

Using XTrans for Broadcast Transcription

A User Manual

Version 3.0

Linguistic Data Consortium

October 11, 2007

Developer: Haejoong Lee <haejoong@ldc.upenn.edu>
Manager: Stephanie Strassel <strassel@ldc.upenn.edu>

<http://www.ldc.upenn.edu/Projects/Transcription>

1	Introduction	3
2	Tool Overview	3
3	Getting Started	6
3.1	Open Audio File	6
3.2	Open Transcript File.....	7
3.3	Associate Transcript and Audio File.....	7
4	Audio Segmentation	7
4.1	Mark a Region and Insert a Segment.....	7
4.2	Modifying Segments.....	8
4.3	Overlapping Segments.....	8
5	Assign Speaker Information	8
5.1	Creating Additional Segments for Existing Speaker.....	10
5.2	Editing SpeakerIDs	10
5.2.1	Correcting Speaker Information	10
5.2.2	Merging Two Speakers	11
6	Overlapping Speech.....	11
7	Assigning Sections (Story Boundary Labeling)	11
8	Assigning SUs (Sentence Unit Labeling).....	12
9	Transcription.....	13
9.1	Language Input Functions.....	13
9.1.1	Changing Input for Windows Systems	13
9.1.2	Changing Input for Unix Systems.....	14
9.2	Transcript Markup Shortcuts	15
9.3	Spell Check.....	15
10	Other Features.....	15
11	Data Format.....	15
12	Saving and Exporting Transcripts	16
	Appendix 1: Recommended Strategy	17
	Appendix 2: Keybindings	18
	Appendix 3: Sample Transcript Output.....	20

1 Introduction

To support the demand for efficient, careful and consistent transcription, LDC has created a next-generation speech annotation toolkit, XTrans, to support a full range of speech annotation tasks including quick and careful and rich transcription and MDE annotation of broadcast audio, telephone speech and meetings.

XTrans is a multi-lingual, multi-platform transcription toolkit. Powered by Qt's international language support, it can be used for transcription tasks in many different languages. Based on Qt, it can be easily ported to most UNIX derivatives, Microsoft Windows and Mac OS X.

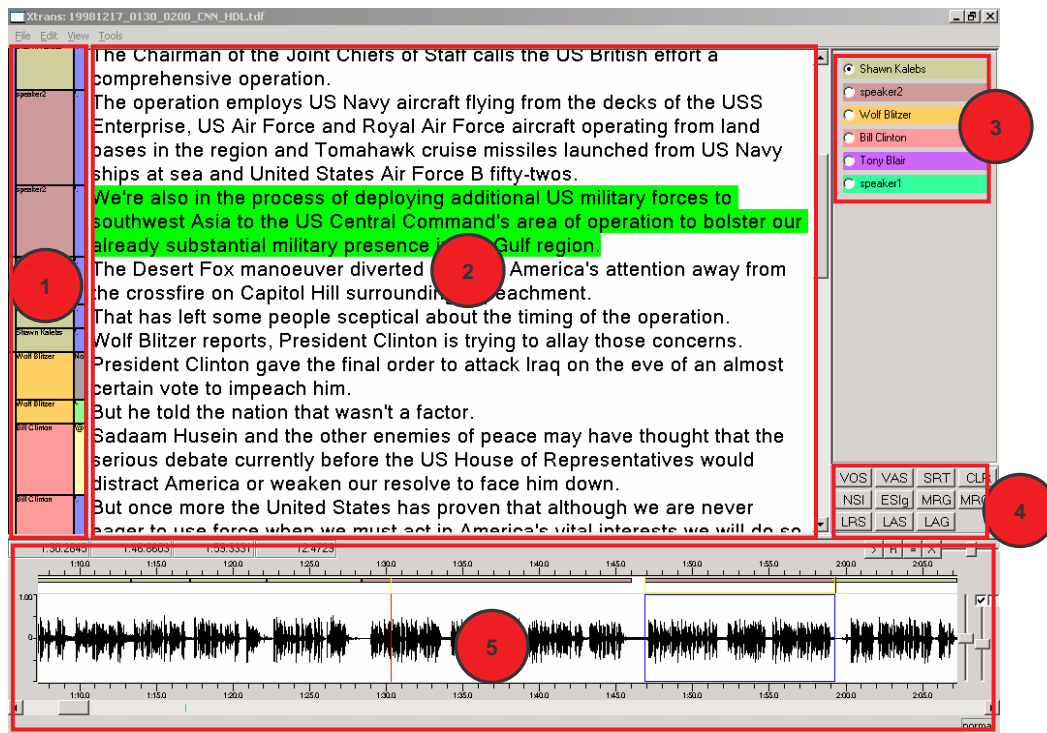
XTrans consists of several re-usable components such as text pane and waveform display. Most of the components are written in Python with some components written in C++.

The current version of XTrans runs on FreeBSD, Linux and Windows platforms.

2 Tool Overview

This section gives a very brief introduction to the tool design and layout.

XTrans is divided into five main areas, labeled on the image below:



(1) The Segment Panel

This region displays three kinds of segments in three separate columns. Each column contains boxes that correspond to segments that have already been created for that type of segment.

The left-most column shows **sections** (i.e., story boundaries)

- Reports (e.g. news stories) are in blue
- Conversational (e.g. interviews) are in green
- Non-news (e.g. commercials) are in light yellow
- Unassigned sections are in grey

The middle column shows **speaker segments**, one color per speaker. The speakerID is also displayed.

