# Translators' Manual

# PiPeNovel: Pilot on post-editing Novels

Antonio Toral

University of Groningen

Sections 1 and 2 are based on PET's user manual by Wilker Aziz and Lucia Specia.

May 4, 2018

# Contents

# Chapter 1

# Running

## 1.1  Requirements

PET only requires a recent version of the Java Virtual Machine (`1.6.0_26` or superior), which is likely to be already installed if you keep your machine up to date.

On Microsoft Windows it is important to check whether the PATH to the Java binary files is properly set. Again, this may be already set. If you try to run the tool 1.2 and it does not work, follow the guidelines in Section 1.3 to set the path.

## 1.2  How to run

Simply execute the file `run.bat`. On Windows double clicking the file will be enough. On Linux or Mac if the "executable permission" is set, double-clicking the file then selecting "run on terminal" will be enough. If not, set it under the tab "Permissions" right-clicking the file, or open a terminal and type `sh run.bat`.

## 1.3  Windows Troubleshooting

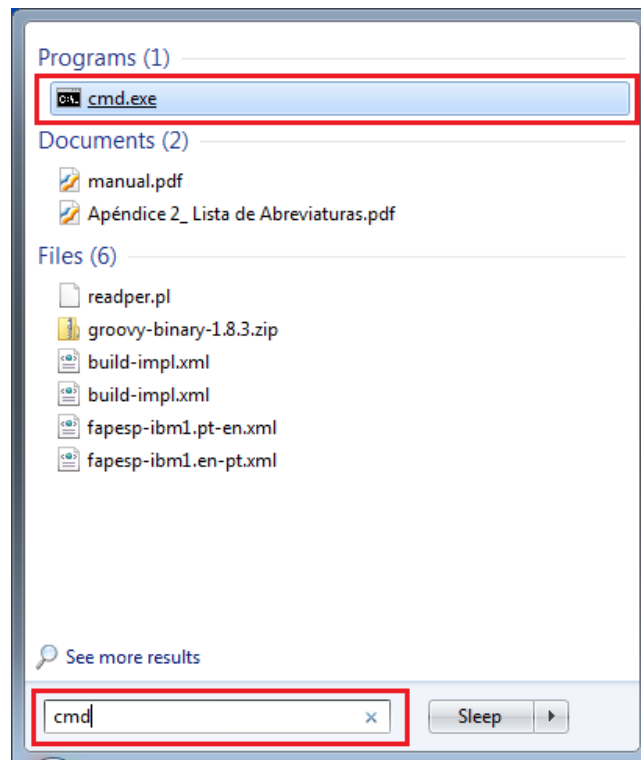1. Do I have the right JVM?

   (a) Open a terminal
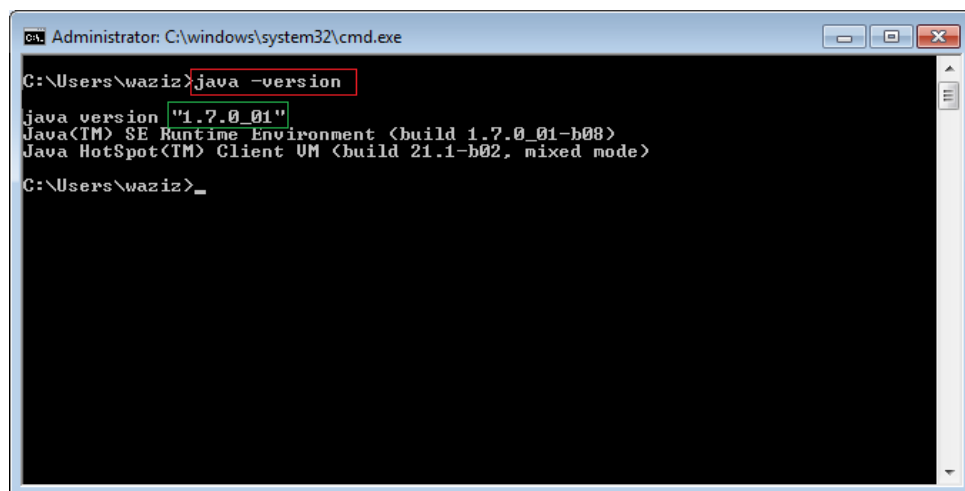
Figure 1.1: Terminal

(b) Type in `java -version`



Figure 1.2: Java version

2. What if my computer does not recognize the command `java`?

   (a) Make sure you have downloaded and installed the JVM(`http://java.com/en/download/index.jsp`).

