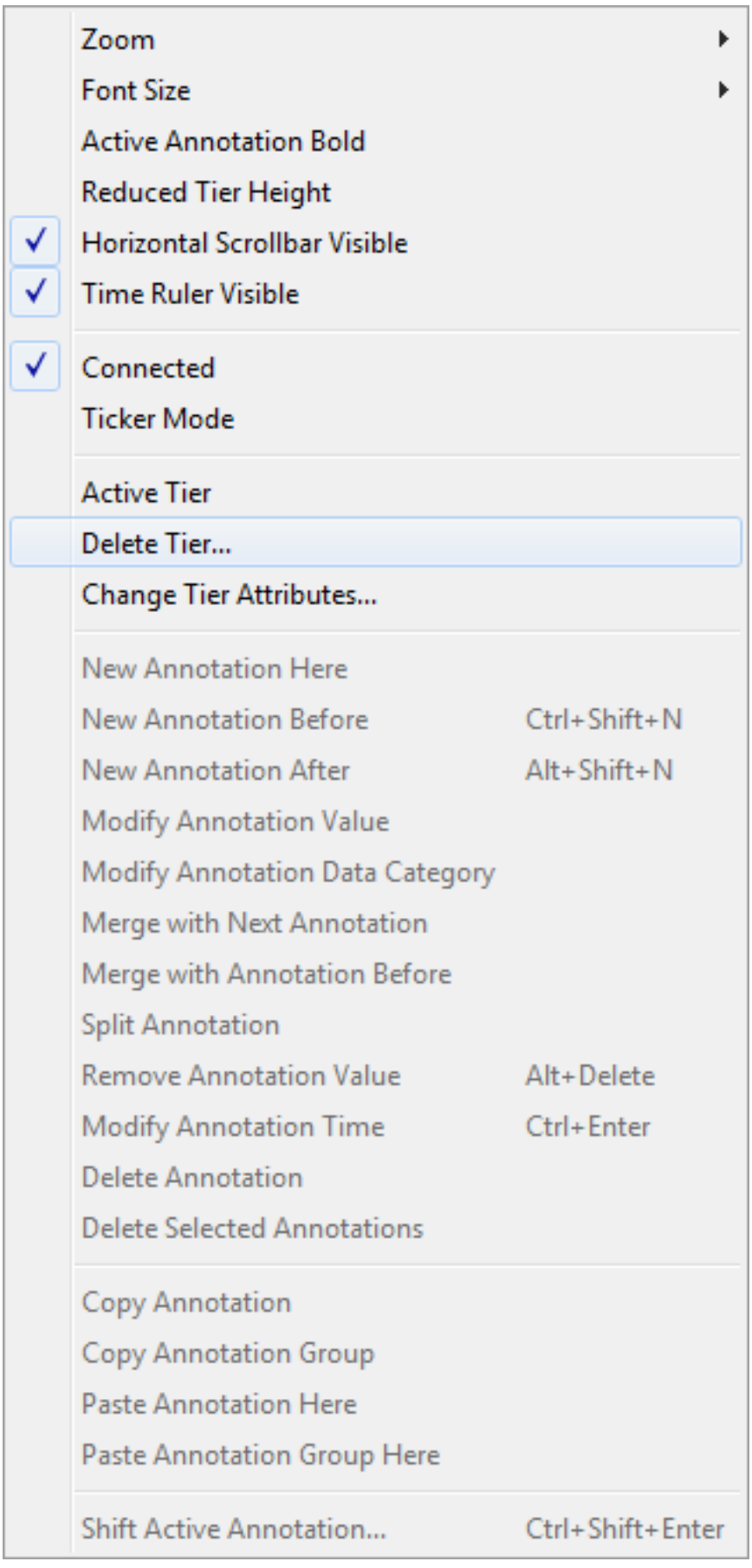


Figure 5.19. Timeline context menu

c. A warning dialog appears asking you to confirm the deletion (see below).

2. Use the Tier menu.

   a. Click on Tier menu.

   b. Go to Delete tier… A dialog window appears.

   c. The labels of all available tiers are displayed a pull down box, e.g.:

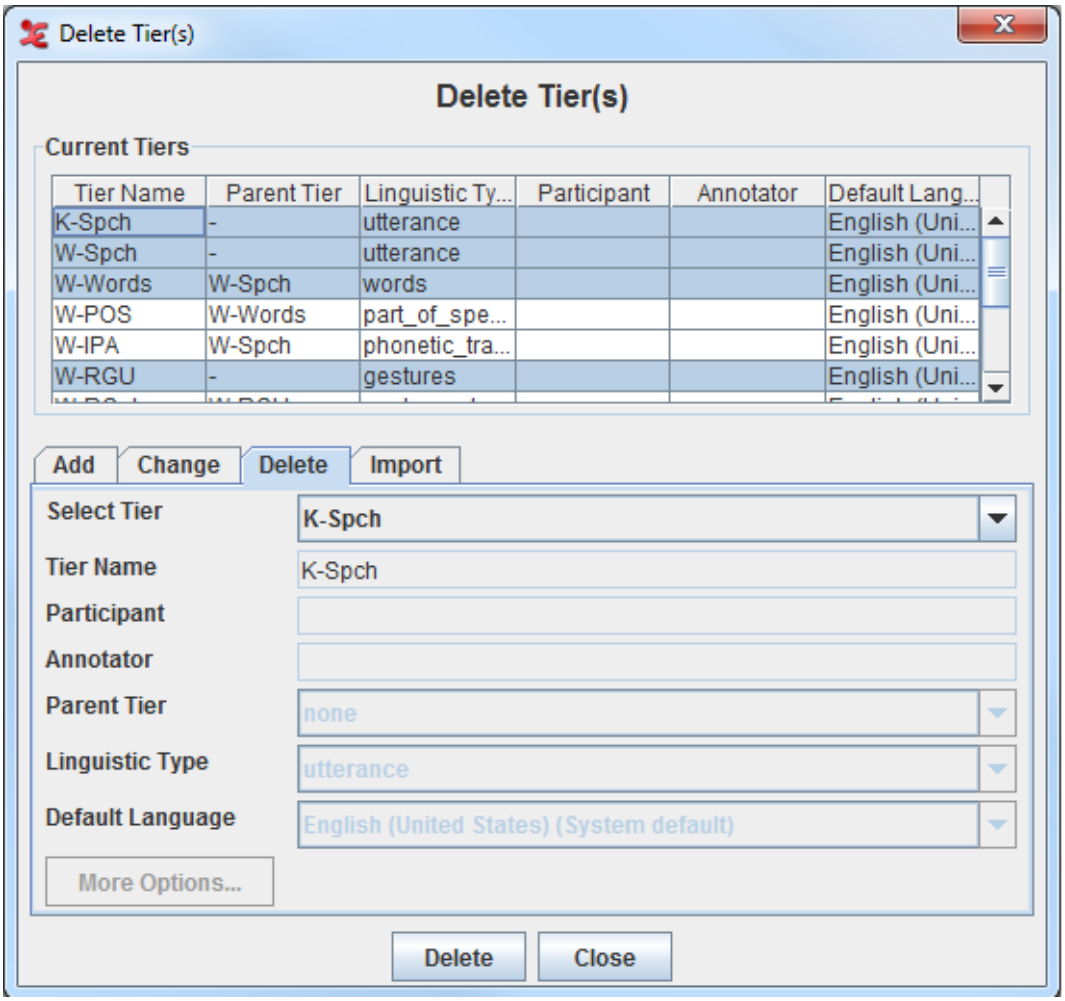


Figure 5.20. Delete tier

   d. 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.:

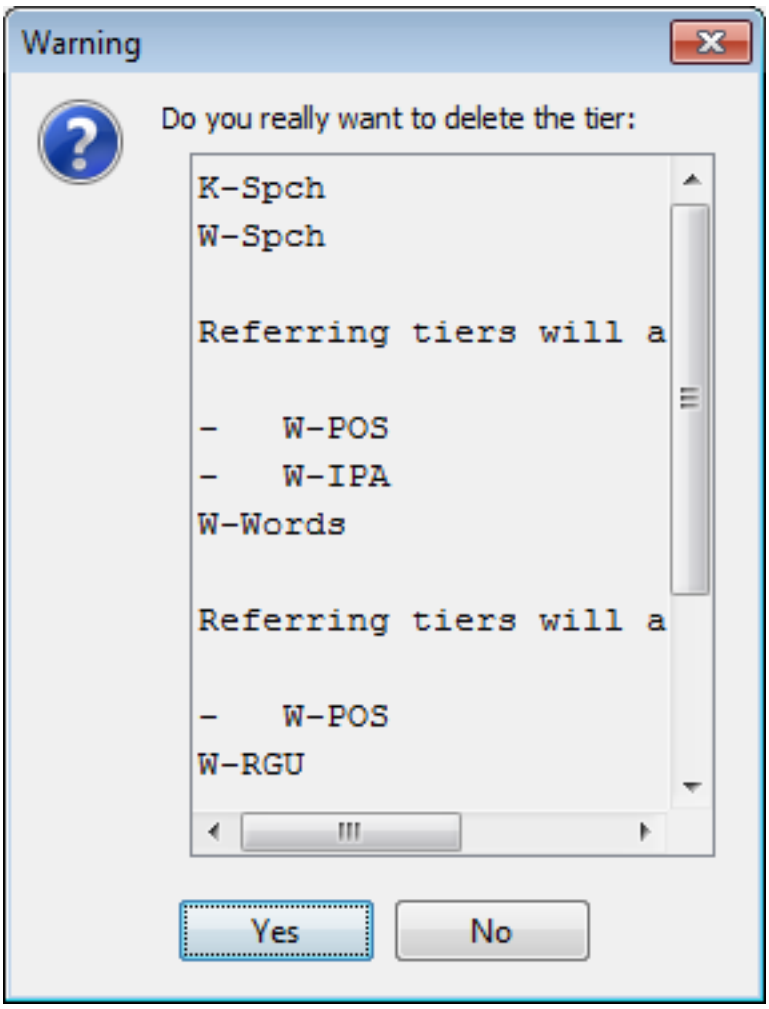


Figure 5.21. Deleting tier confirmation

e. 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 Section 5.4.8.

5.4.6. 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.... 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. Enter a name for the new tier and select the desired linguistic 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 linguistic 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 Section 5.15.2.)

As a final example, consider an audio recognizer (see Section 5.4.3) 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.

5.4.7. 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.

5.4.8. Changing the parent of a tier

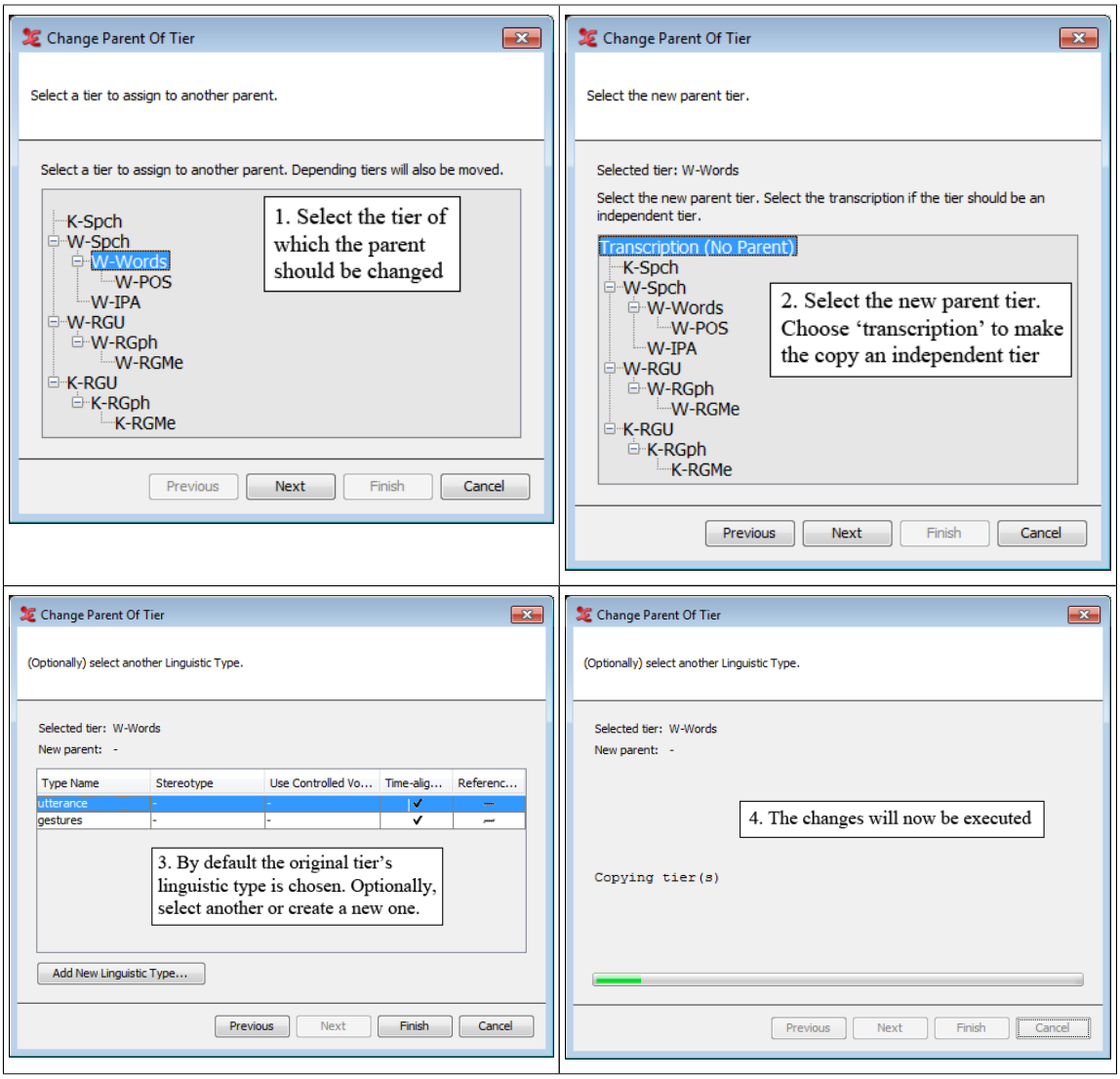
Previous versions of ELAN offered an option to change the parent of a tier. As this operation could cause data loss when not performed carefully, this function has been disabled. However, what can be done now as an alternative is making a copy of a tier and altering that copy. This prevents the possible loss of data, as the original tier stays the same. At the same time the linguistic 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 5.3. 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:

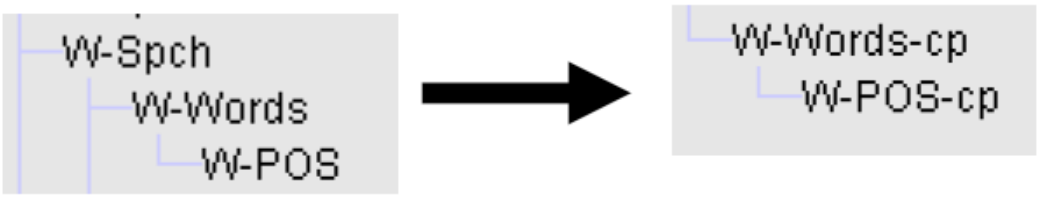


Figure 5.22. Dependent to independent tier

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.

5.4.9. 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 > Add New Participant.... The following *Add new participant* dialog window appears:

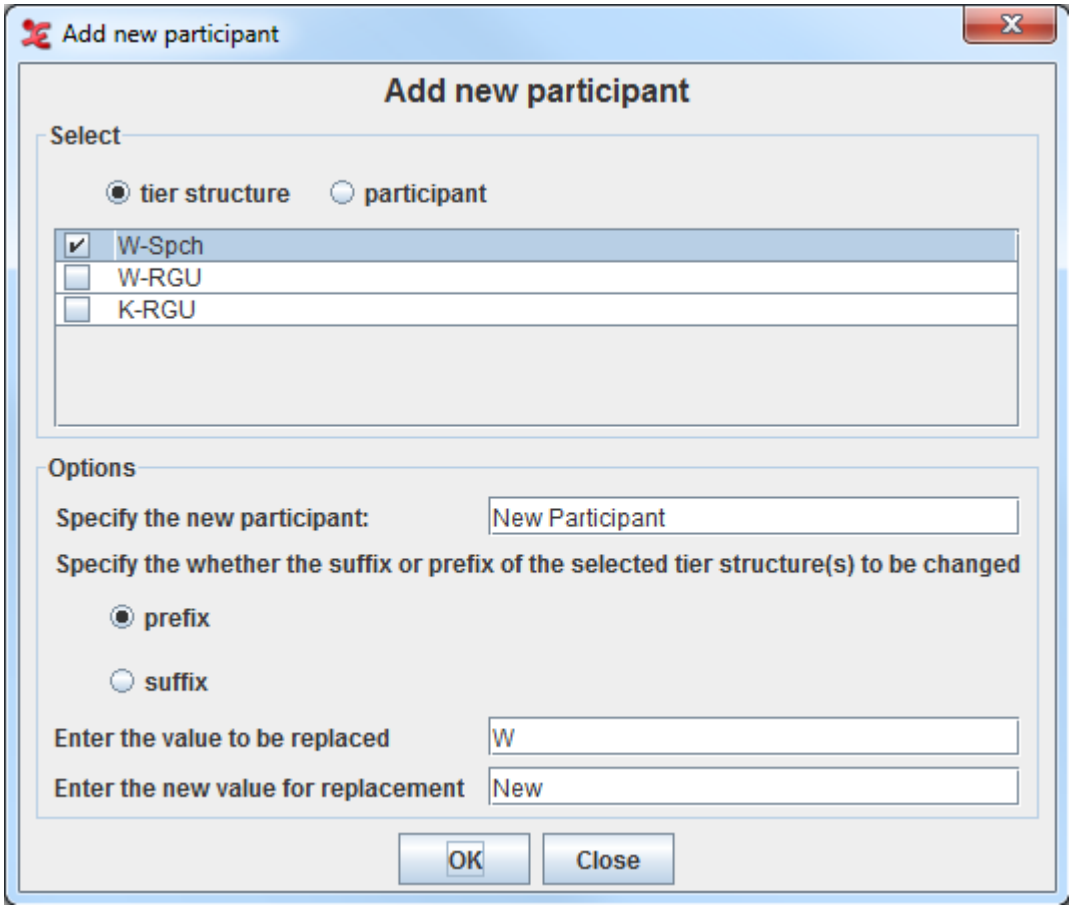


Figure 5.23. Add New Participant

The new set of tiers for the new participant is either based on an existing tier with all its depending tiers or on all tiers with a specific participant attribute. Enter the name for the new participant in the appropriate field. Then, specify whether the suffix or prefix of the selected tier to be changed. Finally, enter the selected tier value you want to replace and the new value for the replacement. The latter will replace the prefix or suffix (it depends on the choice you have made) of the existing tiers and the participant attribute will be set to the new participant.



Note

- The new prefix/suffix value cannot be null.
- The value to be replaced can be null. 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.

5.5. Controlled Vocabularies

5.5.1. The creation of a CV

When you frequently use a certain linguistic 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:

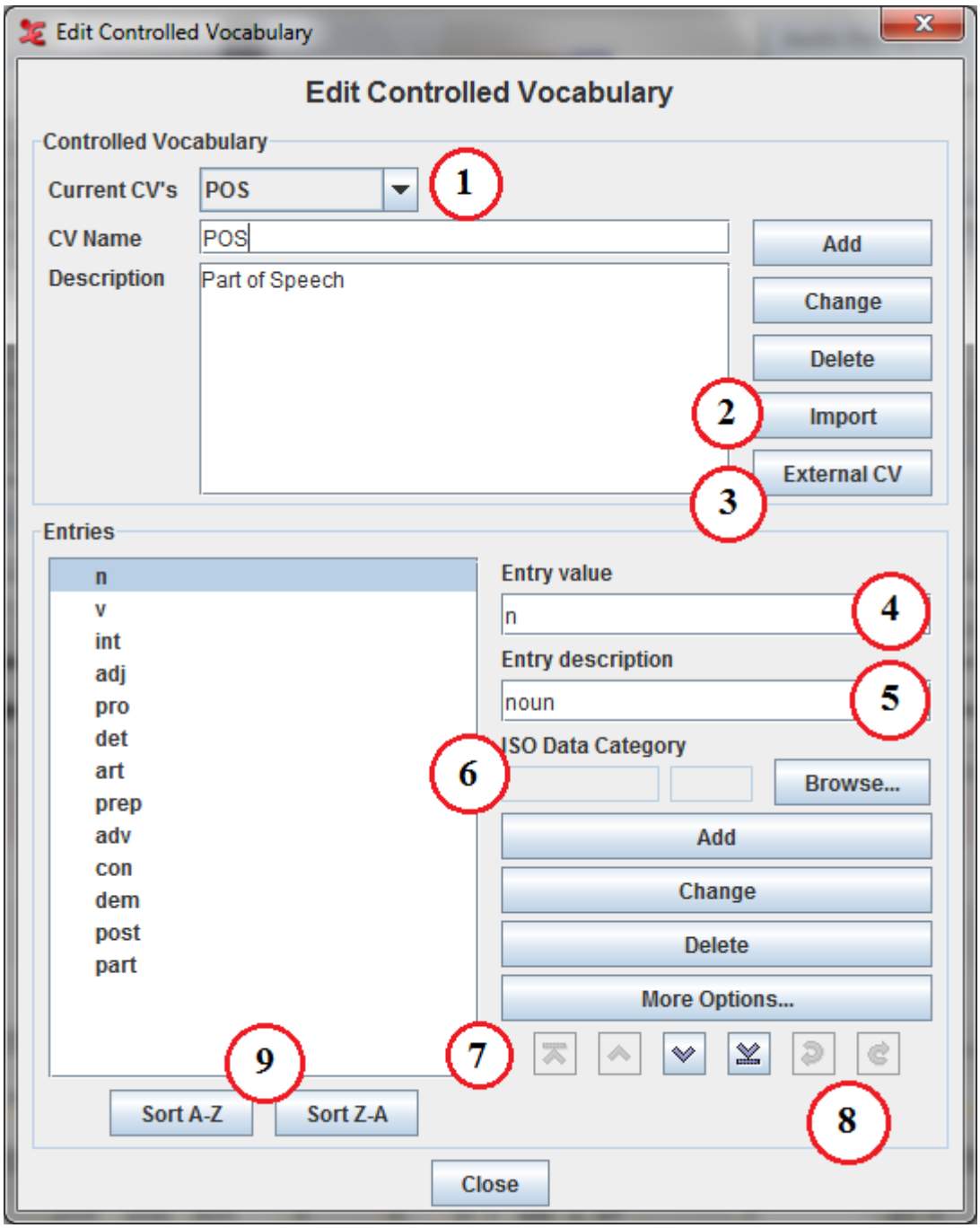


Figure 5.24. Edit controlled vocabulary

1. *The CV being edited*
2. *Import option to retrieve a CV from a template*
3. *Option to link to an External CV*
4. *Annotation value as it is going to be filled in*
5. *A loose description of the value*
6. *Association with a ISO Data Category*
7. *Buttons to move the selected entry in the list (respectively to the top, one place up, one place down and to the bottom).*
8. *Undo/redo changes to the CV*
9. *Buttons to sort the entries in the list*

To create a new CV, do the following:

1. Enter a CV Name and a description
2. Click on the Add button
3. Now enter each CV entry and its description.
4. Confirm every entry addition by clicking on the Add button or by hitting Enter



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.

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.

5.5.2. Associating a CV entry with a Data Category

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 (Figure 5.25), first select a profile in the left panel and then select a data category in the middle panel. Finally click on Apply to associate the selected data category to the CV entry.

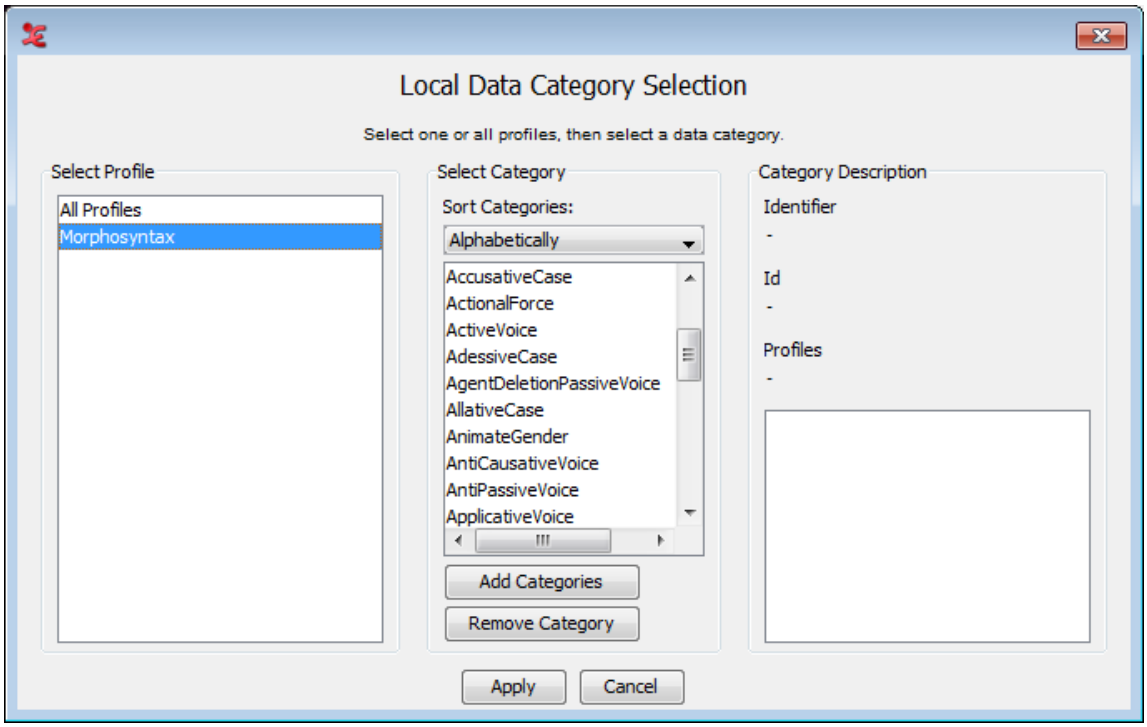


Figure 5.25. Local Data Category Selection

More information about the ISO DCR and how to use it can be found in Section 5.2.

5.5.3. 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 Section 4.2.1) 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 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:

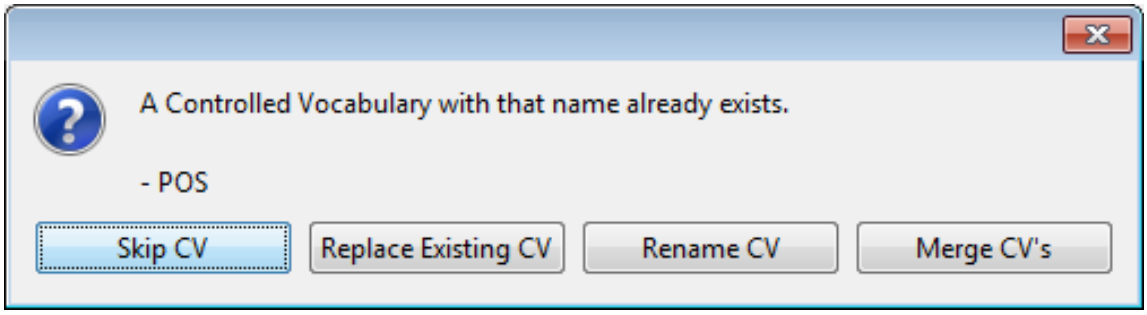


Figure 5.26. Controlled vocabulary already exists warning

- 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.

5.5.4. 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 5.26).

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.

5.5.5. Exporting a CV

A CV can be exported by saving the current file as a template. (See Section 4.2.8)

5.5.6. Using CV's

When editing an annotation that belongs to a linguistic type associated with a CV, a drop down box appears, the suggest panel. Just select the value to be filled in:

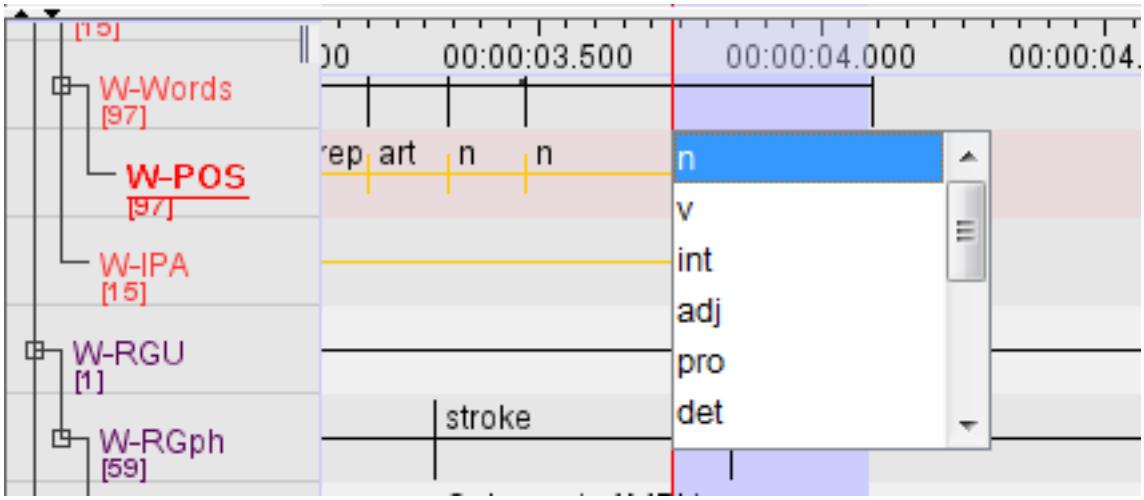


Figure 5.27. Controlled vocabulary drop down box



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 Section 5.5.2), the annotation is also associated with that data category.

More information about entering annotations from a CV and working with the Suggest Panel, see section 5.8.3.

5.5.7. Associating a linguistic type with Controlled Vocabularies

Open the Change the Linguistic type dialog (via Type > Change linguistic type) and make a choice from the Use Controlled Vocabulary dropdown box:

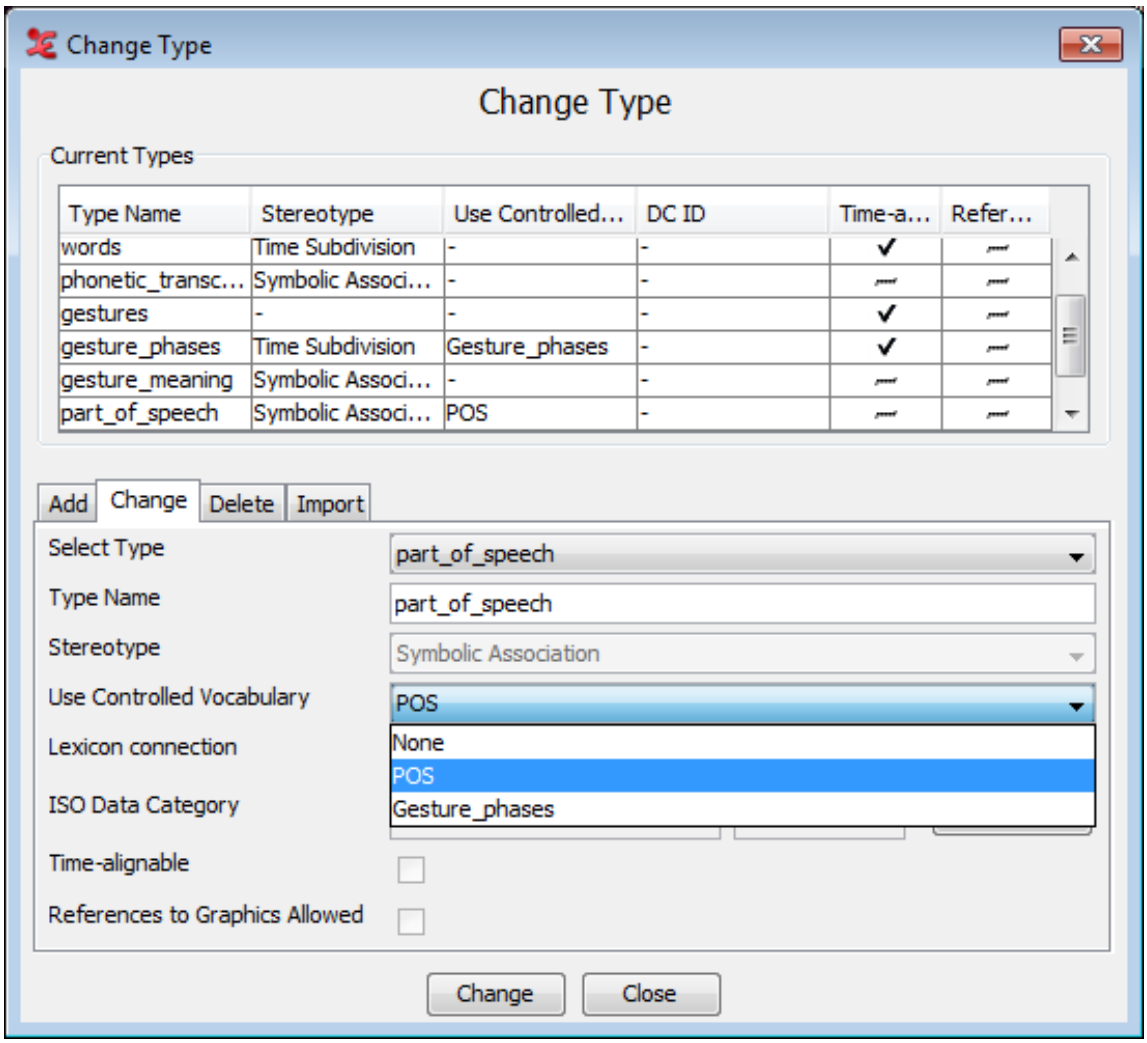


Figure 5.28. Change linguistic type

5.6. Consulting a Lexicon

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 lexica through the following steps:

1. Set up a new Lexicon Service (see Section 5.6.1).
2. Enhance a Linguistic Type so that it's tiers can do a lexicon lookup (sec Section 5.6.2).
3. Perform the lookup (see Section 5.6.3).

5.6.1. 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 lexica. 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

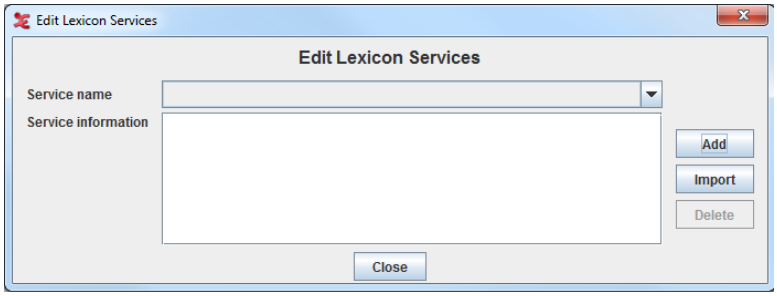


Figure 5.29. Edit Lexicon Services



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 http://tla.mpi.nl/tools/tla-tools/elan/.

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.