The right-most column shows **SUs** (sentence units)

- Statement SUs are shown as dark blue
- Question SUs are shown as red
- Incomplete SUs are shown as green
- Backchannel SUs are shown as yellow
- Unassigned SUs are shown as dark grey

(2) The Transcript Panel

This region shows transcribed text. If you select a region of the transcript (by clicking your mouse somewhere within that segment), it will be highlighted in green and the corresponding audio segment will be highlighted in the waveform display with a blue box surrounding it.

(3) The Speaker Panel

This region shows the list of all speakers assigned to segments in the file. Each name is color-coded to match that speaker in the speaker column of the segment panel display.

(4) The Widget Panel

This region features several shortcut buttons for doing speaker verification and other kinds of quality control. Each button (or widget) has an abbreviation representing its functionality. If you hover over the widget, XTrans will display a longer description of each button's function.

- **MRG** *Merge*: Reassign speaker to another existing speaker for just the current segment
- **MRGg** *Global merge*: Merge two speakerIDs throughout entire file
- **VOS** *View one speaker*: Display only segments for selected speaker
- **VAS** *View all speakers*: Display segments for all speakers
- **SRT** *Sort*: Sort speakers listed on the speaker panel, toggling between chronological and alphabetical display (helpful for doing speaker ID verification and double-checking spelling of speaker names)

- **CLR** *Clear*: Clear the check boxes on the speaker panel
- **NSI** *New speaker information*: Assign new speakerID for just the current segment
- **ESlg** *Edit speaker information - global*: Change speakerID information for the selected speaker throughout the whole file
- **LRS** *Listen random segment*: Listen to a random segment from this speaker (helpful for speakerID verification)
- **LAS** *Listen all segments*: Listen to all segments from this speaker (helpful for speakerID verification)
- **LAG** *Listen all gaps*: Listen to the unsegmented audio "gaps" (helpful for doing quality control, to catch unsegmented speech)

(5) The Audio Panel

This region includes the waveform display and audio control buttons.

The audio control functions at the upper righthand portion of the audio panel include (from left to right): play, repeat, pause/restart, and stop.

Next to the audio control functions is a horizontal slider bar that controls playback speed. The default playback speed is realtime playback. Sliding the bar to the right speeds up playback; sliding it to the left slows down playback¹. Beside the slider bar is a button that displays the playback speed. Clicking this button restores the playback speed to realtime.

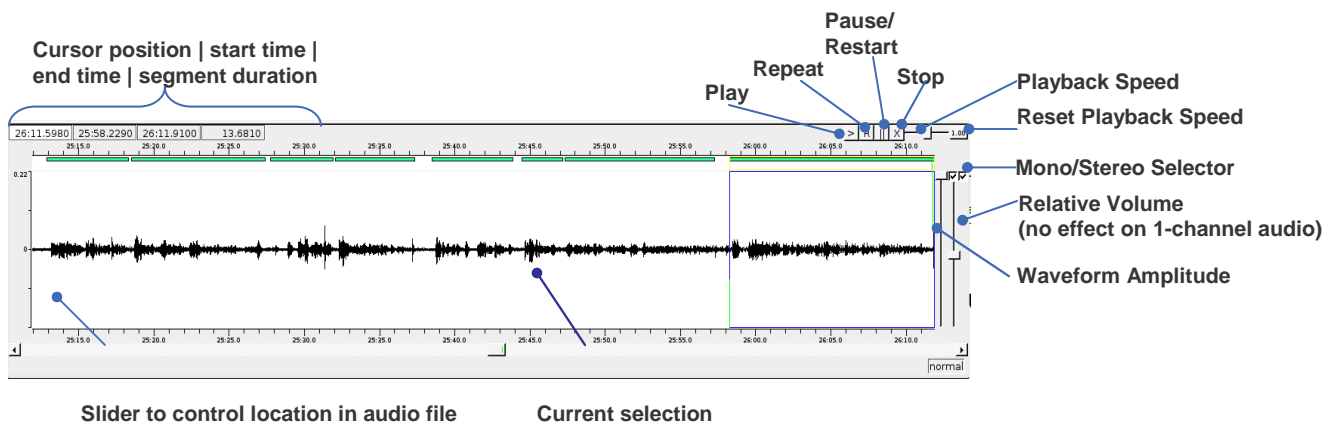
To the right of the waveform display are two vertical slider bars. The first of these controls waveform amplitude display. The second bar controls "relative" audio playback volume. It is relative because it doesn't change the overall volume, but changes the distribution of loudness over the input audio channels. Sliding this bar to the very bottom will mute the channel. Above this second vertical bar are two checkboxes that control mono vs. stereo playback. The default setting for broadcast (single-channel) audio is mono (one box checked), but transcribers can check the second box to enable stereo playback – helpful if you are using headphones.

In the upper lefthand corner of the waveform display you will notice a series of four numbers. The first number shows your current cursor placement. The second and third numbers indicate the start and end time of your current selection, and the fourth number indicates the duration of your current selection.

The long horizontal slider bar running along the bottom of the waveform display adjusts your position in the audio file. Numbers displayed at the top and bottom of the waveform indicate time in minutes and (thousandths of) seconds.

¹ We are currently investigating existing technology for adjusting playback speed without changing sample rate.

The image below summarizes these features.