3. I am sure I have installed the JVM, yet I don't get it working, what can I do?

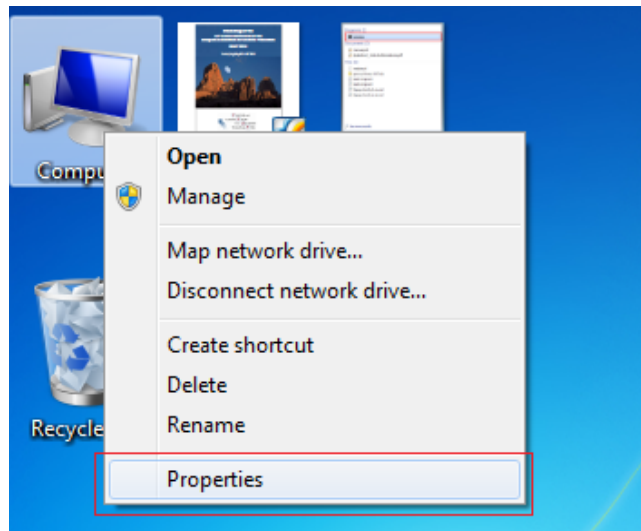   (a) Open your computer's properties

Figure 1.3: Properties
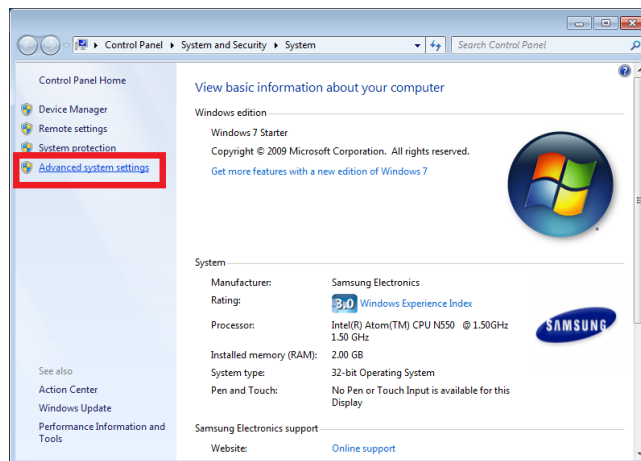
(b) Open the 'Advanced Settings'



Figure 1.4: Advanced settings

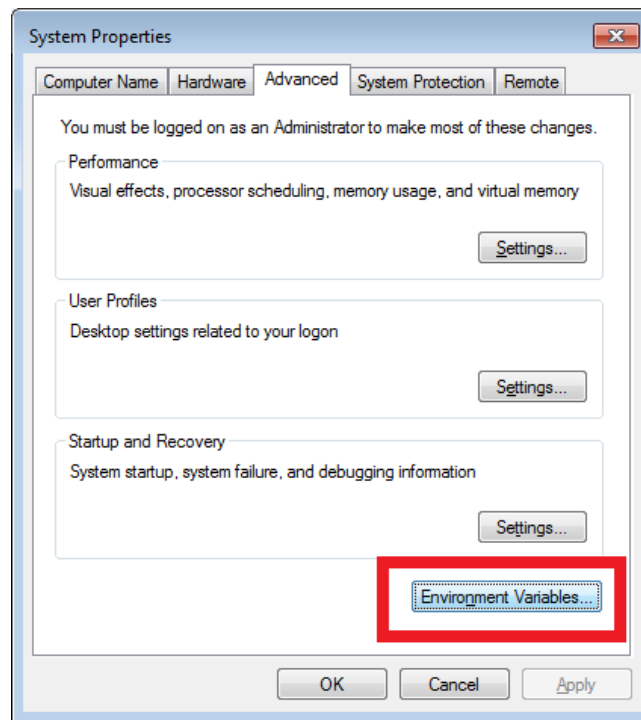(c) Go to 'Environment variables'