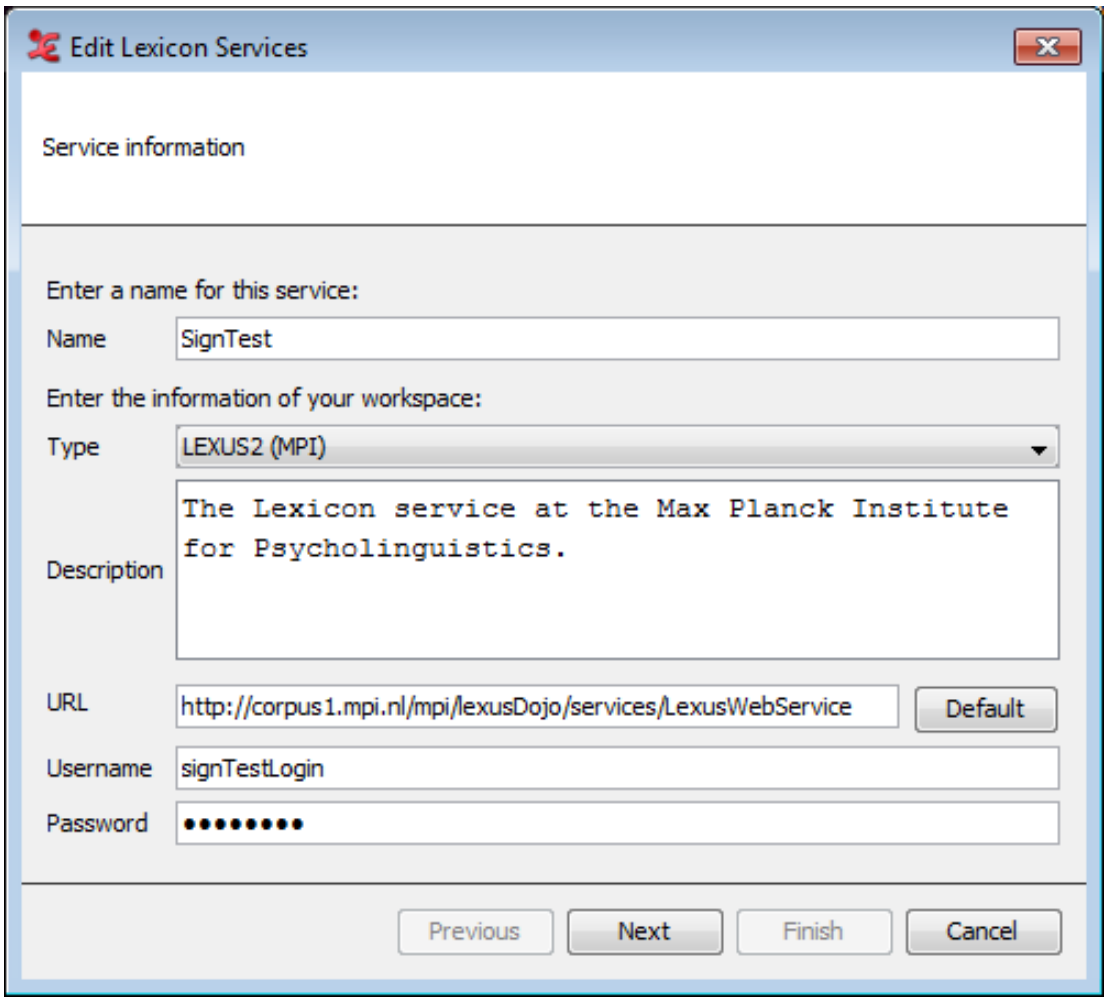


Figure 5.30. Name and Lexicon Service Info

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.

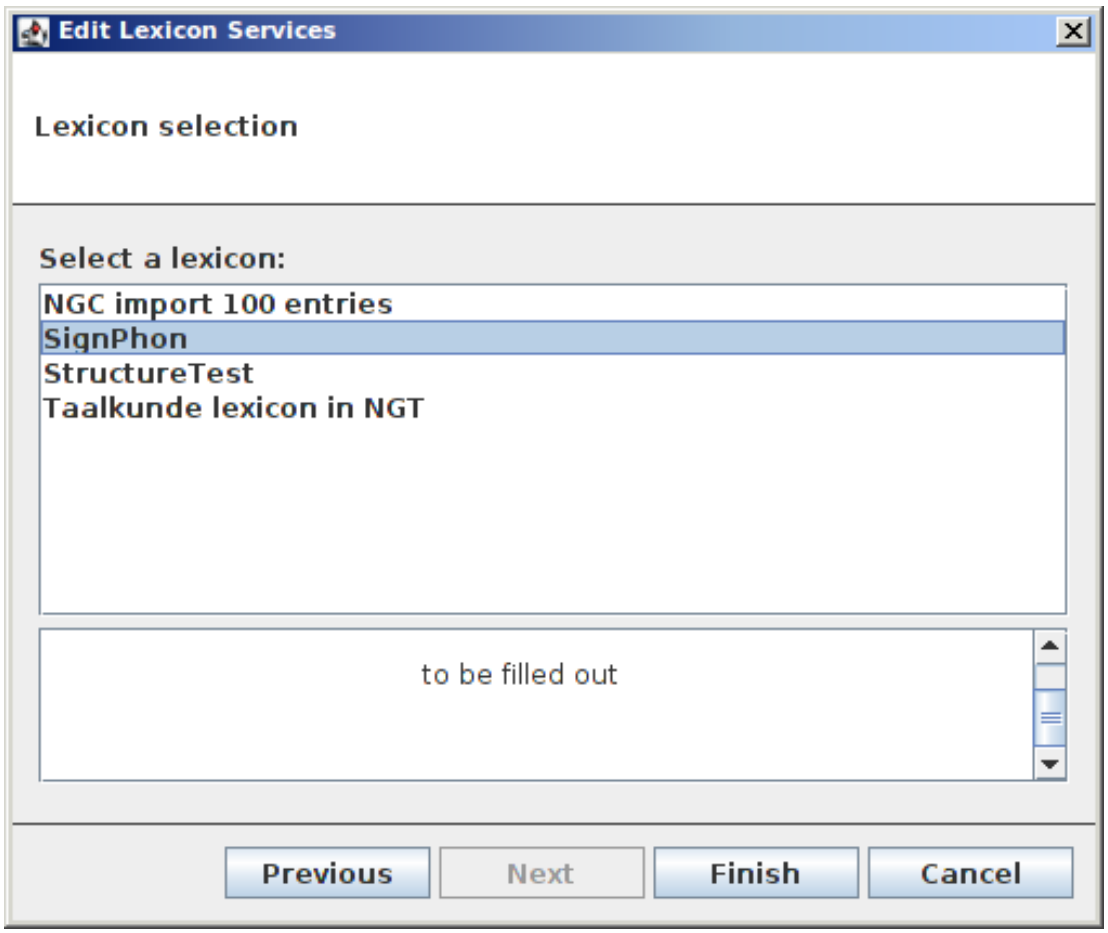


Figure 5.31. Lexicon Selection

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:

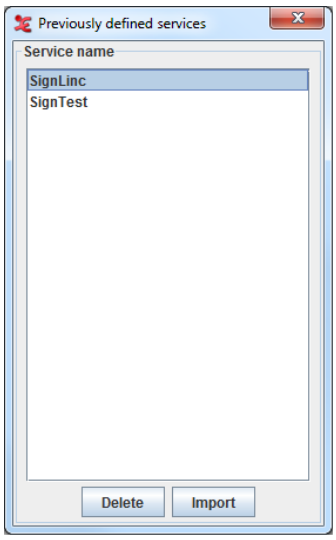


Figure 5.32. Previously defined services

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).

5.6.2. Set the Lexicon Service and entry field information for a Linguistic Type

Before an annotation can be used for a lexicon lookup, the linguistic 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 linguistic type, do the following:

1. In the main menu select Type > Change Linguistic Type...
2. Select the a linguistic 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

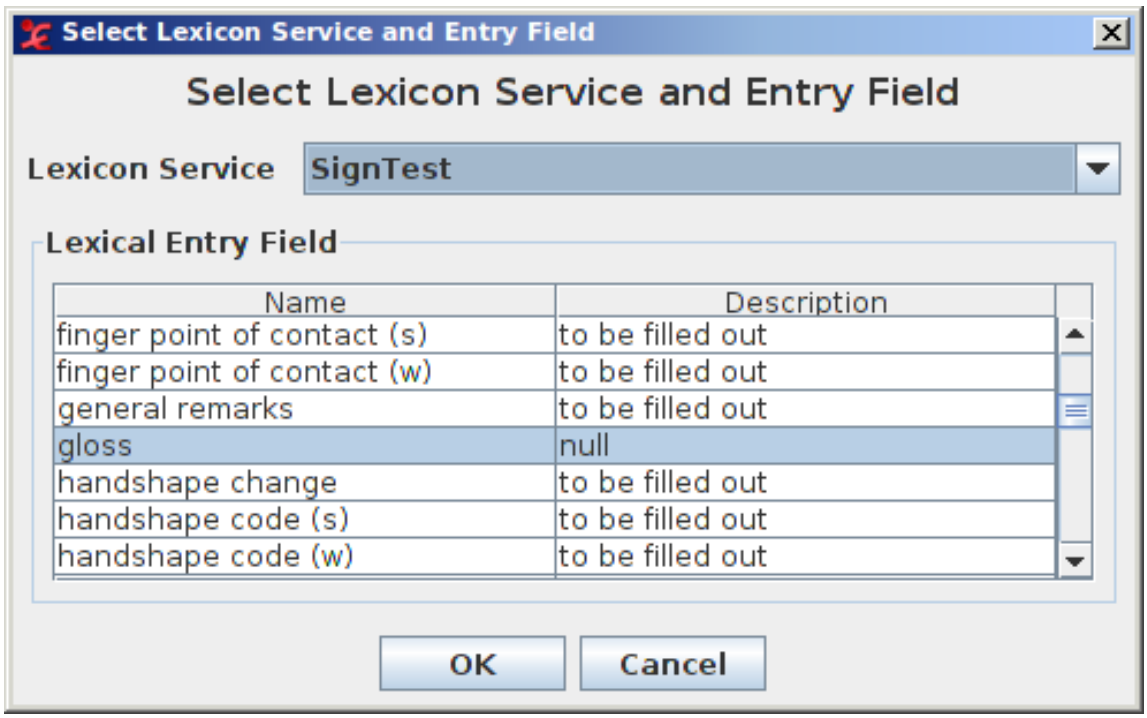


Figure 5.33. Select a Lexicon Service and Entry Field

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.

5.6.3. Perform a Lexicon Lookup

If the linguistic 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 linguistic 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.

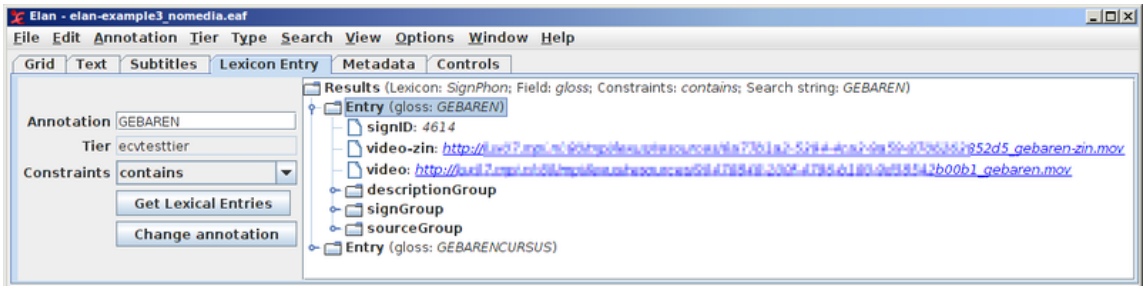


Figure 5.34. The Lexicon Entry tab

5.7. 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 (Section 5.7.1);
- making and saving a selection on an independent tier while playing (Section 5.7.2);
- making and saving a selection on a referring tier (Section 5.7.4);
- deselecting a selection (Section 5.7.5);
- changing the time alignment of an existing selection (Section 5.7.6);
- activating and deactivating the Bulldozer mode (Section 5.7.8).

5.7.1. 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.

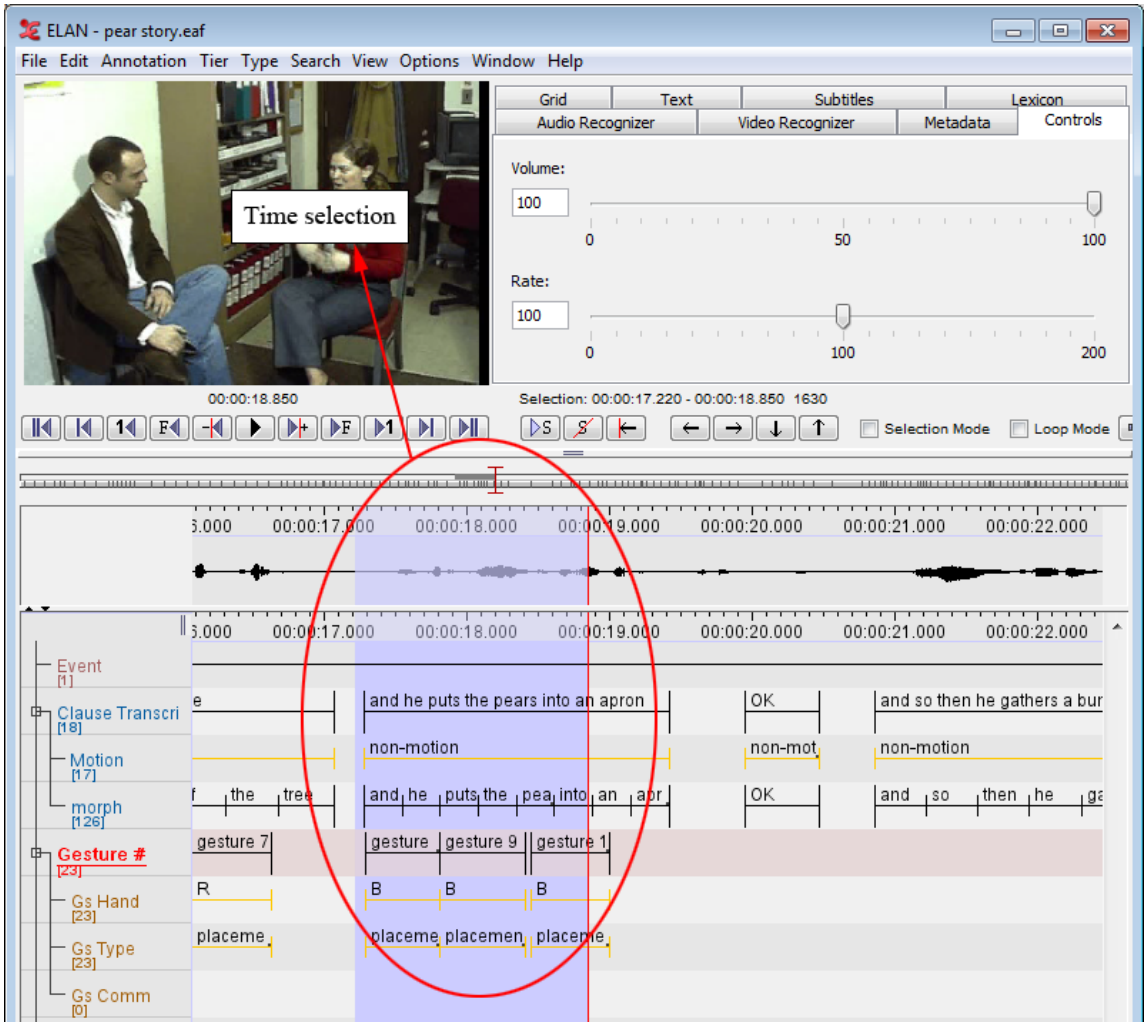


Figure 5.35. Making a selection

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 Section 4.4.14.

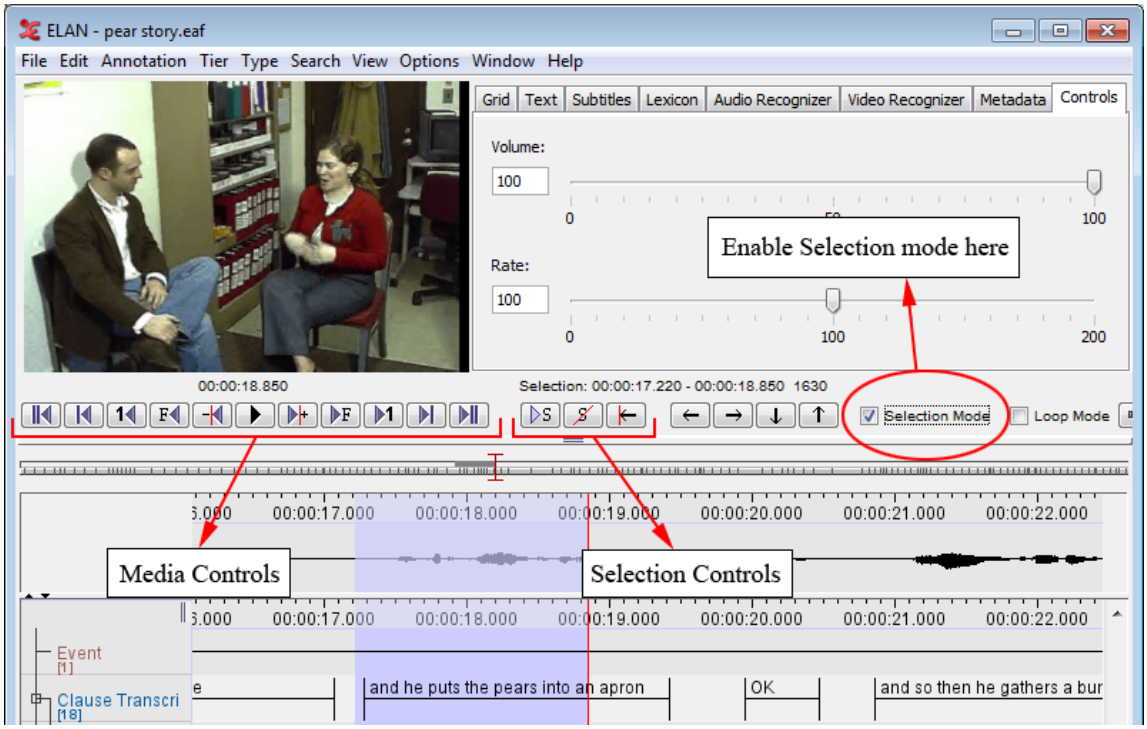


Figure 5.36. Change selection

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.

5.7.2. Using the selection controls

The selection controls allow you to navigate through or to change the active selection. For their use, see again Section 4.4.14.

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

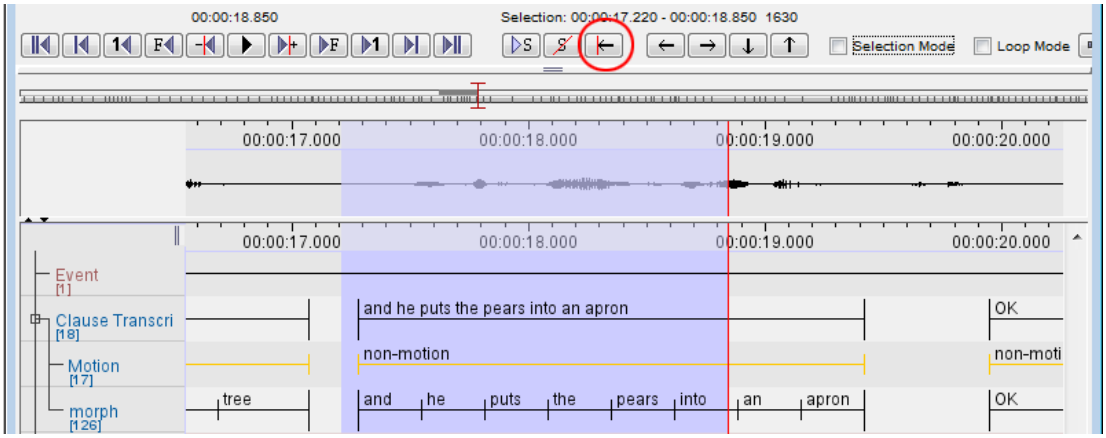


Figure 5.37. Crosshair on begin of selection

2. Enable Selection Mode

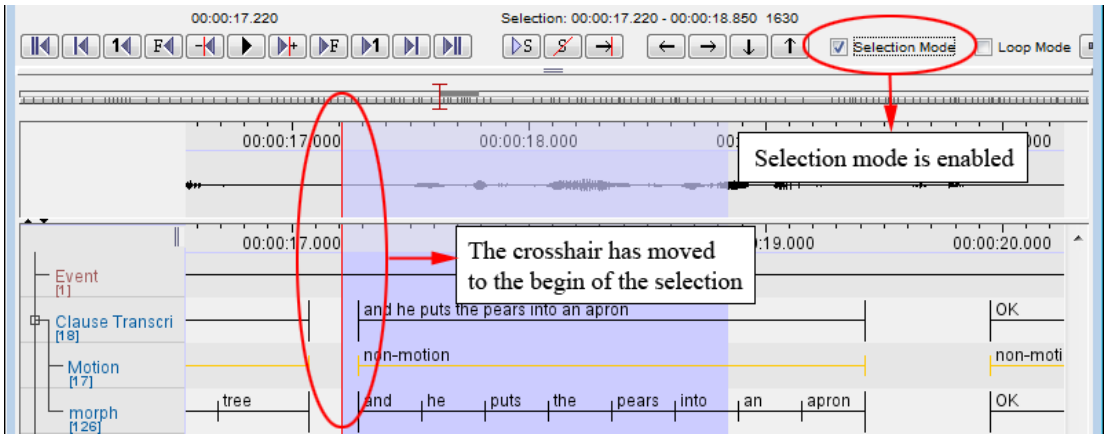


Figure 5.38. Selection mode

3. Go back one second by clicking the corresponding button from the media controls.

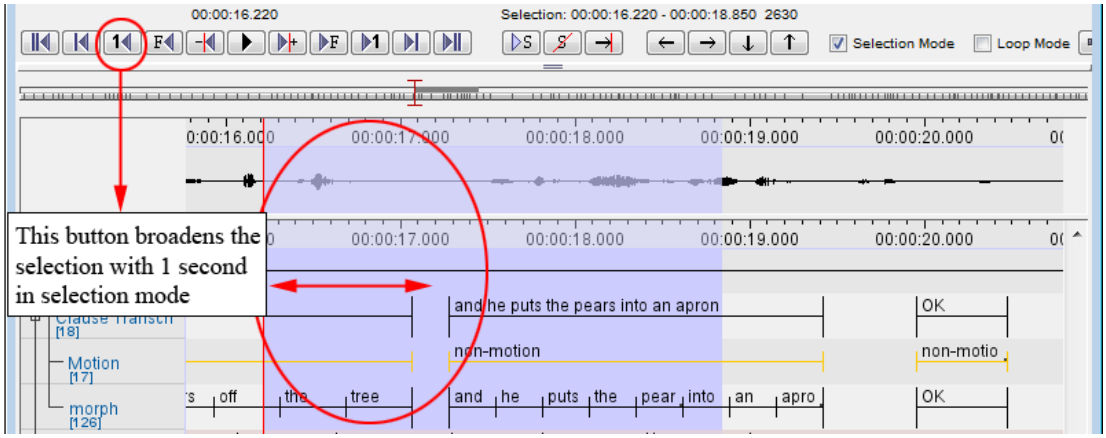


Figure 5.39. Back one second

4. Turn off the selection mode and enter an annotation for the selection.

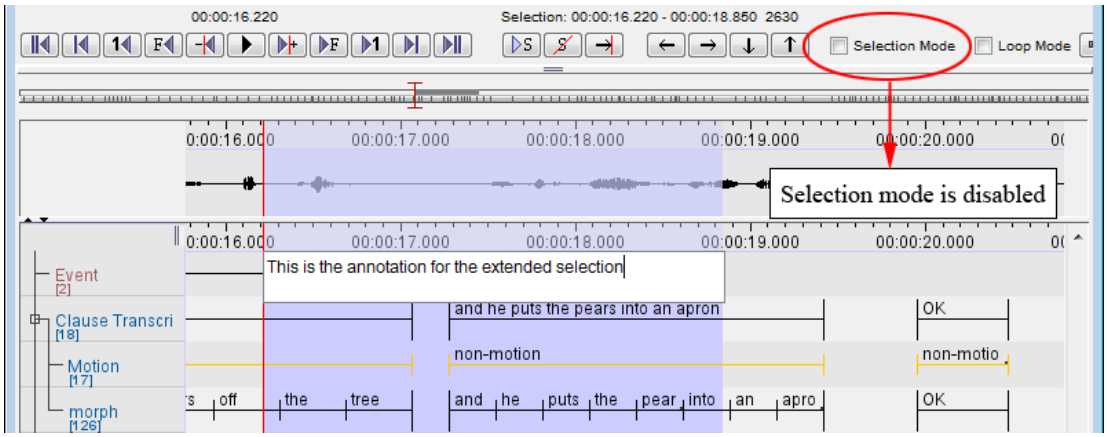


Figure 5.40. Selection mode off

5.7.3. 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 Section 5.7.1).

2. Double-click on the selection.

3. Optionally enter the content of the annotation unit. Press the keys CTRL+ENTER. The selection is saved.

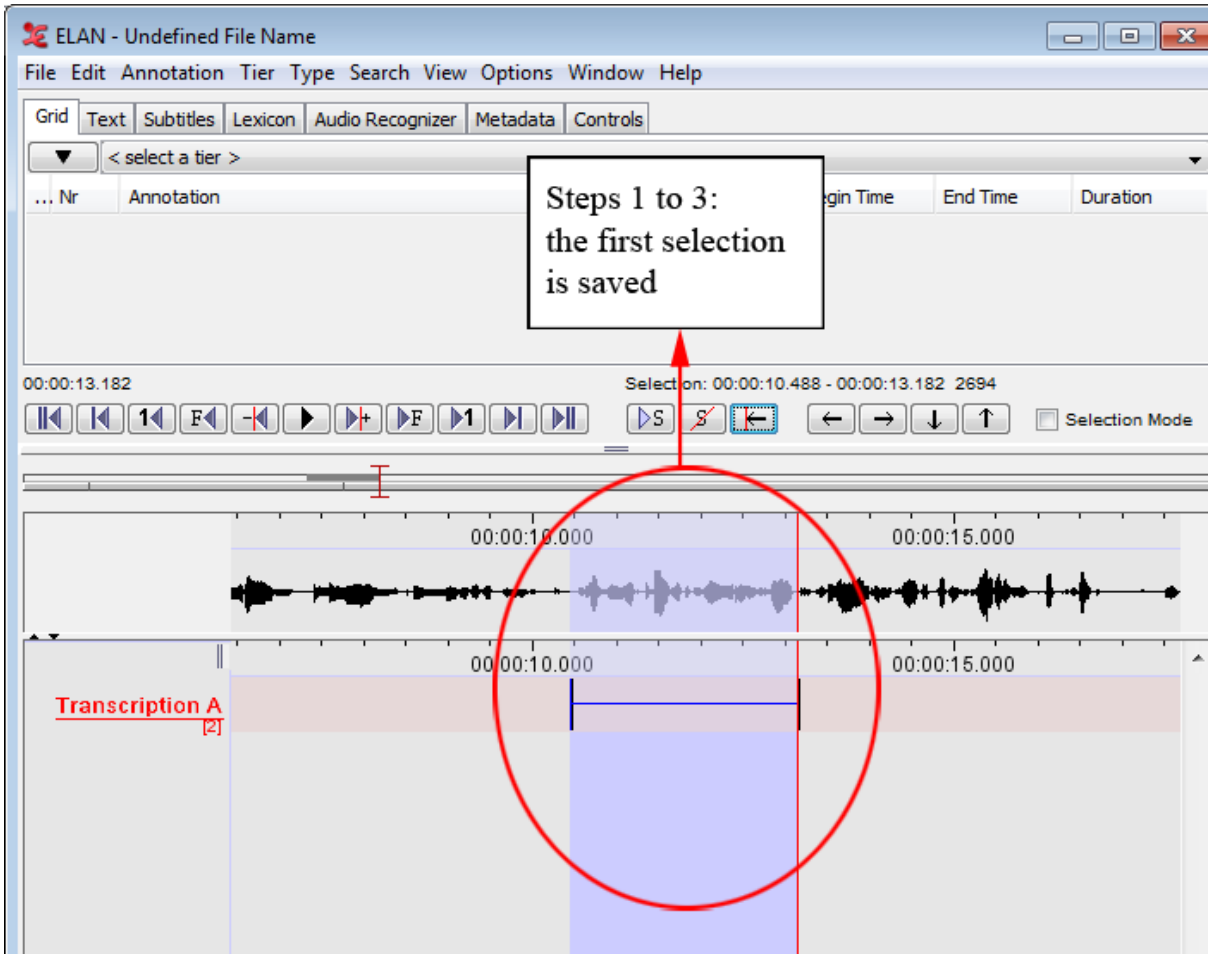


Figure 5.41. Annotation without gaps (steps 1-3)

4. Press the keys ALT+SHIFT+C (or ALT+C) or click on the clear selection icon to deselect the selection (see Section 5.7.5 for deselecting a selection).

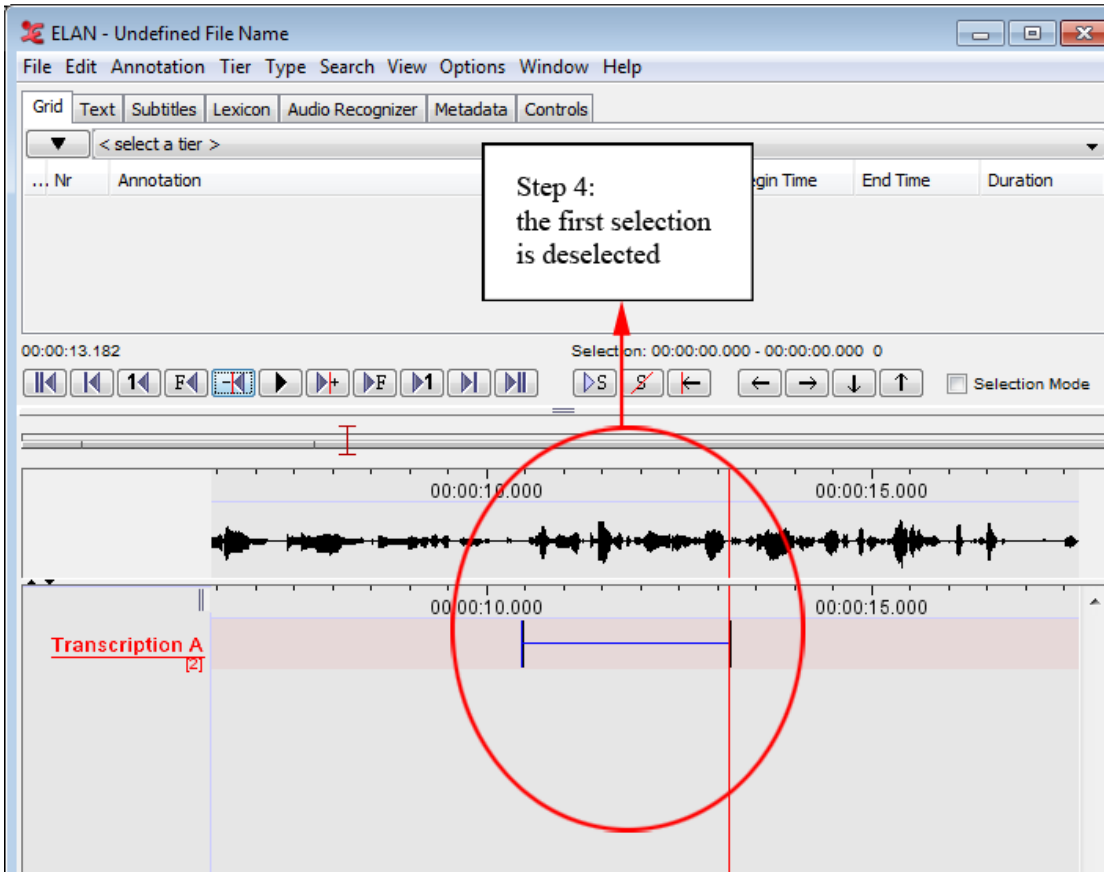


Figure 5.42. Annotation without gaps (step 4)

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.

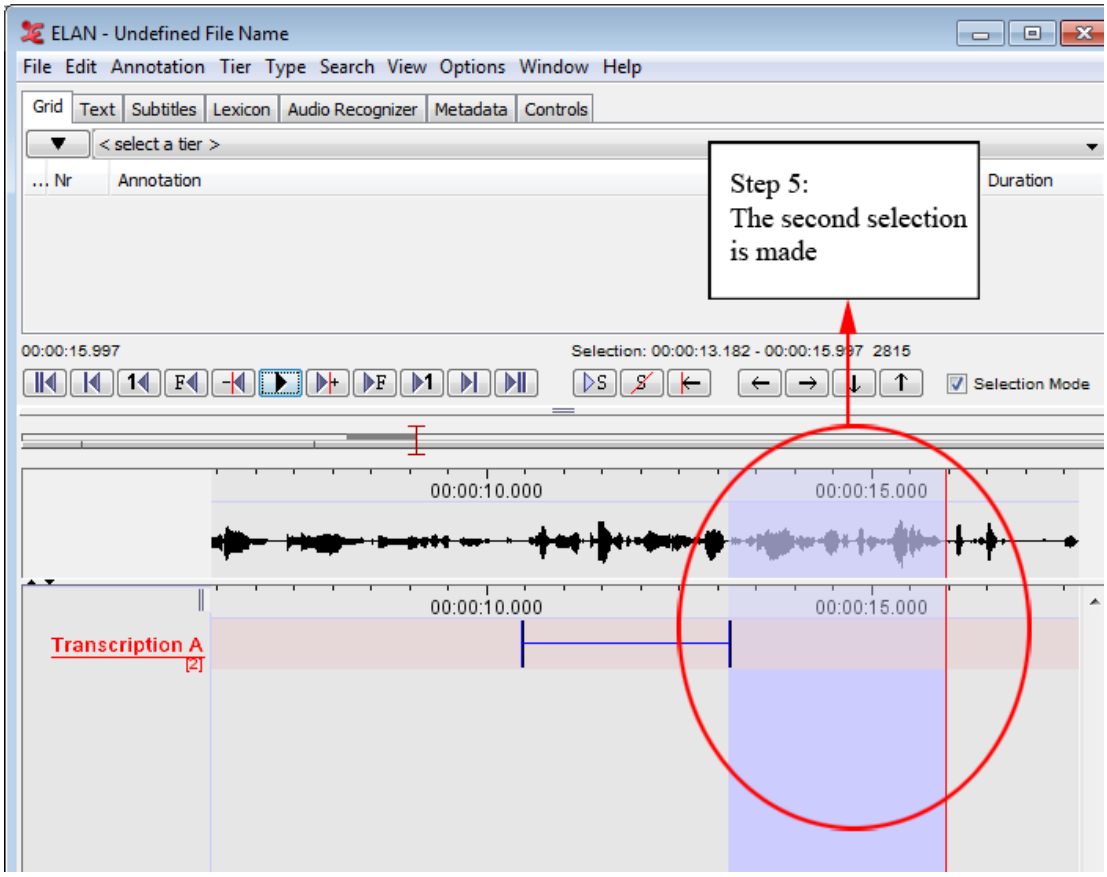


Figure 5.43. Annotation without gaps (step 5)

5.7.4. 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 Section 5.7.1 and Section 5.8).

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.