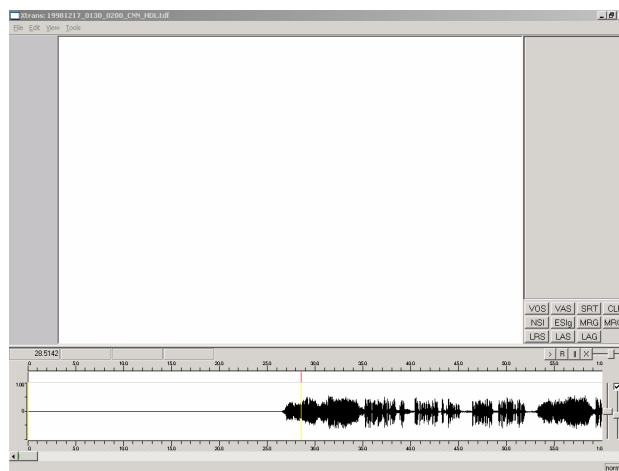


Before you begin transcription, you should also familiarize yourself with the most common playback and waveform display and manipulation functions of XTrans. While all of these functions can be achieved using the mouse in the standard way, transcribers will find it faster to use the keyboard only for most tasks. Default keybindings for XTrans are described in Appendix 2.

3 Getting Started

3.1 Open Audio File

On the file menu, pull down "Open audio file". Currently XTrans accepts .sph and .wav audio format². After you load an audio file, your display will look something like this:



² To be precise, XTrans supports any audio formats that libsndfile supports, which includes more than dozen of file formats. Unfortunately, mp3 is not included; however, we plan to add mp3 support in future versions.

Multiple time-aligned audio files can be loaded (e.g., for meeting or telephone speech transcription). Note that XTrans will display audio file information (file name, path and sampling rate) when you right click on the waveform display.

3.2 Open Transcript File

If you are starting a new transcript, select "New" from the file menu.

If you are returning to an existing transcript, select "Open" from the file menu. Currently XTrans can directly open files in its native .tdf format.

To open an existing transcript in Transcriber (.trs) format, select "Import" > "Transcriber" from the file menu.\

3.3 Associate Transcript and Audio File

Before you begin transcription you must first "link" the audio with your transcript. To do this, select "Edit" > "Blindly associate transcript to audio".³ You cannot proceed without doing this step.

If you have opened an existing transcript, you may also select "Edit" > "Blindly associate transcript to audio". However, it's also possible to use the "Edit" menu to manually "Associate transcript to audio", to specify which speaker is associated with which audio channel. This is most useful when working with meeting recordings with more than two speakers.

4 Audio Segmentation

The first stage of transcription is segmentation, which refers to the process of virtually chopping a long audio file into smaller meaningful units, e.g. sections, speaker turns, sentences or other salient units⁴.

Segmentation can be done with the keyboard only, with the mouse only, or with a combination of both. After you've become familiar with basic XTrans functionality, you will find that using only the keyboard is both faster and more intuitive than using the mouse.

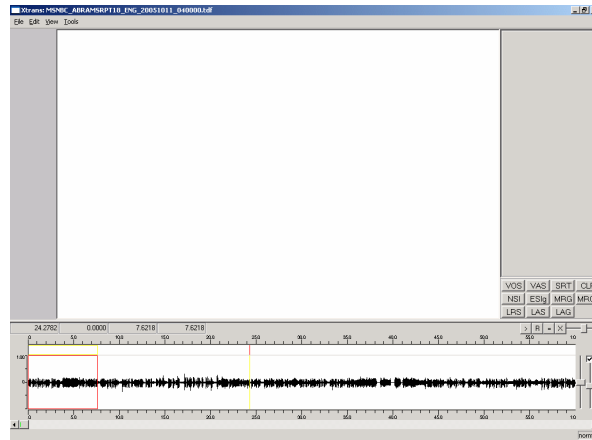
4.1 Mark a Region and Insert a Segment

To begin, place your cursor at the start of the waveform and type `Alt+M`. This starts playback and marks the start time of your selection. When you reach the end of a segment you wish to mark (for instance, the end of a sentence), type `Alt+M` again⁵. This will stop playback and also mark the region you have just played, surrounding it with a red box in the waveform display:

³ This step will be done automatically in future versions of XTrans.

⁴ For Quick Transcription of broadcast data, we recommend starting with sentence-based segmentation.

⁵ Selecting a region to segment can also be done with the mouse (holding down the left mouse button and swiping over the selected region, but it is strongly recommended that transcribers use keyboard functions for segmentation instead of the mouse.



After you've marked your segment, you must insert it into the transcript. To do this, hit `Ctrl+N` (or on *NIX, `Ctrl+Insert`). (To proceed with segmentation, go to Section 5).

4.2 Modifying Segments

To delete an existing segment, hit `Ctrl+D` (or on *NIX, `Ctrl+Delete`).

To join or split two existing segments, type the following commands:

<code>Ctrl+J</code>	Join the current and following segments ⁶
<code>Ctrl+Return</code>	Split segment into two at cursor point

You can also correct or change the start or end boundary points of an existing segment in several ways. With the mouse, hold down the `Shift` key then place your cursor on the start or end boundary you wish to modify, and drag it to the correct location. XTrans automatically adjusts the segment start and/or end time to your new selection.

Segments can also be adjusted without the mouse, using the standard XTrans keybindings described in the Appendix.

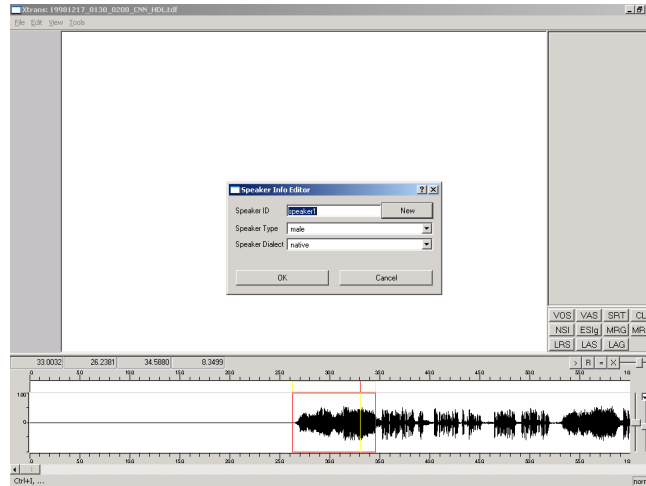
4.3 Overlapping Segments

XTrans prohibits the user from creating overlapping segments for a single speaker. However, segments from two different speakers can overlap. See Section 6 for a full treatment of overlapping speech segmentation.

