

# eCoLoTrain

## Translation Memory Module

### Exercise II

## Translating using an empty TM and a terminology database

### Prerequisites

For being able to complete the following exercise you need to have gone through courses II, III and IV, included in our Translation Memory Module.

You also need to have installed on your computer *DéjàVu X*, *Microsoft Office Word* and *Microsoft Office Excel*. The screen shots used to illustrate several steps of this exercise were taken from *DéjàVu X*.

### 1. Outline of this exercise

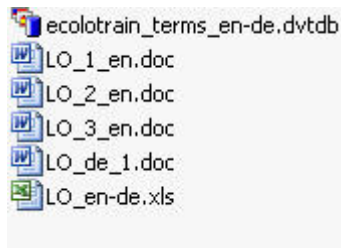
On completion of these exercises, you should be able to:

- Create a new translation project in *DéjàVu X*
- Create a new Translation Memory and select an existing termbase during the creation of a translation project
- Translate a text in *DéjàVu X* using the empty TM and the selected populated termbase
- Analyse and pre-translate an update version of the previous text using the newly created and populated TM
- Finish translation of the update text
- Export the translation back to its original format

### 2. Preparing your working environment

Before starting to work with *DéjàVu X*, prepare your working environment by downloading the necessary files and by creating the folder structure you will need to save your translation project and your new memory database.

- First, create under **My Documents** a folder called "**CAT tools**"
- Download from the eCoLoTrain webpage the compressed folder called "**Translating.rar**". When you decompress it you will see that it contains six sub-folders named by language-pairs. Each sub-folder includes a termbase file (with the *DéjàVu X* extension "\*.dvtdb"), several source files to choose and an *Excel* table containing terms in both languages

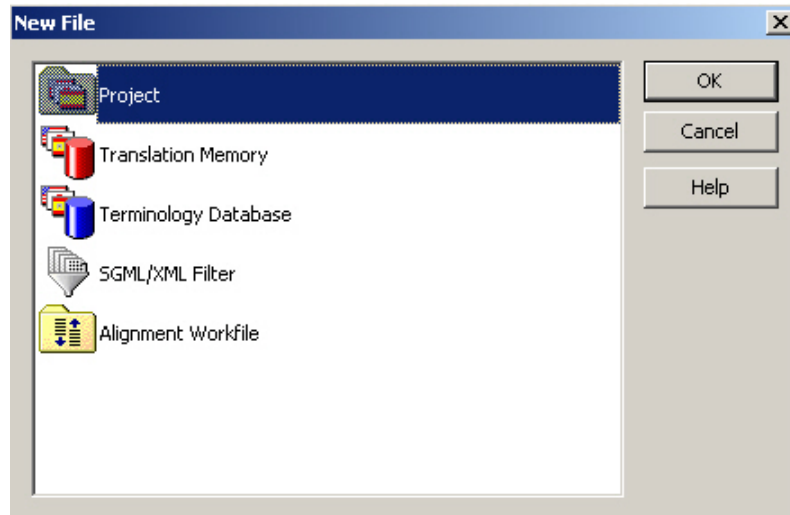


- Create under Translating a sub-folder called "**Project**", in which you will save the project definition file and the TM you will create in *DéjàVu X*

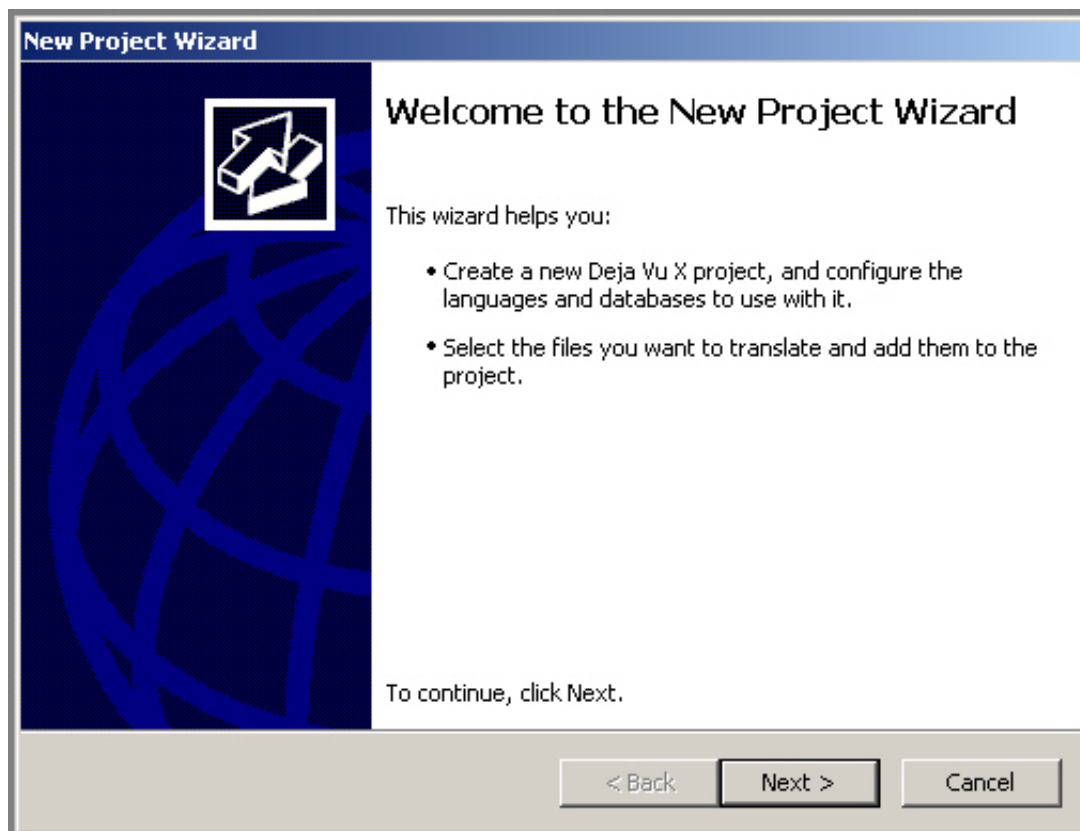
**You are now ready to start!**

## I. Creating a new translation project

- Start *DéjàVu X*
- In *DéjàVu X* choose the option **New** from the **File** menu
- From the **New File** dialog select **Project** and click on **OK**

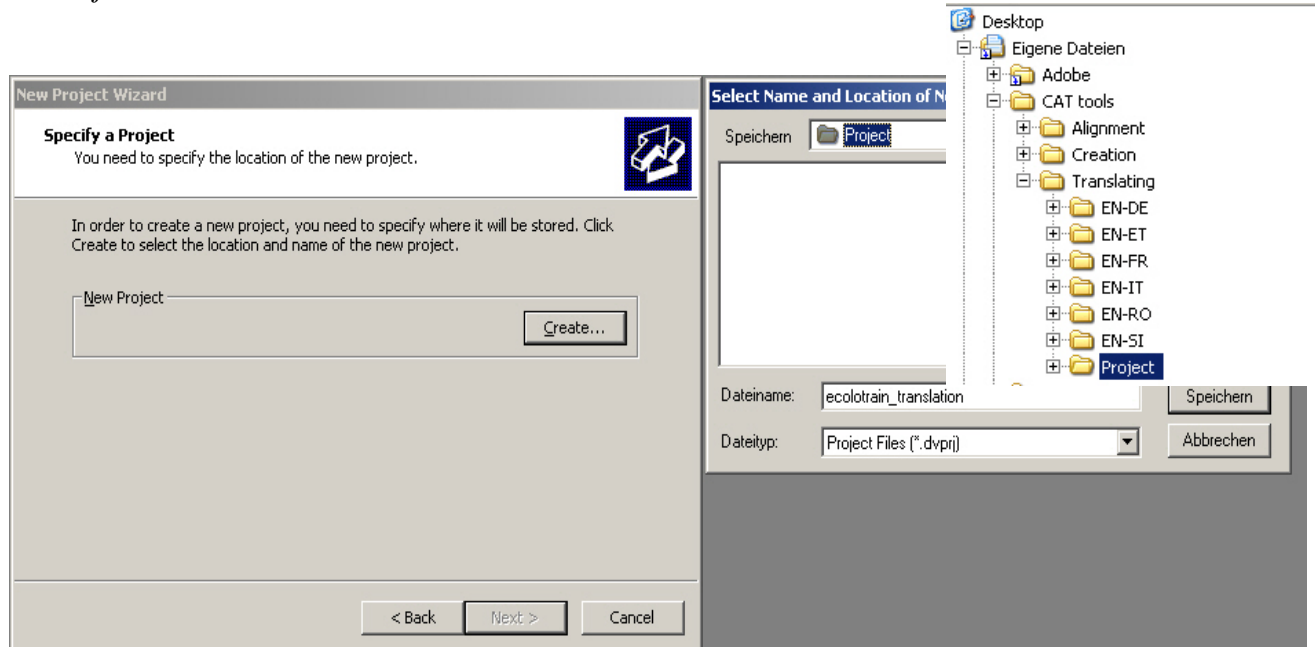


- The **New Project Wizard** opens and guides you through the creation of your translation project



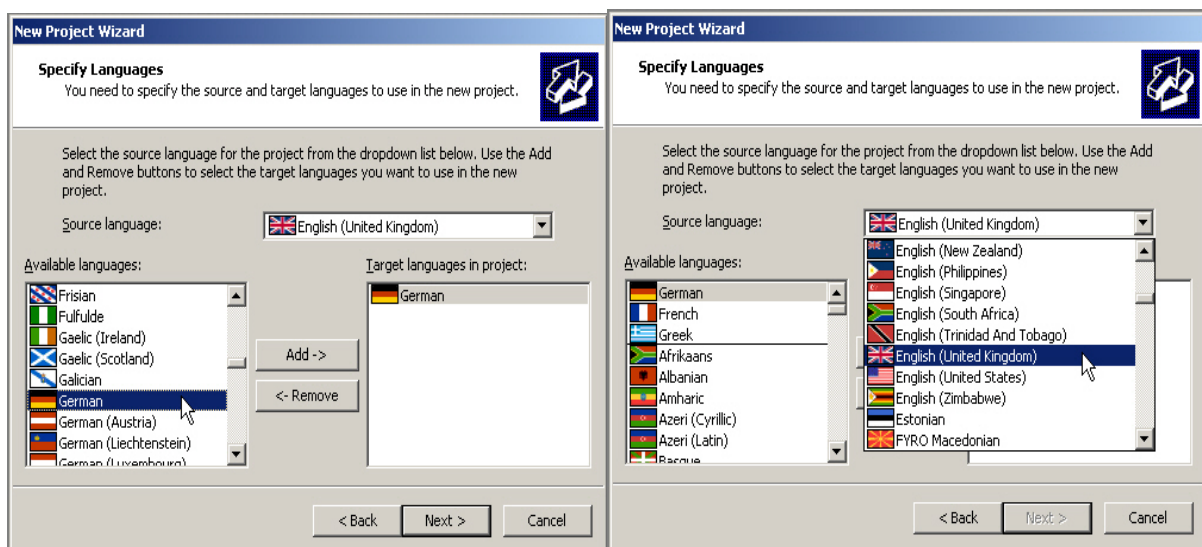
## 1. Specify Project

- Click on **Next** and in the dialog box create your new project, by clicking on the **Create** button and specify a location and a name for your project. For this, choose the folder **Project** you created at the beginning of this exercise and enter a name in the field File name, e.g. "ecolotrain\_translation". The extension ".dvprj" is added by *DéjàVu X*.