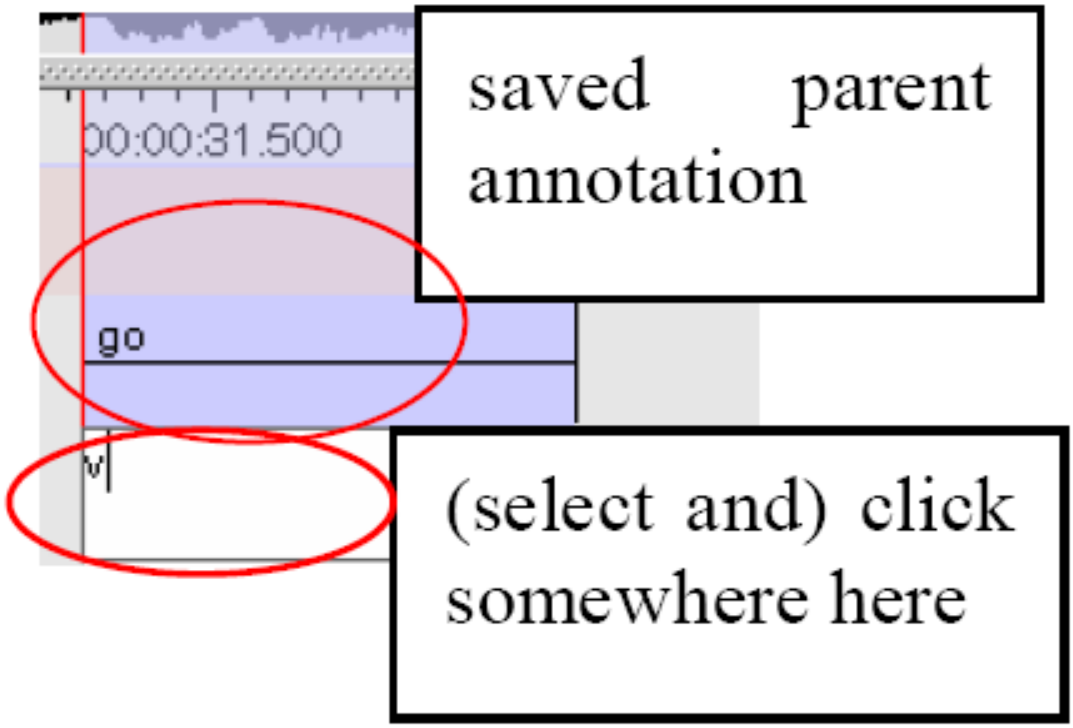


Figure 5.44. A selection on referring tier

3. Do one of the following:
   a. Enter an annotation (see Section 5.8), 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.

5.7.5. 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 Section 5.7.1).



Note

Whenever you select another time interval, the old selection is automatically deselected, unless you enabled Selection Mode.

5.7.6. 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.

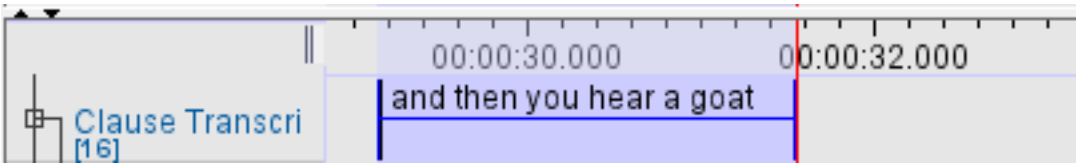


Figure 5.45. Select annotation unit

2. Select the region where you want the modified annotation to be placed.

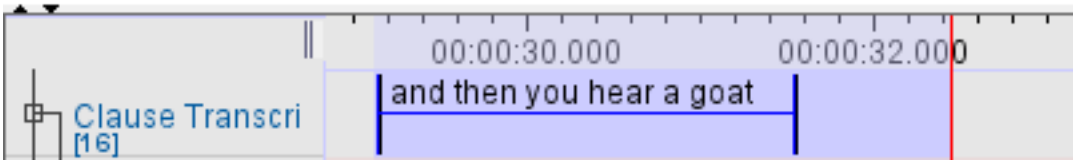


Figure 5.46. Region for modified annotation

3. Right click on the original annotation and select Modify annotation time or press CTRL+ENTER

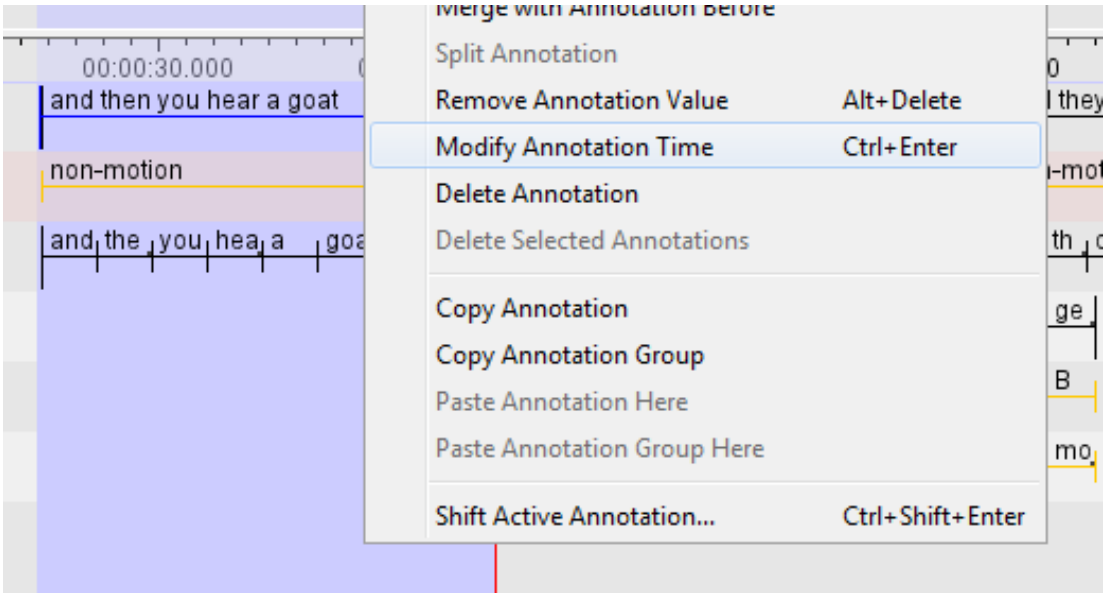


Figure 5.47. Modify annotation time

4. Now the length of the annotation becomes that of the selection from the second step.

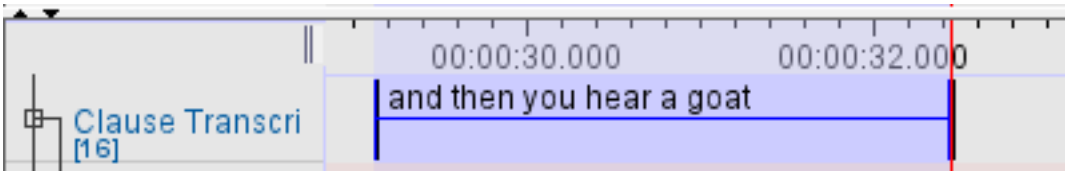


Figure 5.48. Annotation length changed

By dragging 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

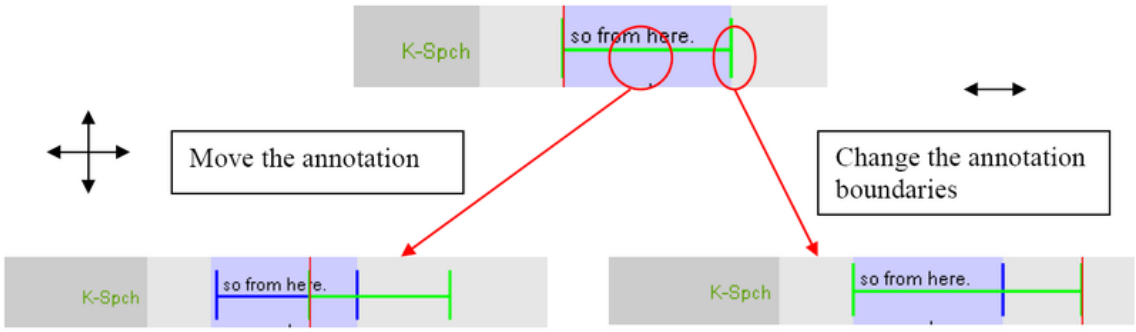


Figure 5.49. Changing selection boundaries by mouse



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 Section 5.1).
- 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 Section 5.7.7.
- If two annotations are adjacent you can snap them by specifying the maximum close-value in ms.

By entering a number of milliseconds

Annotations can be shifted to the left or the right by a number of milliseconds. See Section 5.7.7.

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.

5.7.7. 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 time 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.

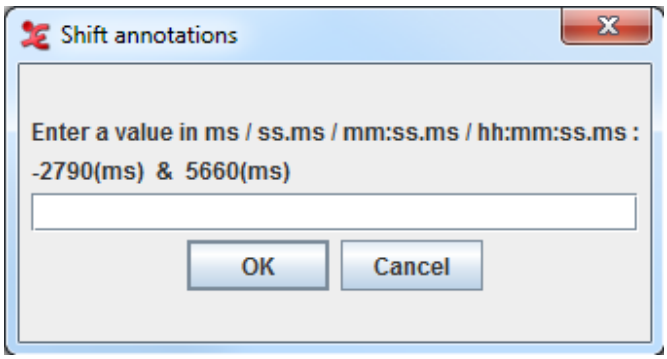


Figure 5.50. Shift annotations

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 5.50. 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.

5.7.8. 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.

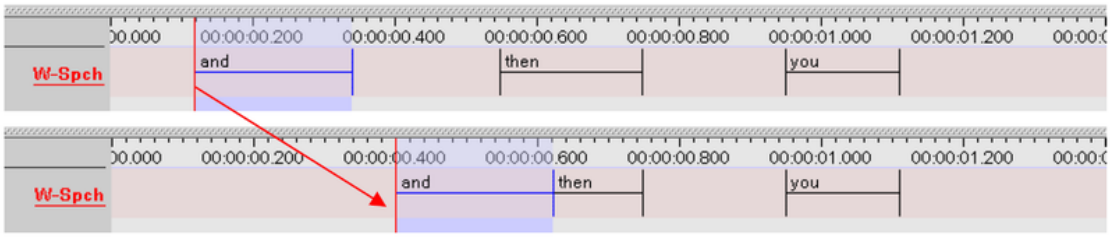


Figure 5.51. Normal mode

- 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.

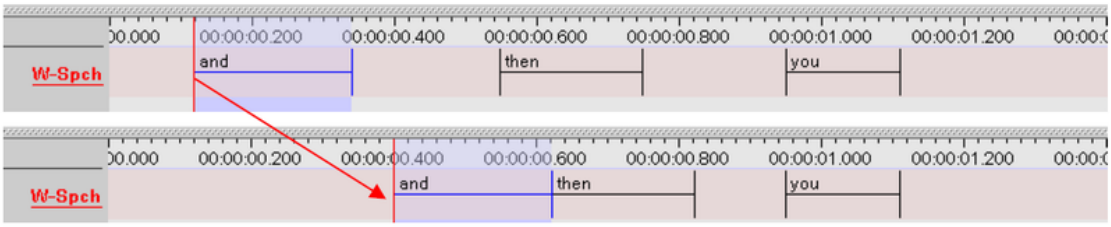


Figure 5.52. Bulldozer mode

- 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 Section 4.2.3).

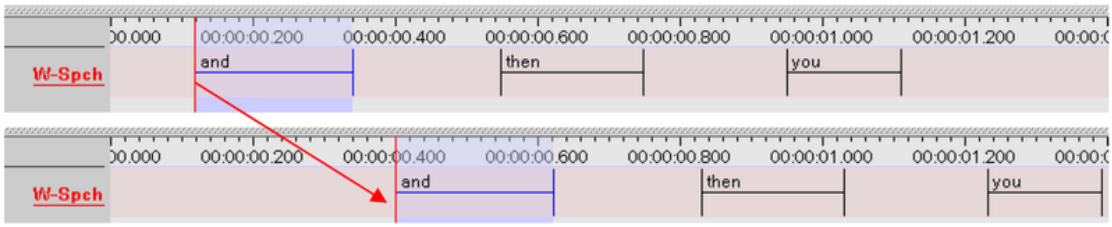


Figure 5.53. Shift mode

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 checkmark appears next to it. To switch back to the Overwrite mode, repeat steps 1 and 2 above. The checkmark disappears.

5.8. How to enter 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 Section 5.8.1 and Section 5.8.2);
- entering annotations before/after other annotations (Section 5.8.4);
- modifying the content of annotations (Section 5.8.12);
- deleting annotations (Section 5.8.16);
- entering annotations in different character sets (Section 5.8.20).

5.8.1. Entering annotations (into the Inline Edit box)

5.8.1.1. 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.:

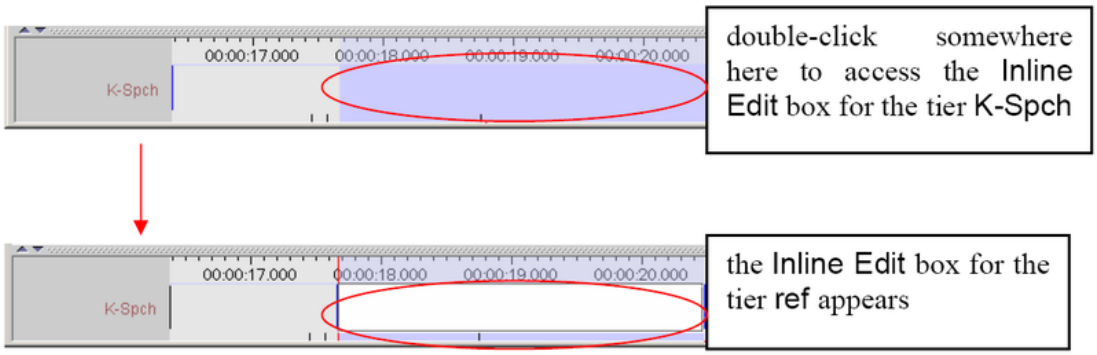


Figure 5.54. Inline Edit box

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.

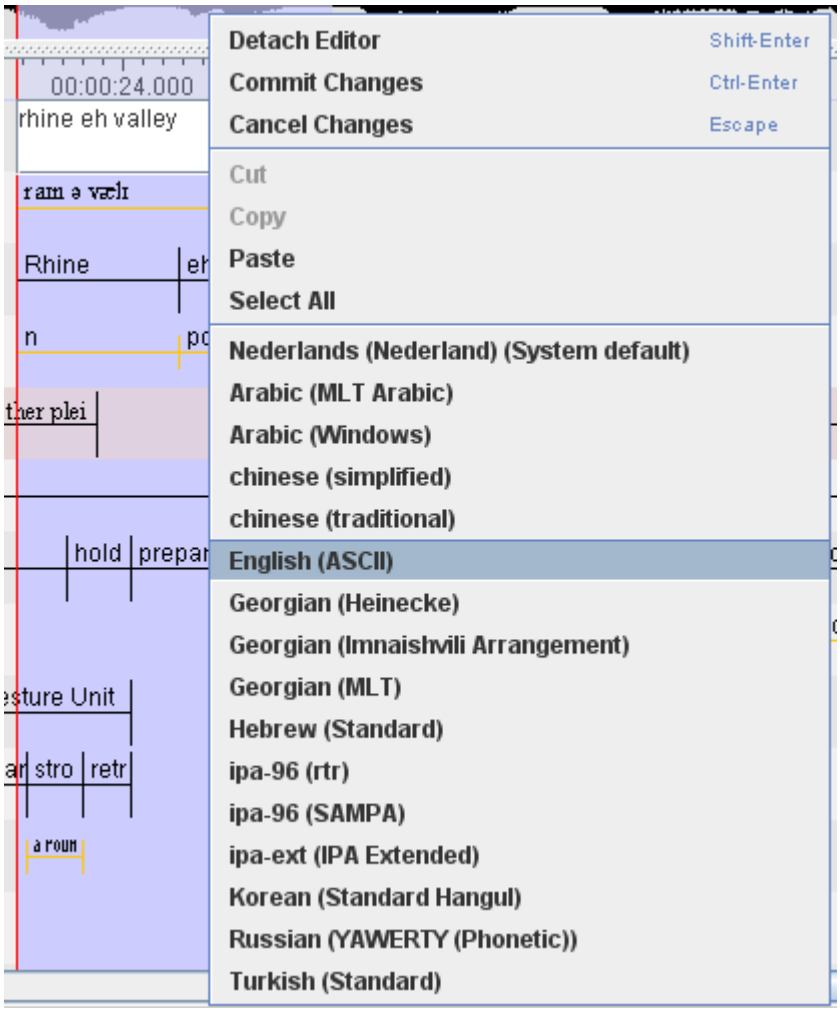


Figure 5.55. Select character set

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 Section 5.7.4.

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 Section 4.4.28

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 Section 4.4.28)

5.8.1.2. 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.

5.8.1.3. 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 Shift+E.

5.8.2. 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 Section 5.7.1), or click on an existing annotation in the Timeline or Interlinear Viewer.
2. Access the Inline Edit box (see Section 5.8.1).
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:

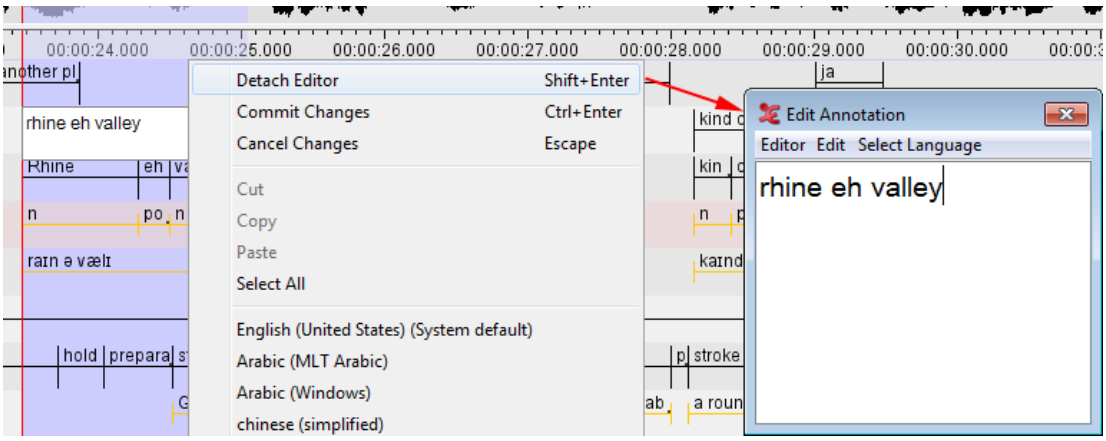


Figure 5.56. Edit annotation

The Edit Annotation box is automatically preconfigured for the default character set of the tier (see Section 5.4.4). 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.:

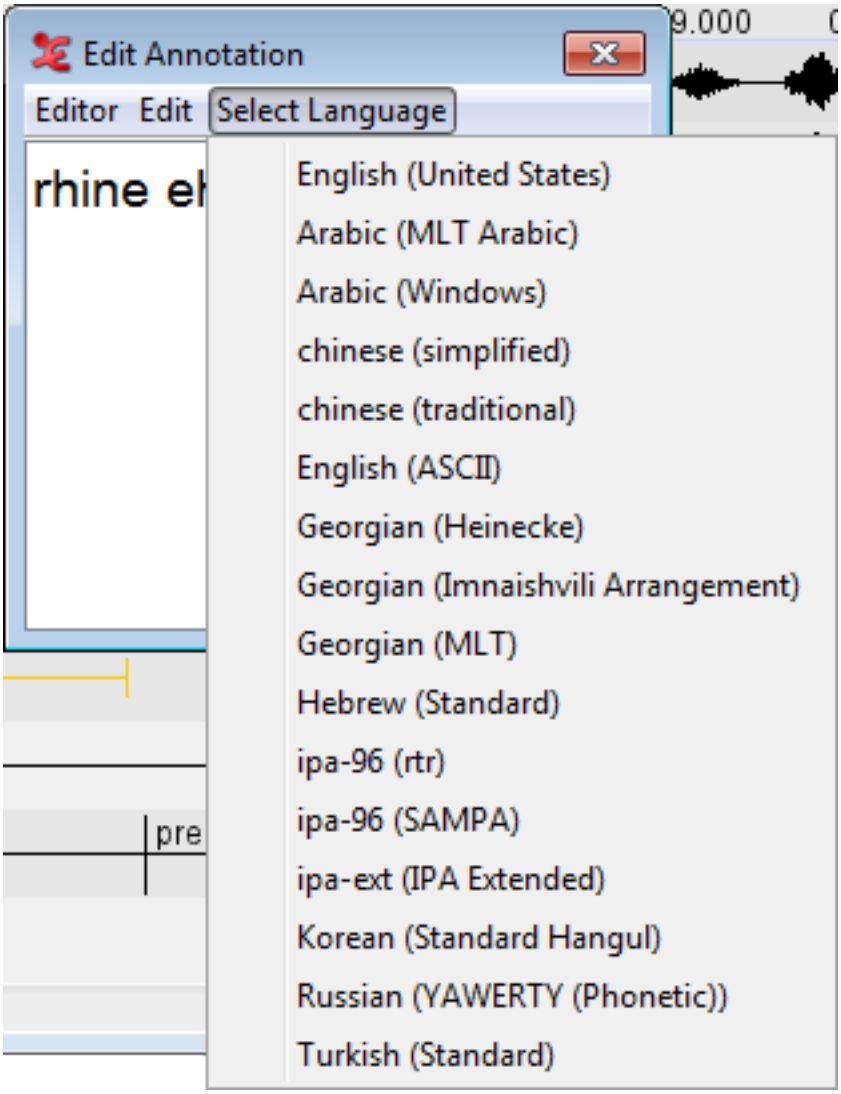


Figure 5.57. Select language

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 Section 5.8.20).

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.

5.8.3. Entering annotations from a Controlled Vocabulary

If the linguistic 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.

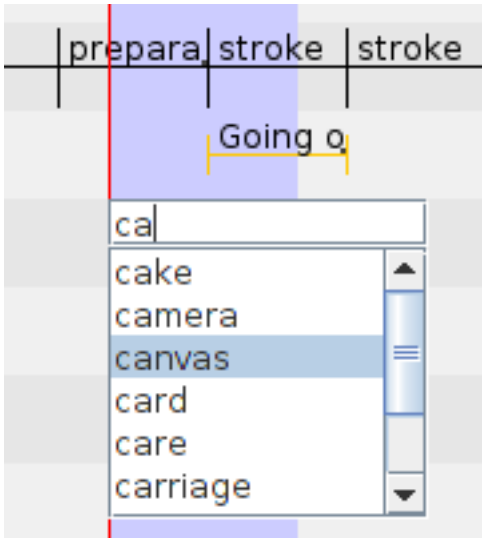


Figure 5.58. The suggest panel

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 section 4.4.28)

5.8.4. 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:

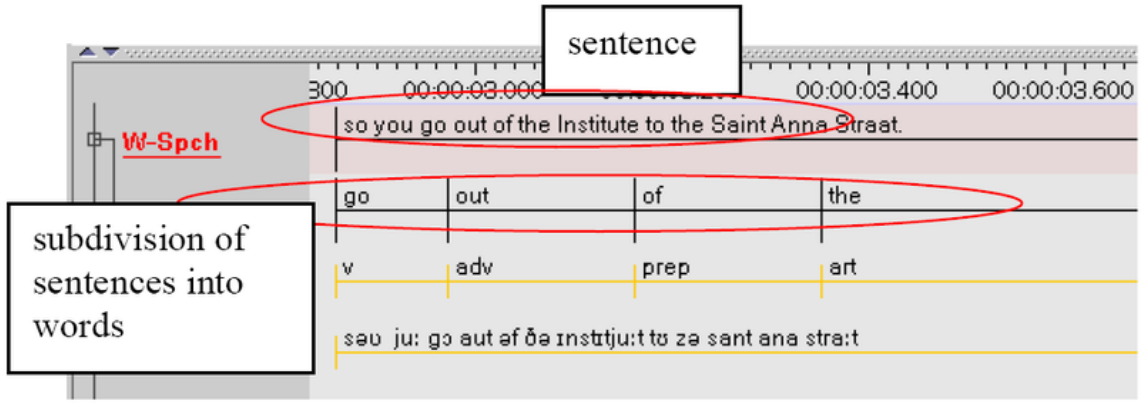


Figure 5.59. Subdividing annotation

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 5.60. New annotation after



Note

This option is only available for those tiers that are assigned to the stereotypes Time Subdivision and Symbolic Subdivision (see Section 5.1).

An annotation is always subdivided into two units. If you need further subdivisions, repeat the steps above.

5.8.5. 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.

5.8.6. 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:

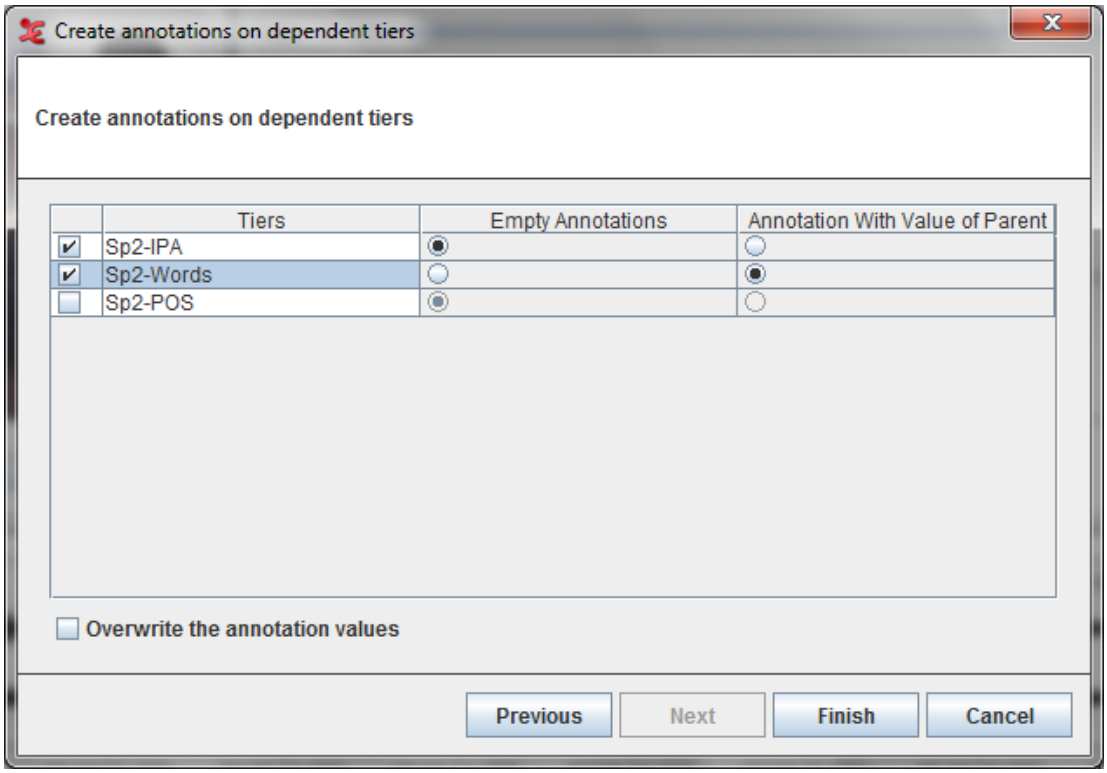


Figure 5.61. Create annotations on dependent tiers

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.

5.8.7. 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.

5.8.8. 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 5.62) 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.

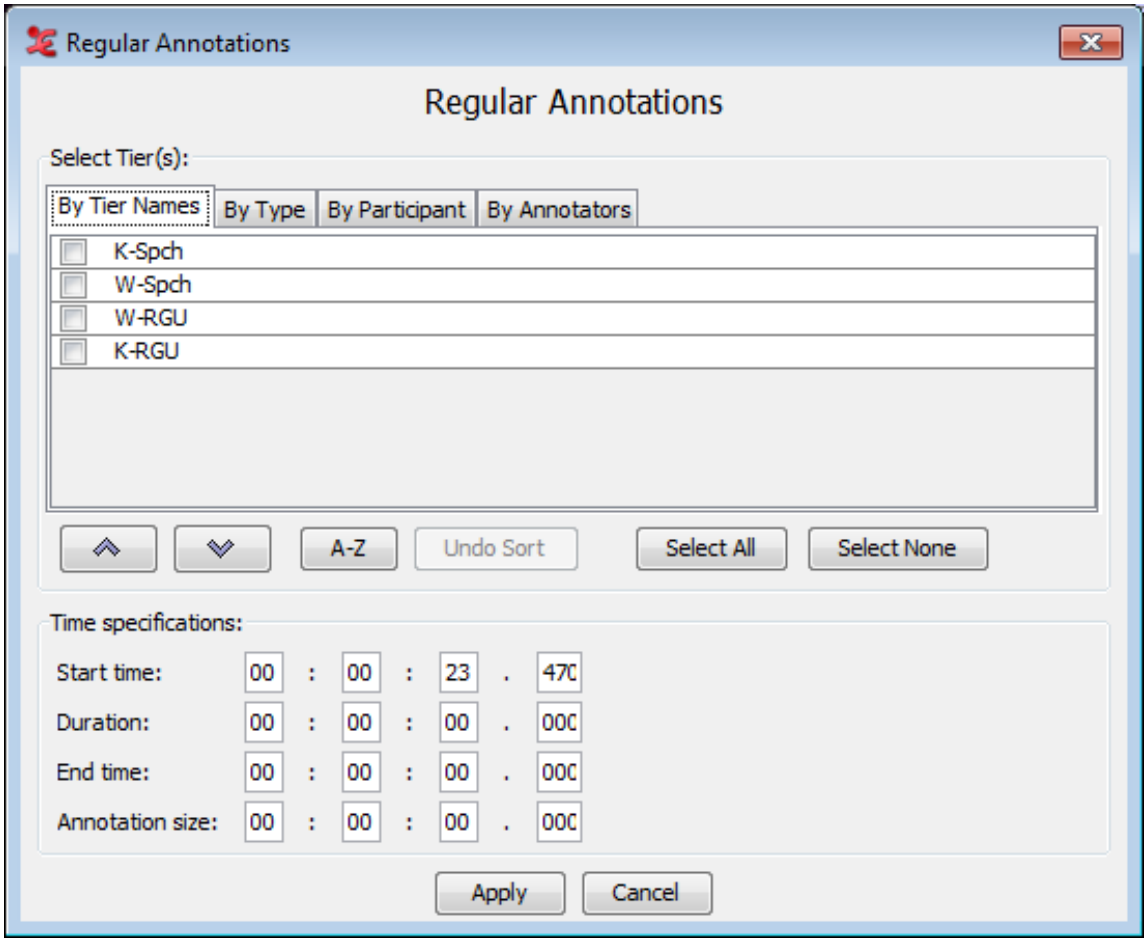


Figure 5.62. Regular Annotations

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.

5.8.9. Duplicating annotation values

If two annotations are adjacent and have the same length (for instance after creating regular annotations, see Section 5.8.8) 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.

5.8.10. 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 Section 4.4.28).

5.8.11. 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 5.63 you can select one or more tiers to provide labels and numbering.

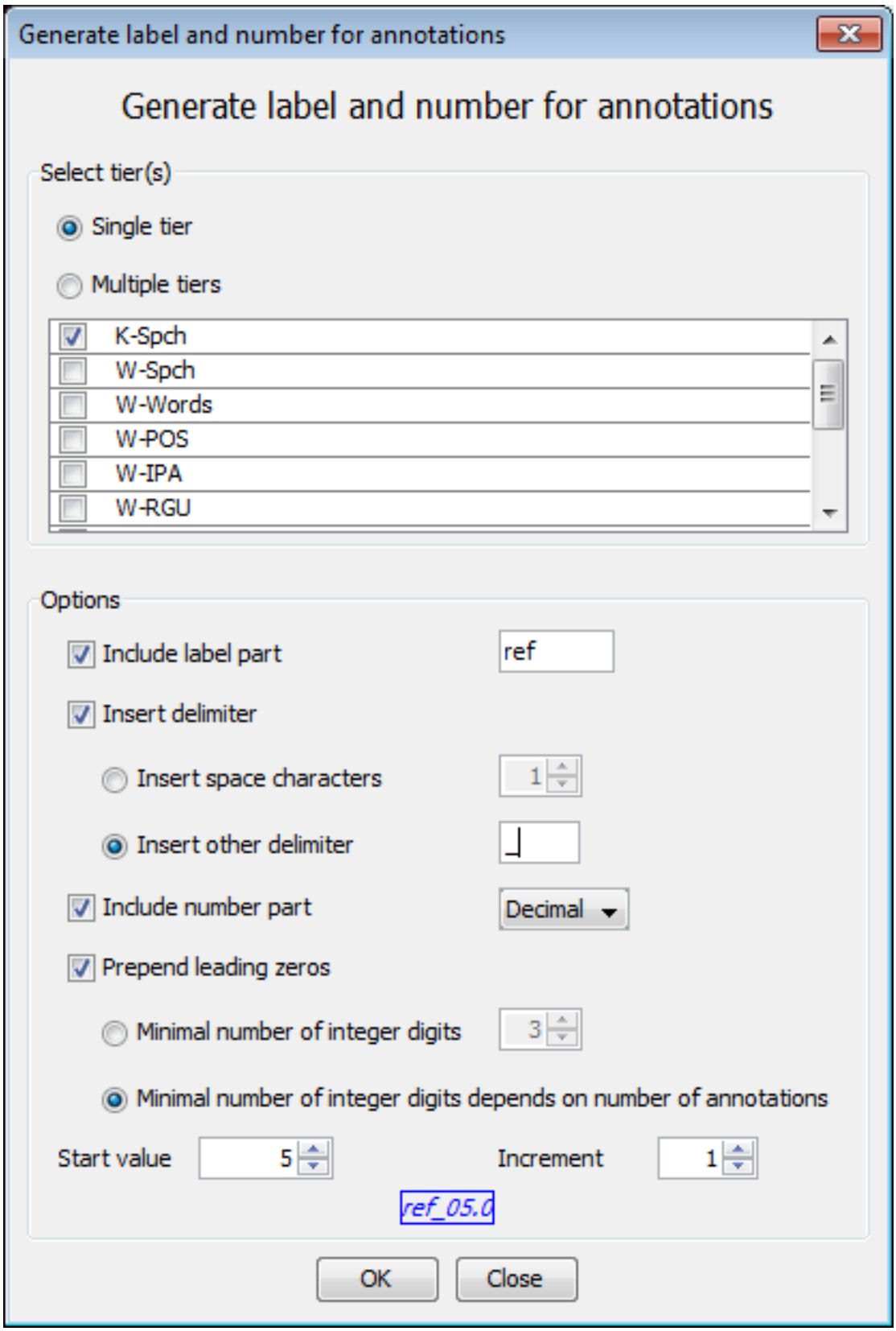


Figure 5.63. Label and Numbering annotations

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.

5.8.12. 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 Section 5.8.1).
2. In the Grid Viewer:
   a. Double-click on the annotation that you want to modify.
   The Inline Edit box appears (see Section 5.8.1).

5.8.13. Moving annotations to another tier

It can happen 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 not be moved along with the parent annotation unit.

5.8.14. Changing the case of Annotations

To change the letter case of all annotation on a tier, click Tier > Change Case of Annotations.... Then select the tier you wish to change, choose the letter case the annotation should get and click OK. If you select Lower-case, you are able to specify that each annotation should begin with a capital.

5.8.15. 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.

5.8.16. 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.