5 Assign Speaker Information

As soon as you insert a new segment, XTrans prompts you for speaker information for that segment.

⁶ Note that this operation joins two consecutive segments **for the same speaker**.



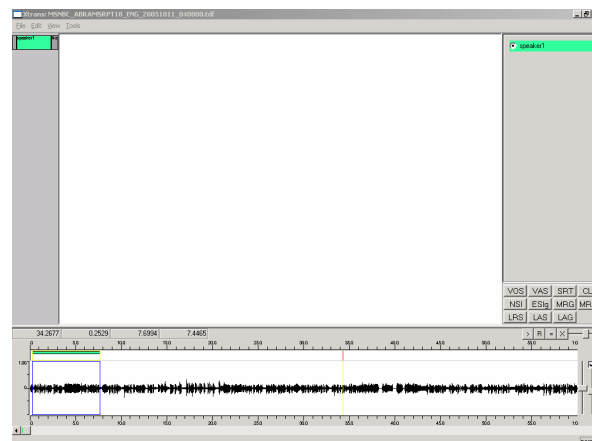
There are three fields to complete:

Speaker ID	name or unique identifier for speaker
Speaker Type	male, female, child or unknown
Speaker Dialect	native, other (for non-native or foreign)

If you know the name of the speaker, type it into the SpeakerID box. Otherwise, use an anonymous ID. To assign a new anonymous SpeakerID, click the "New" button in the SpeakerID dialog box. XTrans will automatically assign the next unused ID, in the form speaker1, speaker2 etc.

As soon as you have completed the speaker information, XTrans will record this information and update the transcript display with three new pieces of information:

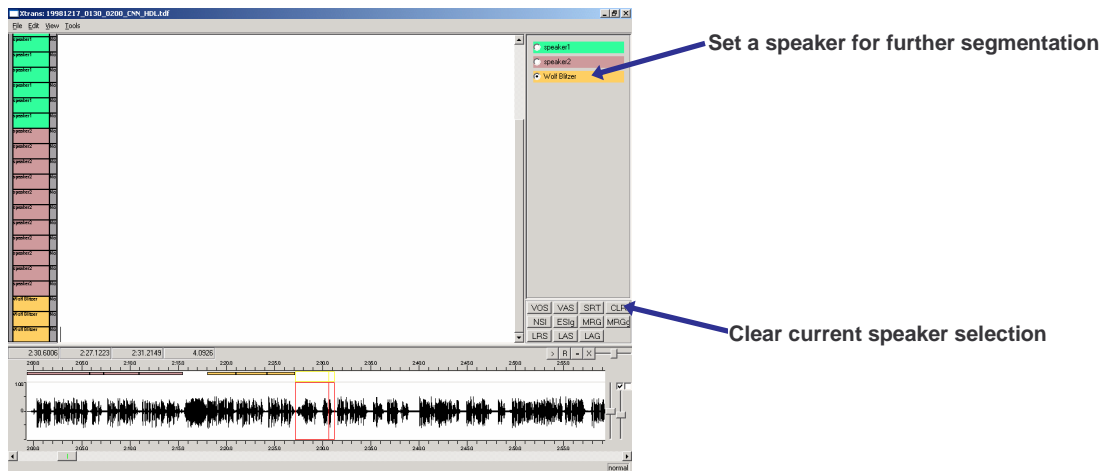
- a box in the speaker turn column that corresponds to this segment, color-coded and displaying the speaker name
- a color-coded horizontal bar in the waveform display that indicates the duration of the segment
- a color-coded speaker entry in the speaker panel



You will notice that if you select a segment by clicking on the horizontal bar for that segment, the segment will be highlighted with a blue box, and XTrans will jump to the corresponding region in the transcript window. Any text in the corresponding transcript segment will also be highlighted in green.

5.1 Creating Additional Segments for Existing Speaker

When you need to create another segment for an existing speaker, you can click that speaker's name in speakerID panel on upper right. This will "set" subsequent segments to this speaker, eliminating the need for you to input information for this speaker again. As soon as you encounter speech from another speaker, you can "clear" the selected speaker by clicking CLR in the widget panel. This will force you to assign a new speakerID for the next segment you create.



5.2 Editing SpeakerIDs

There are several ways to modify speaker information.

5.2.1 Correcting Speaker Information

To change speaker information for a single segment, you can right click on the speakerID box for that segment and select "Assign new speakerID (single)". The speakerID dialog window will appear and you can modify speaker information. You can also change speaker information for a single segment by first placing your cursor within that segment, then hitting the NSI (new speaker information) button on the widget panel.

To change speaker information for all segments from a given speaker, you can right click on the speakerID box for one of the segments and select "Edit speaker information (global)", or you can click the ESG (edit speaker information – global) button on the widget panel.