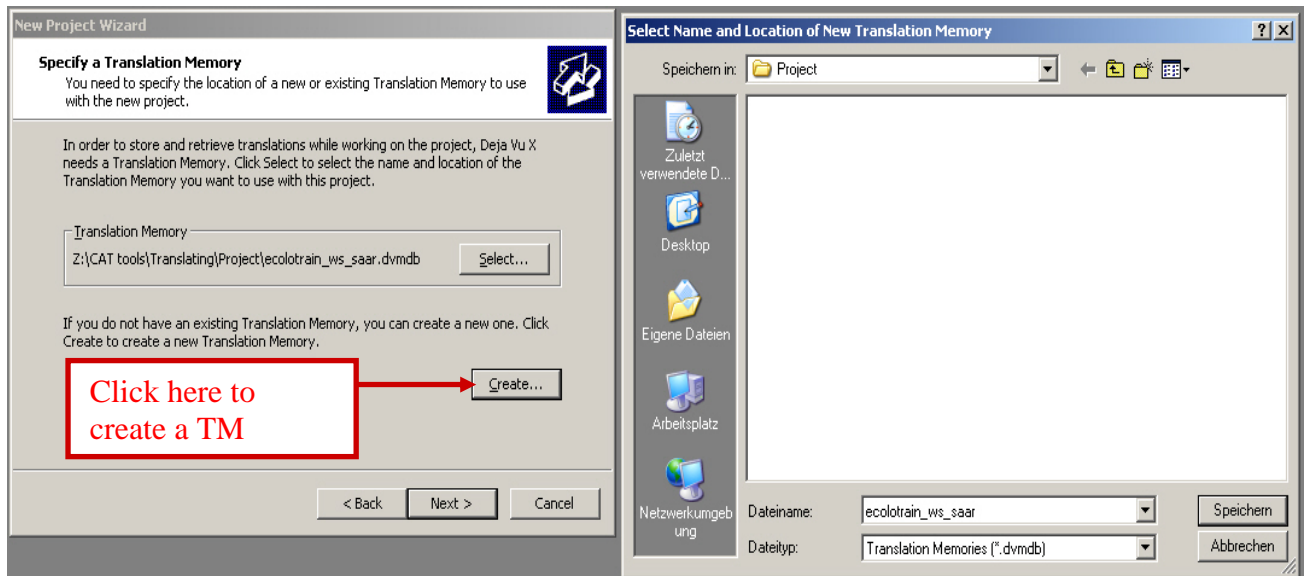
## 2. Specify Languages

- Click on **Next**. In the next dialog box, you have to specify the source and target language of your project. From the source language list select English (United Kingdom) and from the list of the available languages select the target language (select from Estonian, French, German, Italian, Romanian or Slovenian) and click **Next**.



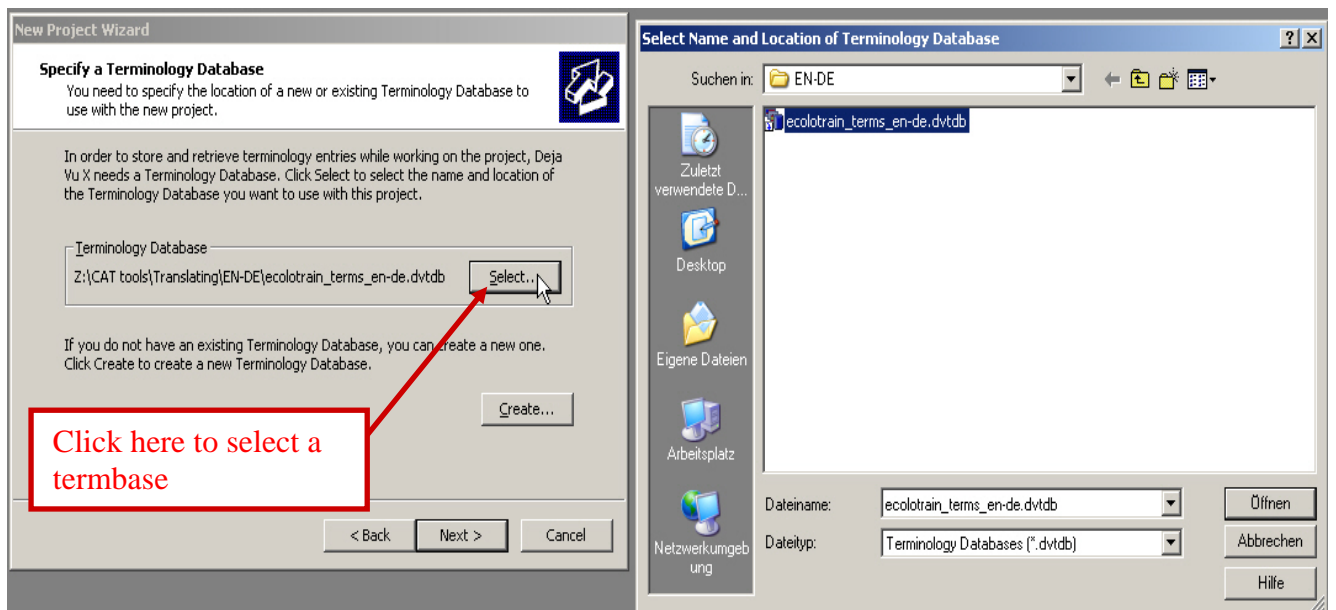
### 3. Create TM

- In the next window you have to create a new, empty translation memory database for our project. Click on **Create** and again specify your folder **My Documents\CAT tools\TranslatingProject** where your translation memory will be stored and enter a name for the new TM in the field **File name** (e.g. "ecolotrain\_ws\_saar"; the file name extension ".dvmdb" is again added automatically). After entering the file name and clicking **Save**, you are taken back to the **Project Wizard**, where your TM is selected. Click **Next**.



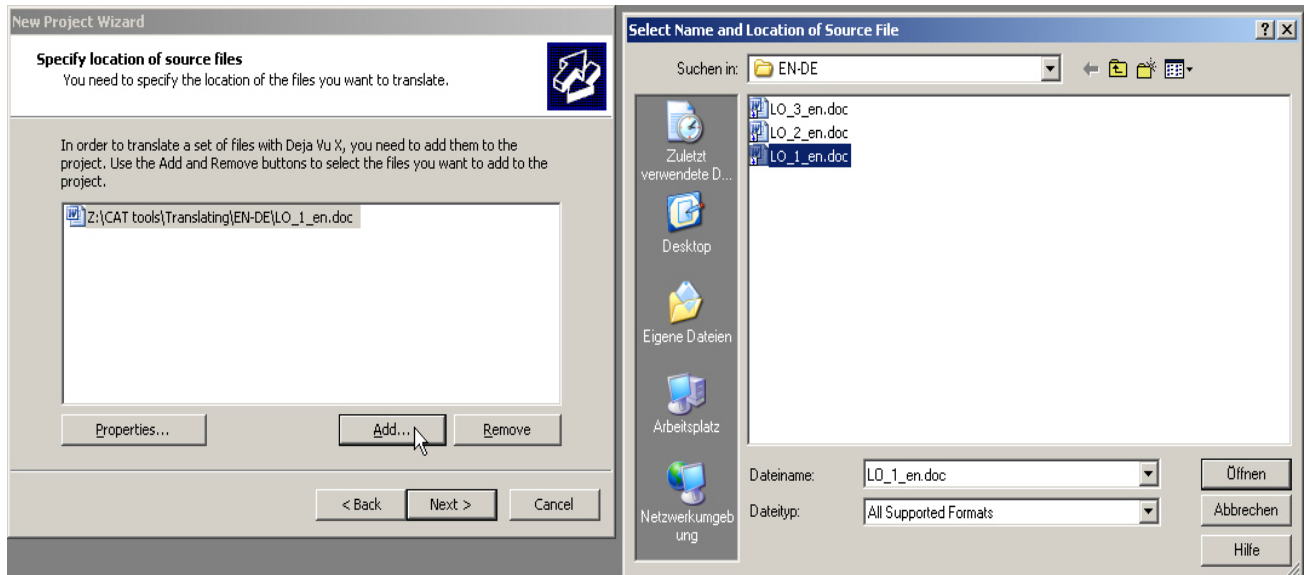
### 4. Select Termbase

- In the next step you have to select the termbase which you want to use for your translation project. Clicking on **Select** opens the dialog box for selecting the location of termbase you want to open. In the folder **My Documents\CAT tools\Translating** you will find several subfolders for different language pairs (EN-ET, EN-DE, EN-FR, EN-SI, EN-RO, EN-IT). Open the subfolder with the target language you selected for your project definition. In there you will find a termbase file with the name "ecolotrain\_terms\_en-xx.dvtdb", where "XX" is the code for your target language (in the example on the screenshot "ecolotrain\_terms\_en-de.dvtdb").



