

Content of Pre-treatment packages for Multilingual Documents and General Instructions for External Translators (with SDL Studio) v13

Directorate for Support and Technological Services for Translation

Euramis Pre-Translation Unit

Contents

1. Content of pre-treatment zip packages	3
1.1 Content of a pre-treatment zip package for a multilingual (XM) document.....	3
1.2 Content of the individual Source Language (SL) - Target Language (TL) zip packages	3
1.2.1 Translatable files in the individual SL - TL zip packages	5
1.2.2 Translation memories in each individual SL - TL zip package	7
1.2.3 Studio project file in each individual SL - TL zip package	8
1.2.4 TMXs subfolder in each individual SL - TL zip package	8
1.2.5 Support files.....	9
2. General instructions for external translators	10
2.1 Setting-up the TL document with DocEP	10
2.2 Handling DocEP tags in the SDL Studio Editor	11
2.3 Multilingual documents without two-column amendments	12
2.4 Multilingual documents with two-column amendments.....	13
2.5 Handling of documents with two-column amendments which contain non-contractual source languages (for which a Pivot is required)	13
2.5.1 For the translation using the original zip package and multilingual document	13
2.5.2 For the translation using the PIVOT document and package	15
2.5.3 Generation of complete TL document.....	16
2.6 Delivery of the TL document	16

1. Content of pre-treatment zip packages

1.1 Content of a pre-treatment zip package for a multilingual (XM) document

The pre-treatment package for a multilingual document contains individually zipped packages for each source to target language combination, as well as support files.

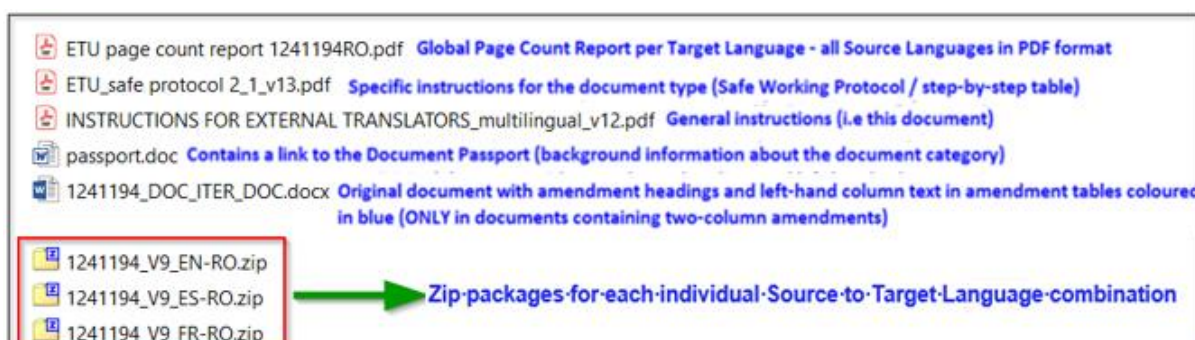


Figure 1: Content of zip package for a multilingual (XM) document.

1.2 Content of the individual Source Language (SL) - Target Language (TL) zip packages

The content of the individual SL-TL zip packages is different depending on whether:

- a) the source language is EN or any other language and
- b) the document contains two-column amendments or not.

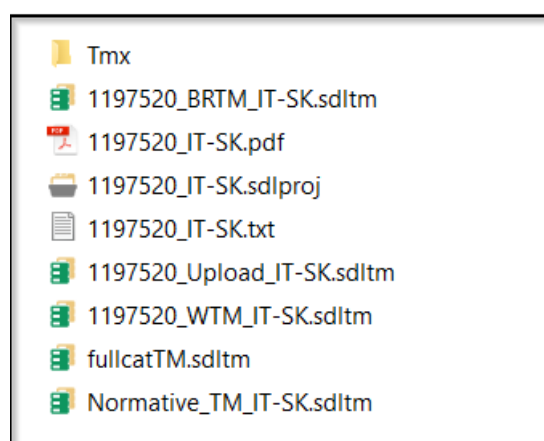


Figure 2: Content of a Source Language (other than EN) - Target Language zip package for documents with two-column amendments (general case).

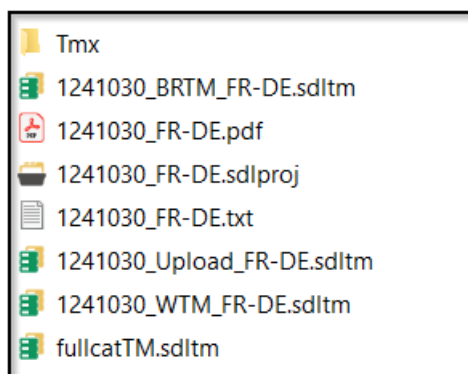


Figure 3: Content of a Source Language (other than EN) - Target Language zip package for documents without two-column amendments (general case).