5.2.2 Merging Two Speakers

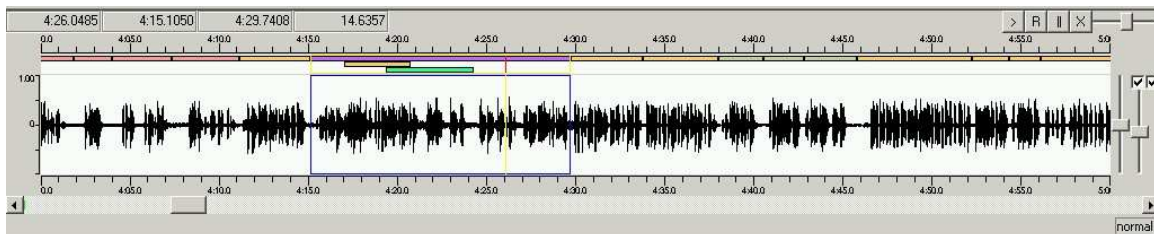
To reassign a speakerID to another (existing) speaker for a single segment, you can right click on the speakerID box for that segment and select "Merge into another speaker (single)". The speakerID dialog window will appear and you can select the new speaker assignment. You can also change speaker information for a single segment by first placing your cursor within that segment, then hitting the MRG (merge current speaker into another speaker) button on the widget panel.

To reassign speaker information for all segments from a given speaker (i.e., merge two speakerIDs into one for the whole file), you can right click on the speakerID box for one of the segments and select "Merge into another speaker (global)", or you can click the MRGg (merge – global) button on the widget panel.

6 Overlapping Speech

As noted above, XTrans prevents users from creating overlapping segments for a single speaker. However, transcription of overlapping or simultaneous speech is very straightforward. Although broadcast audio is single channel, XTrans creates a separate *virtual* channel for each unique speaker. There is then no prohibition on overlapping segments from two or more *different* speakers.

To create an overlapping speech segment, transcribers simply perform segmentation as described in Section 4: mark the start and end of the segment, associate it with the correct speaker, and insert the segment. Overlaps of two, three or more speakers are possible. XTrans visually displays the overlapping speech regions in the audio panel by showing overlapping color-coded horizontal bars for each segment, one per speaker.



Clicking on any of these horizontal bars will select that segment in the waveform and in the transcript.

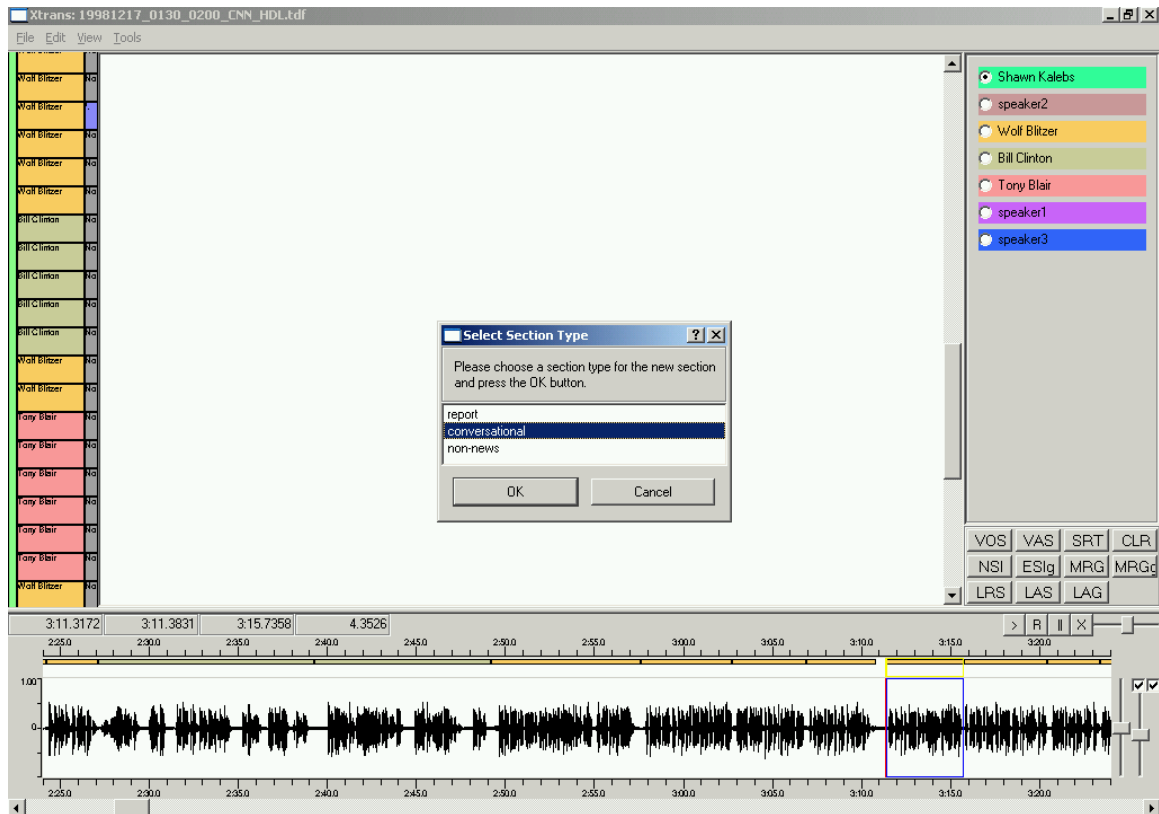
Overlap is not explicitly indicated in the transcript display, but can easily be extrapolated from the transcript file itself by examining segment start and end times.

7 Assigning Sections (Story Boundary Labeling)

To create a section (story, conversation, or commercial) boundary, click on the horizontal bar of the speaker segment in the waveform that corresponds to the

beginning of the new section. The reverse action – clicking inside the transcript to select a segment – also works to select the segment.

Then type `Ctrl+I Ctrl+S` to insert the new section boundary. You will be prompted to select the appropriate section type (report, conversational or non-news).



After you've assigned the section type, the transcript itself will not change, but you will notice that the box in the section column that corresponds to this region of the program now has a color-coded section value assigned, whereas regions without section assignments remain grey.⁷ A horizontal line also appears in the section column at the segment corresponding to the start of this section.