## 5. Specify Source Files

- Now you have to specify the source file(s) you want to translate in your project. Click on **Add** and again go to the folder **My Documents\CAT tools\Translating** and select the subfolder with your language pair. From this folder please select the English Word file "**LO\_1\_en.doc**":



## 6. Client and Subject

- Clicking **Next** brings you to the dialog box where you can specify the client for whom you are making the translation and a subject field. Since in our case the client list is still empty, we leave the **Client** field empty, but choose from the list of subject fields the Subject "**004 Computer Science and technology. Computing**". *DéjàVu X* provides a predefined list of subject fields, which can be changed by the user. For the moment, we accept the default list.

## New Project Wizard

### Specify Client and Subject

You need to specify the client and subject information for the new project.



You can optionally specify client and subject information for the new project. This information will be used when reading from and writing to the databases. Use the dropdown lists below to select a client and a subject.

Client:

Add/Remove...

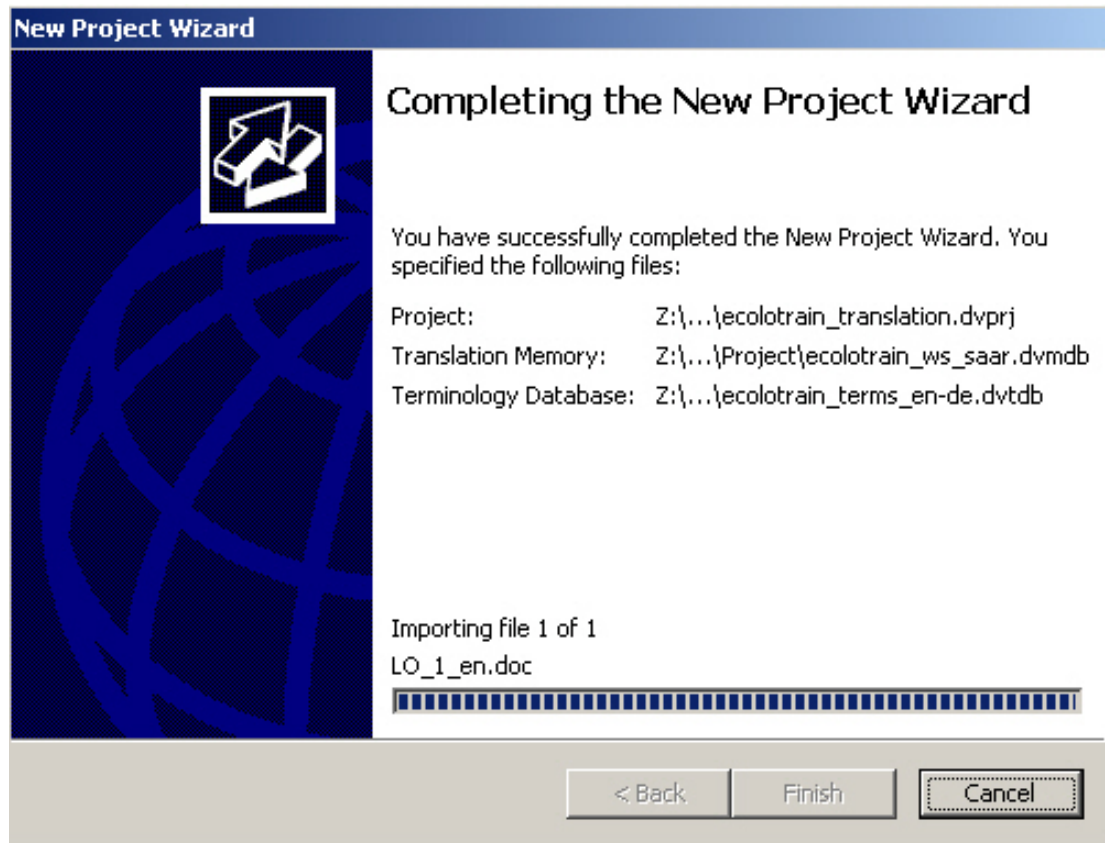
Subject:

- 0 - Generalities
- 00 - Prolegomena, Fundamentals of knowledge and culture, Computer science
- 001 - Science and knowledge in general, Organization of intellectual work
- 002 - Documentation, Books, Writings, Authorship
- 003 - Writing systems and scripts, Including: signs and symbols
- 004 - Computer science and technology, Computing
- 0042 - Computer architecture
- 0043 - Computer hardware

Cancel

## 7. Completing the New Project Wizard

- Clicking **Next** again brings us to the last dialog of the Project Wizard. Click **Finish**. The process of importing the source language file into the editor of *DéjàVu X* begins.



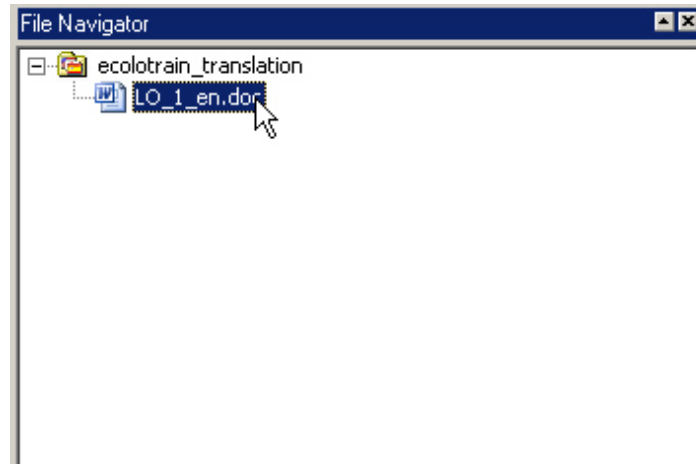
You have successfully created a new translation project in Déjà Vu X!

## II. Translating in DVX

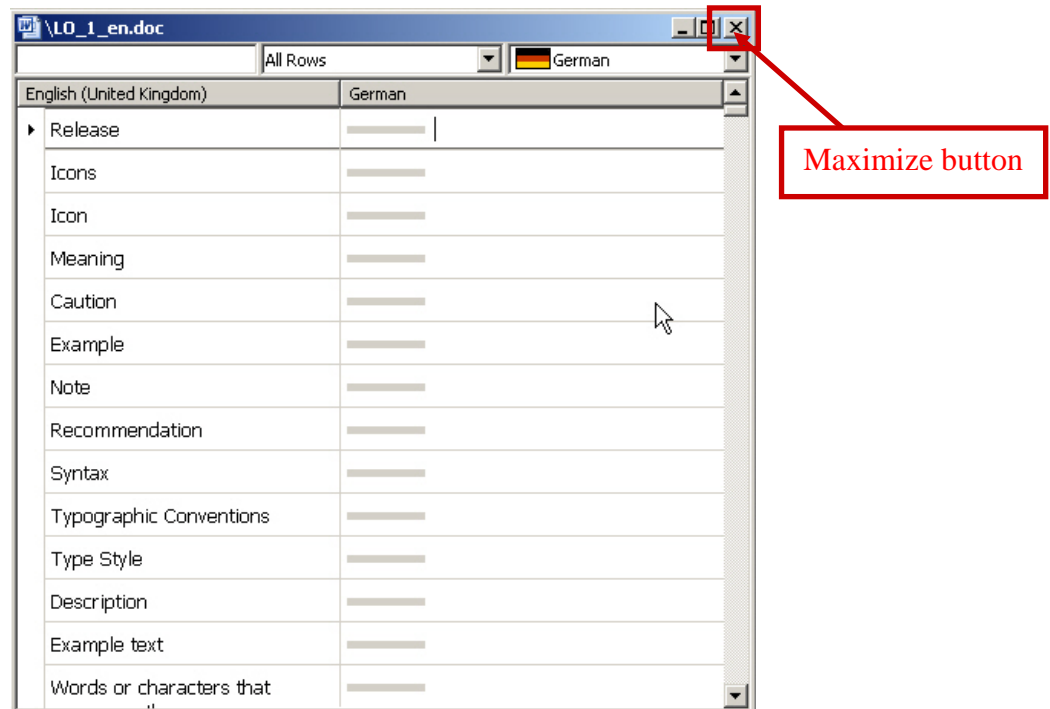
### 1. Getting Started

- After having completed the project definition, you will see on the right hand side of your screen the File Navigator window listing the name of your project and the source file(s) belonging to this project.



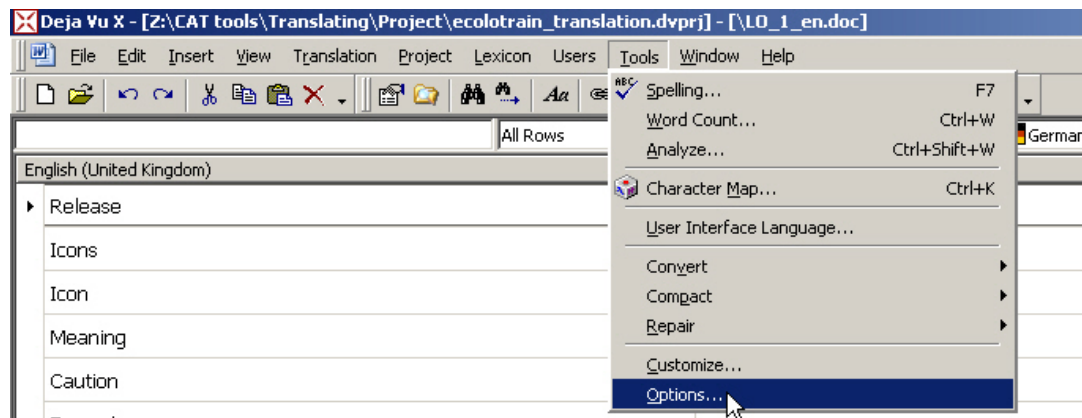


- Double clicking on the name of the file you want to translate opens the editor window of *DéjàVu X* with the source text in the first column and an empty column for the translation. Click on the **Maximize** button on the upper right corner of the editor window, in order to maximize it.