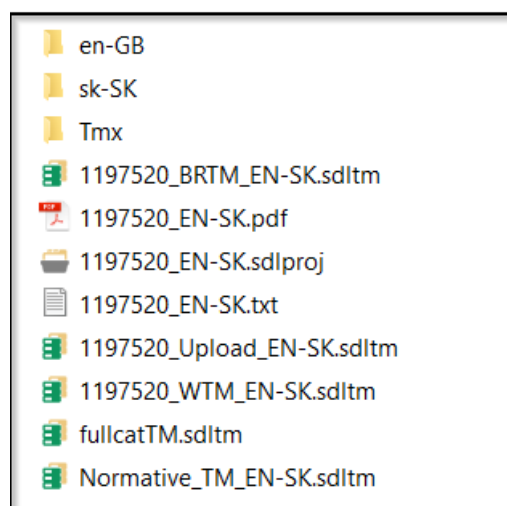


Figure 4: Content of an EN - Target Language zip package for documents with two-column amendments (exception).

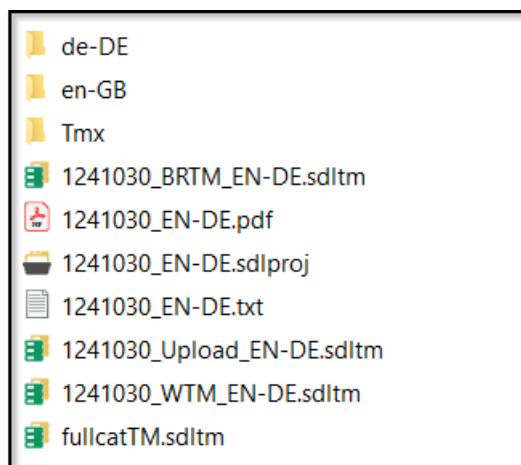


Figure 5: Content of an EN - Target Language zip package for documents without two-column amendments (exception).

1.2.1 Translatable files in the individual SL - TL zip packages

General rule (Figures 2 and 3) - any source language other than EN

There are NO translatable files in the Studio projects. Therefore, there are no SL or TL folders either.

Exception (Figures 4 and 5)

The only exception concerns the the **EN** language - Target language zip package which contains several files, as described below.

Case 1: documents with two-column amendments:

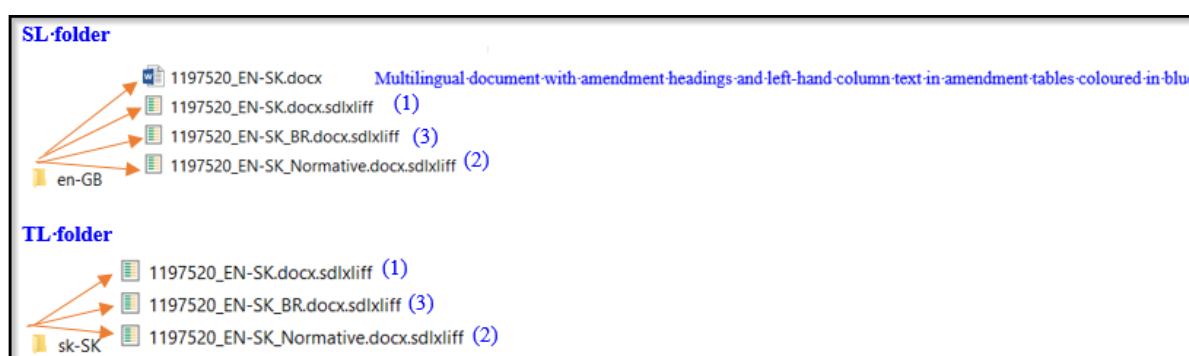


Figure 6: Translatable files in the SL and TL folders - documents with two-column amendments

The **original** multilingual document in its original format used to create the pre-treatment package is included in the source language (SL) folder.

N.B: For the documents containing two-column amendments, this is the so-called **Coloured version** in which the amendment headings and the left-hand

column text in the amendment tables have been coloured in blue in order to facilitate the identification of the corresponding segments in the Studio Editor.