5.8.17. 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:

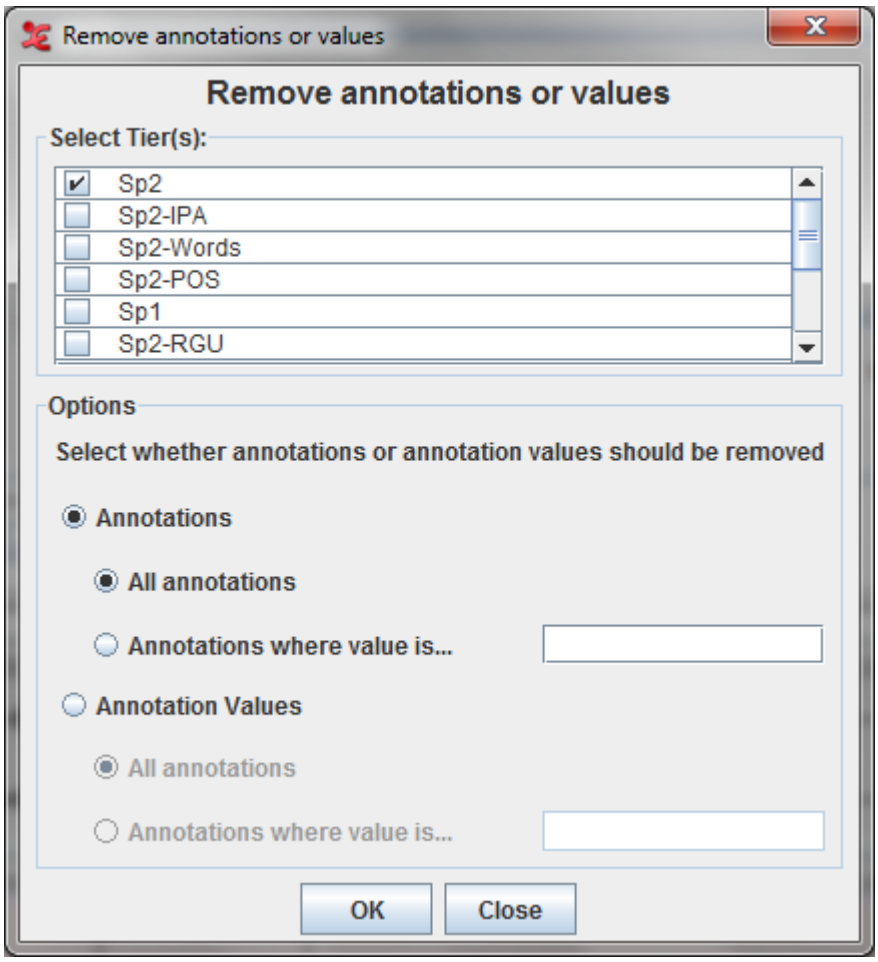


Figure 5.64. Delete annotations or annotations values

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 or select Annotations where value is... to delete only annotations/ values where annotation value matches the given value and finally click on Ok.

5.8.18. 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.

5.8.19. 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 5.65 illustrates this.

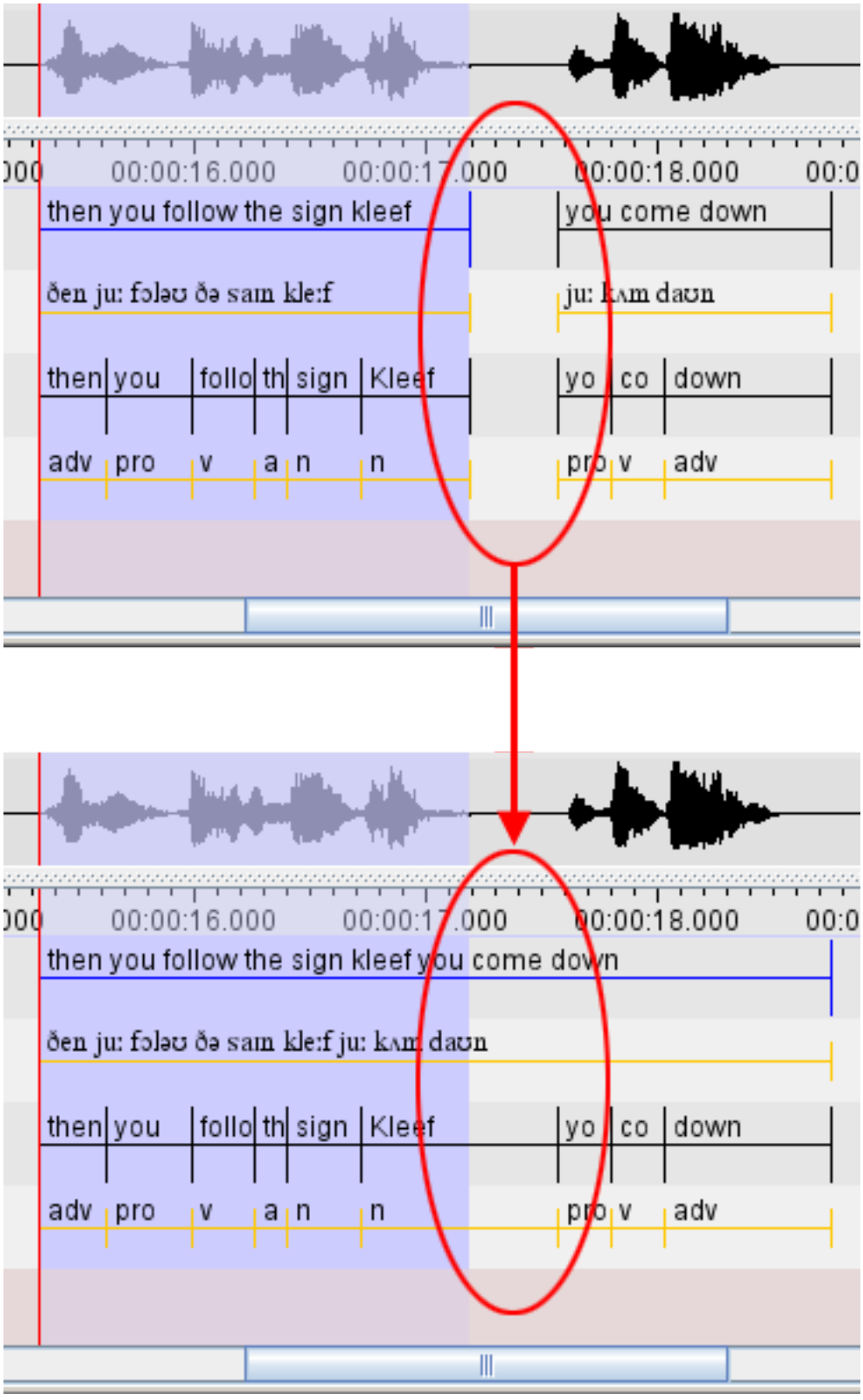


Figure 5.65. Merge annotations

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.

5.8.20. 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 5.66. ipa-96 keyboard map



Figure 5.67. 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 “#;” is matched to the key “@”, i.e., in order to get “#”, 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 “#”. To get “#”, 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:

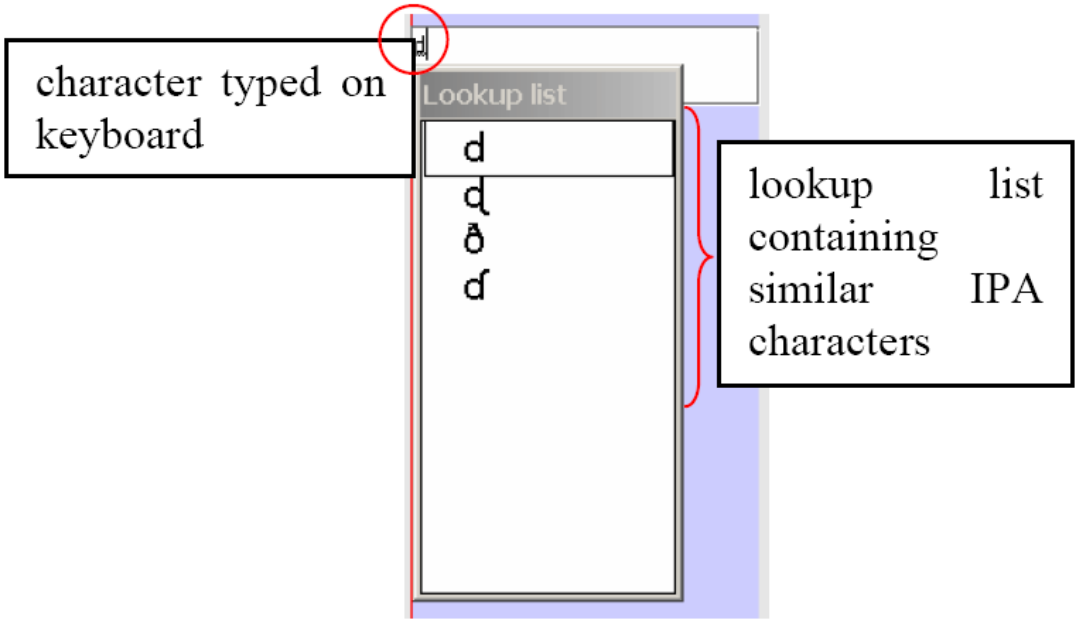


Figure 5.68. Roman typographic root

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.:

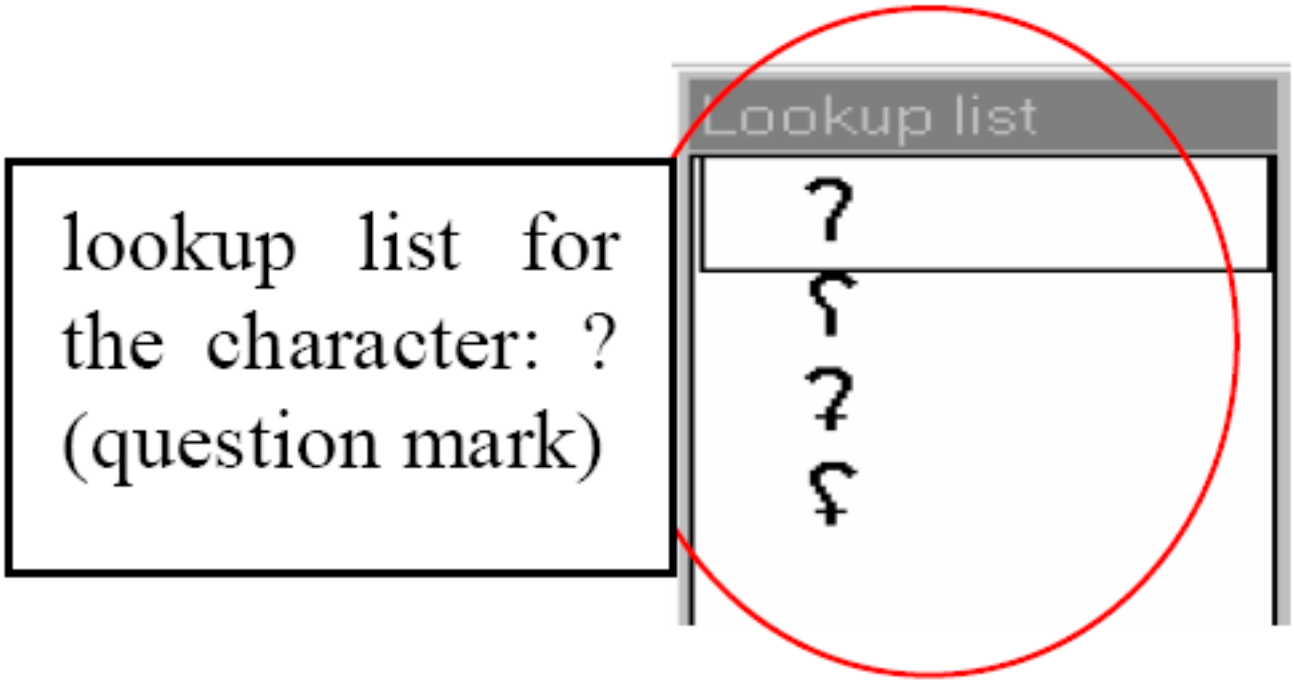


Figure 5.69. Non-letter symbols

- Diacritics are entered with the help of the following keys:

Table 5.4. 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:

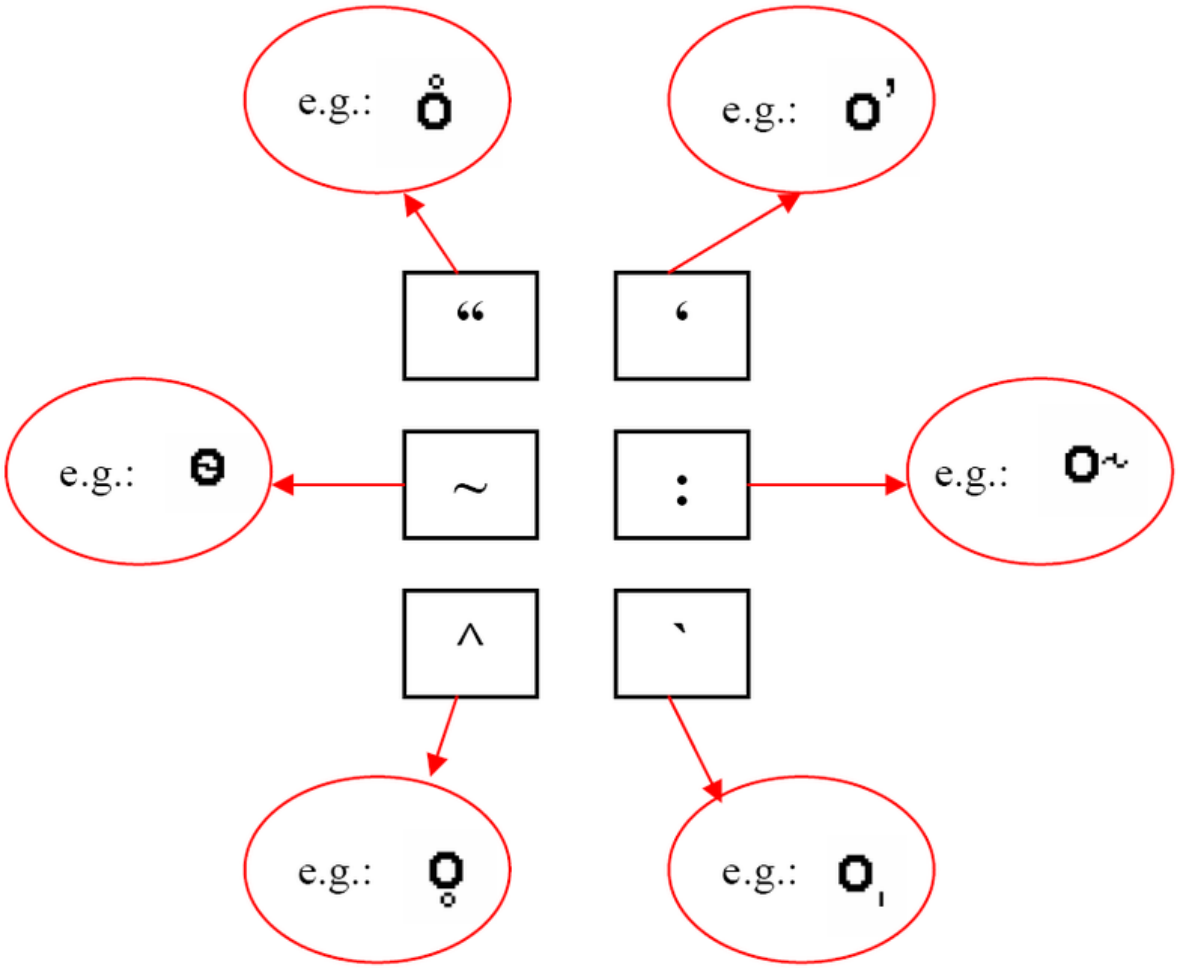


Figure 5.70. Diacritics



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.:

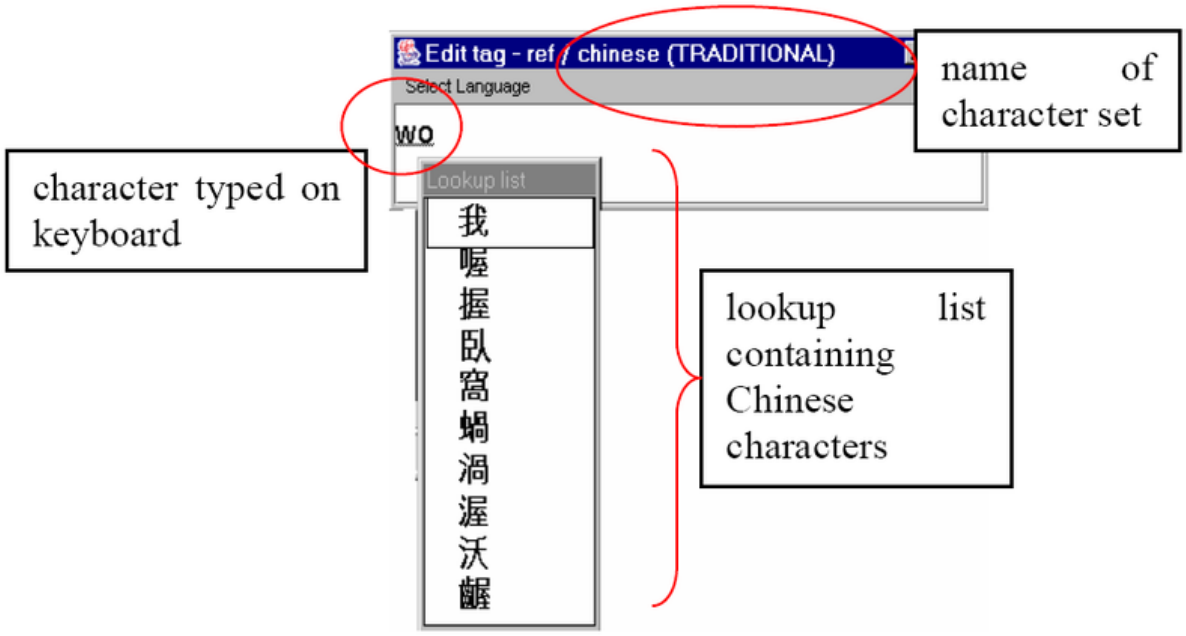


Figure 5.71. Chinese characters

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 5.72. Keyman

Ukelele (Mac), http://scripts.sil.org/ukelele



Figure 5.73. 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).

5.8.21. Associating an annotation with a Data Category

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 5.74), first select a profile in the left panel and then select a data category in the middle panel. Finally click on Apply to associate the selected data category to the annotation.

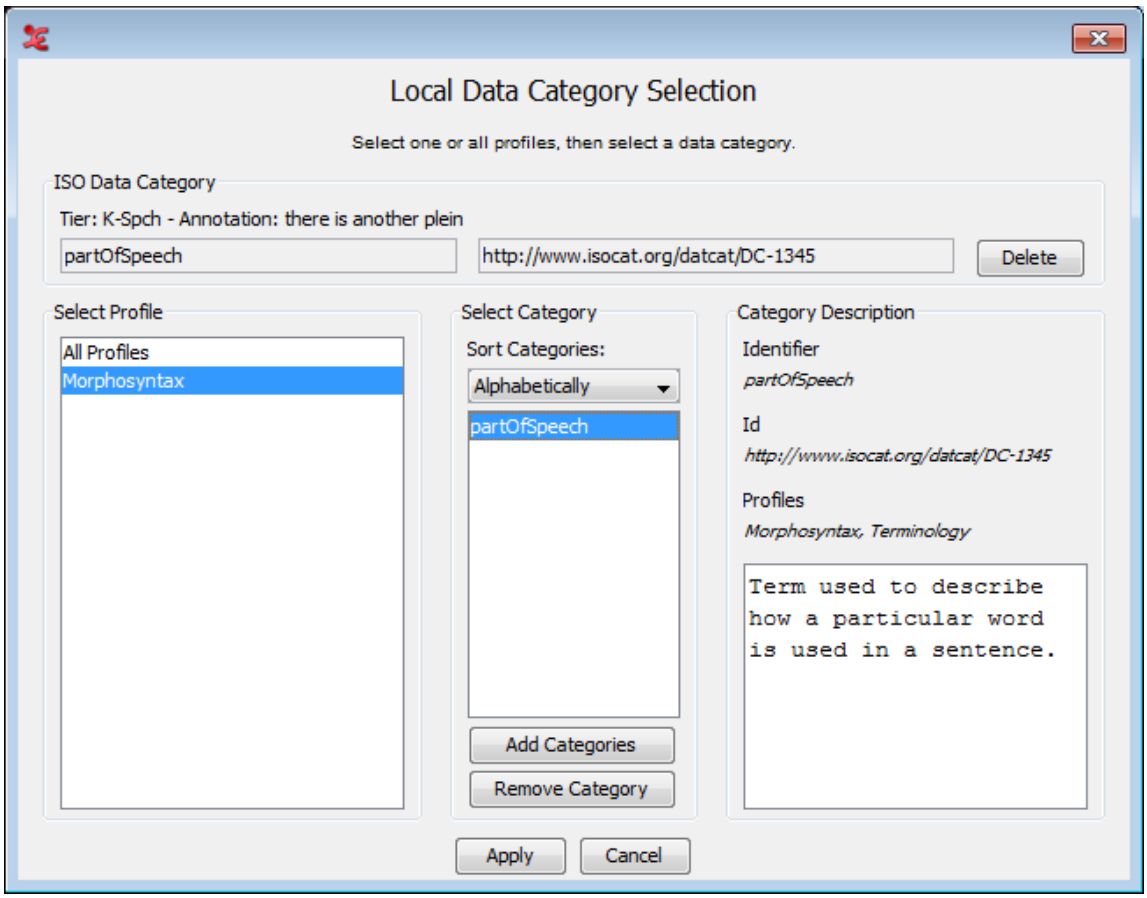


Figure 5.74. Local Data Category Selection: Annotation

More information about the ISO DCR and how to use it can be found in Section 5.2.

5.9. How to create annotation units on dependent tiers

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 a child tier. 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 (= child) 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 the destination tier already contains annotation units, choose between overwriting or preserving them. If its still empty you can ignore this option.
6. Select Create destination annotation for empty source annotation if you want to create for every source annotation a destination annotation, even if its empty.

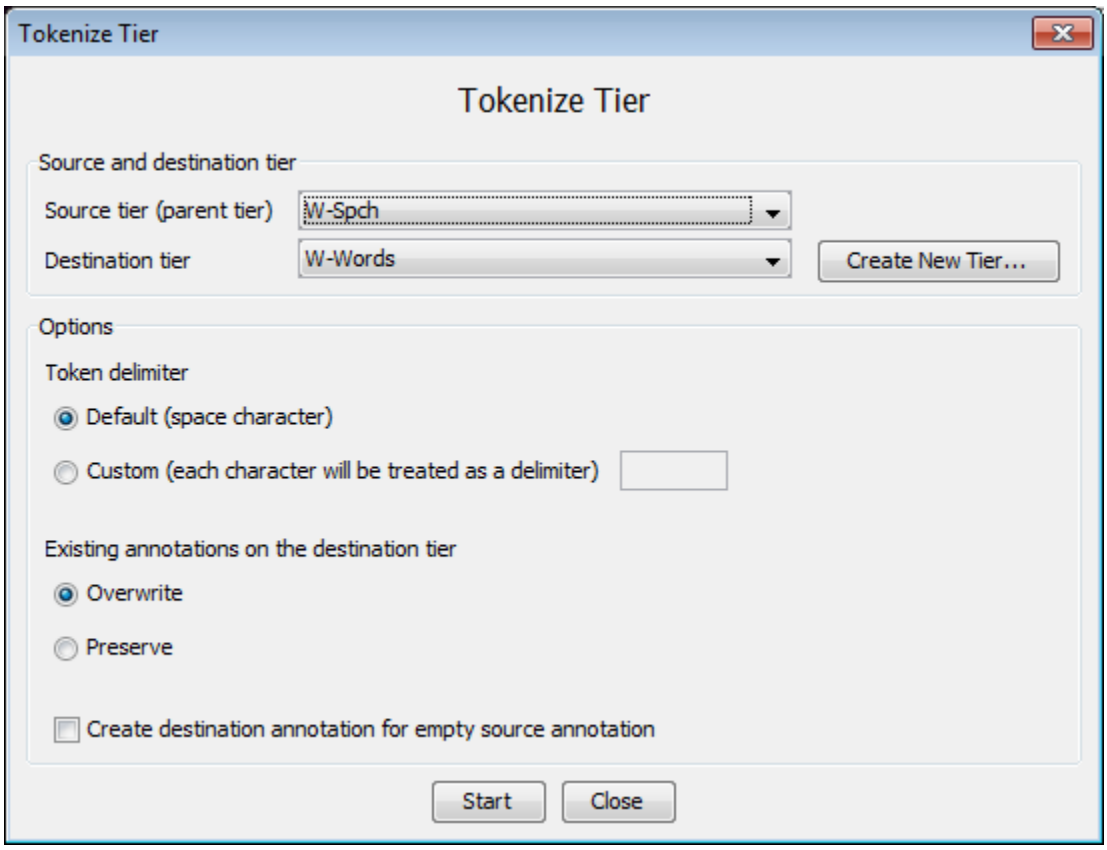


Figure 5.75. Tokenize tier

7. Click on Start, the tokenization will begin:

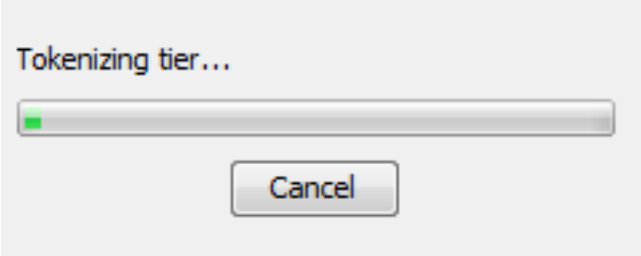


Figure 5.76. Tokenization

8. When it is finished, you will see that every annotation unit from the source tier has been tokenized on the destination tier:

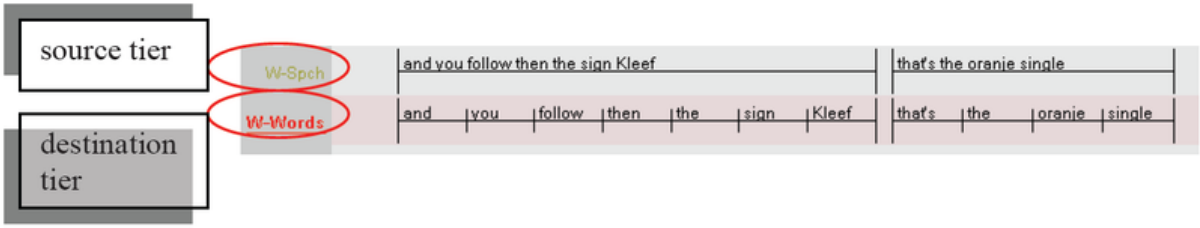


Figure 5.77. Source and destination tier

Even when tokenizing to a tier from the type time subdivision, all tokens (words in this example) on the destination tier have the same size (i.e. duration). In that case you probably want to adjust their length, as described in Section 5.7.6.

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 Section 4.4.6. If you want to copy or filter the contents from 1 tier to another symbolic associated tier, have a look at Section 5.10.

Tabs and newline characters are always treated as delimiters.

5.10. How to filter and copy tiers

5.10.1. 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.

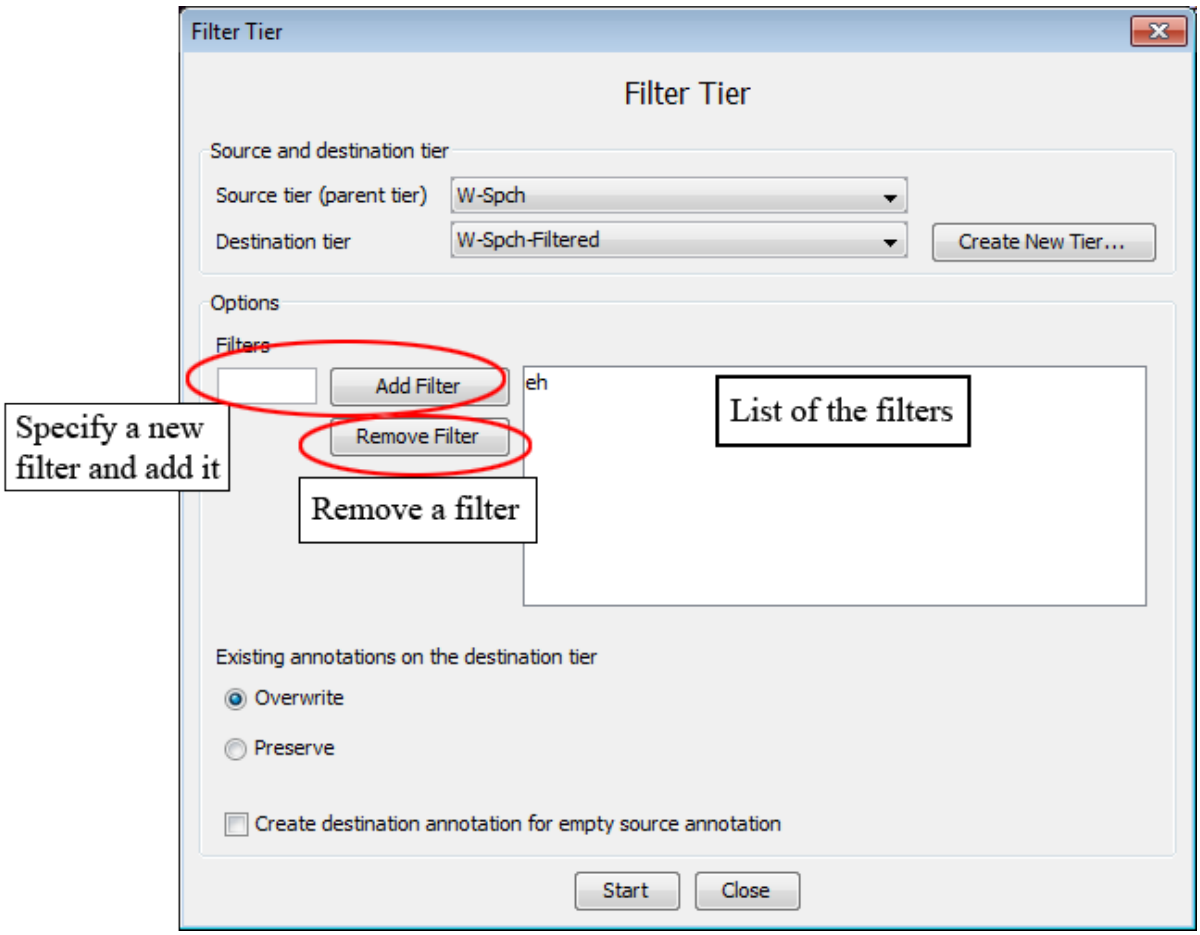


Figure 5.78. Filter tier

5.10.2. 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 linguistic 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.

Note that this is similar to the change parent tier functionality (see Section 5.4.8). However it differs in 2 aspects:

- It is not mandatory to copy the child tiers.

- The parent for the copy can any be a tier in the transcription, including the tier itself.

5.11. 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 Section 5.11.1) and the second specifically aimed at comparing annotators (see Section 5.11.2).

5.11.1. 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.