## 2. DVX Options: General

- You have to set up your *DéjàVu X* system to support an efficient way of translating your text. Open the **Tools** menu and choose **Options**.



- The **Options** dialog box opens. In the **General** tab you should set the options as shown in the screenshot below. The meaning of the options is:

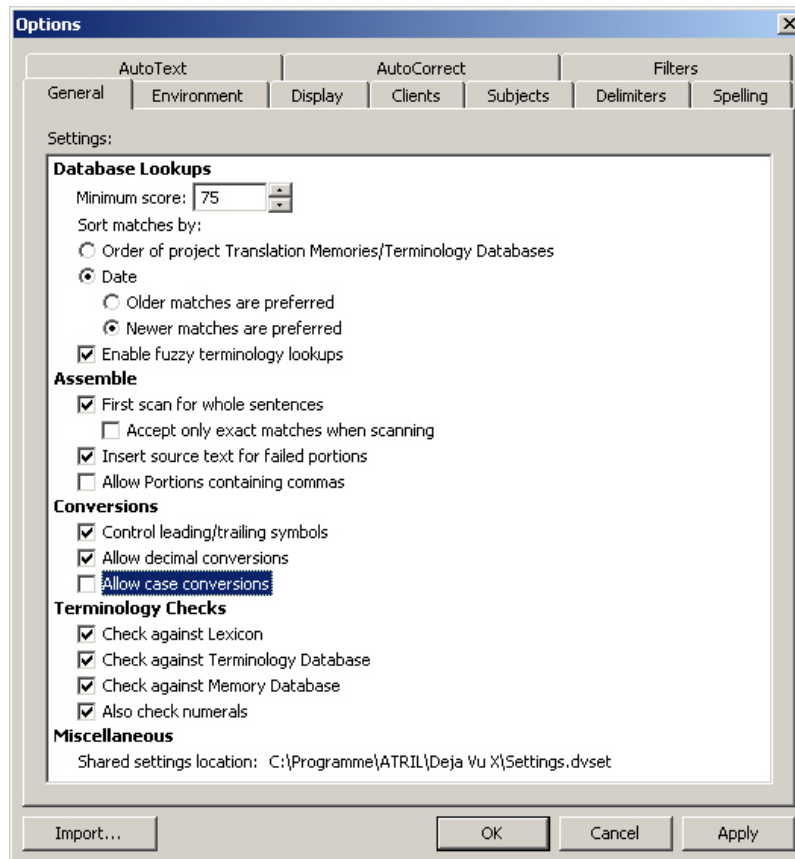
**Database lookups:** here you can set the minimum fuzzy match value (default 75%), the order of lookup in TMs and termbases, preference for older or newer matches if more than one match is found for a segment fuzzy terminology lookup, which is necessary because you want to find not only exactly identical terms but also e.g. inflected forms.

**Assemble:** to assemble translations from portions for segments which have no exact or fuzzy match in the TM, that is to combine results from both terminological and translation memory databases within one segment. You can also specify for DéjàVu X to first scan for whole sentences before assembling and that the source text shall be inserted for the portions which are not found in the termbase or TM.

**Conversions:** here you specify whether leading or trailing symbols (e.g. spaces, punctuation etc.) are controlled by the system and whether you want to allow conversion of number formats (e.g. between English and German/French ) and conversion of cases (i.e. convert the case of target word to the case of the word in the source text, which is, of course, not very helpful for translating from English into German).

**Terminology Checks:** in this part of the dialog just select all options, although we do not use a Lexicon in our exercise. Terminology will be checked against all possible sources containing terminological information.

The last setting in this dialog is showing the location of the file in which all these settings are stored on your computer.



### 3. Options: Environment

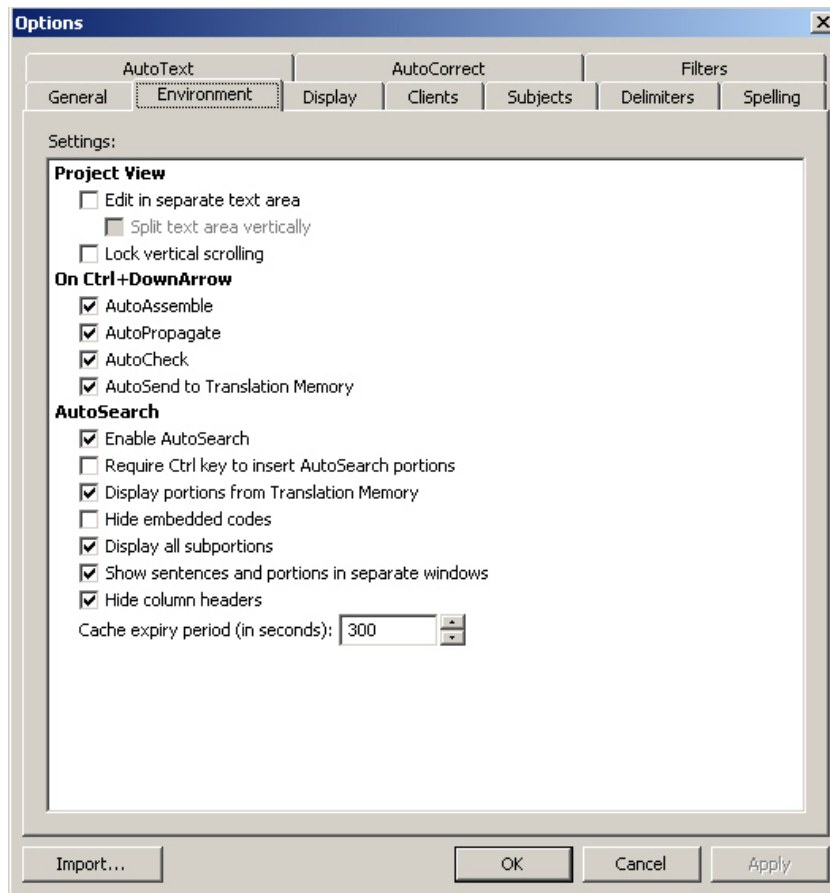
After having made these settings go to the **Environment** tab in the same **Options** window and set the environment settings as shown in the screenshot below. The most important of these settings are:

- **On Ctrl+DownArrow:** here you specify the actions to be taken, when pressing **Ctrl+DownArrow** in any translation segment:


  - AutoAssemble:** assemble a translation from portions of a segment found in the TM or termbase
  - AutoPropagate:** insert the current translation in all other occurrences of the same segment in the project file(s)
  - AutoCheck:** check the translation for consistency with the TM and the termbase
  - AutoSend to Translation Memory:** automatically send the translation of a segment to the TM when pressing **Ctrl+DownArrow**
- **AutoSearch**

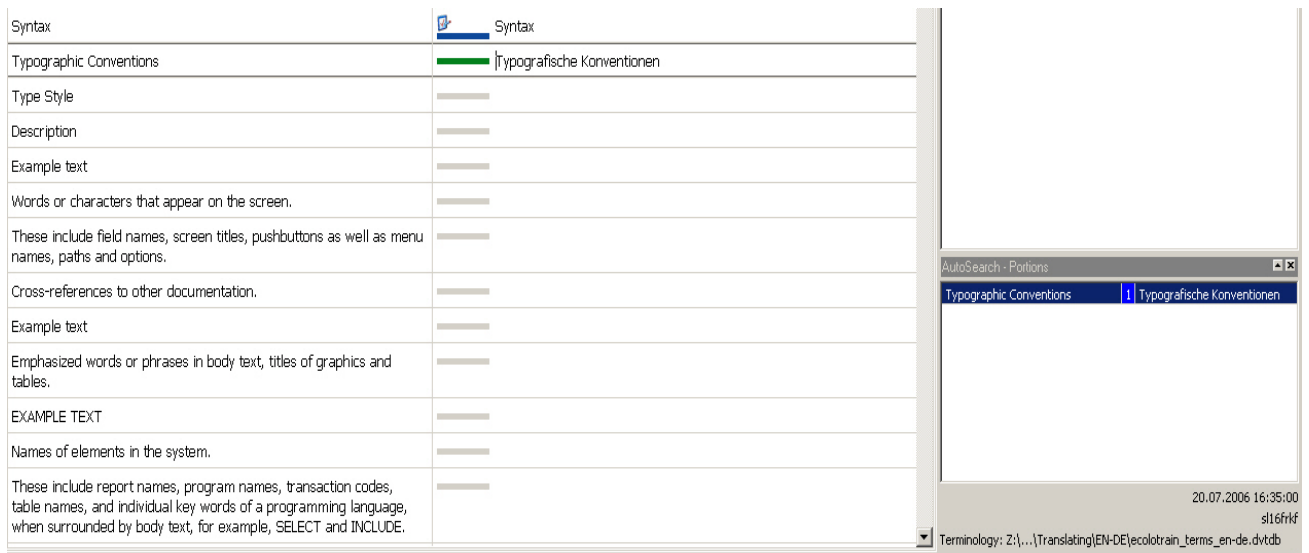
  - Enable AutoSearch:** automatically searches for segment and portions of segment in TM and termbase
  - Display portions from Translation Memory:** not only portions of segments found in the termbase but also those found in the TM are displayed in the AutoSearch window
  - Display all portions:** display larger portions as well as their subportions
  - Show sentences and portions in separate windows:** split AutoSearch window in two parts, one displaying only whole segments at the bottom of the *DéjàVu X* window and another one displaying portions from TM and termbase on the right hand side of the

window below the File Navigator window  
**Hide column header:** the column header in the AutoSearch windows containing the source and target language names can be displayed or hidden

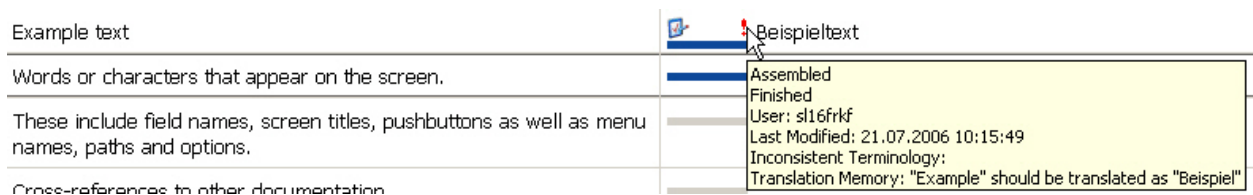


#### 4. Translating

- After having selected the settings in both the **General** and the **Environment** tab, click **OK** and you start translating. Click in the target cell of the first segment - since we have just selected the **AutoAssemble** option and nothing was found for the first segment in the termbase (the TM is still empty) the text of the source segment is inserted in the target cell and can be replaced by the correct translation. After having translated a segment, press **Ctrl + DownArrow** which brings you to the next segment, again the source text is inserted and can be replaced. Every time you confirm a translation in this way the small column between source and target segment is changed to  which means that *DéjàVu X* tried to assemble this translation and that the translation is now "finished". When you reach the first segment for which *DéjàVu X* finds a translation in the termbase, the AutoSearch - Portions window will show this translation with a number on a blue background between source and target term (blue means: portion found in termbase). Since we enabled AutoAssemble, the translation is automatically inserted in the target segment, marked with a green line between source and target, which means that this was 100% match in the termbase. You can confirm this translation using **Ctrl + DownArrow** and go to the next segment.