The following translatable SDLXLIFF files are included:

- An SDLXLIFF file (1234567_EN-TL.**docx.SDLXLIFF**) named after the original document to be translated with the text of amendment headings and of the left-hand column in the amendment tables coloured in blue **(1)**.
- An SDLXLIFF file (1234567_EN-TL_**Normative.docx.sdlxliiff**) resulting from the pre-translation at 98% with the Normative TM **(2)**.
- An SDLXLIFF file (1234567_EN-TL_**BR.docx.sdlxliiff**) resulting from the pre-translation at 98% with the Normative TM and a consecutive pre-translation at 100% with the BRTM **(3)**.

Case 2: documents without two-column amendments:

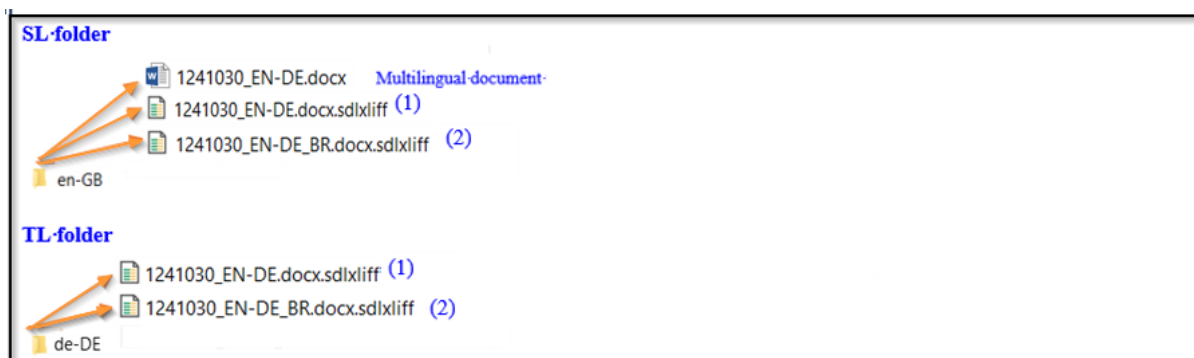


Figure 7: Translatable files in the SL and TL folders - documents without two-column amendments

The **original** multilingual document in its original format used to create the pre-treatment package is included in the source language (SL) folder.

The following translatable SDLXLIFF files are included:

- An SDLXLIFF file (1234567_EN-TL.**docx.SDLXLIFF**) named after the original document to be translated **(1)**.
- An SDLXLIFF file (1234567_EN-TL_**BR.docx.sdlxliiff**) resulting from the pre-translation at 100% with the BRTM **(2)**.

1.2.2 Translation memories in each individual SL - TL zip package

Depending on the document type, each individual SL-TL package may contain:







 1197349_BR2TM_EN-RO.sdltm	a Basic Reference 2 Memory
 1197349_BRTM_EN-RO.sdltm	a Basic Reference Memory
 1197349_Upload_EN-RO.sdltm	an Upload Memory (can be disregarded in your workflow)
 1197349_WTM_EN-RO.sdltm	a Working Memory
 fullcatTM.sdltm	a fullcat Memory (can be disregarded in your workflow)
 Normative_TM_EN-RO.sdltm	a Normative Memory

Figure 8: Translation Memories

The **Normative TM** contains the tmx file of the standard segments in the amendment headings preceding each amendment. This TM is created for documents with two-column amendments only.

The **BRTM** contains

- either the Basic Reference tmx file only, in the case of documents with two-column amendments (e.g.: for protocol 2.1 the BRTM contains the .tmx of the COM document);
- or the Basic Reference tmx file and the Normative tmx file containing standard text specific for that particular document type, for documents without two-column amendments.

This memory is used for pre-translating the original document at 100% in order to get a) the Basic Reference document text into the left-hand column or b) the Basic Reference document text and the translation of the standard text.

The **BR2TM** contains the Secondary Basic Reference tmx file(s) only (for example: for protocol 2.6.1 the BR2TM contains the .tmx of the draft report and its amendments). This memory is used for pre-translating the .BR.docx.sdlxliff file at 100% in order to get the text from previous versions into the right-hand column.

The **Working TM (WTM)** contains all the material, i.e. all the tmx files for the Basic Reference document(s), the Retrieval, the Normative tmx (in case of documents without two-column amendments) and the other reference documents belonging to the procedure.

The **Upload TM** and the **fullcat TM** are for EP internal use only, they should be disregarded in your workflow.

The zip package contains the General Instructions and the ETU Safe Working Protocol (i.e. a table in Word format, describing step-by-step how to work with the memories) to be used for that document. If no protocol is available, please

use the WTM in each individual SL-TL package as a resource for interactive translation in Studio or a Studio-compatible translation tool.