Figure 1.5: Environment variables

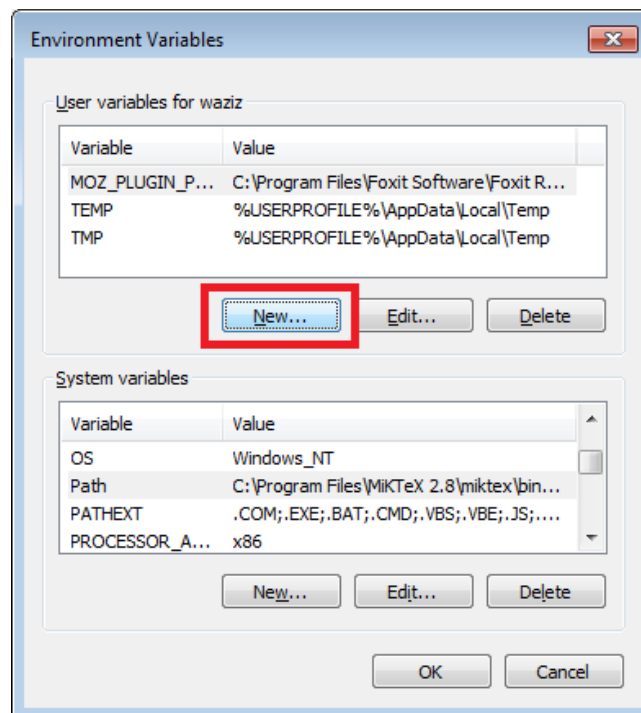(d) Add a new variable



Figure 1.6: New user variable

(e) Set variable's name and value to 'PATH' and '<path-to-java>' respectively
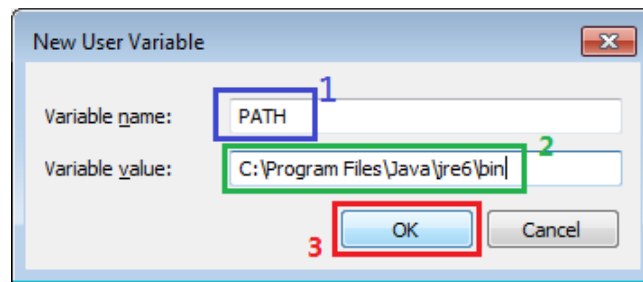
Figure 1.7: Java's path

(f) Conclude
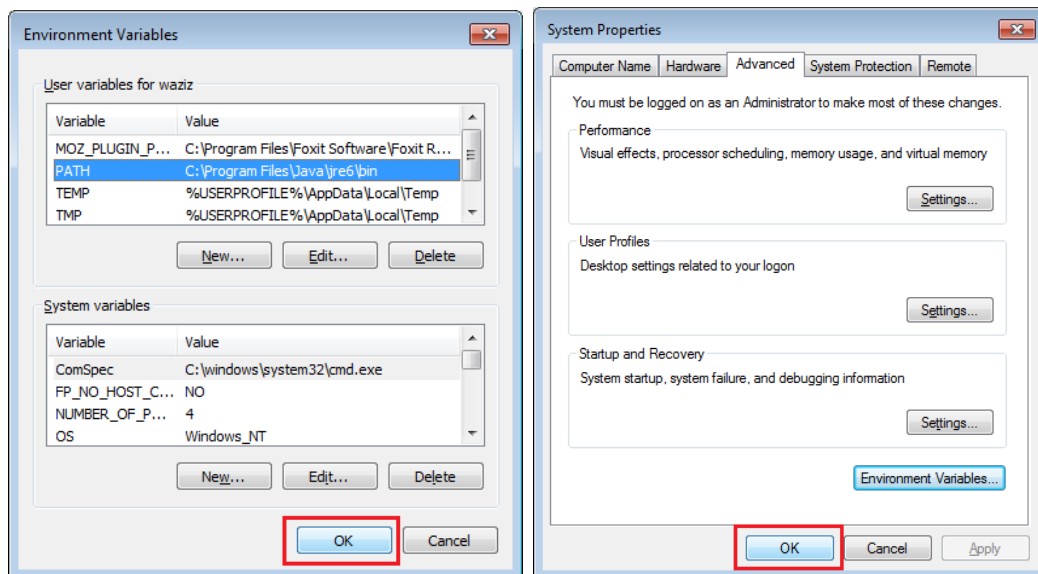


Figure 1.8: Conclude

4. How do I find the path to Java binaries?

   (a) Open a terminal (see Figure 1a)

   (b) Type in `where java`
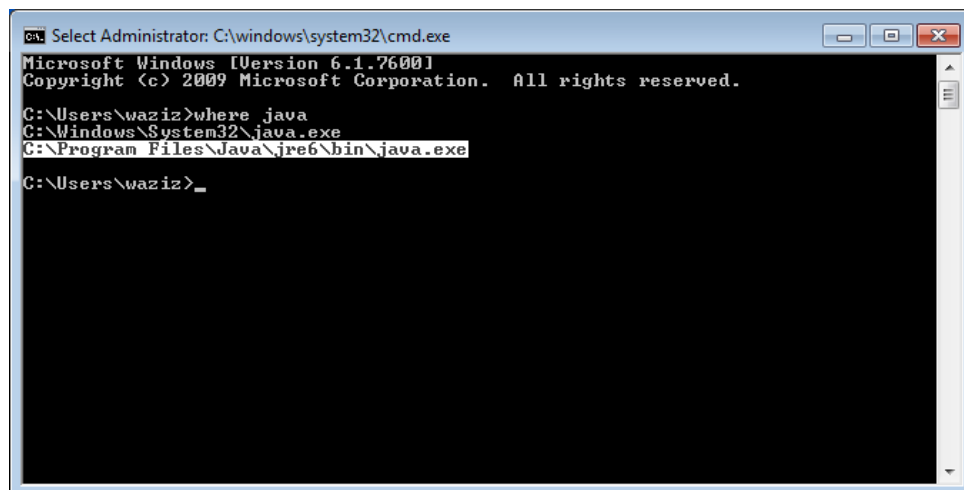


Figure 1.9: Finding Java

If this does not work you may try the following:

(c) Open the Windows Explorer (e.g. double click 'Computer')

(d) Find the directory where you have installed Java (usually `C:\Program Files\Java`)