- The translated segment will again be marked as finished. In your translation, it may happen that the **AutoCheck** option seems to find an "inconsistent" translation, which does not correspond to the translation for a term listed in the termbase. In this case, the source word for which *DéjàVu X* pretends to have found an inconsistent translation is highlighted in red and a red exclamation mark is shown between the source and the target segments. If you point with your mouse on the exclamation mark a quick info appears telling you where the inconsistency has been found.



- Check the translation and, if you think it is OK, just confirm it again using **Ctrl + DownArrow**.


## 5. Codes and Bars

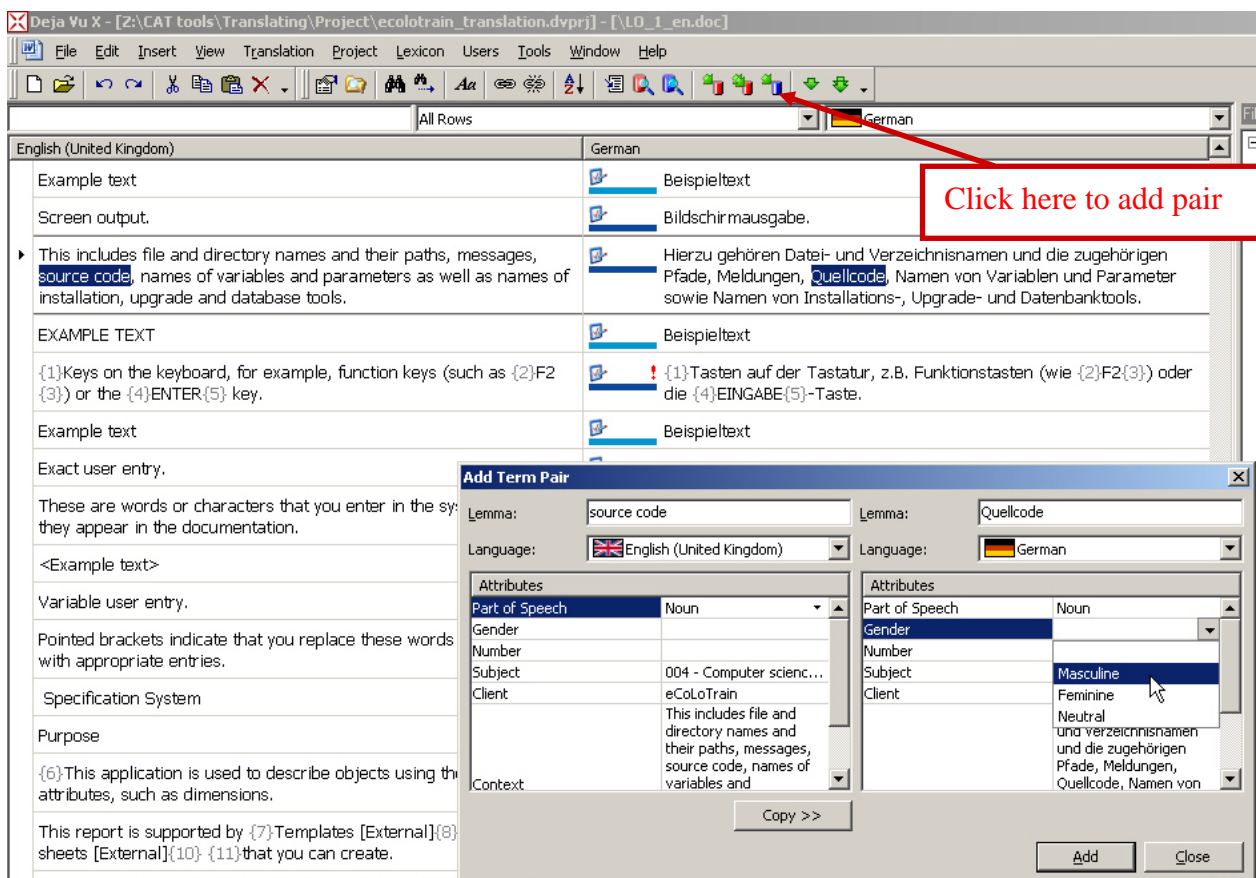
- During translation process you will meet sentences containing "**embedded codes**", i.e. codes which are placeholders for some formatting codes in the original text. In *DéjàVu X* these codes are always displayed as numbers in curly brackets, e.g. {1} throughout the whole text. In *DéjàVu X* you are not allowed to change or delete these codes, you cannot change the order of these codes. If a code is accidentally deleted, a warning symbol (⊗) is displayed between the source and the target segments, as soon as you leave the segment using **Ctrl + DownArrow**. In this case, copy the code using normal Windows copy techniques from the source segment to the correct location in the target segment. Press **Ctrl + DownArrow** again and the warning symbol disappears.
- You will also see segments with a light blue bar between the source and the target segments. These are segments where the translation has been auto-propagated after the

segment was translated for first time in the project. If the translation is also correct in this context, you can confirm it using **Ctrl + DownArrow**.

EXAMPLE TEXT	Beispieltext
{1}Keys on the keyboard, for example, function keys (such as {2}F2 {3}) or the {4}ENTER{5} key.	! {1}Tasten auf der Tastatur, z.B. Funktionstasten (wie {2}F2{3}) oder die {4}EINGABE{5}-Taste.
Example text	Beispieltext
Exact user entry.	Exact Propagated Last Modified: 20.07.2006 17:39:24
These are words or characters that you enter in the system exactly as	

## 6. Add Terms

- During your translation process you might want to **add source and target terms** to your termbase, in order to make them available for other translation projects, too.
- To add a new pair of source and target terms to your termbase, highlight both terms in the source and target segment and click on the **Add Pair to Terminology Database button** () in the tools bar. In the dialog box which opens you see the two terms together with the corresponding segment as **Context**. You can specify further information like Part of Speech, Gender etc. and then click on **Add** to send the entry to the termbase.



The screenshot shows the CAT tool interface with a translation project open. The source text (English) and target text (German) are visible. A red box highlights the 'Add Pair to Terminology Database' button in the toolbar, with a red arrow pointing to it and the text 'Click here to add pair'. The 'Add Term Pair' dialog box is open, showing the source term 'source code' and the target term 'Quellcode'. The dialog box includes fields for Lemma, Language, and Attributes (Part of Speech, Gender, Number, Subject, Client). The 'Part of Speech' is set to 'Noun' and the 'Gender' is set to 'Masculine'. The 'Context' field contains the source text segment: 'This includes file and directory names and their paths, messages, source code, names of variables and parameters as well as names of installation, upgrade and database tools.'

- Go on with your translation as far as you want. If you want to finish translating, just choose the Option **Close** from the **File** menu in *DéjàVu X*.

### III. Update Scenario

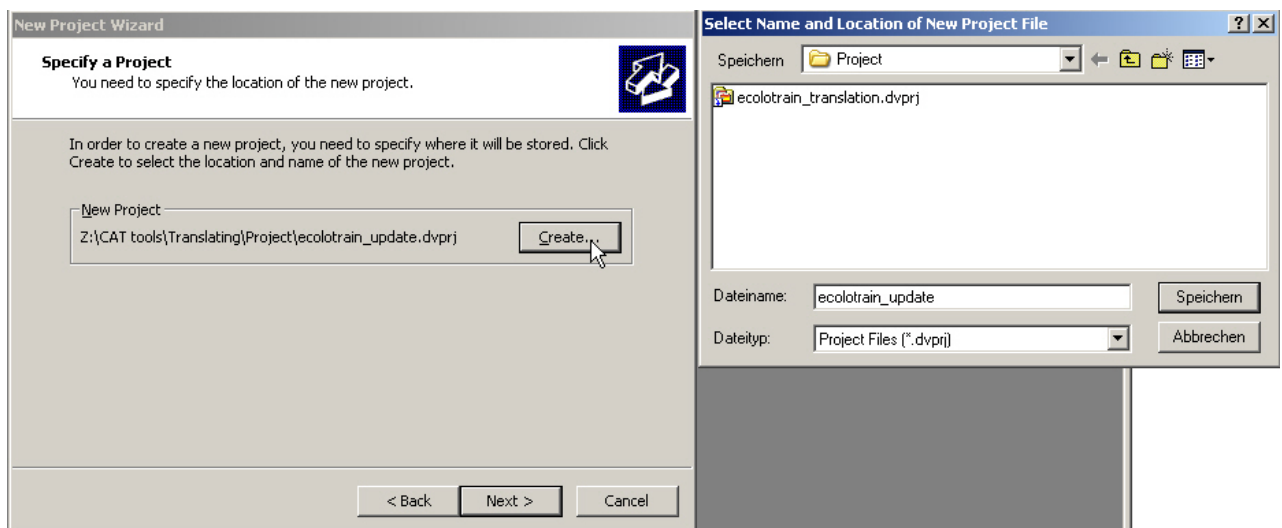
In this part of the exercise, we will assume that we finished our first translation project and will now go on to a project, where we have to translate an update version of the document we have just translated.