1.2.3 Studio project file in each individual SL - TL zip package

Each project is named after the original FdR number of the document plus the source and the target languages (**1234567_SL-TL.sdlproj**)

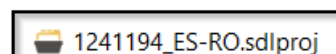
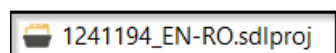


Figure 9: Studio Projects

1.2.4 TMXs subfolder in each individual SL - TL zip package

The folder contains the tmx files that have already been imported into the TMs according to the Safe Working Protocol.

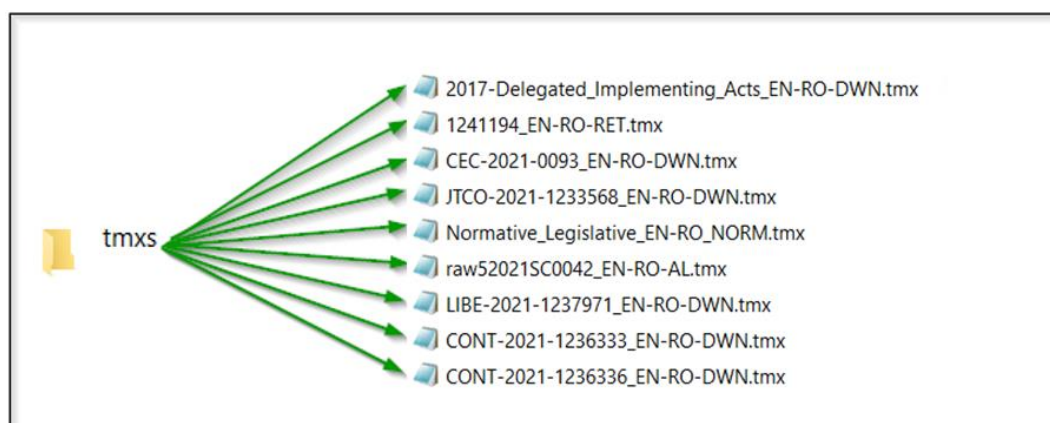


Figure 10: Tmx files in the tmxs subfolder

1.2.5 Support files

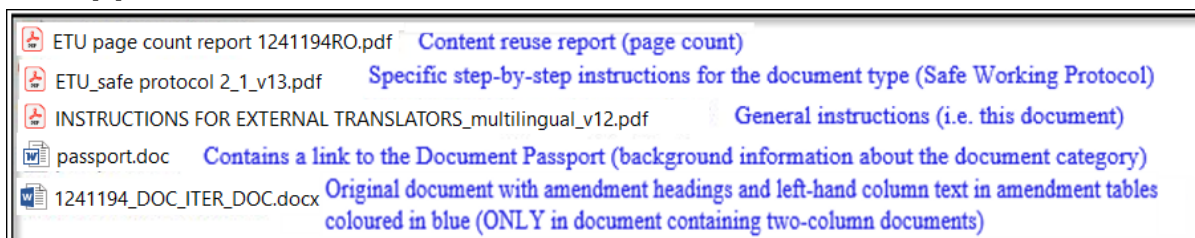


Figure 11: Support files in the zip package for the multilingual document.

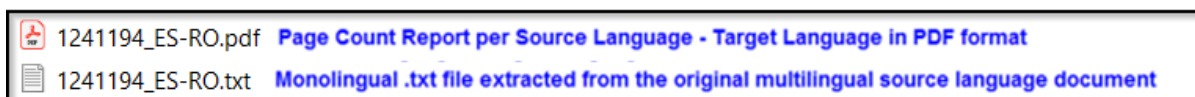



Figure 12: Additional support files in the individual SL - TL zip packages.

2. General instructions for external translators

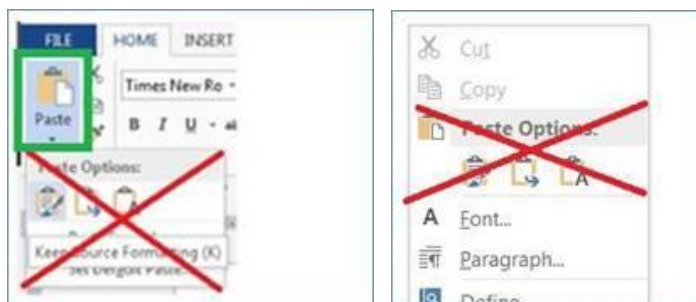
Always refer to the specific Safe Protocol for the type of document concerned before starting to work on the document.