(e) Go to `jre6\bin`
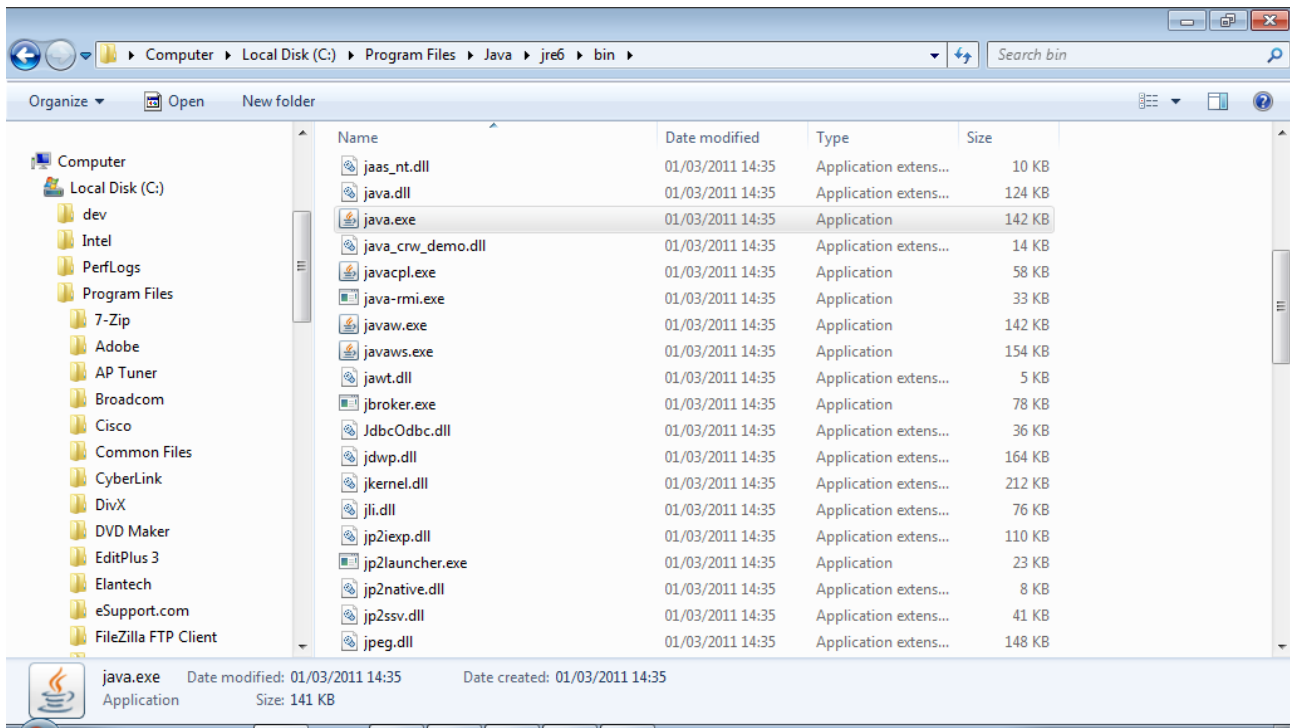


Figure 1.10: Java installation folder

(f) Copy the path from the address bar
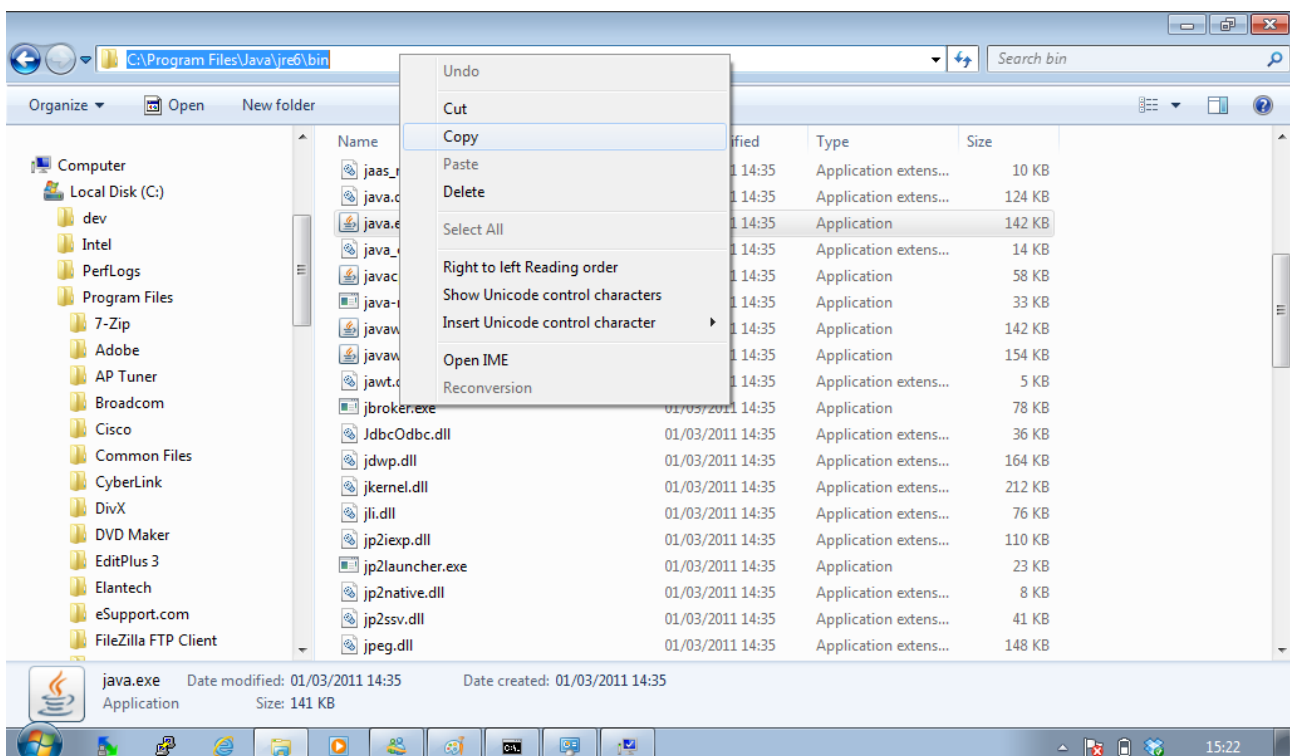


Figure 1.11: Path to binaries