In this update scenario, we will use the TM populated with the translations of the first project and also the termbase we used for the first project.

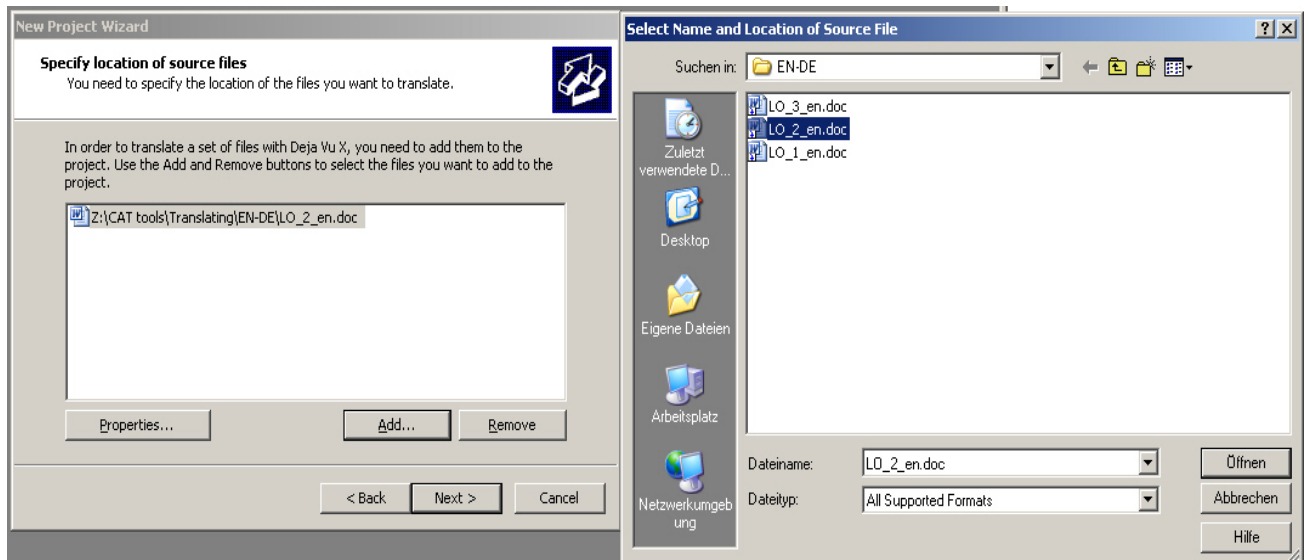
We will have to translate a document which is quite similar to the first one.

#### 1. Getting Started

- In the first step, you need to define a new project in the same way as before by choosing **New** from the **File** menu. In the New Project Wizard create a new project and save it in the folder **Project** in **My Documents\CAT tools\Translating** (name it e.g. *ecolotrain\_update.dvprj*).



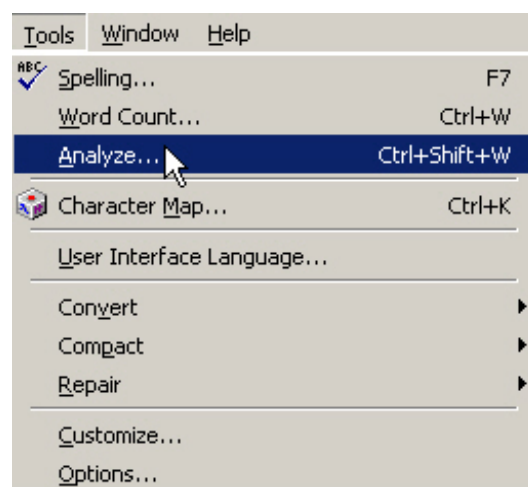
- Click on **Next** and choose the same source and target language as in the first project. When you are asked to specify the TM and then the termbase to be used, choose the TM and termbase created and/or used in first project (**ecolotrain\_ws\_saar.dvmdb** as TM from the **Project** folder and **ecolotrain\_terms\_en-xx.dvtdb** as termbase from your language pair folder, with EN as source language).
- As source text to be translated, choose from your language pair folder the file *LO\_2\_en.doc*.



- In the window where you have to specify client and subject field, leave again the client field empty and choose as **Subject** "004 - Computer Science and Technology. Computing".
- Clicking **Next** completes the definition of the project, clicking **Finish** starts the import of the file into *DéjàVu X* and shows the file in the **File Navigator**.
- Double clicking on the file name in the **File Navigator** opens the file in the editor again.


## 2. Analysing

- Before starting to translate this update document, analyse it using the analysis feature of *DéjàVu X*. This will provide information on the number of exact and fuzzy matches this document contains. It will also show how many segments are repeated twice or more in the document. To do this, choose the Option **Analyze** from the **Tools** menu.



- In the Analyze dialog box click on the **Analyze** button. *DéjàVu X* analyses the document on the basis of the TM and displays the results.



Analysis				
Language:  German				
	Lines	Words	Characters	Codes
Guaranteed	0	0	0	0
Exact	26	94	611	7
Duplicate	0	0	0	0
99%	2	4	24	0
98%	0	0	0	0
97%	0	0	0	0
96%	0	0	0	0
95%	0	0	0	0
90%-94%	0	0	0	0
80%-89%	1	14	101	1
70%-79%	4	71	432	4
60%-69%	0	0	0	0
50%-59%	0	0	0	0
40%-49%	0	0	0	0
30%-39%	0	0	0	0
20%-29%	0	0	0	0
10%-19%	0	0	0	0
0%-9%	0	0	0	0
No Match	27	421	2711	36
<b>Total:</b>	<b>60</b>	<b>604</b>	<b>3879</b>	<b>48</b>

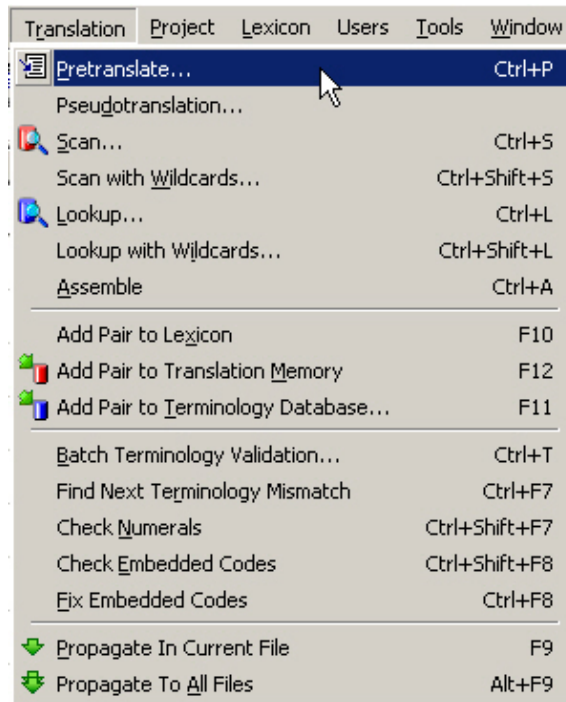
Global  
Internal Repetition: 6%

Analyze      Copy      Save      Close

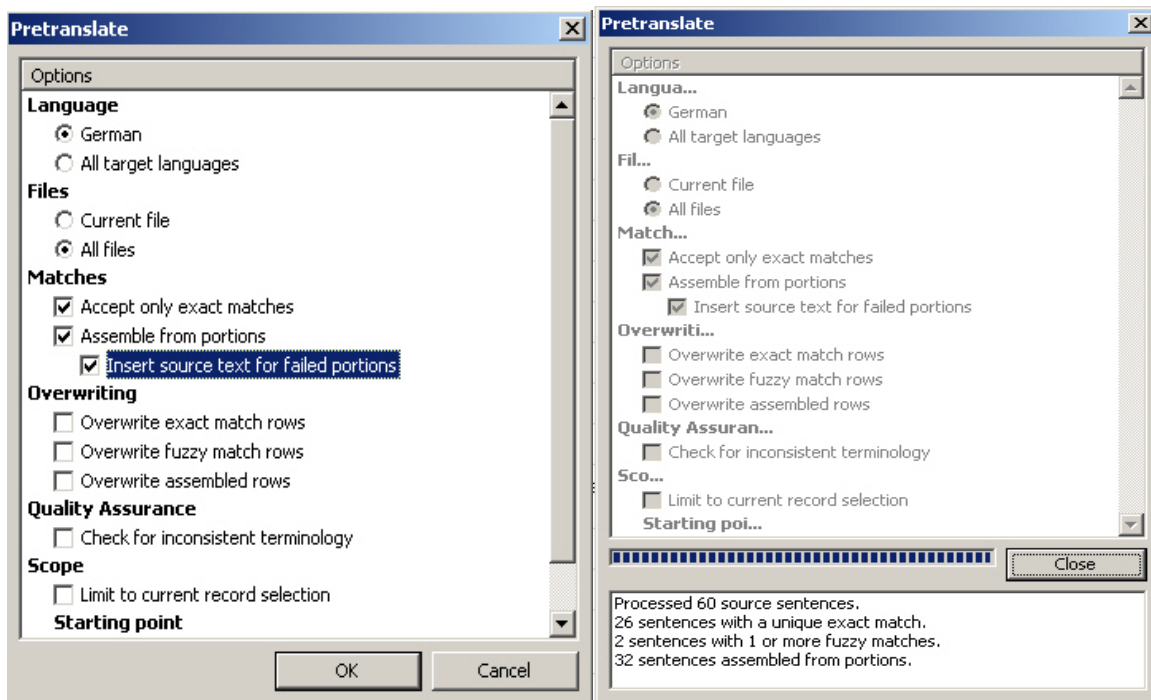
In this example, there are 26 segments ("Lines" in the Analysis window), which are exact matches, two 99% fuzzy matches, one fuzzy match between 80-89% four 70-79% fuzzy matches and 27 segments with No Match at all. You can close the Analysis window or save the result to a file. If you do so, choose your Project folder as location for the text file containing the result.