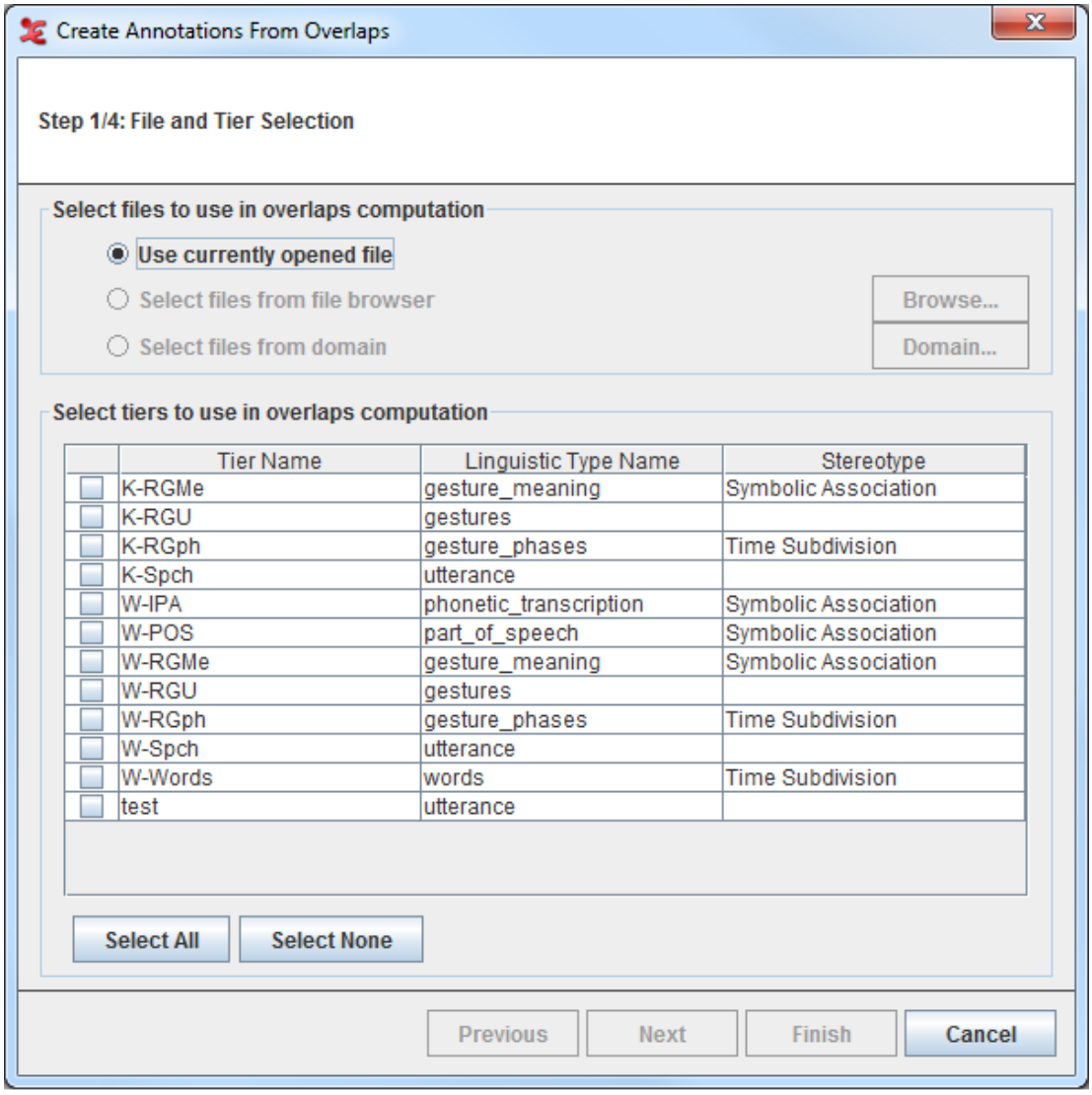


Figure 5.79. Create annotations from overlaps



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 Section 4.7.4.4

2. Step 2/4: Overlaps Computation Criteria

   In this step you can define the overlaps computation criteria in the *Overlaps Computation Criteria* window

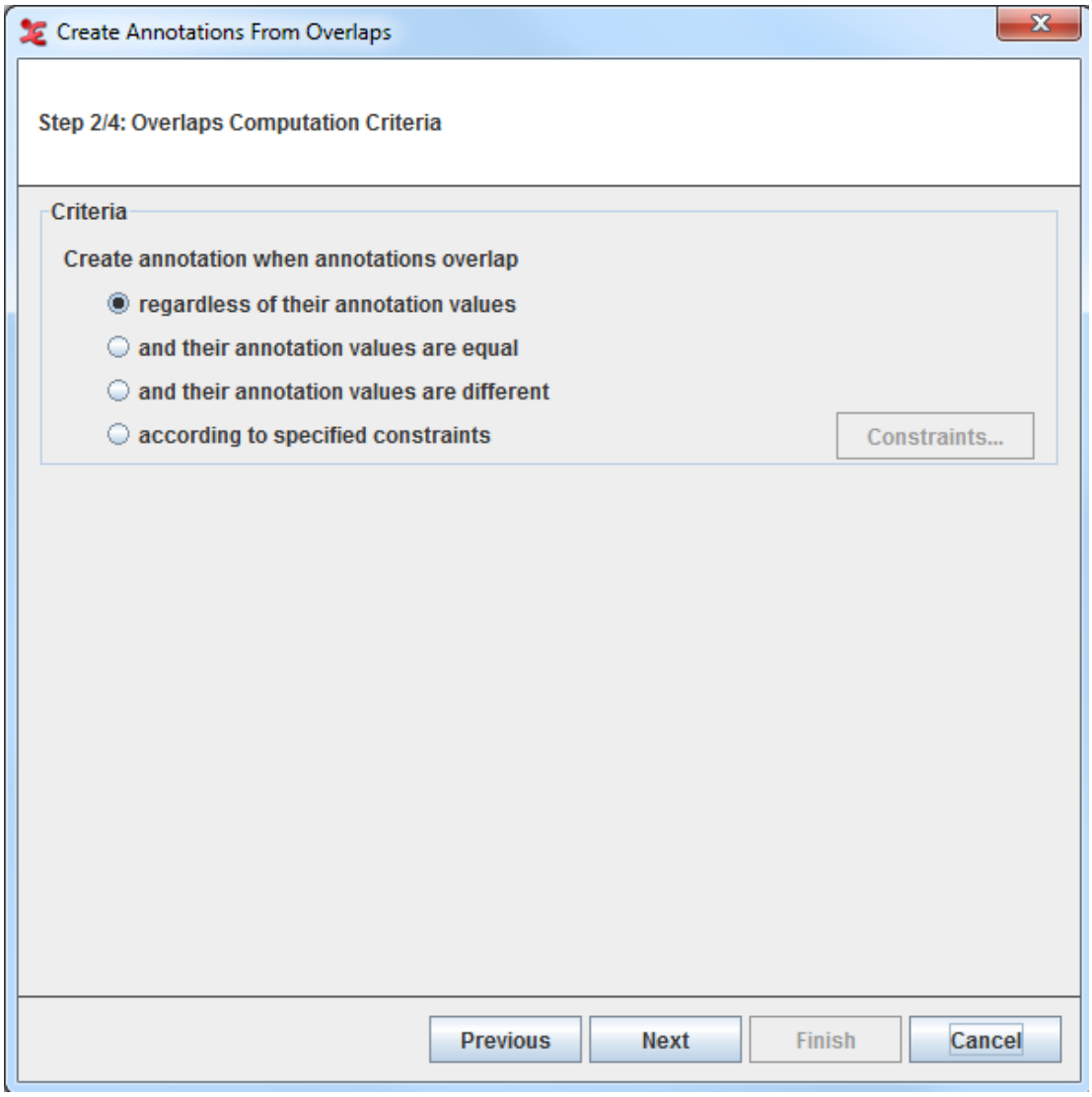


Figure 5.80. Overlaps Computation Criteria

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:

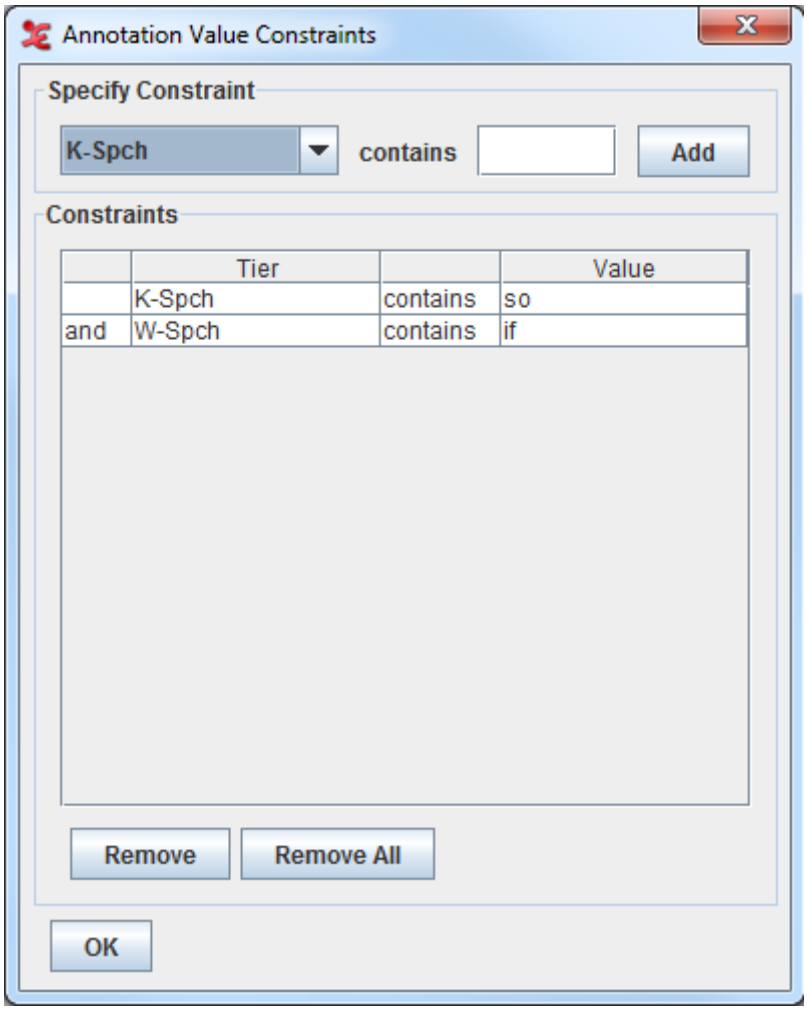


Figure 5.81. Annotation Value Constraint

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 Linguistic 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. You can now go to the next and last step, by selecting Next.

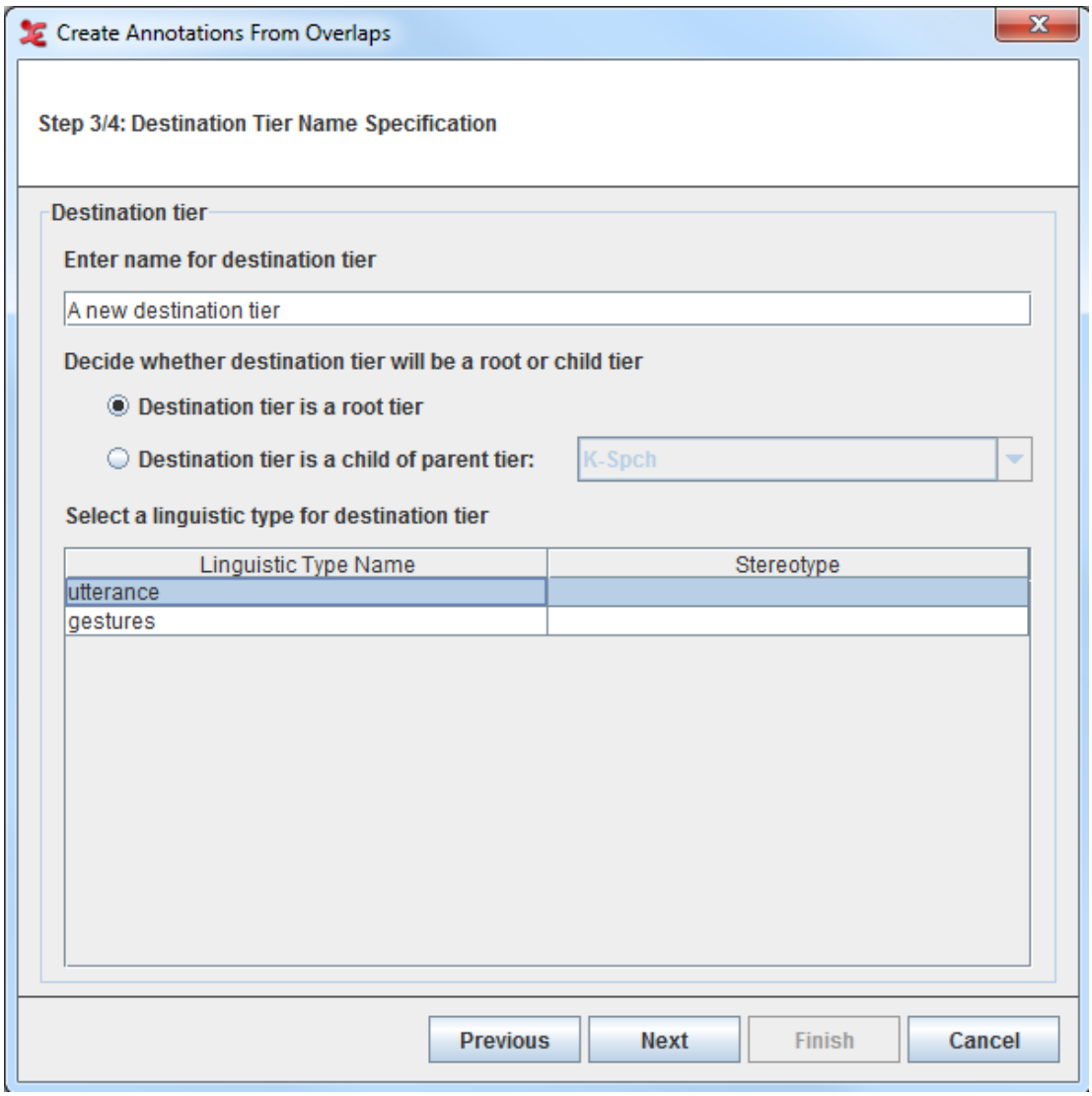


Figure 5.82. Destination Tier Name Specification

4. Step 4/4: Destination Tier Value Specification

   Here you can specify the value for the destination tier.

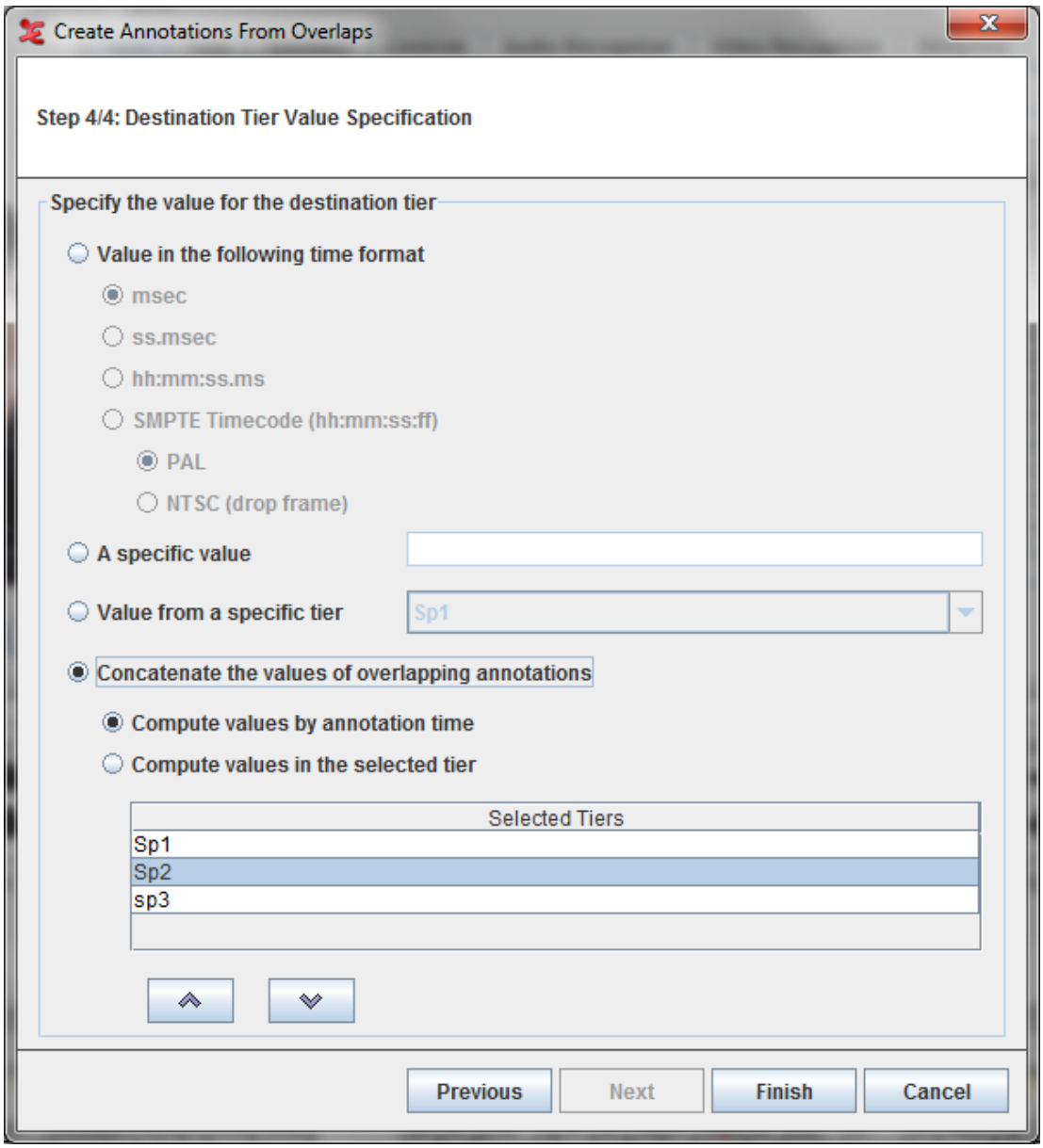


Figure 5.83. Destination Tier Value Specification

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.

5.11.2. Comparing annotators

Sometimes it is necessary to compare the work of annotators to be able to improve annotation quality. The function giving the essential information for this purpose can be accessed via Tier > Compare Annotators.... Select a tier for each of two annotators and click Compare. In the table that now appears overlapping annotations are in the same row. The begin time and end time of each annotation is given as well as the amount of overlap, the total extent (time from lowest begin time till highest end time) and a value that indicates the agreement: overlap / extent. The average agreement is printed above the table. By clicking Save the information in the table can saved into a text file.

5.12. 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 5.79)

2. Step 2/4: Subtract Computation Criteria

   In this step you can choose the subtraction computation criteria in the *Subtract computation criteria* window

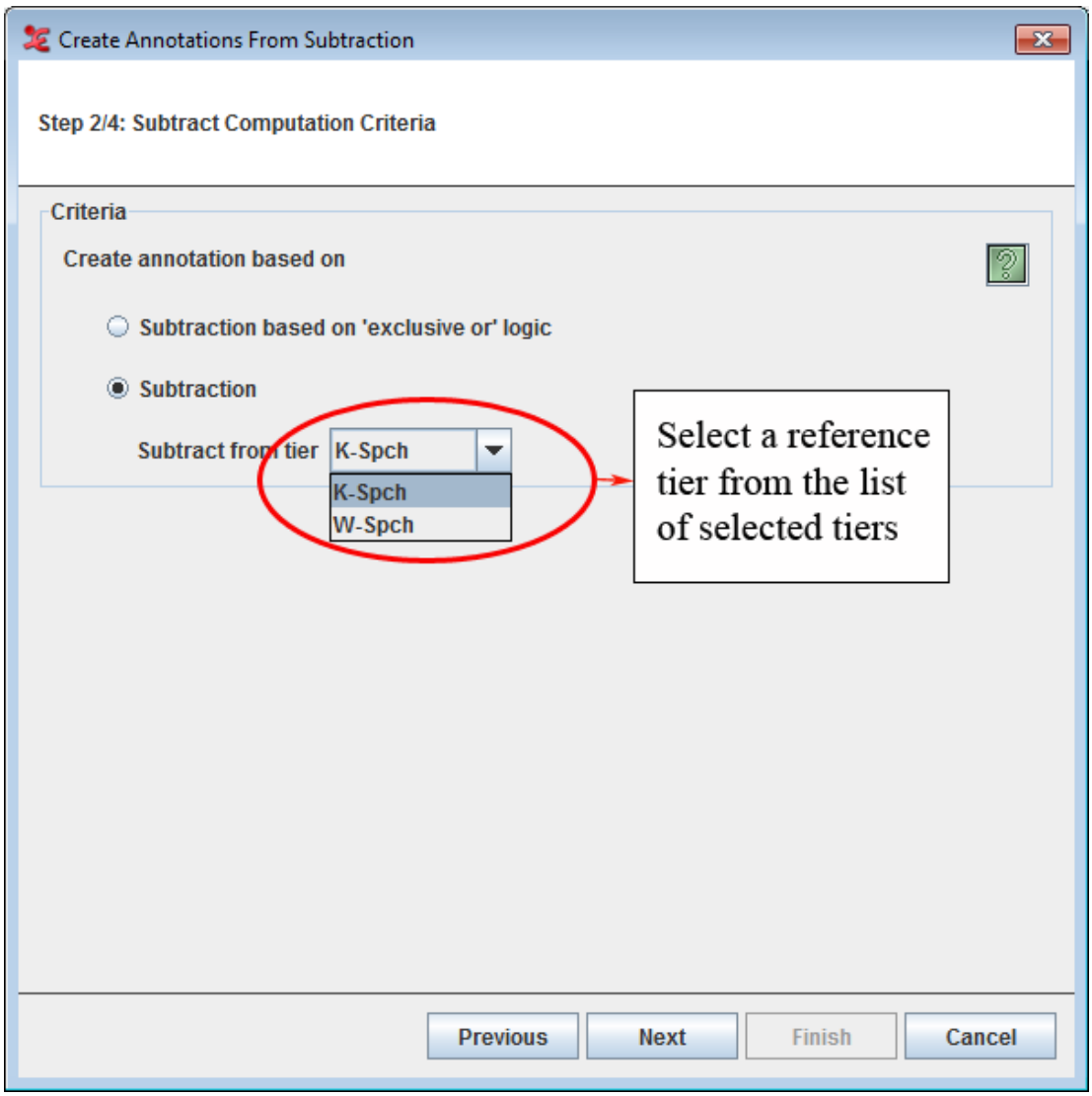


Figure 5.84. Subtract Computation Criteria

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 5.82)
4. Step 4/4: Destination Tier Value Specification

   Here you can specify the value for the destination tier.

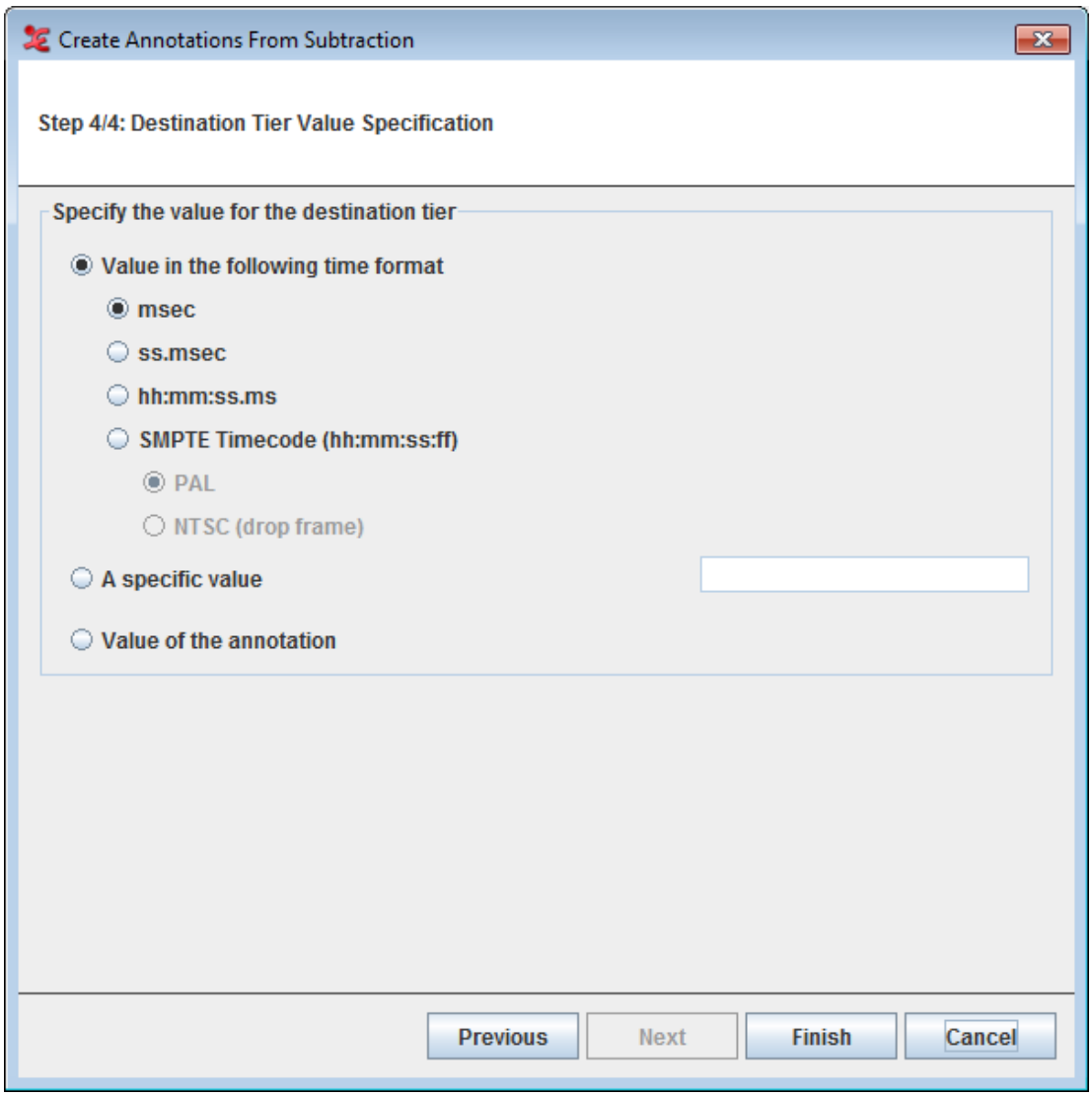


Figure 5.85. Destination Tier Value Specification

For detailed explanation on the options see Figure 5.83.

Finally, you can click on Finish. The new tier will be created and populated.

5.13. Copy and paste annotations

5.13.1. 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

5.13.2. 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 linguistic type! If you want to copy the annotation to a tier of a different type, use *Duplicate Annotation* (see Section 5.13.5).

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 Section 9.1.3 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 Section 5.13.3).

5.13.3. 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 linguistic 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.

5.13.4. 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 linguistic 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

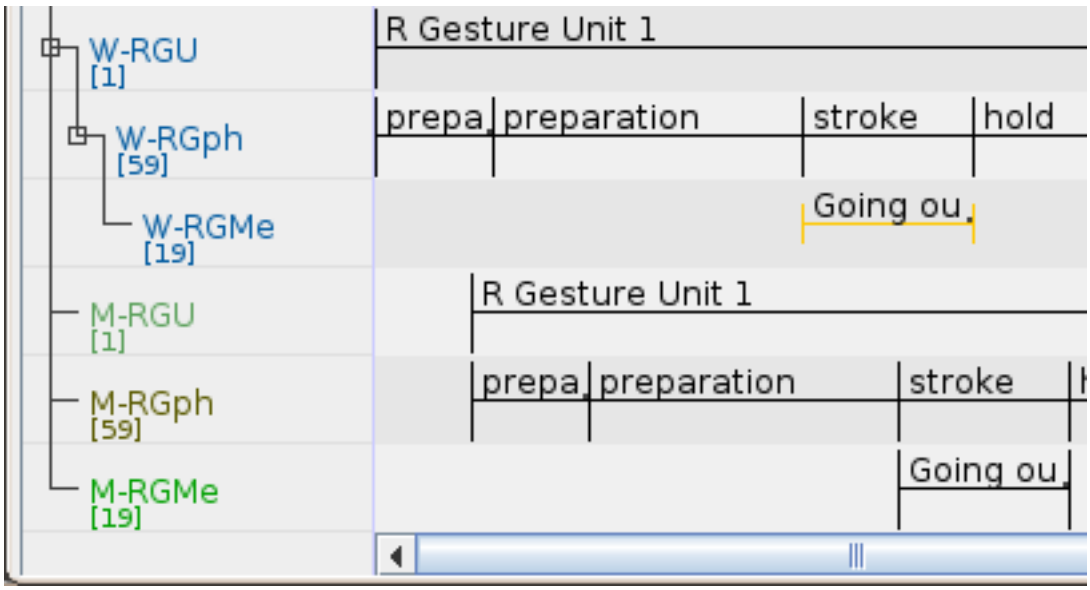


Figure 5.86. Copying a annotation group

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.

5.13.5. 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 linguistic type.

5.13.6. Synopsis

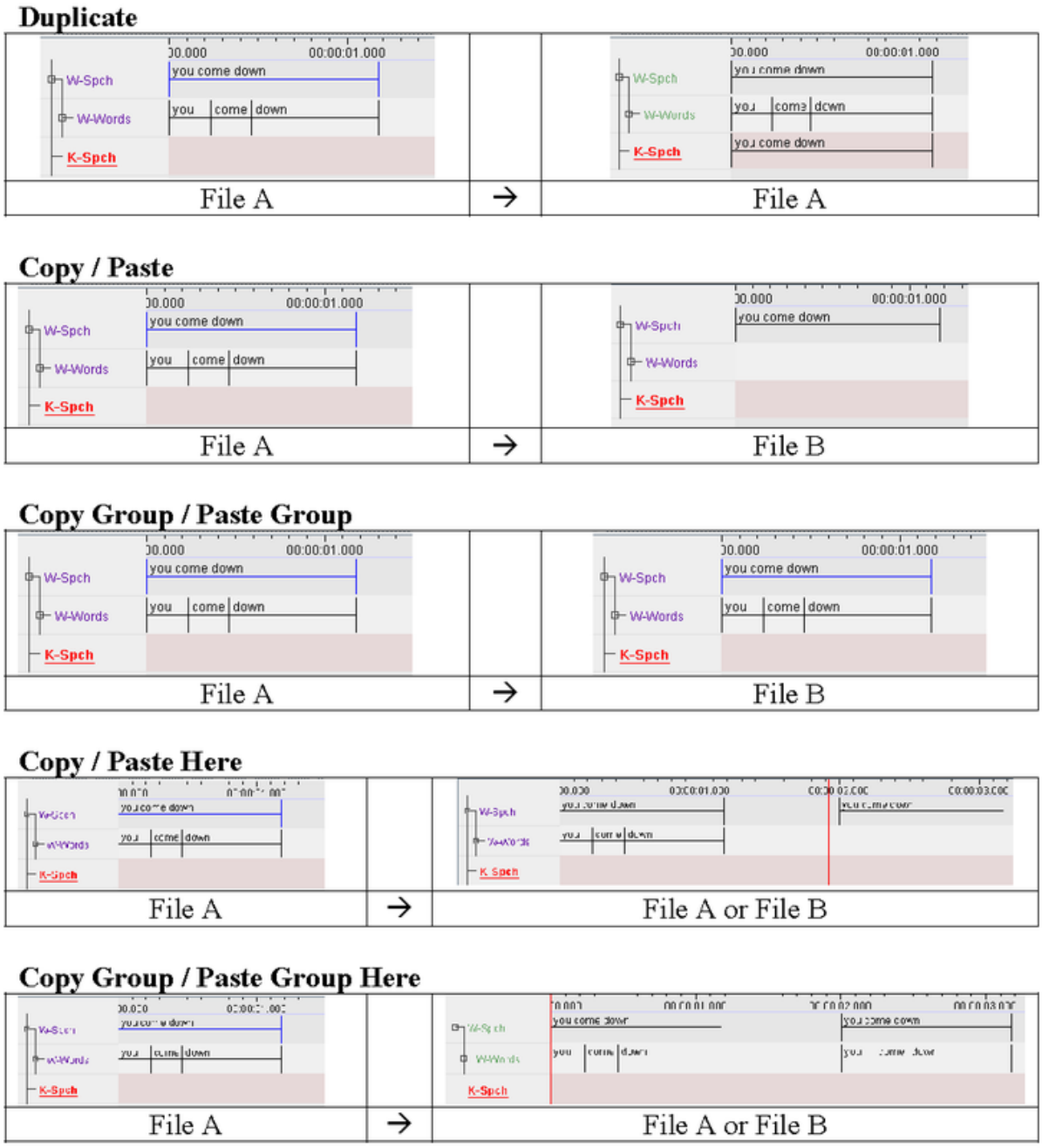


Figure 5.87. Synopsis copy and paste annotations

5.14. 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 / linguistic type
- Copy tier / change parent tier
- Insert annotation before / after

- Change tier attributes / linguistic type
- Clear selection (through the icon or keyboard shortcut)
- Delete annotation / tier / linguistic type
- Modify annotation / annotation time
- Create new annotation
- Copy / Paste / Duplicate annotation
- Play around selection time setting
- The tokenizer
- Add / Change / Edit controlled vocabularies

5.15. Information about annotations

5.15.1. 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.

5.15.2. 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:

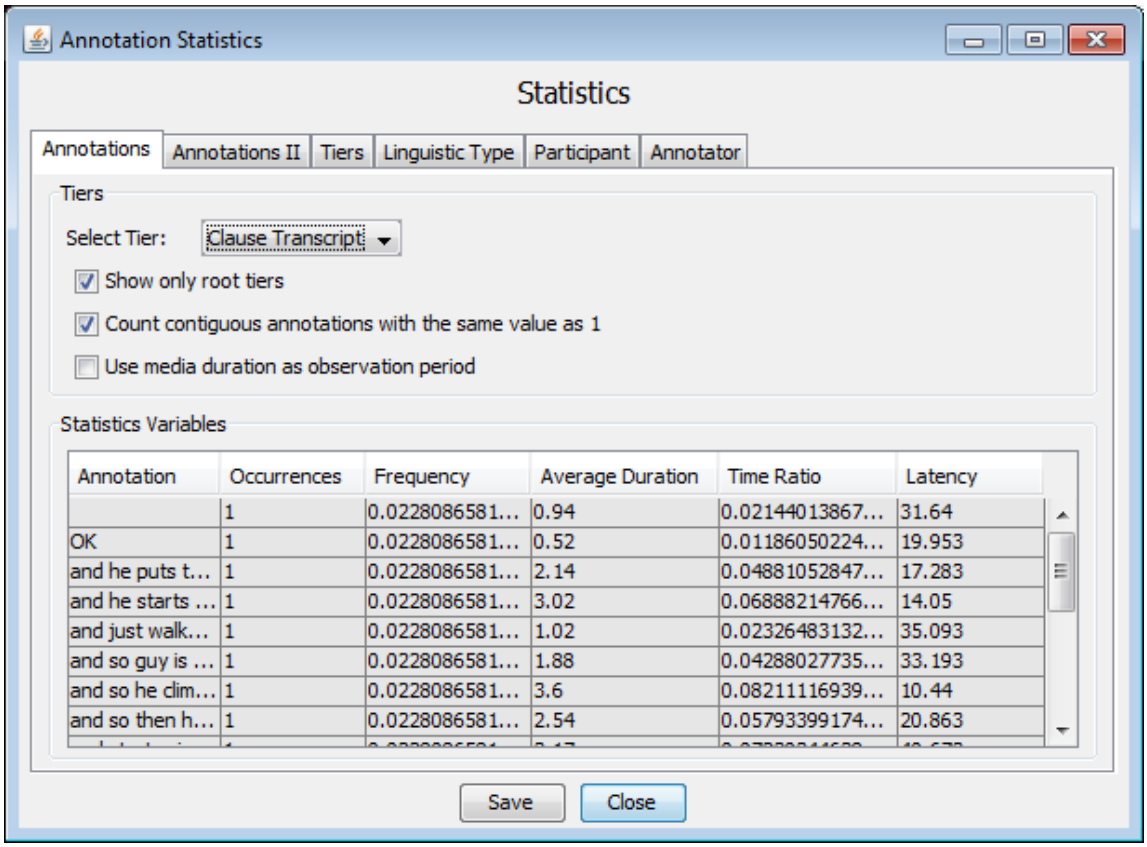


Figure 5.88. Annotations Statistics

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:

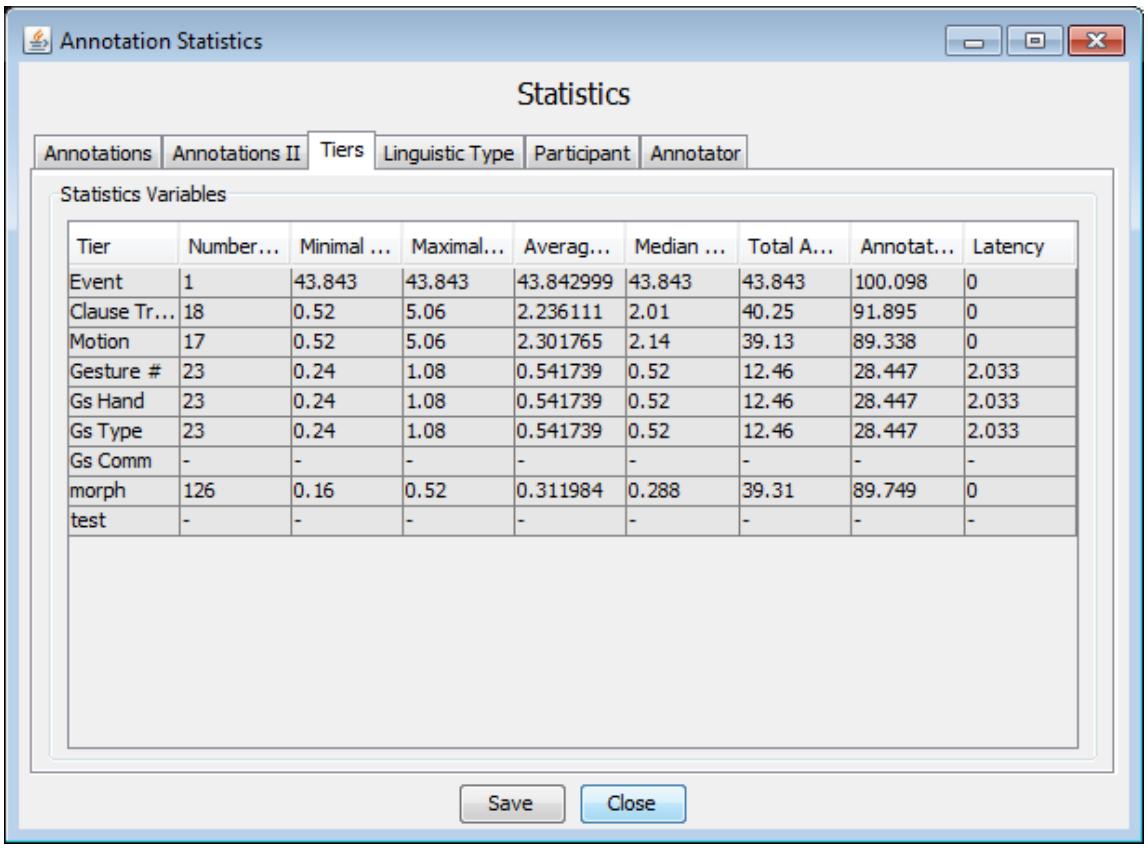


Figure 5.89. Tiers statistics

- 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.
- 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 Linguistic Type, Participant, Annotator are similar to the statistics on the Tiers tab except that they apply to linguistic type, participant and annotator respectively.

Chapter 6. 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.

6.1. 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 4, Chapter 5 and Chapter 7).

6.2. Synchronization mode

The Synchronization mode is described in Section 4.2.3, 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.