# Chapter 2

# Interface

## 2.1  Main

The main page (see Figure 2.1) is the page that you see when PET is loaded. Observe that the title displays the default *user* of the tool ("demo" in this example), as set in the configuration file.
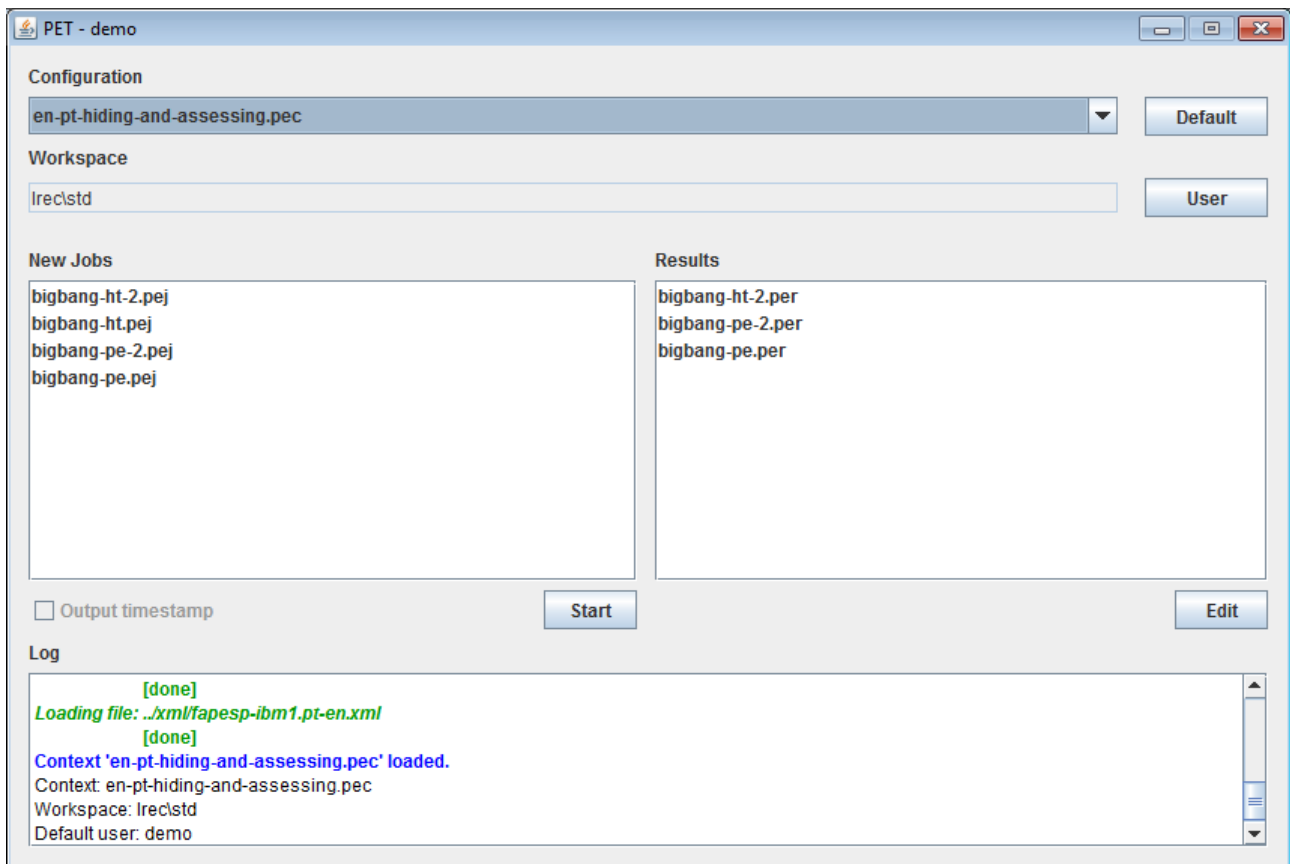