Note that all section boundaries must correspond to the beginning of a speaker segment.

8 Assigning SUs (Sentence Unit Labeling)

SU assignment is very simple if you follow the practice of creating one segment per sentence during the segmentation task. To label the SU type for an existing segment, simply enter one of the following commands while your cursor is in the segment (either on the waveform or in the corresponding transcript region):

⁷ Alternatively, you can right-click on the box in the section column to assign or re-assign section for a region. You can also type `Ctrl+E Ctrl+S` to edit type information for an existing section.

Ctrl+I Ctrl+U Ctrl+S	Set "statement" SU
Ctrl+I Ctrl+U Ctrl+Q	Set "question" SU
Ctrl+I Ctrl+U Ctrl+I	Set "incomplete" SU
Ctrl+I Ctrl+U Ctrl+B	Set "backchannel" SU
Ctrl+D Ctrl+U	Unset SU markup

The transcript itself will not change, but you will notice that the box in the SU column that corresponds to this segment now has a color-coded SU value assigned, whereas segments without SU assignments remain dark grey.⁸

9 Transcription

There are a handful of features in XTrans to facilitate transcription itself. The basic transcription process involves typing the words that correspond to an existing segment. Recall that the transcript and waveform display are linked in that you can click on a segment in the waveform and XTrans will find that segment in the transcript, and vice versa.

9.1 Language Input Functions

By design, XTrans supports multilingual text input. The default input mode is English.

9.1.1 Changing Input for Windows Systems

You can change input methods on a Windows system by accessing the "Language Bar" functions of Windows.

WindowsXP

For instance, to enable the input of Arabic text on Windows XP:

1. Make sure you have an XP installation disk
2. Go to Regional and Language Options
3. Select the Language tab
4. Select "Install files for complex script and right-to-left languages" This will copy files from somewhere to somewhere else and add Arabic to the lists of available languages to add.
5. Hit Apply and reboot the machine at the prompt.
6. Go back to Regional and Language Options.
7. Go to the Advanced tab and select Arabic from the "Languages for non-unicode programs" list.
8. Go back to the Language tab and make sure that Arabic is listed as an installed service under Details and "Text Services and Input Languages".

The result of this process should be a little blue box in the lower right corner of the screen that displays "EN" when English is the selected input language and

⁸ Alternatively, you can right-click on the box in the SU column to assign or re-assign SU for a segment.

"AR" for Arabic. The user can set their desired input language by selecting from that blue box.

Windows2000

To enable the input of Arabic text on Windows2000:

1. Go to the control panel
2. Select Regional options
3. Select Input locales
4. Hit the change button
5. Add required languages

Note that it may be necessary to reboot your machine for changes to take effect.

9.1.2 Changing Input for Unix Systems

There are multiple ways to change the language input mode on a *NIX system.

Method 1: Keyboard Mapping

With this method the user sets the keyboard mapping via a separate terminal window (xterm). For instance:

Change input mode to English (\$ is a shell prompt):

```
$ setxkbmap us
```

Change input mode to Arabic (\$ is a shell prompt):

```
$ setxkbmap ar
```

Note that the input mode is set to English first, then Arabic, because once the input mode has been changed to Arabic, you will be unable to type any English command on the terminal. In bash (GNU Bourne-Again Shell), if you set it to English first, even though you cannot type, you can still use the up-arrow key to recall the command line that was used to set the input mode to English when you need to type in English.

Method 2: SCIM

Another option is to use an input method like SCIM (Smart Common Input Method). SCIM provides a user friendly, full featured input method user interface for POSIX-style operating systems (including Linux, FreeBSD and other Unix), as well as a development platform to make input method development easier. SCIM runs as a separate program and registers itself as an input method of X Windows system. Then, users change their environment setting to select their input method, e.g. in bash,

```
export XMODIFIERS="@im=SCIM"
```

Any tool that understands X Input Method (XIM) Protocol will get keyboard input from the designated input method. For more information on installing and using SCIM, please visit the SCIM webpage: <http://www.scim-im.org/>

Other Methods

For future versions of XTrans we are researching input methods that can be directly incorporated into the tool, for instance UIM.

9.2 Transcript Markup Shortcuts

The following shortcuts can help the transcriber add required markup to transcripts:

Ctrl+Shift+G	Insert <background></background> in text
Ctrl+Shift+B	Insert <breath/> in text
Ctrl+Shift+H	Insert <cough/> in text
Ctrl+Shift+F	Insert <foreign language="..."></foreign> in text
Ctrl+Shift+N	Insert <lname/> in text
Ctrl+Shift+L	Insert <laugh></laugh> in text
Ctrl+Shift+P	Insert <lipsmack/> in text
Ctrl+Shift+S	Insert <sneeze/> in text
Ctrl+Z	undo the last change
Ctrl+Y	redo the last undo

9.3 Spell Check

Aspell-based spell checker is implemented, but only enabled in the Linux package. Users must install aspell and the dictionaries themselves, since the current XTrans package doesn't include those components.

Spell check will be added to the Windows version in the future.

10 Other Features

To be added.