### 3. Pretranslating

- From this analysis result, we can conclude that it might be worthwhile to let *DéjàVu X* make a pretranslation. As such it will automatically translate all segments with exact matches in the TM. Choose the option **Pretranslate** from the **Translate** menu in *DéjàVu X*.



- In the **Pretranslate** dialog box select the options like the following screenshot.
- In the **Language** option choose your target language, since your project does not contain any other languages. Since you are only translating one file, the setting in the **Files** option is irrelevant. Under **Matches** select **Accept only exact matches**, so that you do not get pretranslations for fuzzy matches, because there might be the danger that you forget to check translations of fuzzy matches afterwards. Under this option also select **Assemble** from portions and **Insert source text** for failed portions. Thus, *DéjàVu X* will try to assemble translations for segments with fuzzy matches or with no matches and will insert the source text in those cases where there are no portions available. In this exercise, use this option to have text inserted in all target segments, even if time will not allow the completion of the entire translation. At the end of the exercise you will export the translation work back into Word format again, without getting too much error messages from *DéjàVu X* on missing target segments or embedded codes. The remaining options are only relevant, if pretranslation is carried out in a project where there were already some parts of the document translated and you want to control whether certain existing translations should be overwritten or not.
- When you have completed the settings of the **Pretranslate** options just click on **OK** and the pretranslation process is started. After completion, the result is displayed at the bottom of the Pretranslate options window.

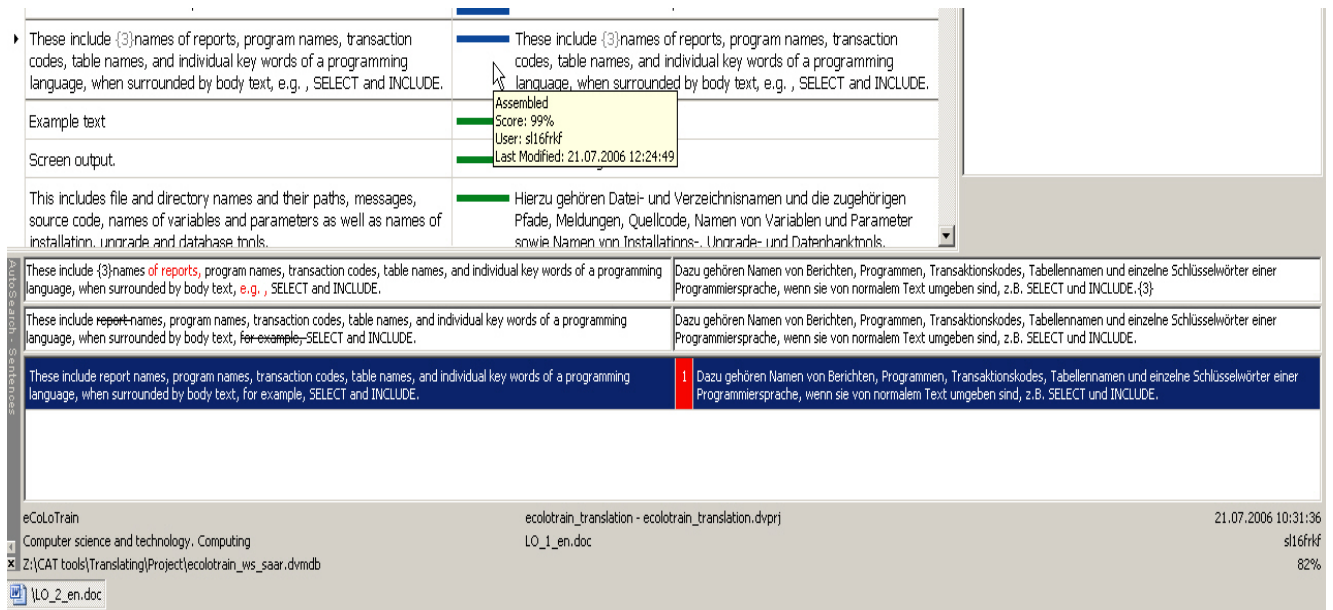


#### 4. Pretranslation Results

- The *DéjàVu X* editor shows the result of the pretranslation process. The status of each segment is colour coded. Segments with a green bar are 100% matches from the TM(s) or the termbase(s), segments with a blue bar are segments which have been assembled because they do not have an exact match. The segments without a match have to be translated manually by the translator using the TM.

English (United Kingdom)	German
▸ Caution	■ Achtung
Example	■ Beispiel
Note	■ Anmerkung
Recommendation	■ Empfehlung
Syntax	■ Syntax
Typographic Conventions	■ Typografische Konventionen
Type Style	■ Format
Description	■ Beschreibung
Example text	■ Beispieltext
{1}Words or characters appear{2}ing on the screen.	■ {1}Words or characters appear{2}ing on the screen.
These include field names, screen titles, pushbuttons, menu names as well as directory paths and program options.	■ These include field names, screen titles, pushbuttons, menu names as well as directory paths and program options.
Cross-references to other documentation.	■ Querverweise zu anderen Dokumenten.
Example text	■ Beispieltext
Emphasized words or phrases in body text, titles of graphics and tables.	■ Hervorgehobene Wörter oder Phrasen im Text, Titel von Grafiken und Tabellen.
EXAMPLE TEXT	■ Beispieltext
Names of elements in the system.	■ Names of elements in the system.
These include {3}names of reports, program names, transaction codes, table names, and individual key words of a programming language, when surrounded by body text, e.g. , SELECT and INCLUDE.	■ These include {3}names of reports, program names, transaction codes, table names, and individual key words of a programming language, when surrounded by body text, e.g. , SELECT and INCLUDE.
Example text	■ Beispieltext
Screen output.	■ Bildschirmausgabe.
This includes file and directory names and their paths, messages, source code, names of variables and parameters as well as names of installation, upgrade and database tools.	■ Hierzu gehören Datei- und Verzeichnisnamen und die zugehörigen Pfade, Meldungen, Quellcode, Namen von Variablen und Parameter sowie Namen von Installations-, Upgrade- und Datenbanktools.

- You can now look through the pretranslated segments using **Ctrl + DownArrow** to go from one segment to the next and confirm the translation. When you come to a segment where you only see the source segment in the target cell, you have to enter a translation manually. Another possibility is to look in the **AutoSearch-Sentences** window at the bottom of the *DéjàVu X* screen to view the fuzzy match(es) found in the TM for this segment. This information is accompanied by a match value on the right hand side of the window and some other information, like subject field, file name of the TM, project name of the fuzzy match etc.

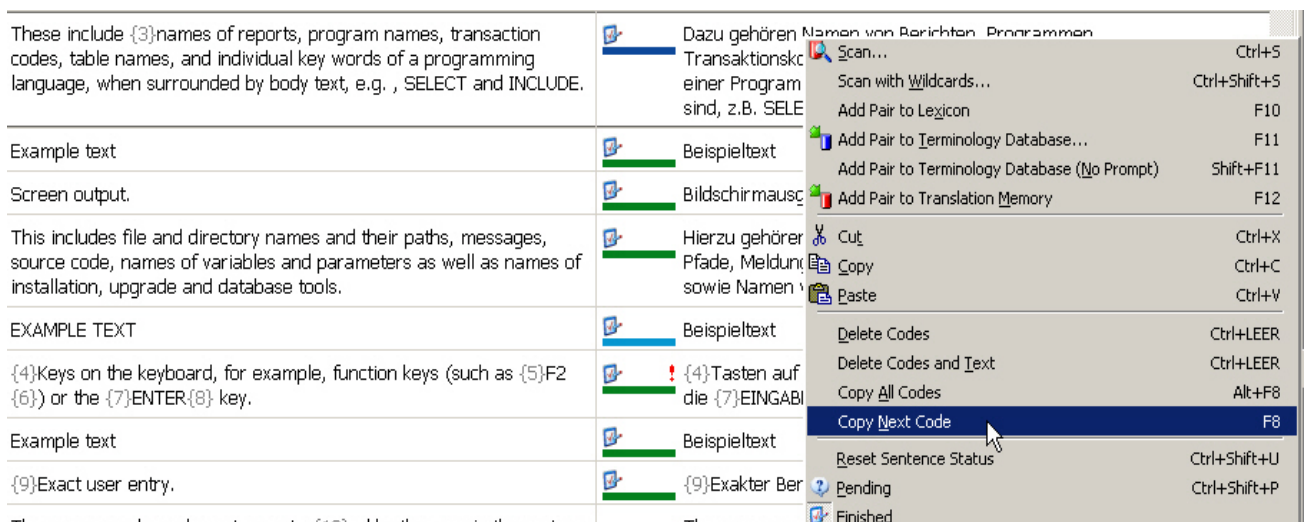


The first four boxes in the AutoSearch-Sentences window contain the source sentence as it occurs in the current document, its translation found in the TM (including embedded codes) as well as the source sentence found in the TM together with the translation. In the source segments, the differences between the old and the new sentences are shown in red characters or marked as deleted.

The box with the blue background and the number (on red background) is the place where you can select the translation to be inserted in the target text part of the editor by double clicking on the number or somewhere else in the segment. If more than one match was found, they appear in this window below each other and you can select the translation you want by double clicking on it.

## 5. Manual Corrections

- Sometimes it might happen that in the target segment an embedded code is missing or is placed at the wrong position. In this case, you have to insert or move it manually. In order to insert a code manually, right click with your mouse at the position where the code has to be inserted and choose **Copy Next code** from the popup menu or press **F8**.



In order to move a code within the target segment, you can use the normal Windows operations for cutting and pasting text elements. Select the code to be moved in the target text, cut it using **Ctrl + X**, move the mouse pointer to the position where the code has to be inserted and paste it using **Ctrl + V**.