Figure 2.1: PET's main page

Figure 2.1 shows the following (note that only the relevant elements for this experiment are discussed):

**Workspace** displays the *workspace*, that is, where the files with the jobs to be performed are stored, and where the edited jobs will be saved;

**Jobs** list of available *jobs* (.pej files), i.e. a collection of segments to post-edit or translate;

**Results** list of (partial) *results* (.per files), that is, the output of jobs that have been started (completed or ongoing);

**Start** starts a *job* from zero;

**Edit** continues a (partial) *result*;

**Important**: if you have started a job and done it partially and would like to resume it, then do NOT click "start" but click "edit". If you do click "start" then all your progress in that job so far will be deleted.

## 2.2   Annotation Page

The annotation page (see Figure 2.2) is the page that you see when a job is loaded.
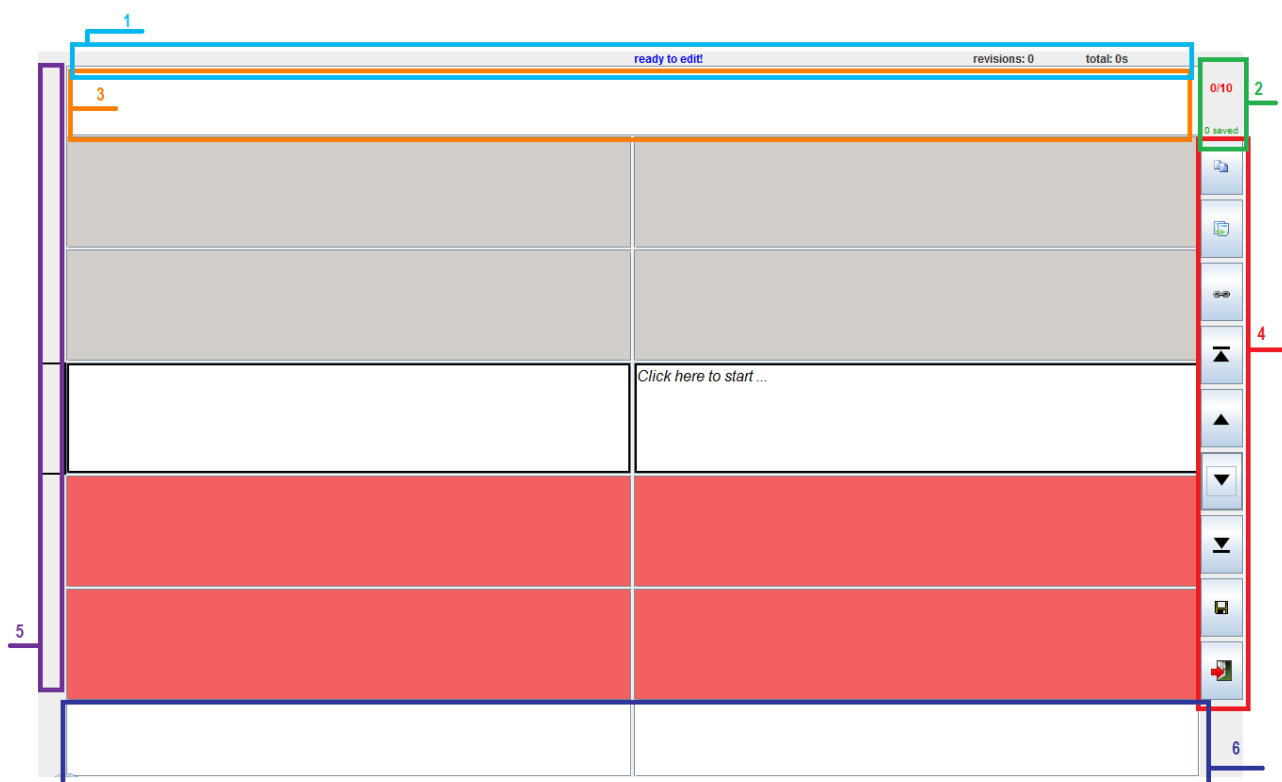


Figure 2.2: PET's annotation page

Figure 2.2 illustrates some important features:

**Labels (1/light blue)** at the top of the window you will find some useful information, such as the status of the current *unit*, the number of revisions and the time spent on that unit;

**Progress (2/green)** on the top-right hand side. The status of the current job is shown in units done out of the total units in the file (in red); the status of the current jobs is also shown in number of units saved to disk; besides, the timestamp of the last time the progress was saved to disk is shown;

**Top box (3/orange)** displays the source text while the active unit is being post-edited;

> **Tool box (4/red)** placed on the right-hand-side, it displays action buttons explained in Section 2.2.1;
>
> **Id box (5/purple)** placed on the left-hand-side, not used in this experiment;
>
> **Bottom box (6/blue)** not used in this experiment;
>
> **Units** the centre of the page contains a grid of 10 lines and two columns (i.e for source and translation). The line in the centre of the screen (highlighted with dark borders) is the active unit, as explained in Section 2.2.2.