11 Data Format

The default XTrans data format is .tdf, or Tab Delimited Format. This is a simple file format in which data is represented as a set of "records", which are in turn a set of "fields" separated by tab characters. The XTrans .tdf format is a set of 13-field records plus some meta-information as described below:

1	file	file name or id	string
2	channel	audio channel	number
3	start	start time	number
4	end	end time	number
5	speaker	speaker name or id	string
6	speakerType	speaker type	string
7	speakerDialect	speaker dialect	string
8	transcript	transcript	string

9	section	section id	number
10	turn	turn id	number
11	segment	segment id	number
12	sectionType	section type	string
13	suType	SU type	string

Each record in an XTrans .tdf file refers to a segment. In addition to the body of segments, there are a few lines of meta-information within the .tdf file. The first line of the format declares the above field specification for segments in the following form:

```
file;unicode    channel;int    start;float    ...
```

The second and third lines specify where the "real" section boundaries are and what types they are. For example,

```
;;MM sectionTypes      [u'report', u'nontrans', None]
;;MM sectionBoundaries [0.0, 425.3, 9999999.0]
```

means that the first section starts at 0.0 second and its type is "report", and that the second section starts at 425.3 seconds and its type is "nontrans", and this is the last section (last ones are ignored). These two lines are optional⁹.

Besides these, comments are allowed anywhere in the file except for the first line. A comment line starts with ";;" and ignoring this line doesn't cause any problem to the data integrity with one exception that lines starting with ";;MM" might have meaningful information as shown in the example above.

12 Saving and Exporting Transcripts

To save your work, type `Alt+F`, `Alt+S` or select "Save" or "Save As" from the File menu. Files are saved automatically at one-minute intervals to a backup file with a .sav extension. If XTrans crashes, users can recover the file by renaming the backup file by removing the .sav extension. In normal conditions, the backup file is removed when the user closes the file.

You can export your transcript file into two other common data formats: .trs (used with the Transcriber tool) and .txt (used with older LDC transcription tools). To do this, select the "Export" function from the File menu¹⁰.

⁹ For instance, telephone speech transcripts usually do not include section information.

¹⁰ A third export format, SGML: Weblog, is used for some specialized LDC annotation tasks and is not intended for general use.

Appendix 1: Recommended Strategy

There are many different ways to interact with XTrans to create a time-aligned transcript. The following is a synopsis of LDC's recommended strategy for creating broadcast transcripts with XTrans. Note that most of these functions are keyboard rather than mouse-based commands. For quick transcription, it is strongly recommended that transcribers choose keyboard over mouse-based functions as much as possible. This takes a little getting used to but you will find it much faster and easier to use the keyboard only rather than switching between keyboard and mouse (and it's easier on your wrists!).

Quick Guide for Quick Transcription

- | | |
|---|---|
| 1. open audio file | File > Open audio file |
| 2. open new transcript file | File > New |
| 3. associate audio and transcript | Edit > Blindly associate transcript to audio |
| 4. begin playback and mark segment start | Alt+M |
| 5. stop playback and mark segment end | Alt+M |
| 6. insert segment | Ctrl+N (Ctrl+Insert on *nix) |
| 7. assign speaker information | dialog box (use tab & arrow keys to select options) |
| 8. create next segment (repeat 4-7). To create segment for same speaker, first select speaker in speaker panel then repeat steps 4-6. | |
| 9. assign section boundary | Ctrl+I Ctrl+S |
| 10. assign SU type | Ctrl+I Ctrl+U Ctrl+___ |
| 11. transcribe the segment ¹¹ | |
| 12. save your work frequently | Alt+F Alt+S |
| 13. repeat steps 4-12 | |
| 14. save and exit | |

¹¹ Some transcribers prefer to fully segment the file the go back and transcribe it; while others prefer to transcribe as they segment.

Appendix 2: Keybindings

The default XTrans keybindings are shown in the table below.

XTrans keybindings are user-configurable by going to the Tools >Edit key bindings menu (or hitting `Alt+T`, `Alt+E`). A dialog box will open. Place your cursor on the second column of the row containing the keybinding you want to change. Double-click on a cell to edit it. Erase the existing text, type the double-quoted string of key combinations you want, then press `Enter`. Click on the `OK` button to commit the change.

If you undo the modification before committing it, use the `Reset` button. The `Reset All` button will cancel all modifications from the current editing session. The `Default` button restores the default setting for the current keybinding; and the `Default All` button does the same for all keybindings. Use the `cancel` button to cancel all changes and close the dialog box.

It is also possible to save the changes as a configuration file using the Tools > Save configuration file menu (or `Alt+T,Alt+S`). To load the saved configuration file, use Tools > Load configuration file menu (or `Alt+T,Alt+L`).

Category	Keybinding	Description
Playback	<code>Alt+M</code>	Toggle between two actions: play/begin_selection vs. stop_playing/finalize_selection
	<code>Alt+S</code>	Stop playback
	<code>Alt+P</code>	Pause/resume playback
	<code>Alt+Space</code>	Play current region
	<code>Alt+Return</code>	Play from the right edge of current selection
	<code>Alt+\</code>	Play just ending part of current selection
	<code>Alt+ </code>	Play from the beginning of current waveform window
	<code>Alt+[</code>	Rewind one second
	<code>Alt+]</code>	Fast-forward one second
Waveform Control	<code>Ctrl+<</code>	Fast-grow current selection; move left edge to the left by a couple of pixels
	<code>Ctrl+></code>	Fast-shrink current selection; move left edge to the right by a couple of pixels
	<code>Alt+></code>	Fast-grow current selection; move right edge to the right by a couple of pixels
	<code>Alt+<</code>	Fast-shrink current selection; move right edge to the left by a couple of pixels
	<code>Ctrl+,</code>	Grow current selection; move left edge to the left by one pixel
	<code>Ctrl+.</code>	Shrink current selection; move left edge to the right by one pixel
	<code>Alt+.</code>	Grow current selection; move right edge to the right by one pixel
	<code>Alt+,</code>	Shrink current selection; move right edge to the left by one pixel
	<code>Alt+A</code>	Zoom in where the cursor is placed
	<code>Alt+Z</code>	Zoom out where the cursor is placed
	<code>Alt+X</code>	Make current selection fit the current waveform window