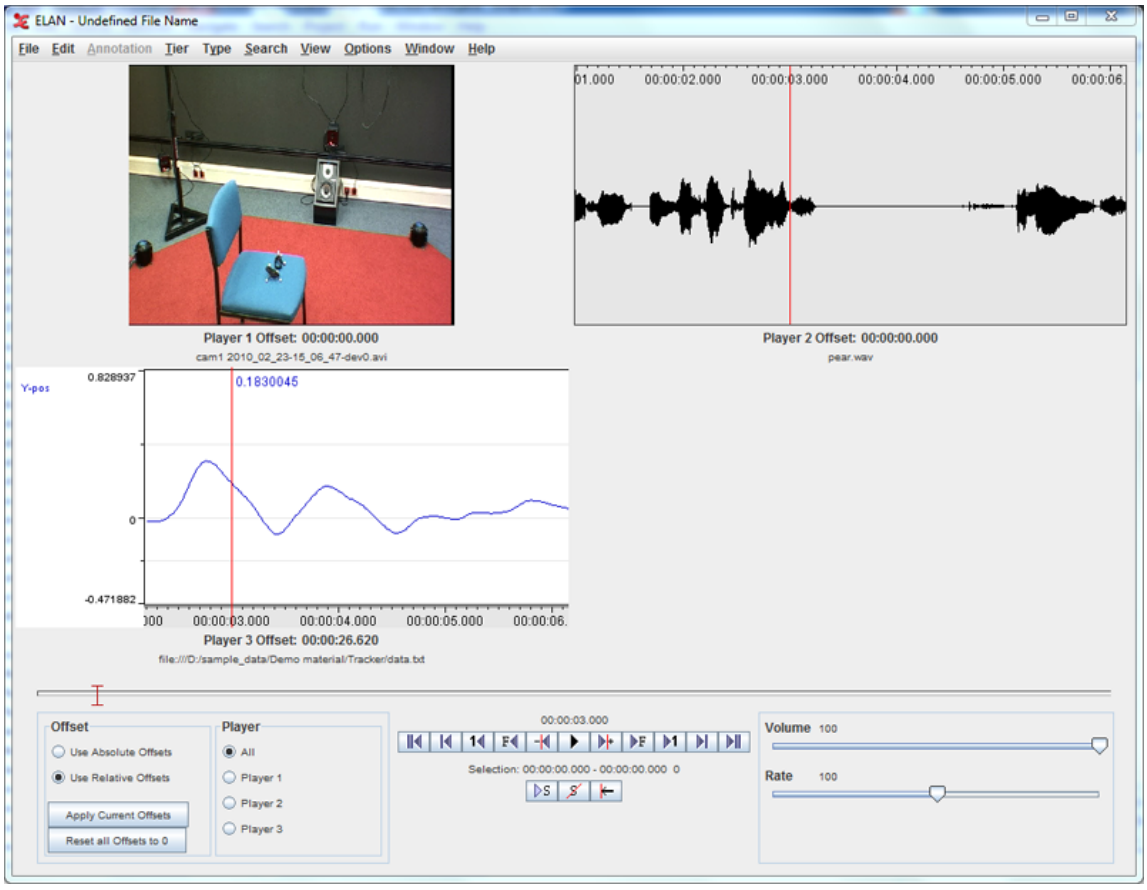


Figure 6.1. Synchronization of video, audio and timeseries files

6.3. 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 6.2) else transcription window is opened with the last used settings (see Figure 6.6).



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 Section 6.1)or in the special purpose Segmentation mode(see Section 6.4)), and then transcribing and translating these turns.

6.3.1. 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 6.6). Here you can select the tier types and the number of columns.

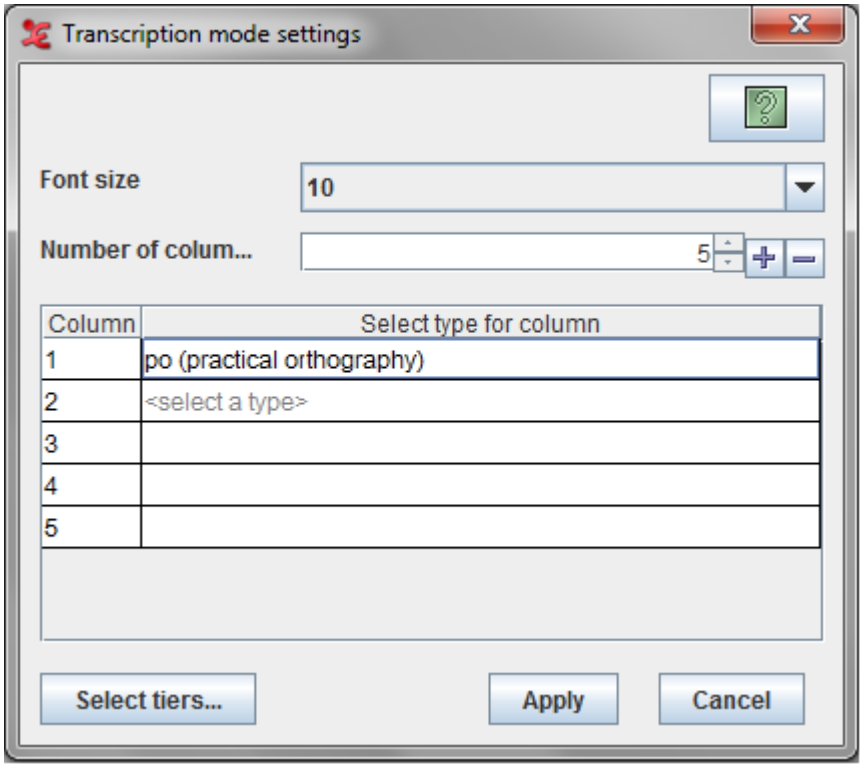


Figure 6.2. Settings Dialog

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 Section 6.3.2.



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 :

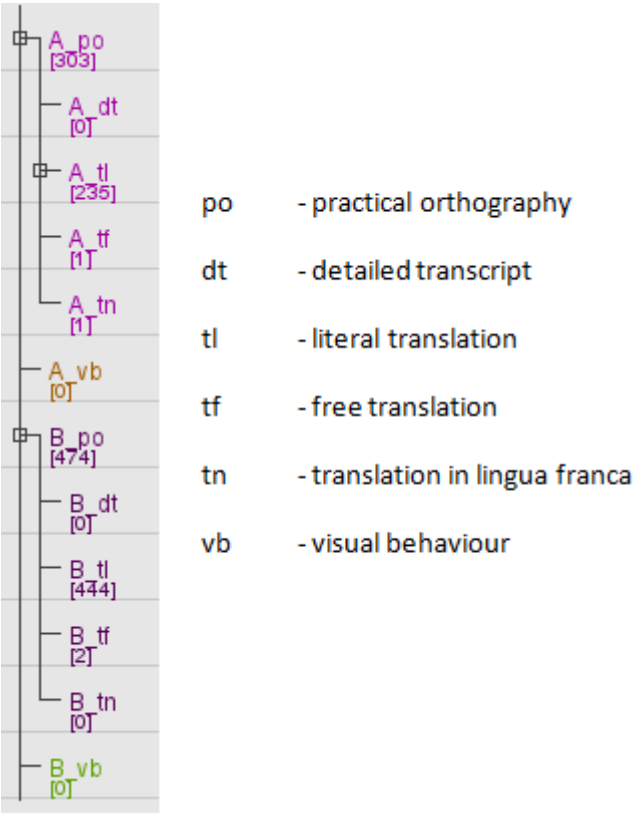
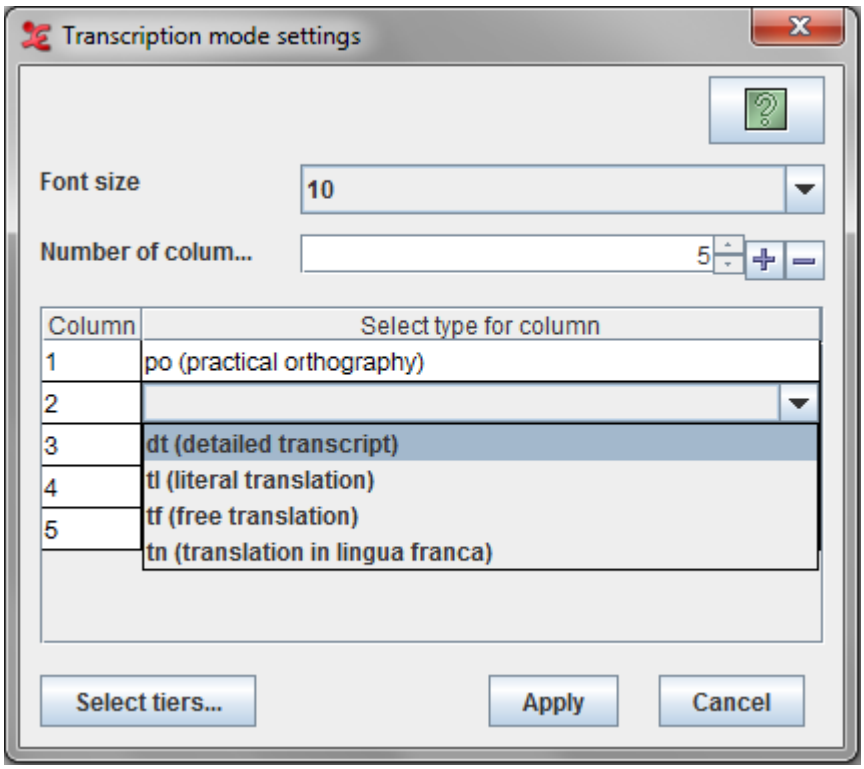


Figure 6.3. Tier Structure

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 5.2). 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 6.6), 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 linguistic types to differentiate the types of information in your tiers. Thus the linguistic type of your free translation tier should be different from the linguistic type of your main transcription.

6.3.2. 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 6.2. This dialog comes up.

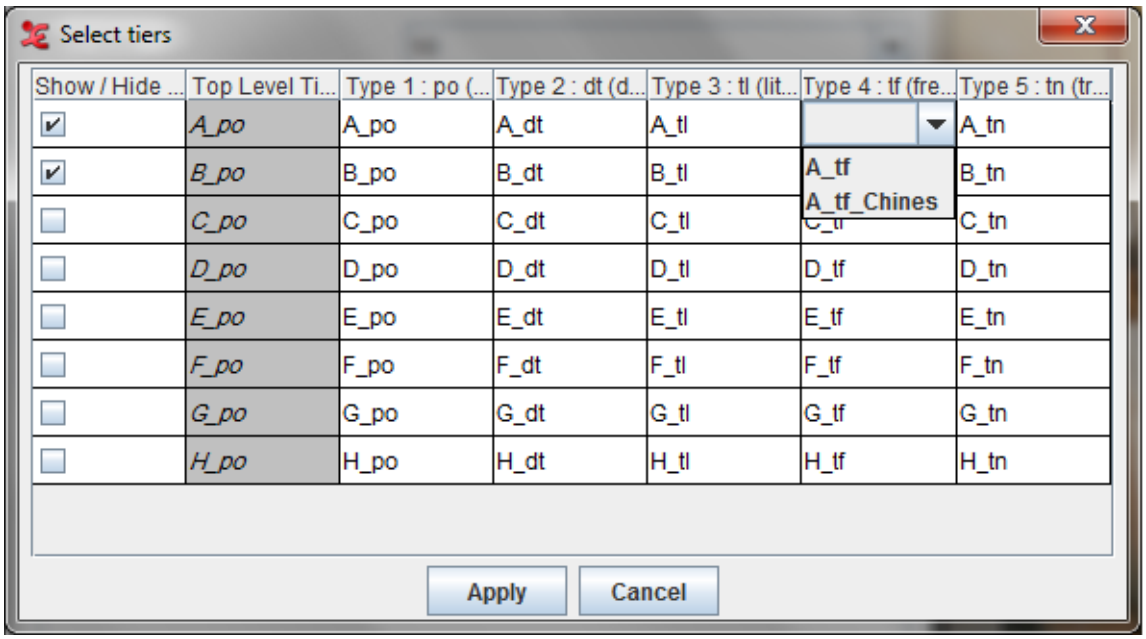
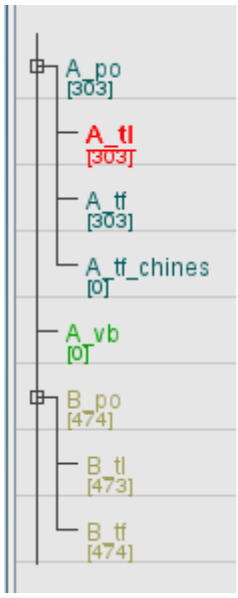


Figure 6.4. Select Tiers Dialog



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.

6.3.3. 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.

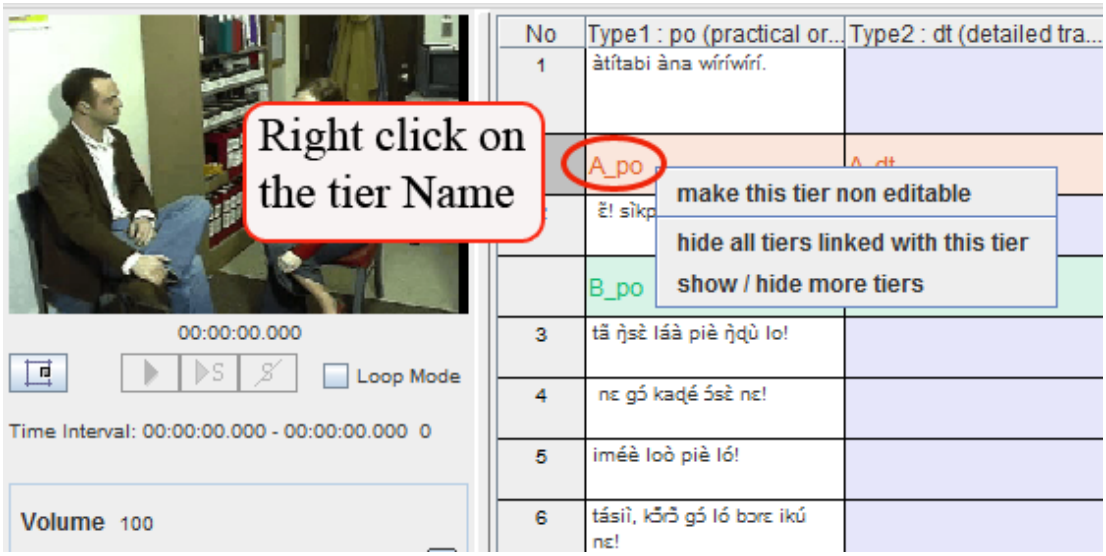


Figure 6.5. Context Menu

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 : Section 6.3.2 . Check all the tiers to show up in the table and uncheck the tiers that are to be hidden.

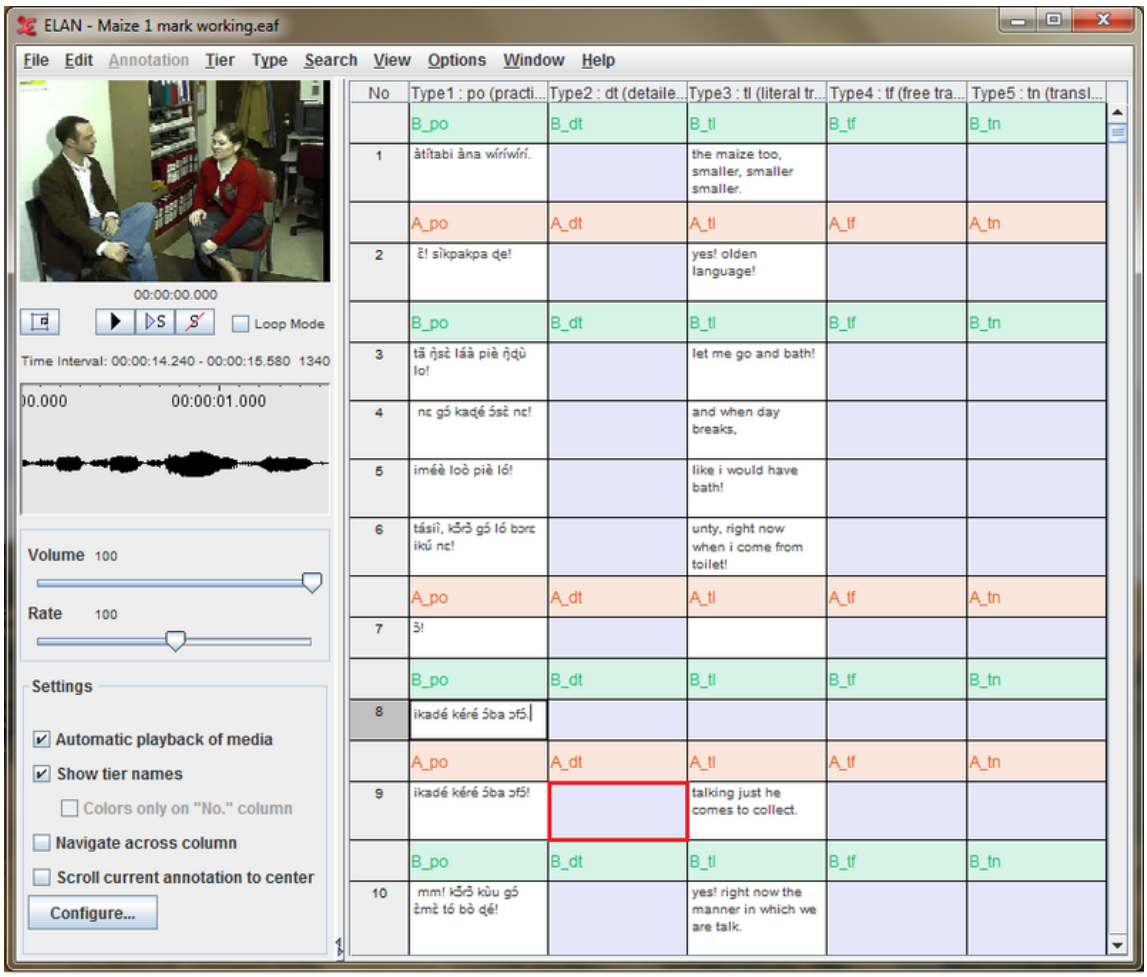
- Another way to open the above dialog is to click Configure... button from the settings panel in the transcription window (see :Figure 6.6). This will popup the transcription mode settings dialog (see : Figure 6.2 and click on the Select tiers... button. This will bring the same dialog (see : Section 6.3.2).

6.3.4. 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 freezed(made non-editable). To freeze/unfreeze the tiers, right click on the tier name (if shown) or right click on any of the freezed or unfreezed annotation of the tier will popup either this menu - see : Figure 6.5 or this menu - see :Figure 6.8. In the popup menu select make this tier non editable to freeze a tier or select make this tier editable to unfreeze the tier.

6.3.5. 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 6.2) will bring this window:



The color filled in this cell with the tier names shown, indicates that there is no actual annotation in the A_dt tier. Putting this cell to edit mode will automatically create an annotation on this tier.

Figure 6.6. 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 othe media player options like, go to next of previous second/pixel/frame are only available as shortcut actions. To see the shorcuts used for this action, goto 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 Section 6.3.8). Default: checked.
  - 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 6.6). If unchecked, colour coding distinguishes different tiers/participants, and hovering over an annotation will give you the tier/participant name (see Figure 6.7). 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.

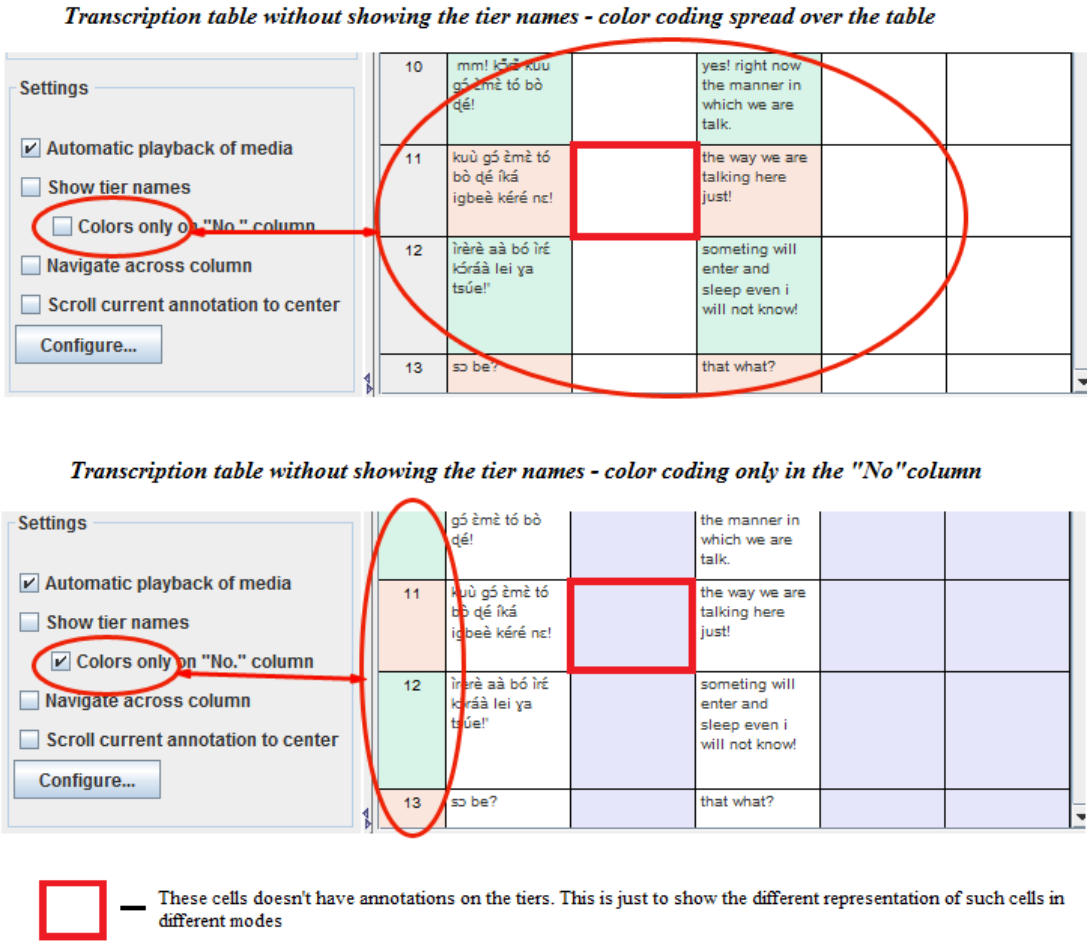


Figure 6.7. Transcription Window without showing the tier names

  - 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 6.2) 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 ENTERto 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 Section 6.3.8.

Other options available in this mode :

- right clicking on annotation will popup this menu.

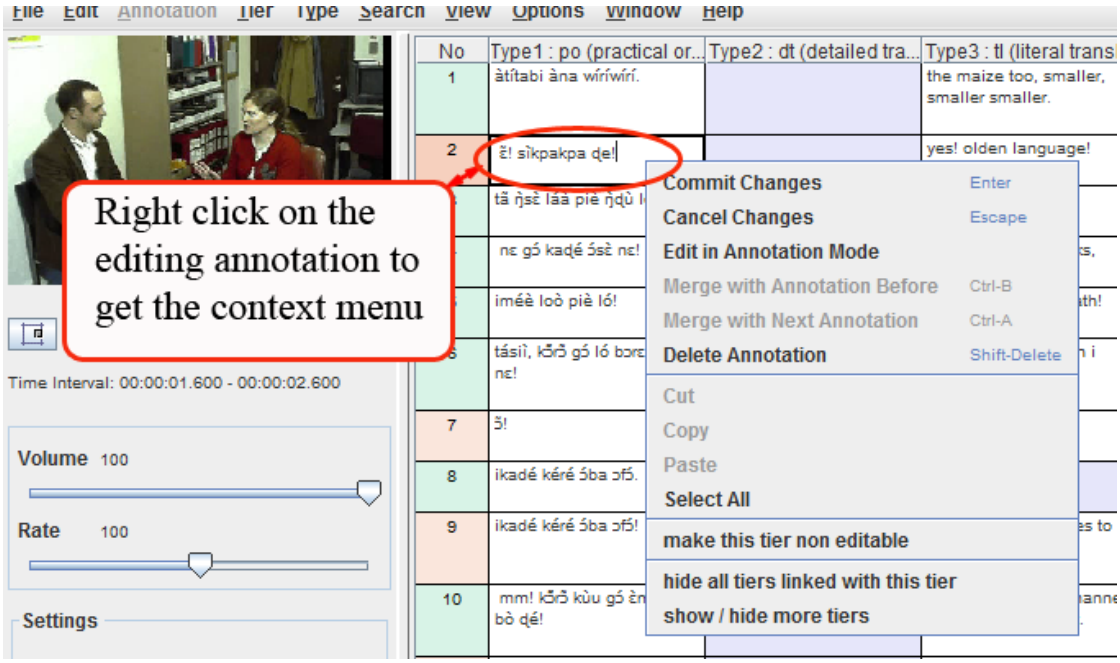


Figure 6.8. Context Menu for a Annotation

– Commit Changes : Saves the changes made to the annotation

– Cancel Changes : Cancels the changes made to the annotation

– Edit in Annotation mode : see Section 6.3.7

– Merge with Annotation Before : see Section 6.3.6

– Merge with Next Annotation : see Section 6.3.6

– 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 Section 6.3.3
- 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.

6.3.6. Merging annotations

To merge two annotations, first make a right click on the annotation. This will pop up a context menu(see Figure 6.8). Then select either Merge with annotation before or Merge with next annotation . For detailed information see Section 5.8.19.

6.3.7. 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 6.8) . 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 Section 5.7.6 and to split annotations see Section 5.8.18

6.3.8. 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 Section 4.4.30

6.4. 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 linguistic 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 Section 9.2.2).

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 linguistic 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 Section 5.5.1.

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).

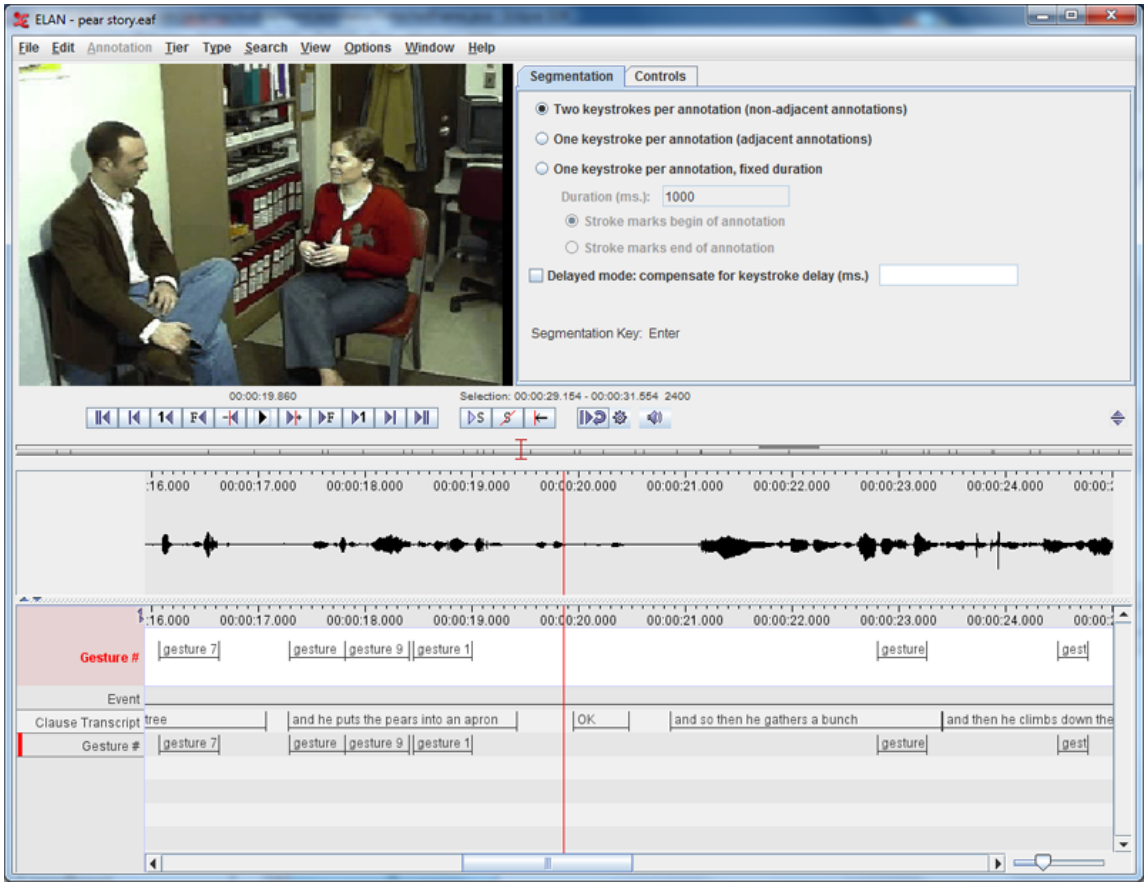


Figure 6.9. Segmentation mode

6.4.1. 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.

6.4.2. 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, or 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.



Figure 6.10. 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.

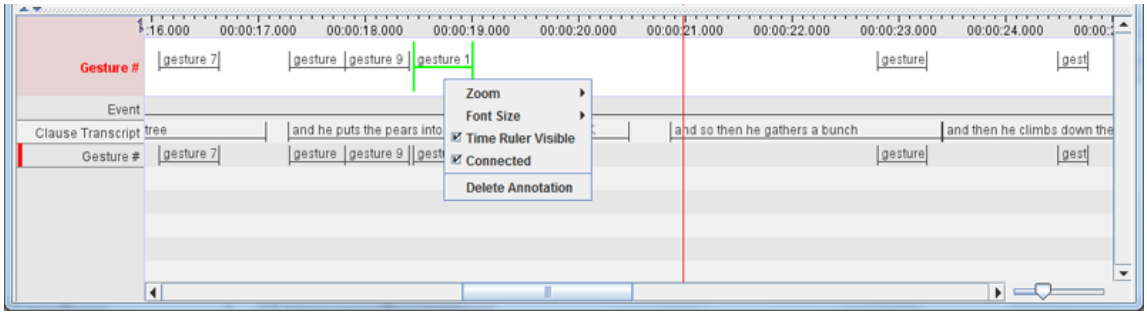


Figure 6.11. The context menu

6.4.3. 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 . Use 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 Section 5.8.19.

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 Section 5.8.18.

6.4.4. 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.

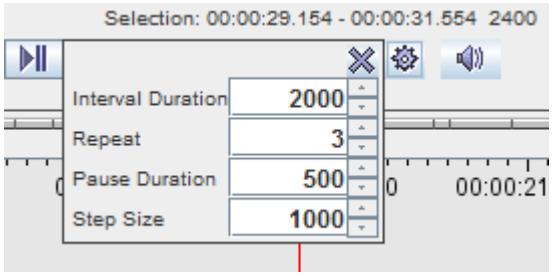


Figure 6.12. Step and Repeat configuration

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.

6.4.5. 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

Chapter 7. Searching

The ELAN tool allows you:

- to search annotations within one document (see Section 7.1)
- to display the results (see Section 7.2)
- to jump to the corresponding annotation the ELAN window (see Section 7.3)
- to search through multiple annotation files (see Section 7.4)
- to do a structured search through multiple annotations files (see Section 7.5)

7.1. 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:

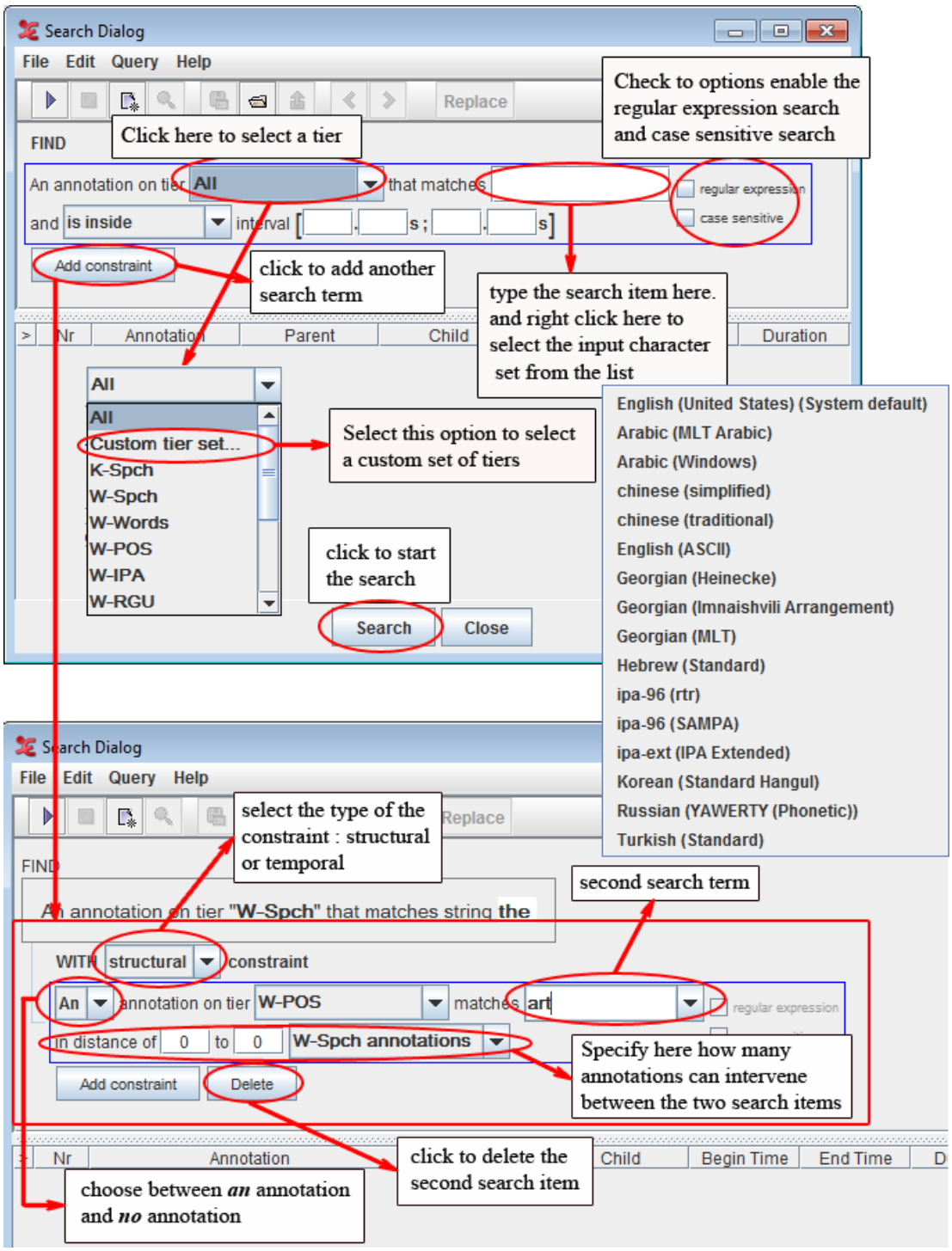


Figure 7.1. Search dialog

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 linguistic type of the selected tier has a Controlled Vocabulary (see Section 5.4.8), 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 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.

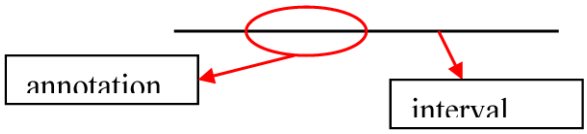
   For example: annotations contained in a structural distance of –1 to 2 tx-annotations from *trees* on the tier tx are *sees*, *trees*, and, *flowers*.