The top half of the `units` area shows already edited units, while the bottom half shows units to come (see Figure 2.2). As the user progresses, some units are displayed in green, some are displayed in red. Green stands for "done", that is, units that have been post-edited/translated, while red stands for "to be done". The unit in white is the active unit (always in the central area), and it turns yellow to indicate that the editing has started. Note that the label at the top will change from "ready to edit" to "editing". The labels at the top also displays the time spent on the current revision ("partial"), the number of complete revisions ("revisions") and the total time spent on revisions that are complete ("total").
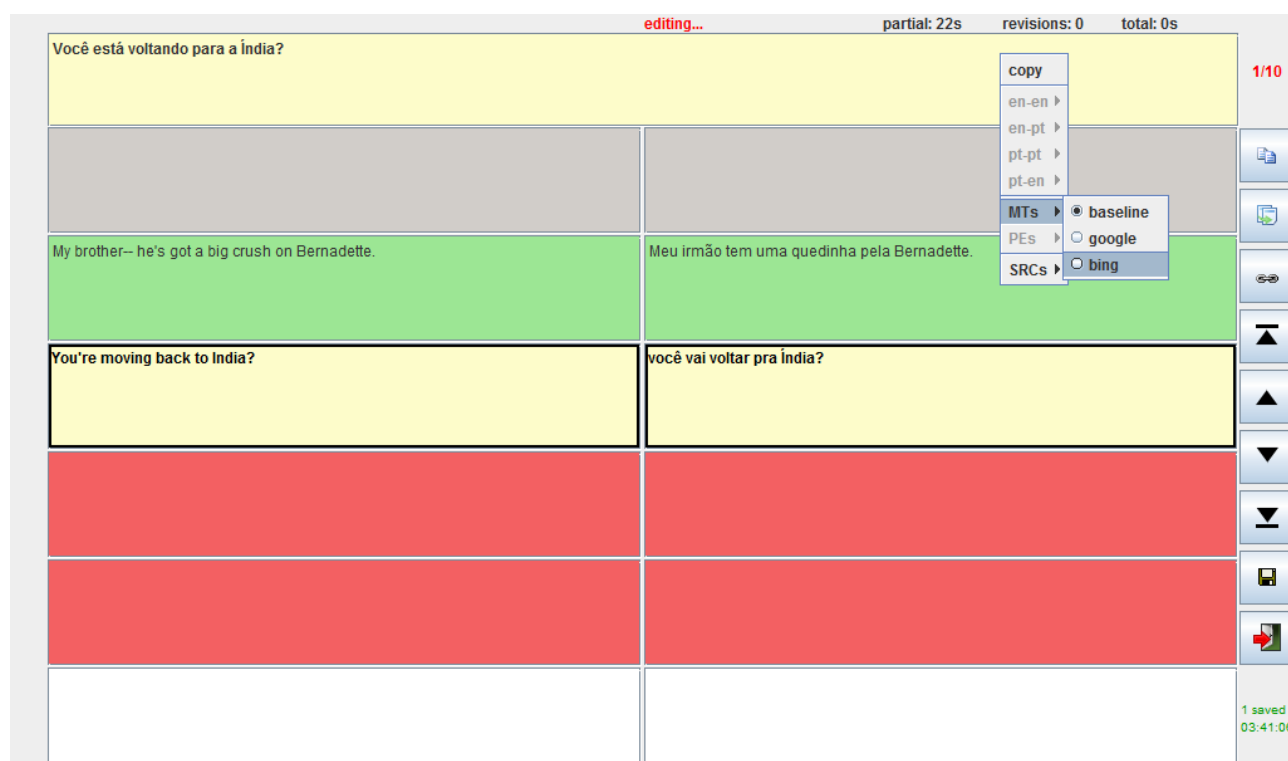


Figure 2.3: PET's annotation page - progress

## 2.2.1 Tool box

The context of the tool box (right-hand side box in Figure 2.2 with the tool's action buttons) displays:

> **Copy** not used in this experiment
>
> **Revert** revert the active unit to its last revision;

**Bind** binds source and target scrolling (useful for long segments);

**Previous undone** searches (backward) for the first undone unit;

**Previous** backs one unit;

**Next** advances one unit;

**Next undone** searches (forward) for the first undone unit;

**Save** saves the current progress to disk;

**Close** closes the job (asks for confirmation and offers a chance to save the progress);

By clicking anywhere on the tool box, it will be focused. Once focused, scrolling down the mouse wheel will move forward and scrolling up will move backward among the units. When the tool box is focused some shortcuts are available. Placing the pointer over a button for a second will show a tool tip text and information about shortcuts for that button.