Segmentation	Ctrl+N	Insert a segment for the selected region (Ctrl+Insert on *nix)
	Ctrl+D	Delete current segment (Ctrl+Delete on *nix)
	Ctrl+J	Join the current and following segments
	Ctrl+Return	Split current segment into two at cursor point
Section Labeling	Ctrl+I, Ctrl+S	Insert section boundary
	Ctrl+E, Ctrl+S	Edit section boundary type
	Ctrl+D, Ctrl+S	Delete section boundary on the left of the cursor
SU Labeling	Ctrl+I, Ctrl+U, Ctrl+B	Set SU type (backchannel) for current segment
	Ctrl+I, Ctrl+U, Ctrl+I	Set SU type (incomplete) for current segment
	Ctrl+I, Ctrl+U, Ctrl+Q	Set SU type (question) for current segment
	Ctrl+I, Ctrl+U, Ctrl+S	Set SU type (statement) for current segment
	Ctrl+D, Ctrl+U	Unset SU type for current segment
Transcript Markup	Ctrl+Shift+G	Insert <background></background> in text
	Ctrl+Shift+B	Insert <breath/> in text
	Ctrl+Shift+H	Insert <cough/> in text
	Ctrl+Shift+F	Insert <foreign language="..."></foreign> in text
	Ctrl+Shift+N	Insert <lname/> in text
	Ctrl+Shift+L	Insert <laugh></laugh> in text
	Ctrl+Shift+P	Insert <lipsmack/> in text
	Ctrl+Shift+S	Insert <sneeze/> in text

Appendix 3: Sample Transcript Output

Meeting recording excerpt

This excerpt shows a .tdf version of a meeting recording transcript. Meetings typically contain 3-8 speakers; this particular recording included 4 speakers. This file is from the Rich Transcription 2007 Meeting Recognition Evaluation training dataset (file NIST_20060223-1416).¹²

```
file;unicode channel;int start;float end;float speaker;unicode
speakerType;unicode speakerDialect;unicode transcript;unicode
section;int turn;int segment;int sectionType;unicode
suType;unicode
;;MM sectionTypes [None, None]
;;MM audioBinding ['NIST_20060223-1416_HM-01_Subj-159_16K16b', 0,
'NIST_20060223-1416_HM-01_Subj-159_16K16b', 0, 'NIST_20060223-1416_HM-02_Subj-
161_16K16b', 0, 'NIST_20060223-1416_HM-02_Subj-161_16K16b', 0, 'NIST_20060223-
1416_HM-03_Subj-162_16K16b', 0, 'NIST_20060223-1416_HM-03_Subj-162_16K16b', 0,
'NIST_20060223-1416_HM-04_Subj-160_16K16b', 0, 'NIST_20060223-1416_HM-04_Subj-
160_16K16b', 0]
;;MM sectionBoundaries [0.0, 9999999.0]
NIST_20060223-1416_HM-01_Subj-159_16K16b 0 0.538744769874
5.62125523013 Subj-159 male native So, I'd like to thank you all for
coming this afternoon to this financial literacy workshop. 0 0 0

NIST_20060223-1416_HM-01_Subj-159_16K16b 0 5.65937238494 10.8506276151
Subj-159 male native What I want to do is try to answer and address
all of your concerns that you have 0 1 1
NIST_20060223-1416_HM-01_Subj-159_16K16b 0 10.9093723849 12.8306276151
Subj-159 male native about money and money matters. 0 2
2

NIST_20060223-1416_HM-01_Subj-159_16K16b 0 13.0593723849 16.3806276151
Subj-159 male native I would be glad to speak to you about all
types of 0 3 3
NIST_20060223-1416_HM-01_Subj-159_16K16b 0 16.9949612403 20.22 Subj-
159 male native products you can use to accumulate money and wealth. 0
4 4
NIST_20060223-1416_HM-01_Subj-159_16K16b 0 20.61 23.52 Subj-159
male native For your retirement for your education concerns 0
6 6
NIST_20060223-1416_HM-01_Subj-159_16K16b 0 23.86 26.83 Subj-159
male native for buying a home reducing -- for reducing debt. 0
7 7
NIST_20060223-1416_HM-01_Subj-159_16K16b 0 26.9949612403 32.0100775194
Subj-159 male native Whatever matters that -- uh whatever
financial matters that you want to address. 0 8 8
NIST_20060223-1416_HM-01_Subj-159_16K16b 0 32.1949612403 38.7150387597
Subj-159 male native What we want to try to achieve in this meeting
is your personal satisfaction. 0 9 9
NIST_20060223-1416_HM-01_Subj-159_16K16b 0 38.8349612403 46.3450387597
Subj-159 male native Is because I want you to be comfortable with a
with asking me any of the financial questions that you have concerns about.
0 10 10
NIST_20060223-1416_HM-01_Subj-159_16K16b 0 46.49 48.17 Subj-159
male native If I'm unable to answer them 0 14 14
NIST_20060223-1416_HM-01_Subj-159_16K16b 0 48.52958159 49.8606276151
Subj-159 male native I will tell you so. 0 15 15
```

¹² For more information about LDC's meeting recording transcription efforts, see <http://projects ldc.upenn.edu/Transcription/NISTMeet/>

NIST_20060223-1416_HM-01_Subj-159_16K16b	0	50.26	53.92	Subj-159
male native As a registered investment adviser the difference that I				
make	0	16	16	
NIST_20060223-1416_HM-01_Subj-159_16K16b	0	53.92	57.97	Subj-159
male native to you is that I am liable for fully disclosing				0
17	17			
NIST_20060223-1416_HM-01_Subj-159_16K16b	0			