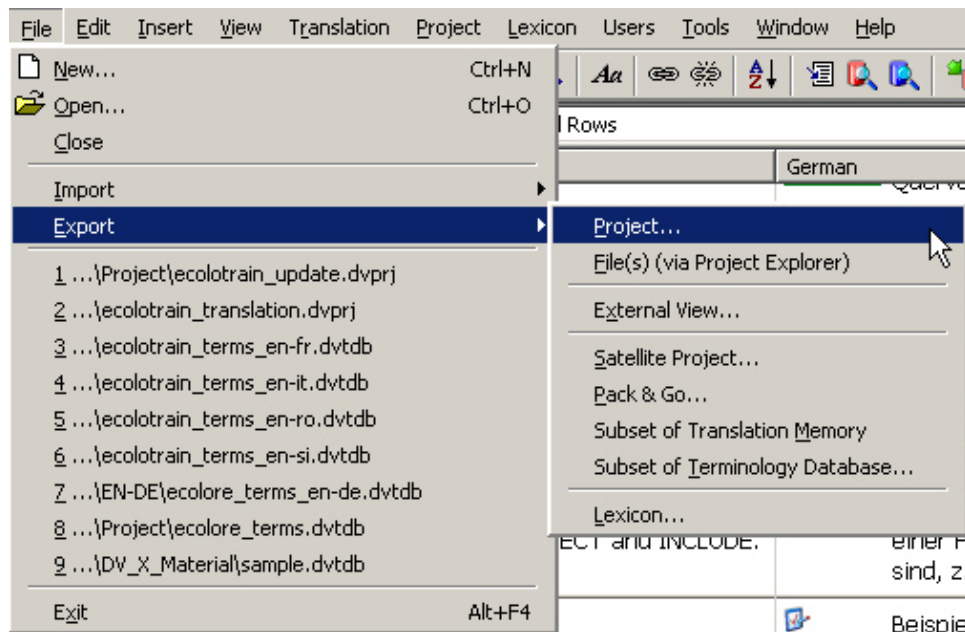
Continue translating in this way.

## IV. Exporting

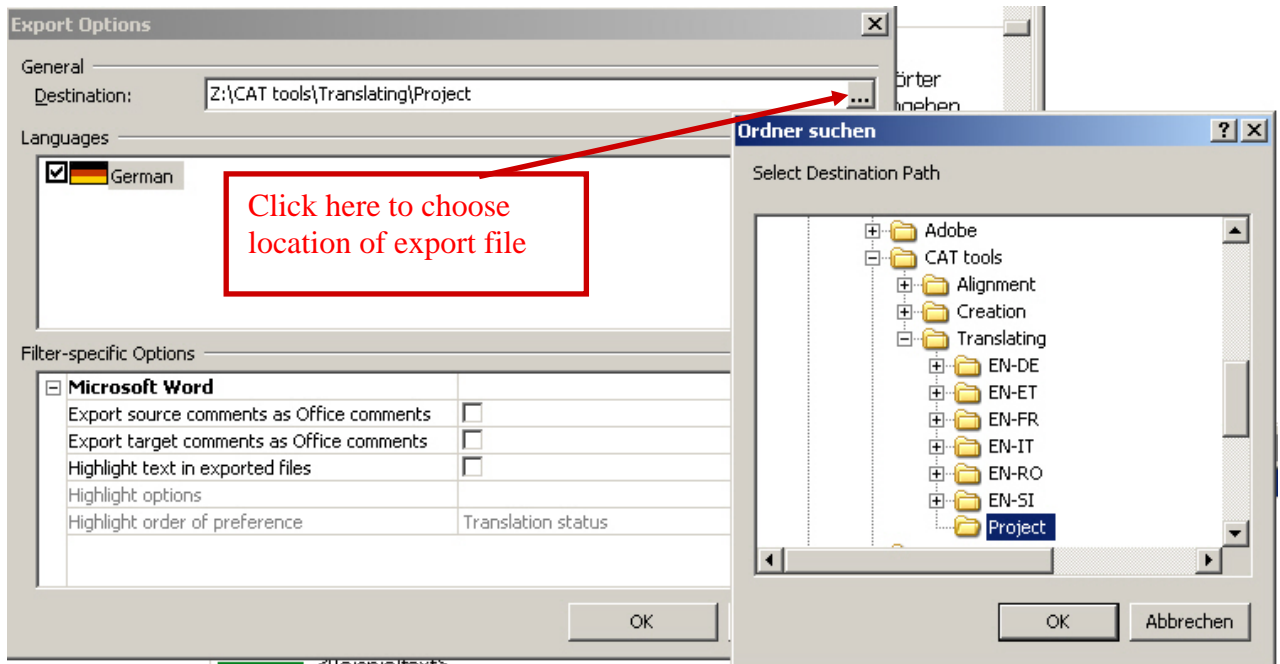
### 1. Getting Started

When the translation of a file or project is finished, you have to export it from *DéjàVu X*, in order to restore it to the original format.

- Choose the option **Export/Project...** from the **File** menu in *DéjàVu X*.



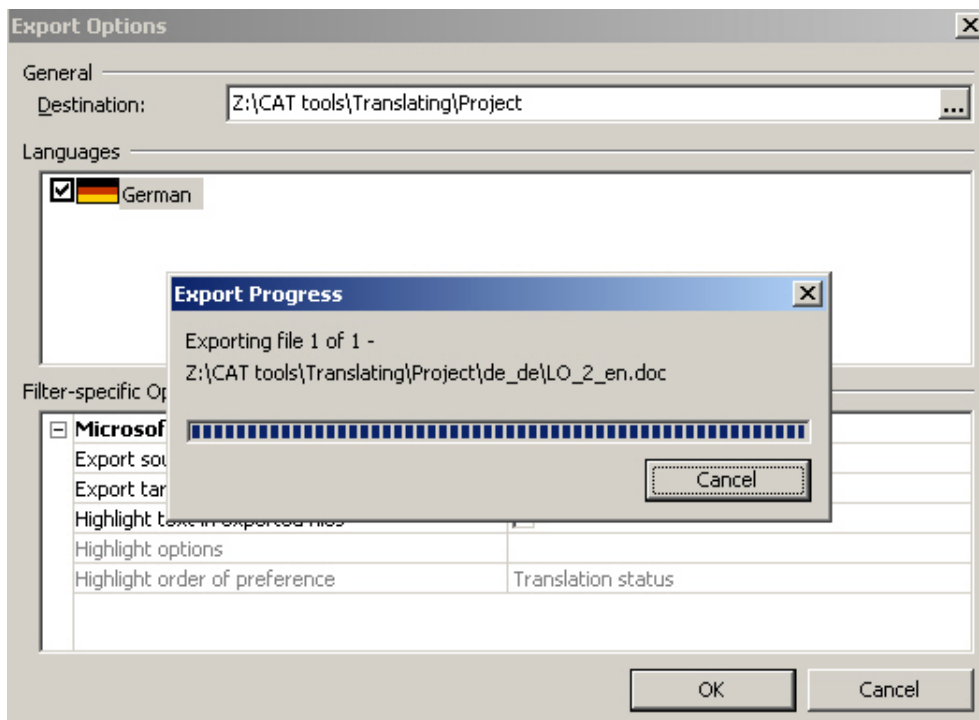
- In the **Export Options** dialog box choose the location where the exported file shall be stored. For this exercise, choose your Project folder in **My Documents\CAT tools\Translating**.



## 2. Export Progress

- After having specified the location of the export file, click on **OK** in the **Export Options** dialog box.

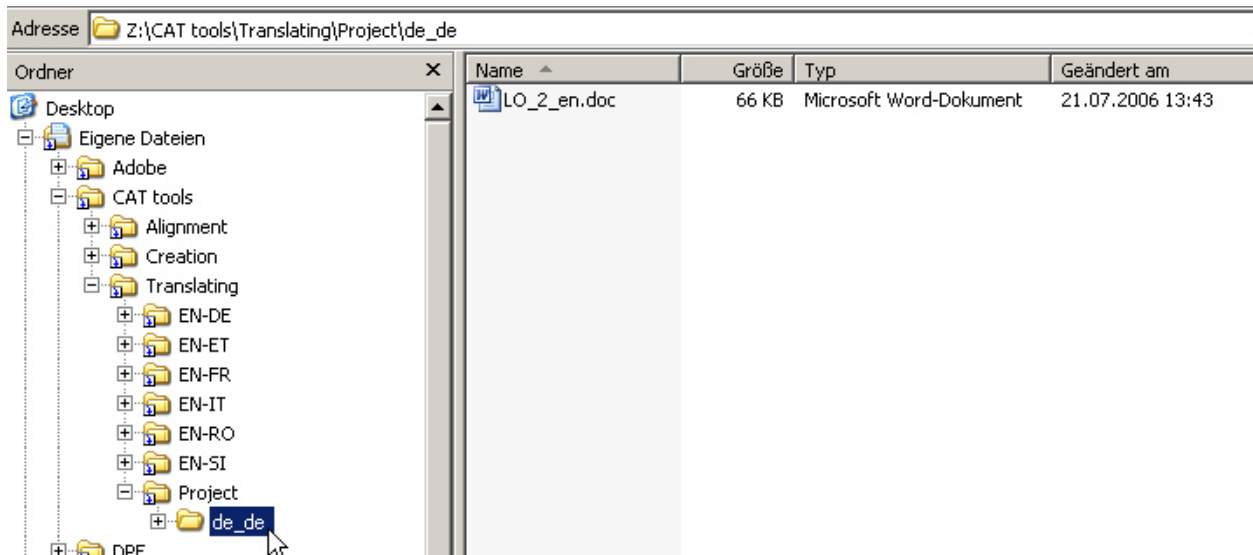
The translation file(s) are now exported from the project.



During the export process *DéjàVu X* creates a new folder for the exported file(s), in order to avoid overwriting the source text, if the translator chooses the same folder where the source

text is located. In the example given new folder "**de-de**" was created for the target language file (German) in the **Project** folder.

During the export phase, an error message will appear for every target segment that does not contain the embedded code(s) of its equivalent source segment. In this case, you will have to end the export, go back in the editor and make sure all the codes from the source segments are copied in their target equivalents. Then, repeat the export steps explained above.



- You can now navigate to the folder where your translation is stored and open the document in *Microsoft Word*. You can verify if the layout of the exported target file is the same its original, which you can still find in the folder for the respective language pair.

## End of exercise II

Congratulations! You have successfully completed exercise II of the eCoLoTrain Translation Memory module.