Shortcuts are not case-sensitive, some of them are:

I enables the *insert* mode, that is, it changes the active unit to the "editing" state;

↓ moves to next unit (if in scrolling state);

↑ moves to previous unit (if in scrolling state);

<END> moves forward until an undone unit is found;

<HOME> moves backward until an undone unit is found;

B binds source and target scrolling;

<F10> saves;

<ALT>+<F4> closes the tool;

### 2.2.2 Active unit

The active unit is the central line in the units area. It is highlighted by a dark border and is the only unit whose target side is editable. By default, the active unit is in the state "ready to edit", when its target box receives the focus the unit will go to the state "editing" (the box turns yellow), which means that effort indicators will start to be logged. The active unit in the "editing" state will be generally referred to as the *editing box*.

The *editing box* offers some shortcuts, some of them are:

<ALT>+↓ finalizes the current unit and moves to next unit;

<ALT>+N finalizes the current editing and moves forward starting with the next unit;

<ALT>+↑ finalizes the current unit and moves to previous unit;

<ALT>+P finalizes the current unit and moves backward starting with the previous unit;

<ALT>+B binds source and target scrolling;

<CTRL>+Z undo editing;

&lt;CTRL&gt;+Y redo editing;

&lt;CTRL&gt;+C copy selection;

&lt;CTRL&gt;+X cut selection;

&lt;CTRL&gt;+V paste;

&lt;CTRL&gt;+R replace/insert selection;

&lt;CTRL&gt;+I insert/replace selection;

&lt;CTRL&gt;+D delete selection;

&lt;CTRL&gt;+S shift selection;

In case a unit is in the state "editing" an attempt to close the job will result in a prompt to finalize or discard the modifications done to that active unit.

### 2.2.3 Drag and drop

Text can be dragged from any text box (source, target, etc.) and dropped into the active unit's target box (if "editing"). Text can also be moved around within the editing box by using drag and drop.

# Chapter 3

# Translation Guidelines

- Fill-in this questionnaire before starting the translation tasks: `https://www.surveymonkey.com/r/TJCPWWM`

- In this experiment you are to complete a number of **tasks**. Each task corresponds to the translation of 10 consecutive sentences.

- The **type of translation** alternates between tasks. There are 2 types:

  1. <u>Translation from scratch</u>. Only the source sentence is provided, you are to write the translation from scratch.

  2. <u>Post-editing</u>. The source sentence is provided alongside a translation produced by an MT system. You are to post-edit this translation. If possible try to fix the translation provided by the MT system. Only if you deem the MT output too time consuming to fix, you can delete it and start from scratch.

     – Two MT systems, called MT1 and MT2, are used in the experiment. Be aware of which MT system you are post-editing, because at the end of the experiment you will be asked questions about these two systems, e.g. how good/bad they are.

- The **aim** is to produce publishable professional quality translations, both for translations from scratch and for post-editing. Thus, please try to do your very best.

- The **time** elapsed to carry out the translations is recorded while the active unit (sentence) is in editing mode (yellow background). Therefore:

  – Do not start performing a translation until you are in editing mode (yellow background). I.e. do not start thinking how you will tackle the translation of a sentence when the active unit is not in editing mode (green or red background).

  – Do not leave a unit in editing mode (yellow background) while you do something else. If you need to do something unrelated in the middle of a translation then go out of editing mode and come back to editing mode when you are ready to resume translating.

- Translations are related to source sentences on a **1-to-1** basis by default. In other words, each source sentence corresponds to 1 target sentence. In the translation of novels it is not that uncommon to have some cases of many-to-1 (more than 1 source sentence translated as 1 target sentence) or 1-to-many (1 source sentence translated as more than 1 target sentence). While producing 1-to-1 translations is <u>preferred</u>, if, for a given input, you deem it very important not to adhere to 1-to-1 equivalence, then proceed as follows:

- If you want to split a source sentence in 2 or more target sentences (1-to-many): enter all the target sentences in 1 target sentence box, separated by a newline (enter).

- If you want to join 2 source sentences (S1 and S2) into 1 target sentence (T2), many-to-1: leave the target box that corresponds to S1 (T1) empty. Enter the translation of S1+S2 in the target box that corresponds to S2 (T2). If you want to translate more than 2 sentences into 1, then the same method applies. For example for 3-to-1 you would leave the target boxes of T1 and T2 empty and produce the translation of S1+S2+S3 in the box T3.

- Please fill-in this questionnaire after completing the translation tasks: `https://www.surveymonkey.com/r/TFT56QW`

- Send back the results by zipping the PET folder and emailing it.

# Chapter 4

# Conditions

By taking part in this experiment you give your active informed consent and agree to the following conditions:

1. Your translation behaviour (time elapsed and keystrokes) is recorded.

2. Your translations might be published under a Creative Commons license, acknowledging your authorship.

3. The log data containing your translation behaviour might be published under a free license for research purposes. Such a dataset would be anonymised.

4. Your answers to the questionnaire might be published under a free license for research purposes. Such a dataset would be anonymised.

5. One or more research papers covering the experiments carried out might be published. Your contribution to the experiments would be acknowledged.

6. You have the right to withdraw your cooperation at any time, and you can do so without any consequences to you.