Table 7.1. Annotation example

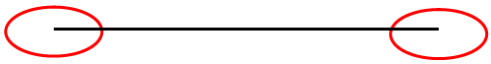
tier	annotations							
st (sentence)	He sees trees and flowers.							
tx (word)	he	sees		trees		and	flowers	
mb (morpheme break)	…	see	-s	tree	-s	…	flower	-s
ps (part of speech)	…	V	SUF	N	SUF	…	N	SUF

2. Constraints based on temporal distance. This means search results are restricted on the basis of the temporal relation between two intervals:

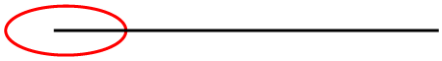
   - *is inside*: the annotation is completely contained within a given interval



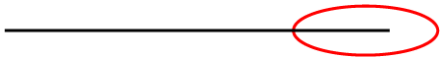
   - *overlaps*: at least a part of the annotation is contained within the given interval



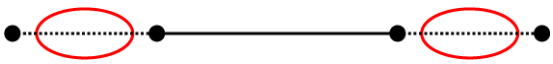
   - *overlaps only begin time of*: the annotation only has its end part in common with the given interval



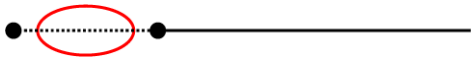
   - *overlaps only end time of*: the annotation only has its begin part in common with the given interval



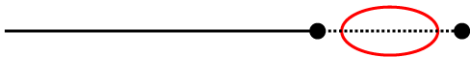
   - *is within … around*: the annotation is contained in an interval around either the begin time or the end time



- *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”.

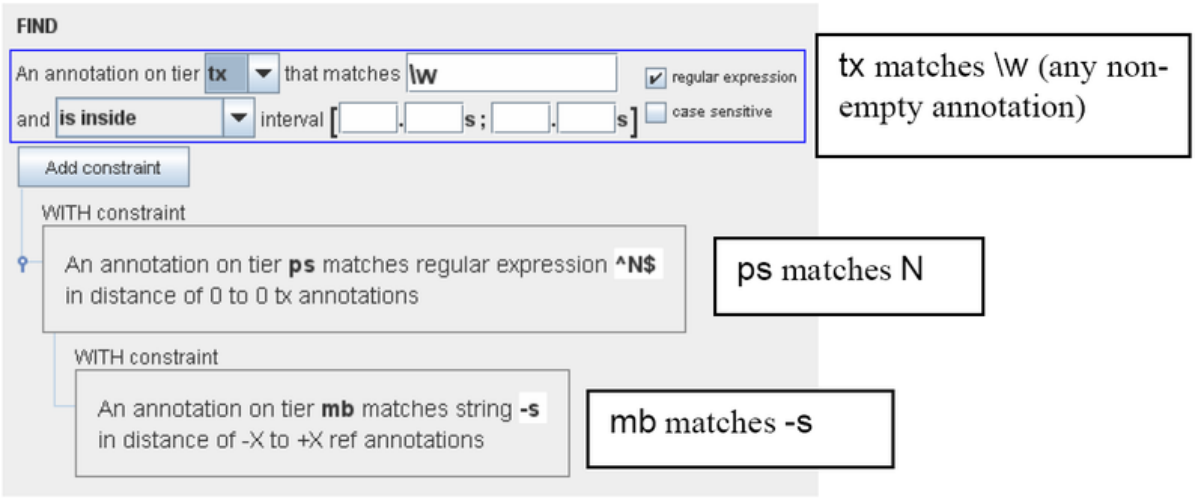


Figure 7.2. Find with constraints

Another option is searching for sequences of utterances, words or other annotations on the same tier, e.g.:

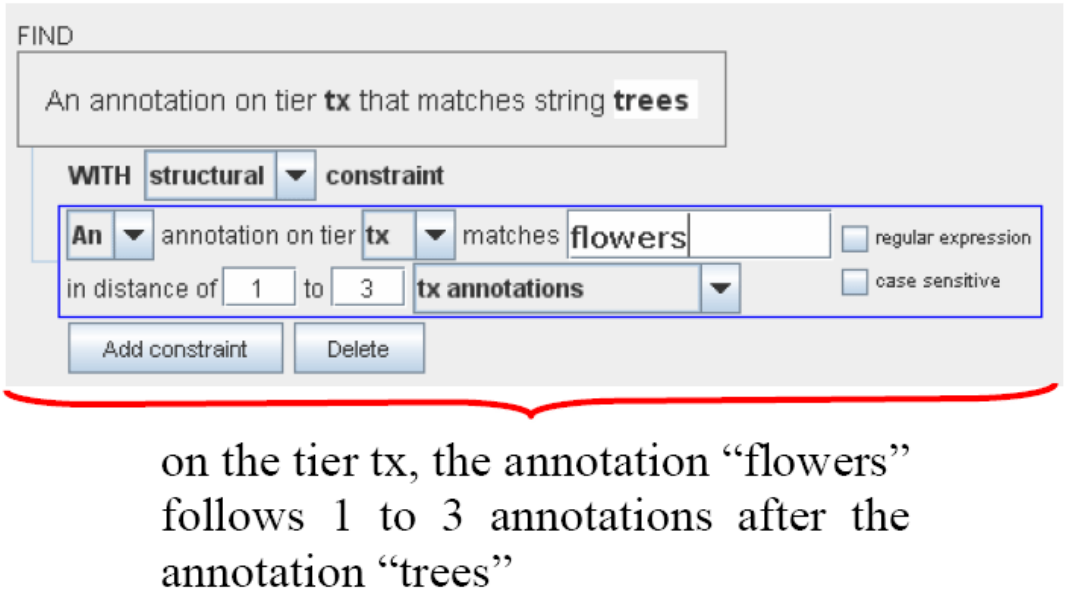


Figure 7.3. Find sequences

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 Section 5.4.4).

   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.

7.1.1. 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:

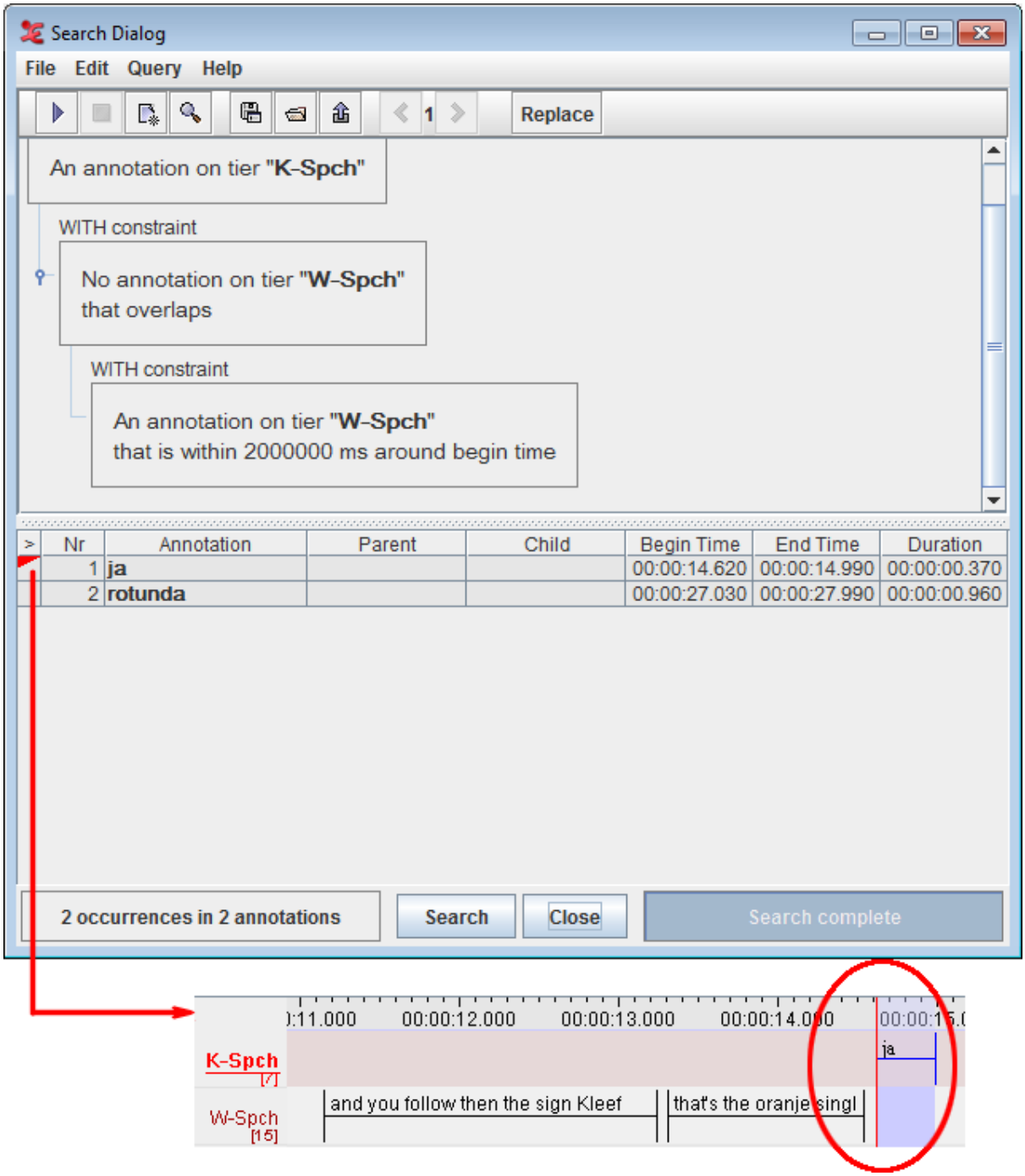


Figure 7.4. Advanced searching example

7.2. 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).

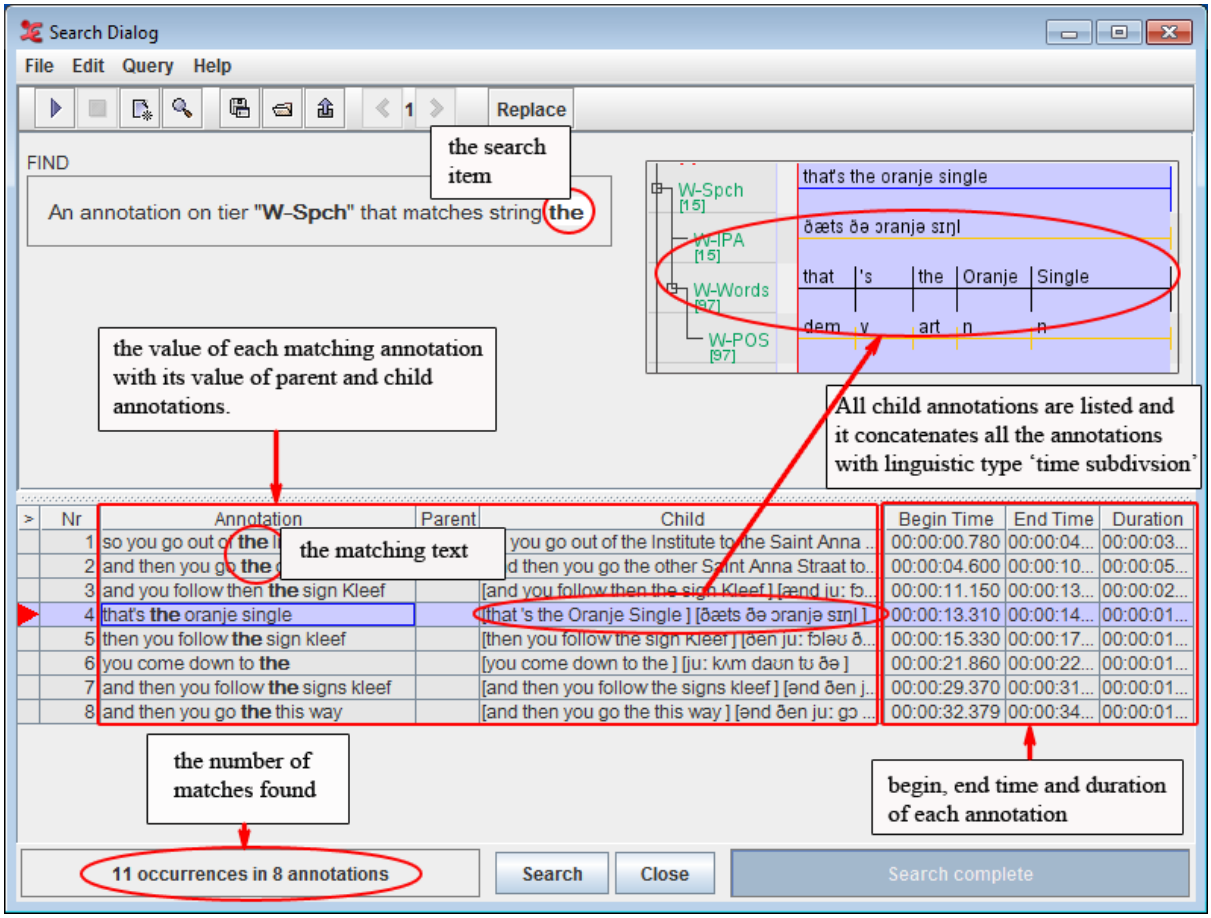


Figure 7.5. Search results

7.2.1. 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.

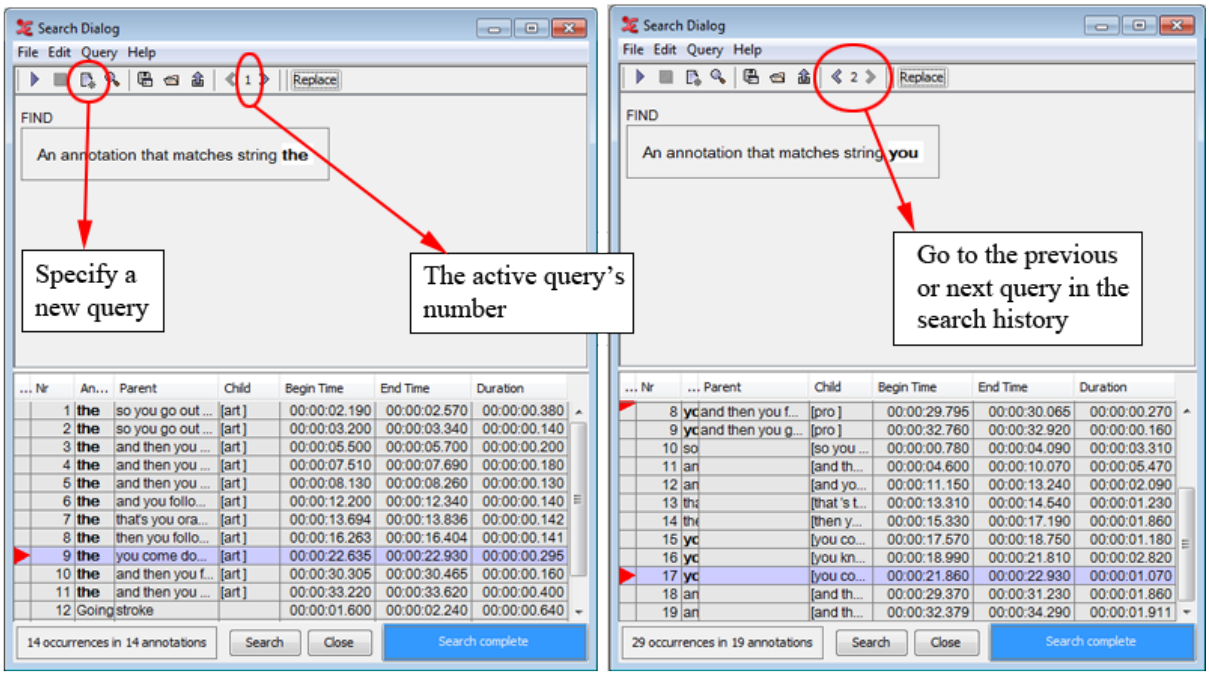


Figure 7.6. Search history



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 Section 7.2.3.

7.2.2. 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:

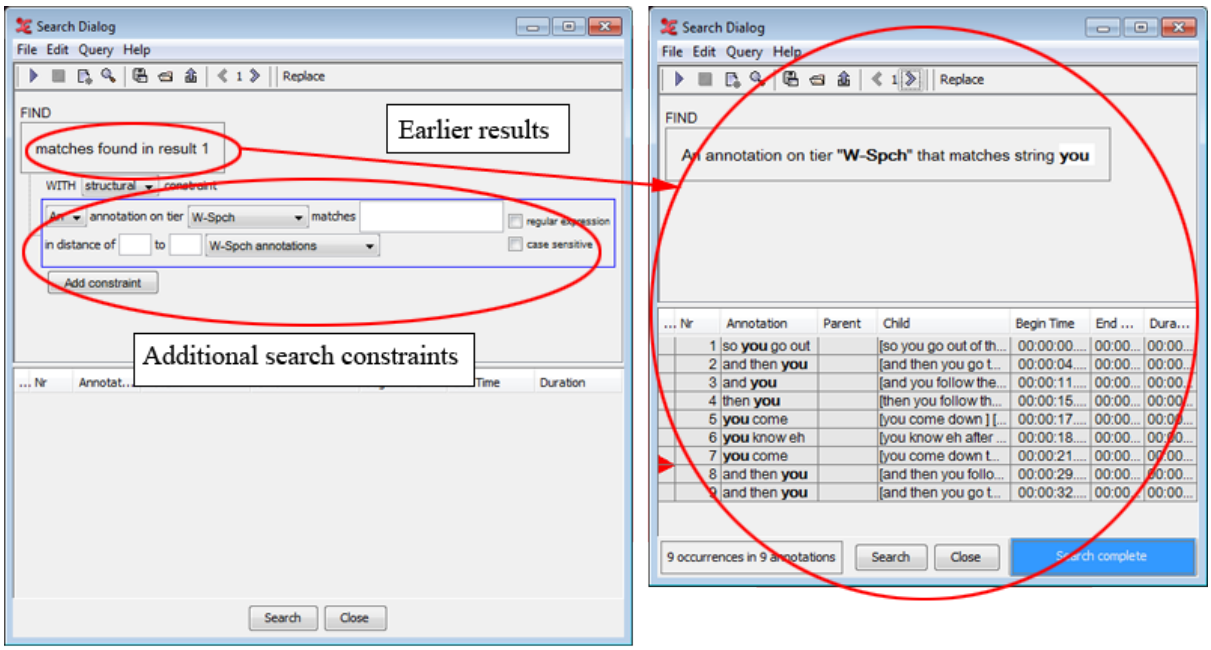


Figure 7.7. Refining a search

By specifying extra search constraints, you can narrow down the results. This is similar to the addition of an extra search constraint.

7.2.3. 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 filename for the active query.
4. Choose save.

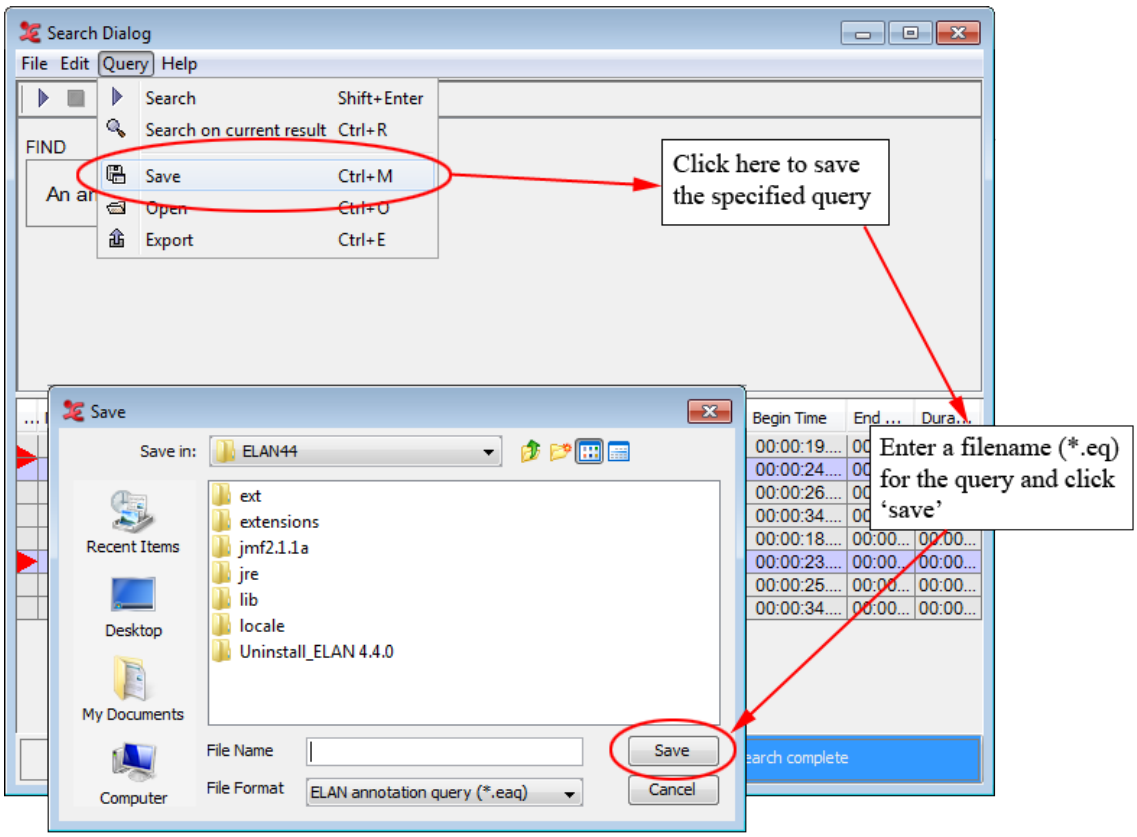


Figure 7.8. Saving a query

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.

7.2.4. 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:

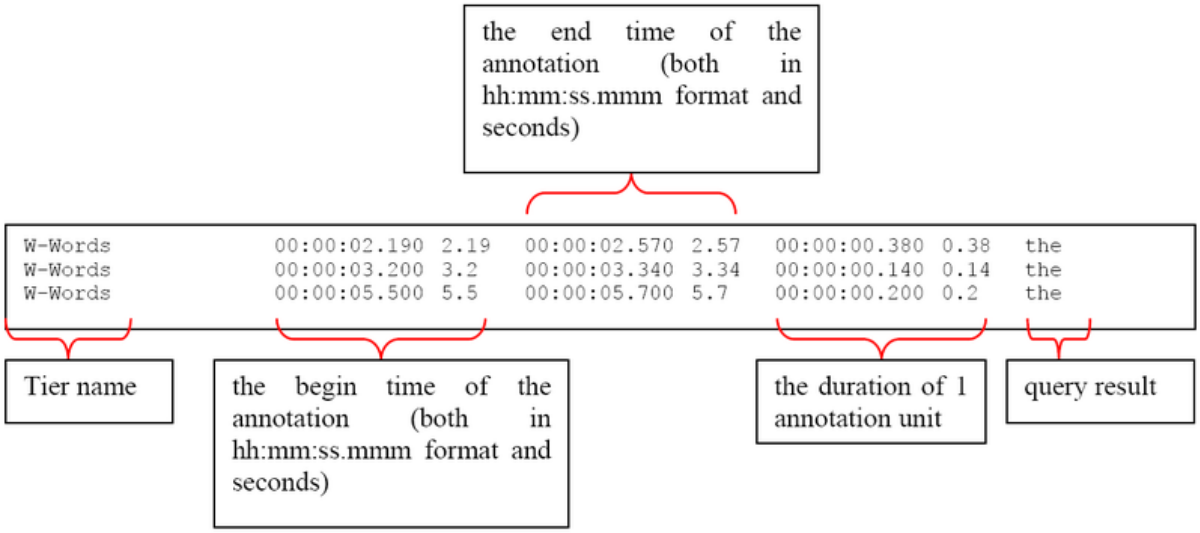
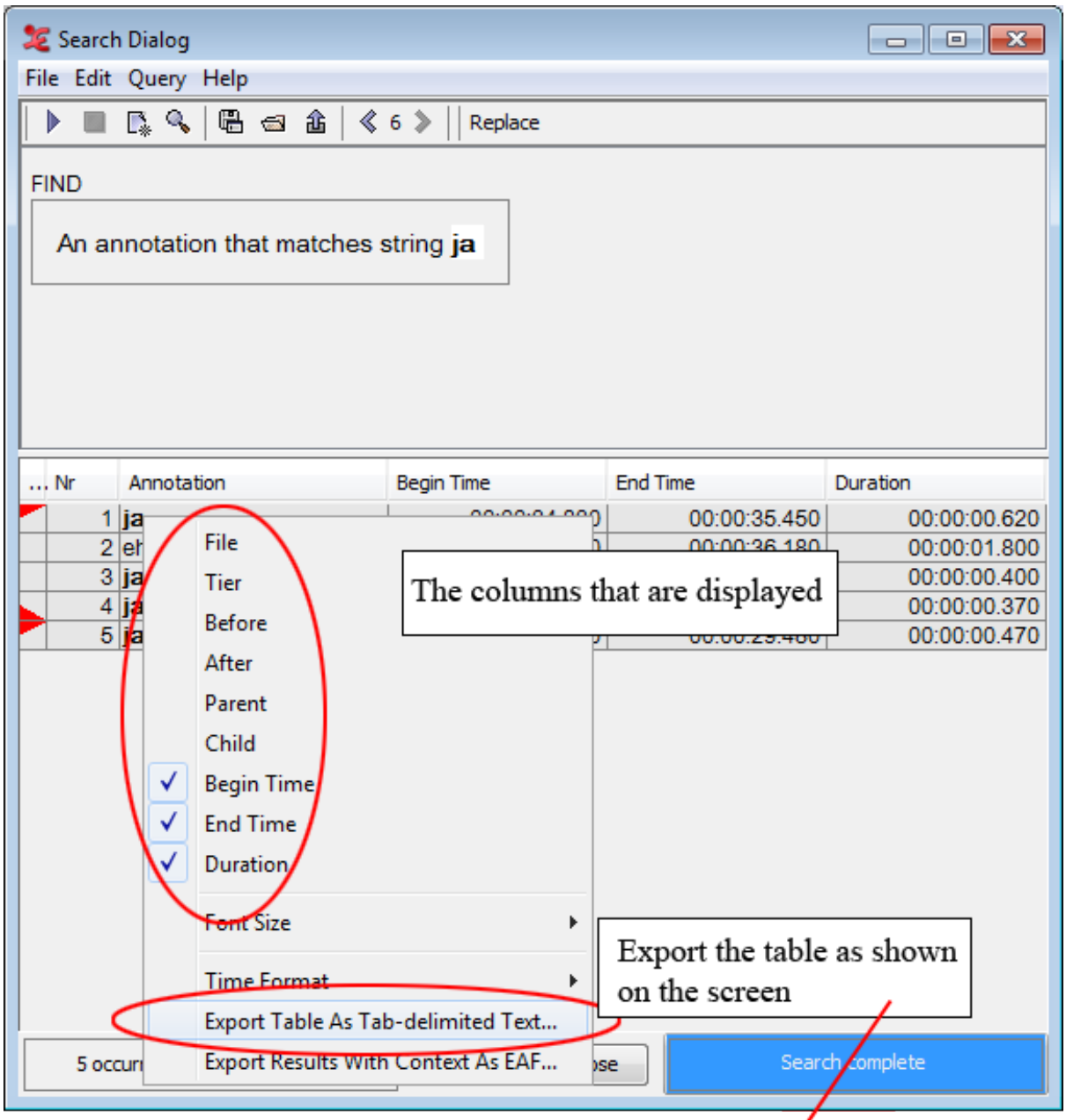


Figure 7.9. Exported search results

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.



Nr	Annotation	Begin Time	End Time	Duration
1	ja	00:00:34.830	00:00:35.450	00:00:00.620
2	eh ja to kleef	00:00:34.380	00:00:36.180	00:00:01.800
3	ja	00:00:10.060	00:00:10.460	00:00:00.400
4	ja	00:00:14.620	00:00:14.990	00:00:00.370
5	ja	00:00:29.010	00:00:29.480	00:00:00.470

Figure 7.10. Export search results

7.2.5. Search and replace

ELAN offers the possibility to replace all of the found search results. Choose Edit > Replace in the Search Dialog:

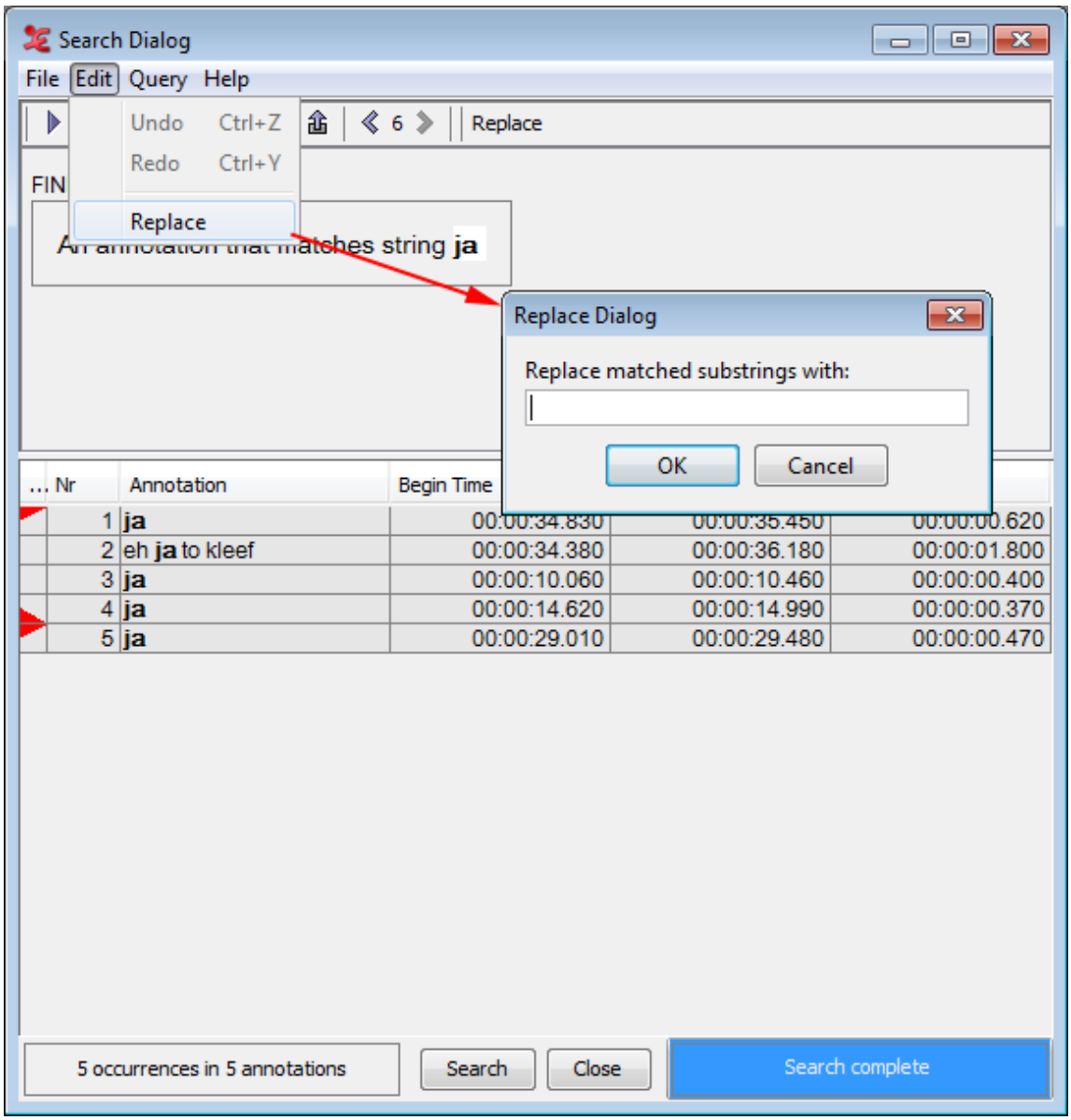


Figure 7.11. Search and replace

Then enter the text that should replace the found results and choose OK

7.3. 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:

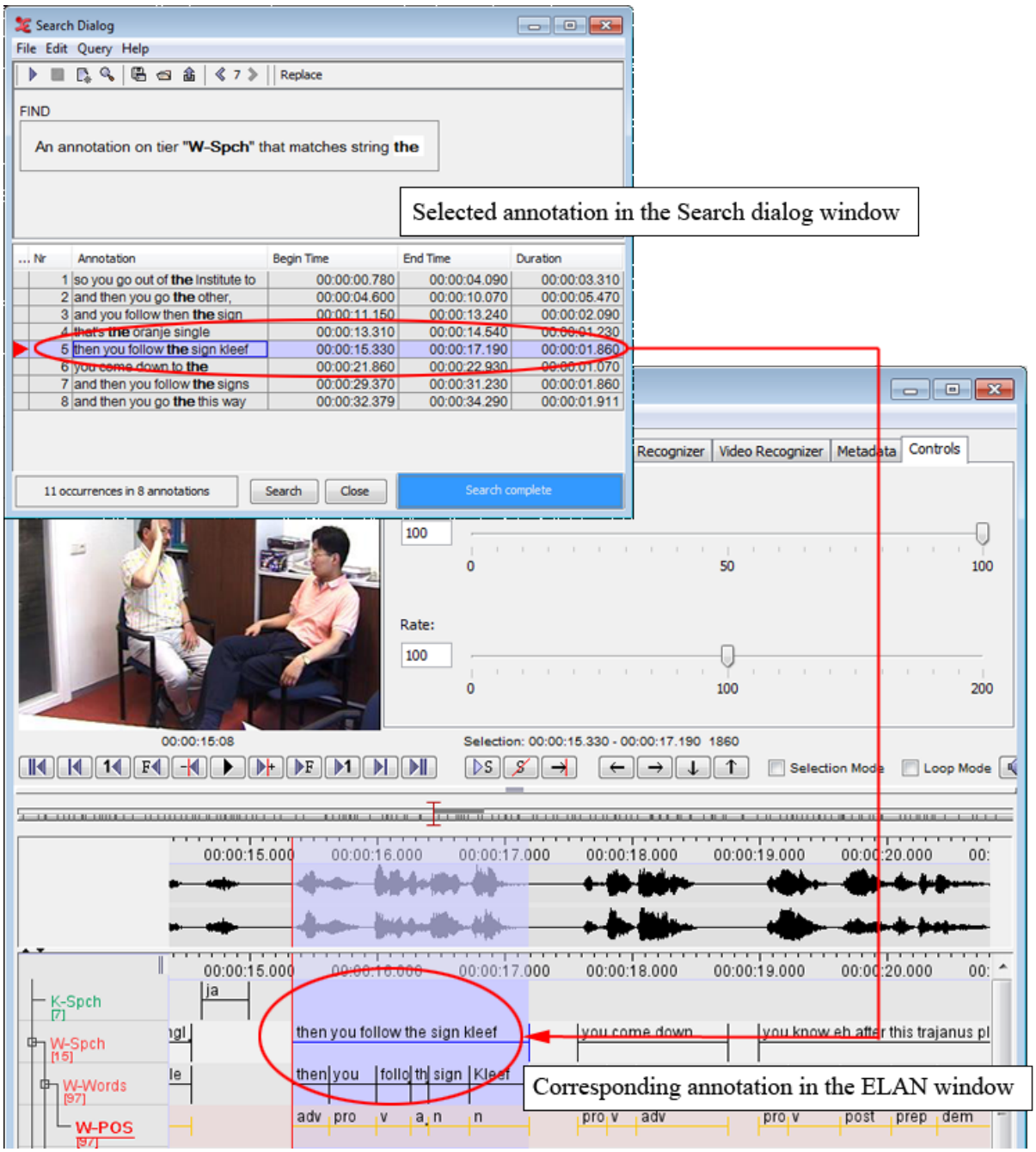


Figure 7.12. From search results to annotations

7.4. Searching through multiple annotation files

Elan offers an option to search for an expression through multiple files. To access it, go to Search > Search multiple eaf… This will open the following dialog box:

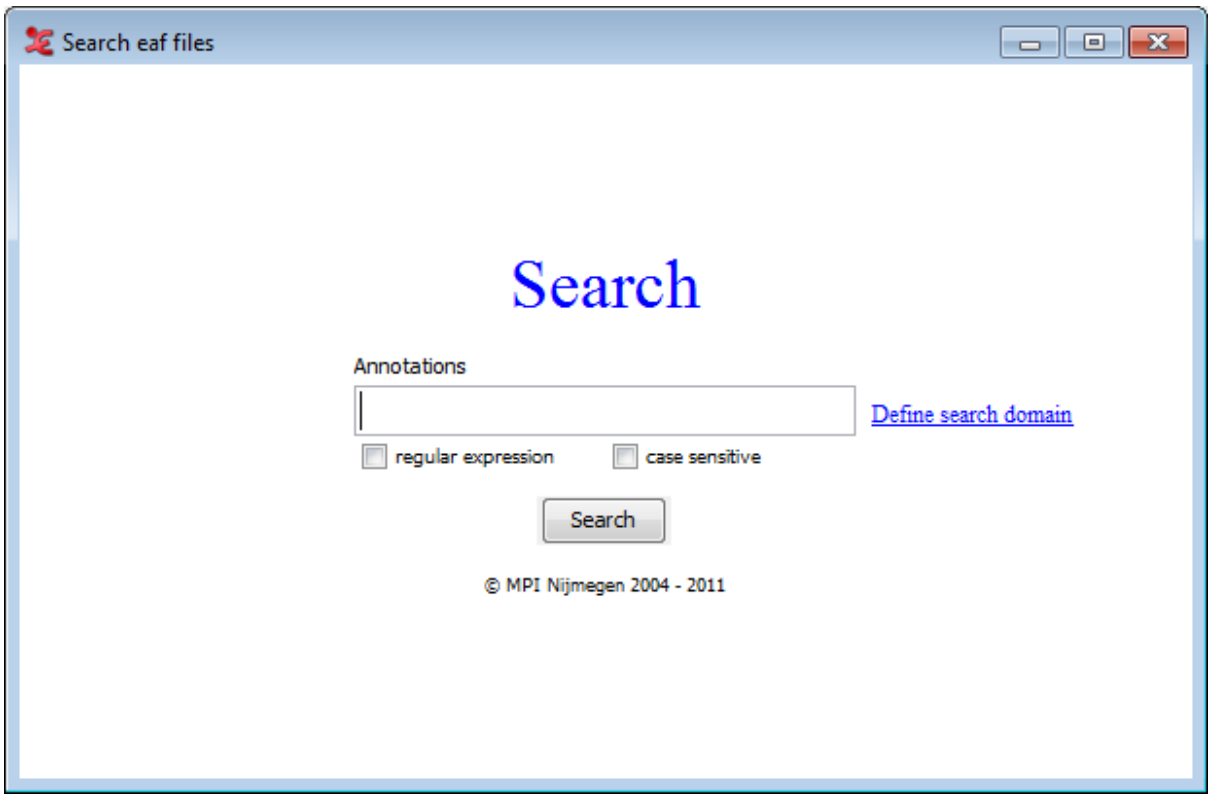


Figure 7.13. Searching through multiple annotation files

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...

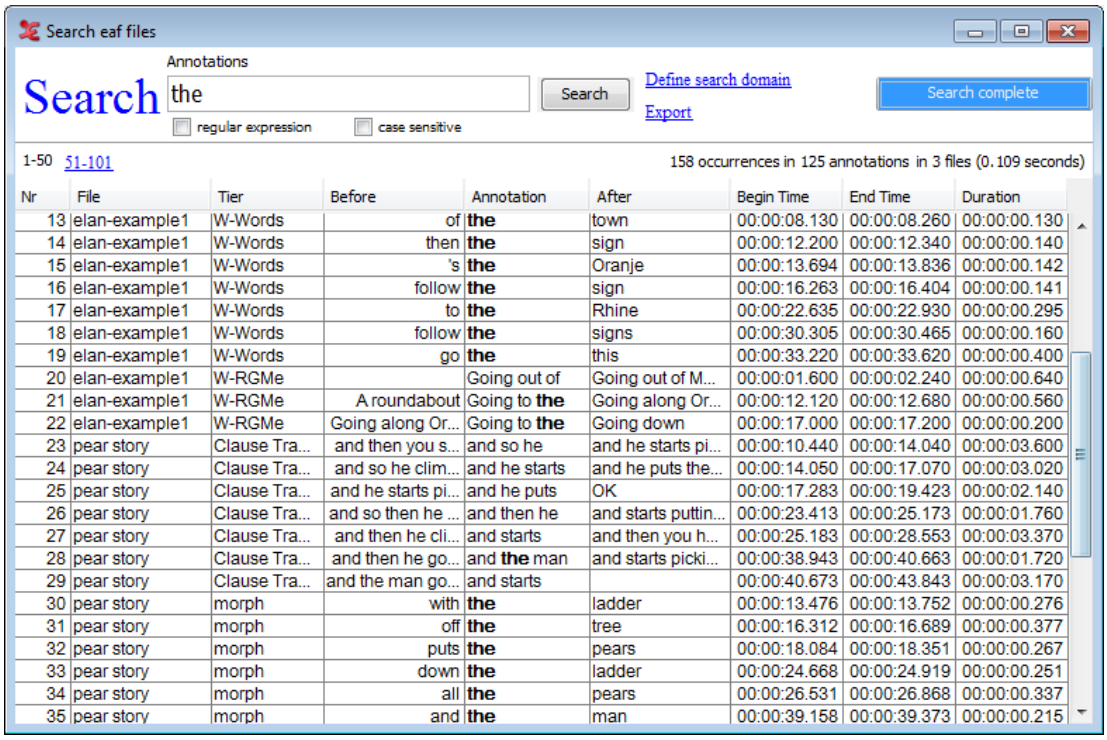
b. Browse to and select an IMDI file that has been exported from a metadata search in the standalone IMDI Browser.

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:



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

Figure 7.14. Searching through multiple annotation files: result

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 Section 7.2.4.

- Just click on Export next to the Search button.
- Enter a filename.
- 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

7.5. 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 Section 7.4) 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.

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:

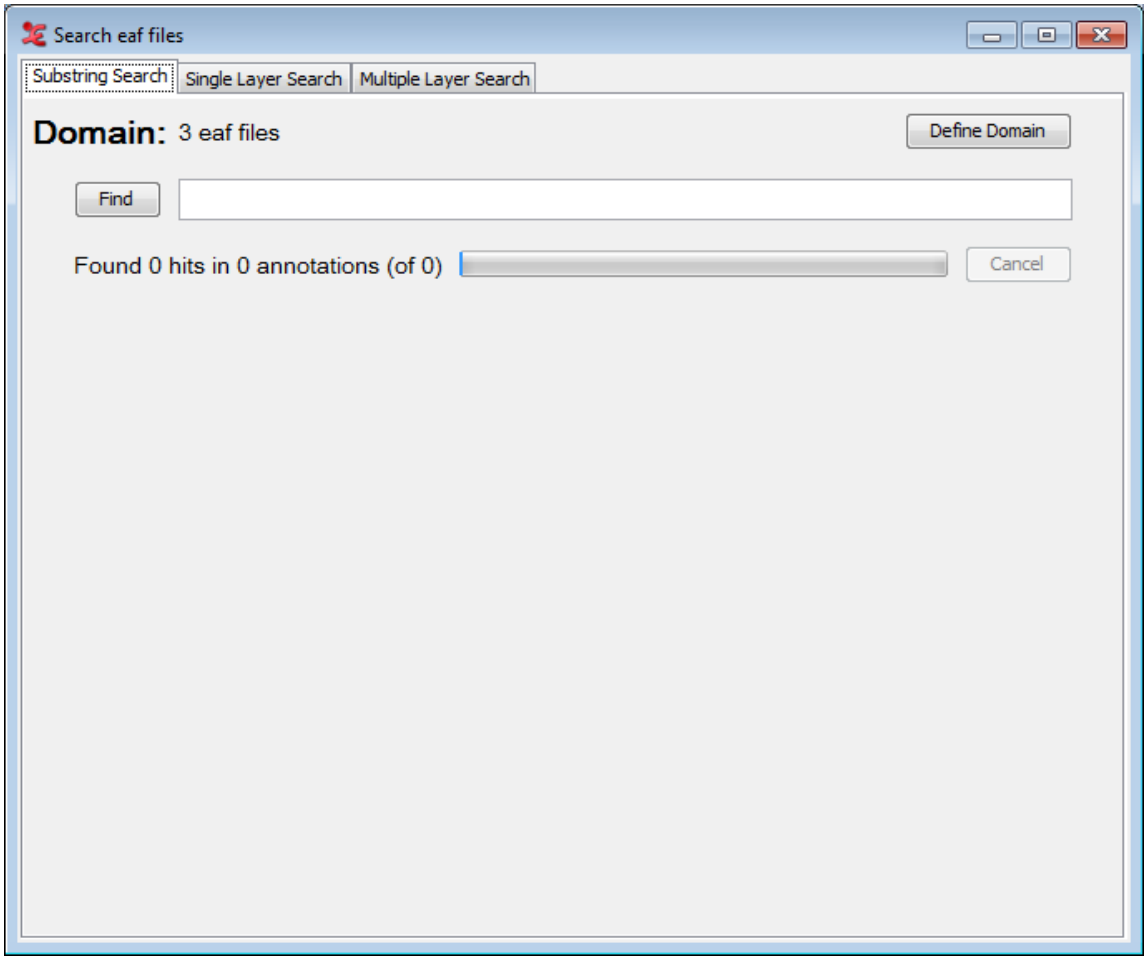


Figure 7.15. Search eaf files

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 Section 7.5.1).

- 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 Section 7.5.2).

- 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 Section 7.5.3).

7.5.1. 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:

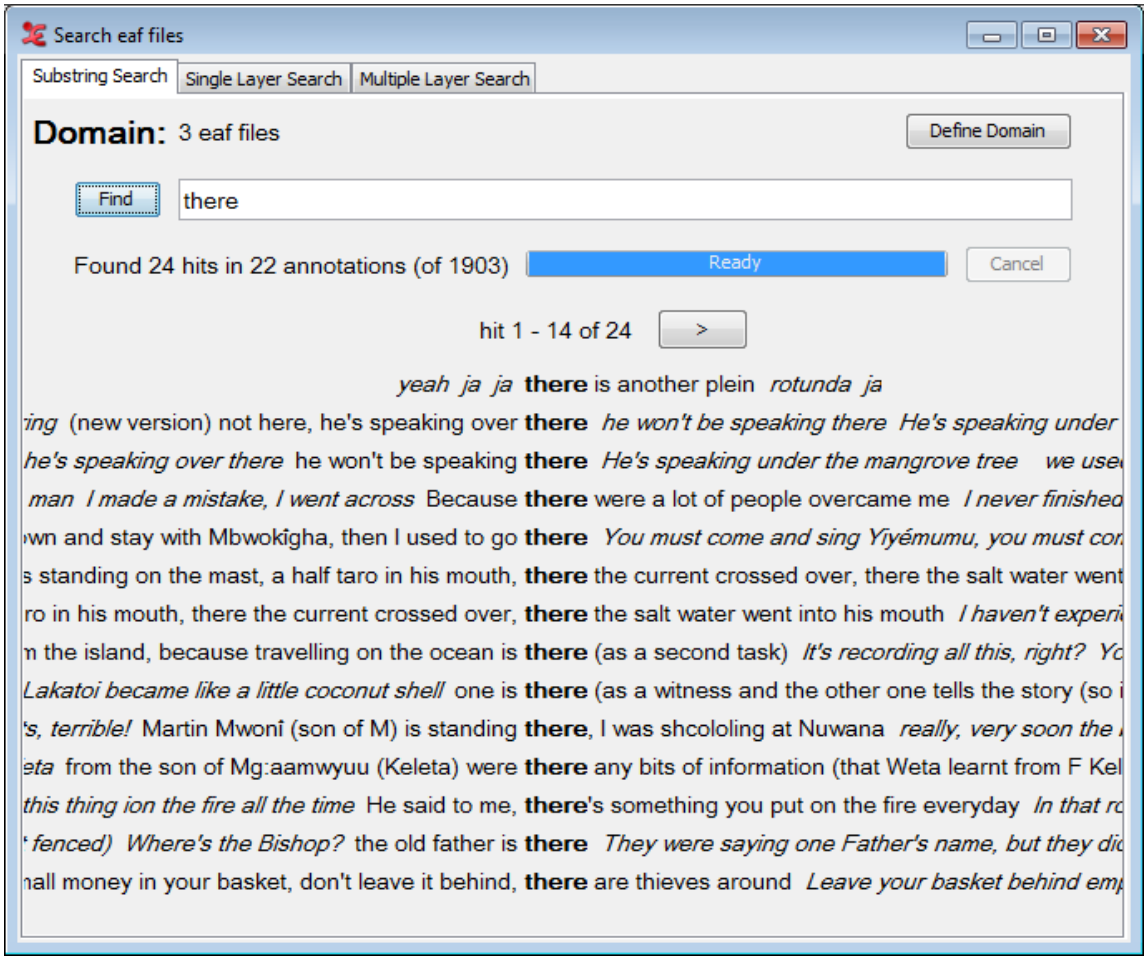


Figure 7.16. Substring Search Results

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:

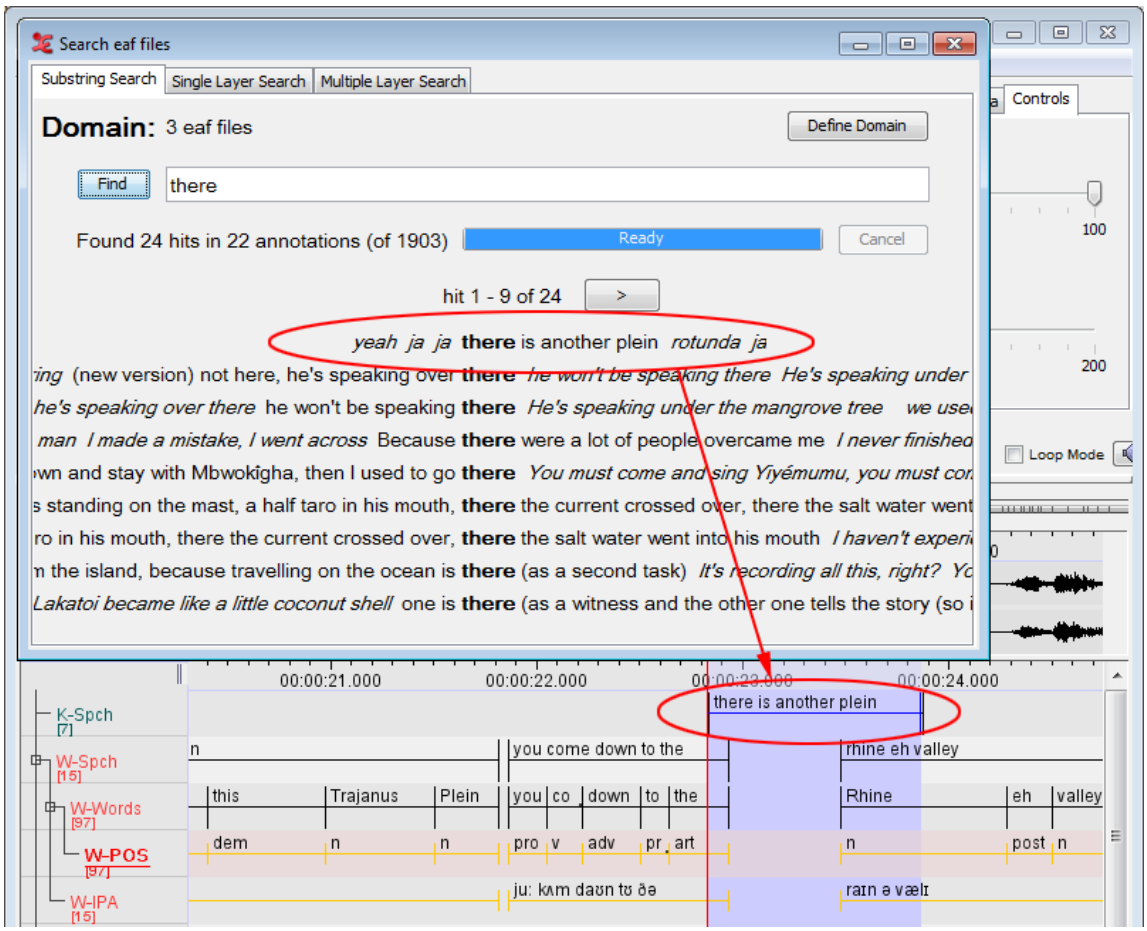


Figure 7.17. Hit in transcription

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 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 filename. 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 filename 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 filename and click Save to save the statistics file.

When you are in the frequency view (Figure 7.18), the context menu 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 filename. The result is a file that contains the following information:

  - Annotation

- Percentage
- Count

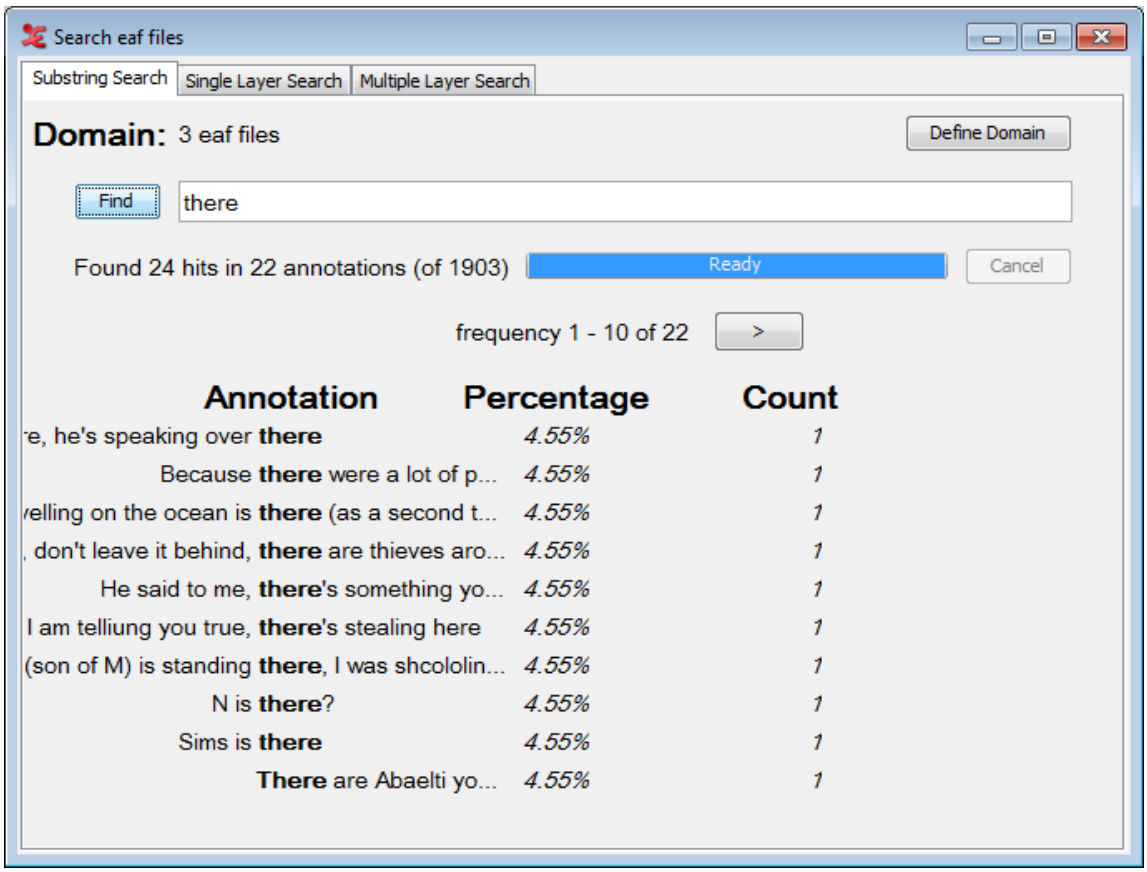


Figure 7.18. Frequency View

7.5.2. 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.

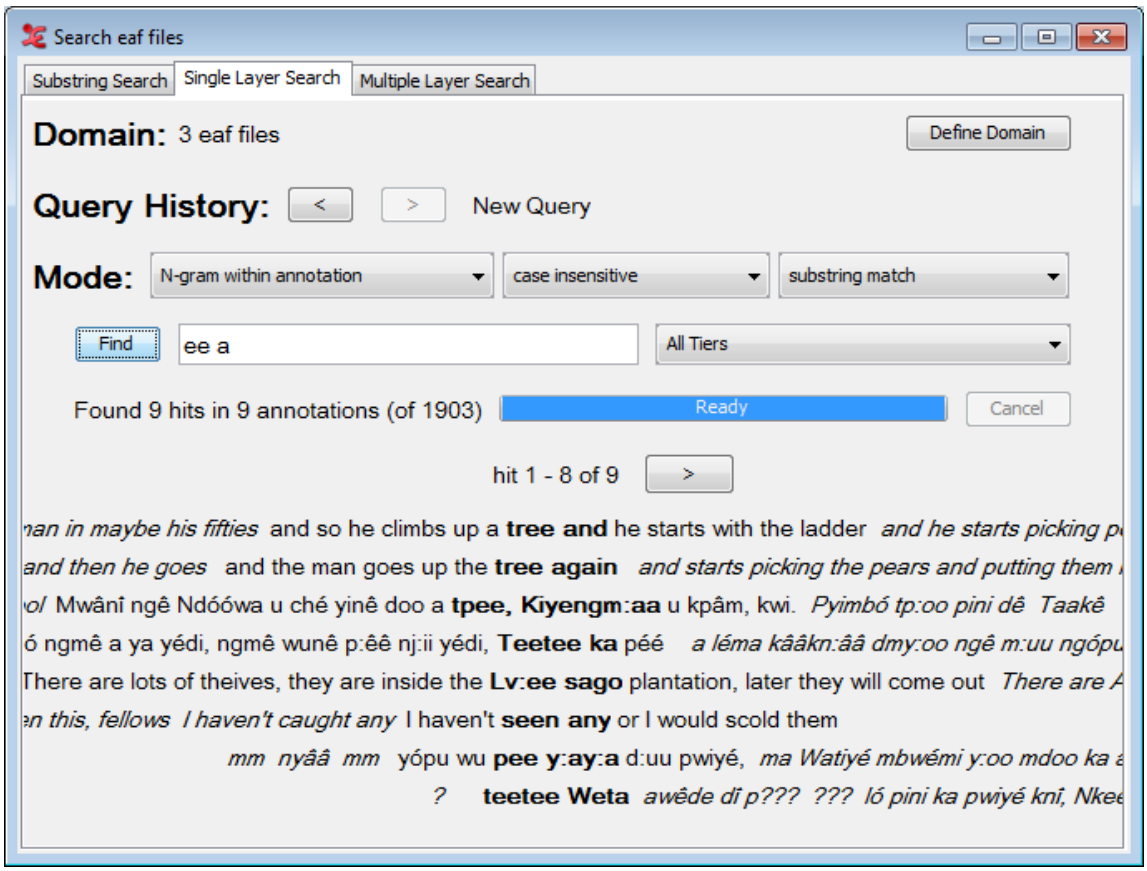


Figure 7.19. Single Layer Search

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.

7.5.3. 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. 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.

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.:

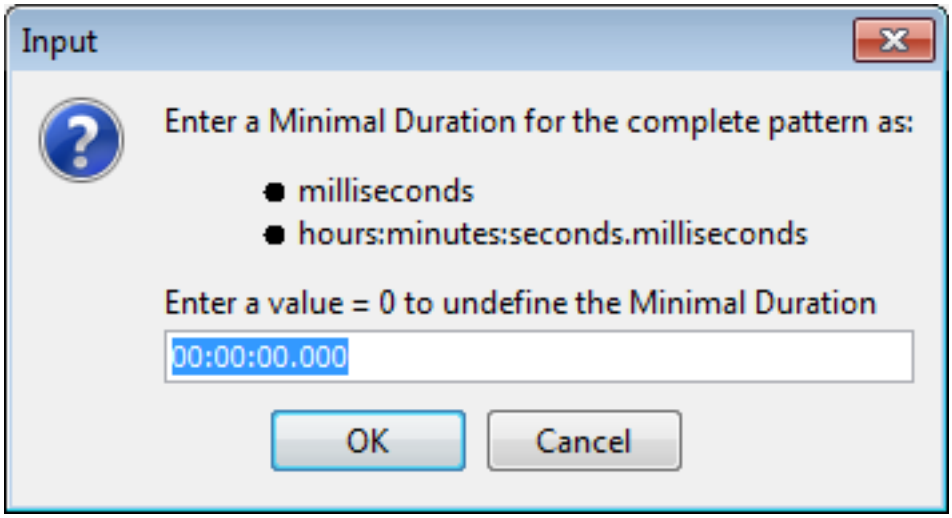


Figure 7.20. Minimal Duration

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 three rows in the query table is that the search engine may find annotations on three 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. Right clicking on the green field gives a context menu offering the following constraints:

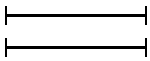
- = N annotations: between the annotations containing the two search strings, there must be exactly N annotations.
- > N annotations: between the annotations containing the two search strings, there must be more than N annotations.
- < N annotations: between the annotations containing the two search 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 exactly 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.

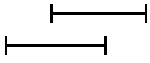
As we saw earlier the search mechanism on this tab has the possibility to construct a query for up to three tiers. Besides the constraints on annotations on a tier, one can also apply constraints on annotations on different tiers. This means that if the search engine has found an annotation that matches 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.

Right clicking the green field between two search strings gives a context menu with the following constraints:

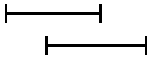
- Fully aligned: the begin time and end time of both annotations are the same:



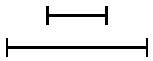
- Overlap: part of both annotations overlap. This includes the other options Fully aligned, Left overlap, Right overlap, Surrounding and Within.
- Left overlap: the begin time and end time of the annotation matching the lower search string lie *before* the begin time and end time of the annotation matching the upper search string:



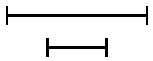
- Right overlap: the begin time and end time of the annotation matching the lower search string lie *after* the begin time and end time of the annotation matching the upper search string:



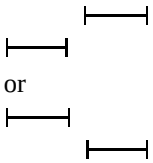
- Surrounding: the begin time of the annotation matching the lower search string lies before the begintime of the annotation matching the upper search string and end time of the annotation matching the lower search string lies after the end time of the annotation matching the upper search string:



- 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:



- 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 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.

- begin time - end 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.

- begin time - end 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 - begin 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.

- end time - begin 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.

- end time - begin 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 - 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 in Figure 7.21:

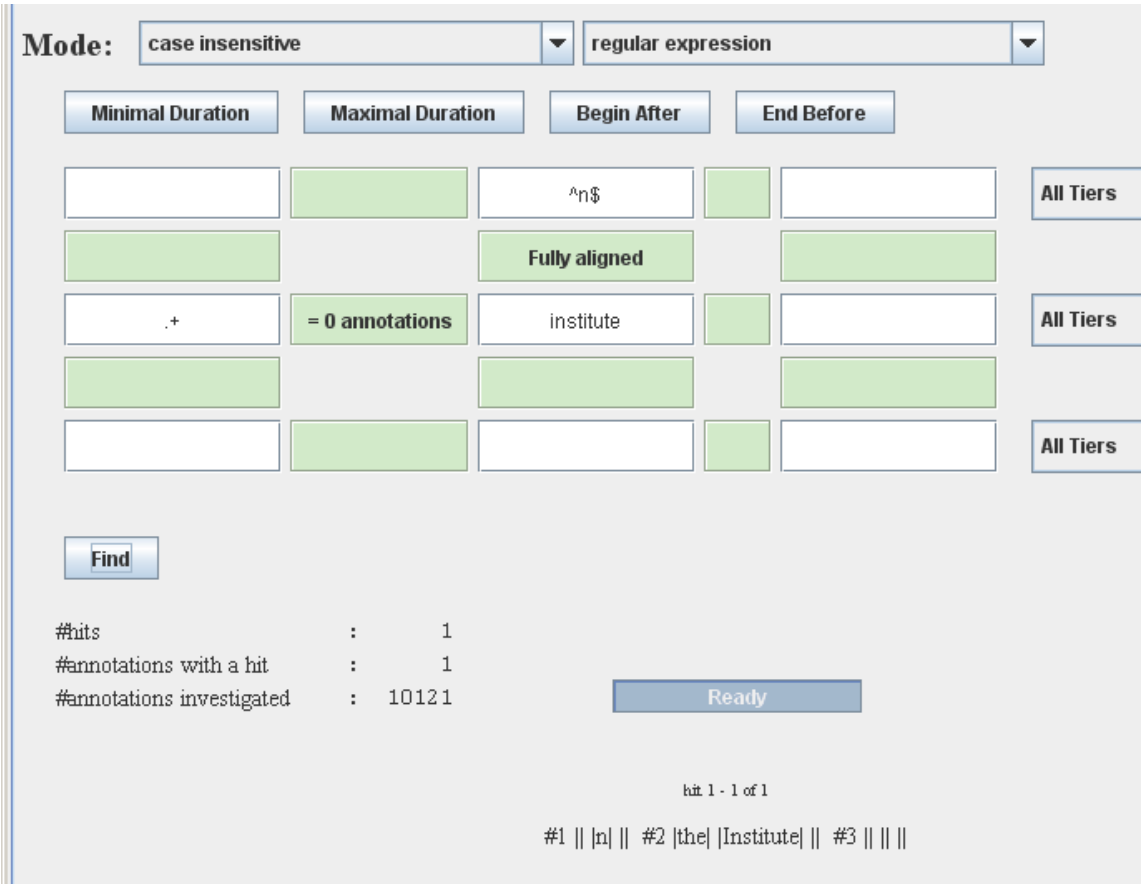


Figure 7.21. Multiple Layer query

Because the search mechanism offers the possibility to search for patterns in three tiers and there are possibly three search strings per tier, the search results also consist of nine elements per hit. Instead of presenting a hit in the form of a table it is presented on one line with indicators for tiers and annotations. Figure 7.21 shows the results of the query above. As you can see the tiers in the result are indicated by #1, #2 and #3 corresponding to the upper, middle and lower query table row respectively. The annotations in a tier are surrounded by vertical bars indicating their start and end.
*Figure 7.21 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 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 Section 7.5.2.

One final but not less important remark concerns the placing of more and less restrictive search strings. Figure 7.21 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$** could also be placed in the lower row and not affect the outcome of the search. While this is perfectly true, we advise you to place restrictive search strings in the left most field on the upper most row possible and the least restrictive search string in the right most field of the lowest row possible. The reason for this is the order in which the search engine considers the search strings in the query. If it finds a restrictive search string it can filter out all the other possibilities, but if it finds a less restrictive search string it has to consider all the matches of this search string. In the example of Figure 7.21 it is clear that if **^n$** 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.

7.6. Search and replace in multiple annotation files

Besides a function to search and replace in the current .eaf file it also possible to search and replace in multiple annotation files. To do so, click Search > Find and replace in multiple files .... Elan first warns you that the changes cannot be undone and that you should have copies of the files you are about to change. If you click Yes, Elan shows the main dialog window. First define the search domain, i.e. the .eaf files to search and replace in, by clicking the Define search domain button. Then 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.

The second step is to select the tiers to search and replace in. You can select either All tiers or Selected tiers. When selecting the latter, click the ... button to specify the tiers. In the new dialog, select the tier(s) in which you want to search and replace and click OK.

Next, enter a search string and a replace string. The search string may be a regular expression (see Appendix A) and may be case sensitive. If it is, check the corresponding checkbox.

Finally, click Find and Replace and start the process by clicking Yes in the warning window. You can watch the progress in the progress bar.

After the process has finished, Elan shows a Process Report that tells you what is replaced in what files.

Chapter 8. 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 Section 4.4.28) . 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 9. 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).

9.1. The mouse options

9.1.1. 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.

9.1.2. 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.

9.1.3. Make a tier the active tier

- Double click on a tier label.
- Right click in the Timeline or Interlinear Viewer and choose Active Tier.

9.1.4. 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.

9.1.5. 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.

9.1.6. 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.

9.1.7. 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.

9.2. The shortcut keys[1]

9.2.1. 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

9.2.2. 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

9.2.3. Working with annotations

ALT+N	Opens the Inline Edit box for inserting a new annotation (only available if a time interval or annotation is selected)
ALT+M	Opens the Inline Edit box for modifying an annotation (only available if an annotation is selected)

Shortcuts in bold mean that they recently have been changed

ALT+SHIFT+M	Opens the Local Data Category Selection for associating the annotation with a data category of the ISO Data Category Registry.
ALT+Delete	Deletes the value of the selected annotation
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 linguistic 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 multiple annotations on multiple tiers.
SHIFT+E	Opens the Inline edit box for the active annotation

9.2.4. 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

9.2.5. Working with selections

SHIFT+SPACE	Play selection

CTRL+SHIFT+SPACE	Play around selection
ALT+SHIFT+C or ALT+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

9.2.6. Searching

CTRL+F	Find
CTRL+SHIFT+F	Search in multiple eaf files

9.2.7. General

CTRL+Z	Undo
CTRL+Y	Redo

9.3. Key concepts of ELAN

9.3.1. 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.

9.3.2. 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.

9.3.3. 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).

9.3.4. 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 linguistic 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[1]

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

Source: http://java.sun.com/j2se/1.4.2/docs/api/java/util/regex/Pattern.html

`.`	Any character (may or may not match line terminators [http://java.sun.com/j2se/1.4.2/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://java.sun.com/j2se/1.4.2/docs/api/java/util/regex/Pattern.html#ubc])
`\p{Lu}`	An uppercase letter (simple category [http://java.sun.com/j2se/1.4.2/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://java.sun.com/j2se/1.4.2/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://java.sun.com/j2se/1.4.2/docs/api/java/util/regex/Pattern.html#cg]

Table A.11. Back references

\n	Whatever the n^{th} capturing group [http://java.sun.com/j2se/1.4.2/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://java.sun.com/j2se/1.4.2/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.

A.1. 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.

A.2. 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://java.sun.com/j2se/1.4.2/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://java.sun.com/j2se/1.4.2/docs/api/java/util/regex/Pattern.html#DOTALL] flag is specified.

By default, the regular expressions `^` and `$` ignore line terminators and only match at the beginning and the end, respectively, of the entire input sequence. If MULTILINE [http://java.sun.com/j2se/1.4.2/docs/api/java/util/regex/Pattern.html#MULTILINE] mode is activated then `^` matches at the beginning of input and after any line terminator except at the end of input. When in MULTILINE [http://java.sun.com/j2se/1.4.2/docs/api/java/util/regex/Pattern.html#MULTILINE] mode `$` matches just before a line terminator or the end of the input sequence.

A.3. 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 ?3.3 [http://java.sun.com/docs/books/jls/second_edition/html/lexical.doc.html#100850] 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.

A.4. 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://java.sun.com/j2se/1.4.2/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://java.sun.com/j2se/1.4.2/docs/api/java/util/regex/Matcher.html] class: Repeated invocations of the `find` [http://java.sun.com/j2se/1.4.2/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://java.sun.com/j2se/1.4.2/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/].