2.1 Setting-up the TL document with DocEP

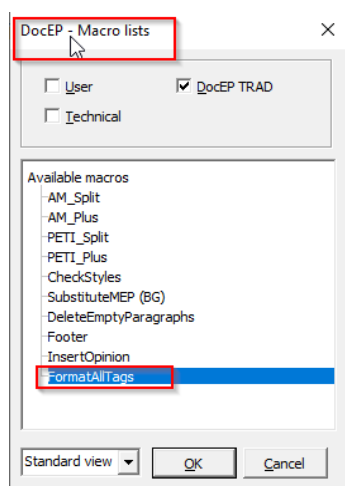
 **docep** is a document creation tool used in the European Parliament. It allows to create documents which contain the official standard phrases encoded in the Recueil de Modèles (RdM) already in the target language.

If the SWP recommends setting up the target language (TL) document to be used as working document in Studio with DocEP 'Create Document', when copying the body of the source language text be careful not to change the DocEP tag-specific style definition (*HideTWBExt* and *HideTWBInt* in dark Blue). Please take the following into consideration:

- ✓ Use the normal MS Word Paste – i.e. either the *Paste* button in the first screenshot below, or *Ctrl+V* – styles will be kept.
- ✓ For amendments use the specific functions in DocEP (*Copy AM*, *Copy all AMs*) which allow you to safely copy text with styles.
- ✗ When you use the additional Paste options in MS Word, the styles will be lost. This is also the case with the option *Keep Source Formatting*: the text looks exactly the same, but all the styles are *Normal*. In the context menu (right click), the regular Paste is not even present and all the available options will result in a faulty Paste.

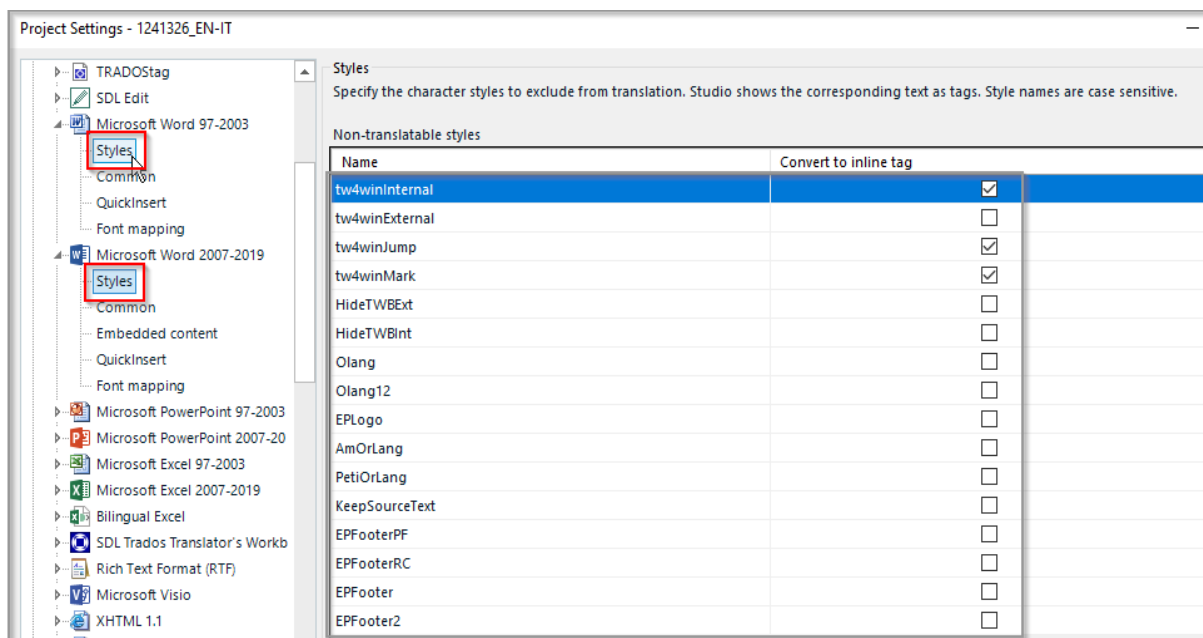


- ✓ Before adding the set-up document to the Studio project always check that all DocEP tags have the two mentioned styles. Should it not be the case (the tag style is *Normal*), use the macro *FormatAllTags* from the DocEP Macro Lists menu. It re-applies the correct styles throughout the document.



2.2 Handling DocEP tags in the SDL Studio Editor

If you use SDL Studio, we recommend that you replicate the settings below in the *File Type* definitions for MS Word 97-2003 and 2007-2019 (under SDL Studio – Project Settings – File Types). DocEP tags will then be treated as structure tags, which by definition remain outside the segments, thus eliminating the risk of damaging them while translating. The settings shown on the images below are also the EP's default in-house settings.





If these settings are not used in SDL Studio, the DocEP tags may be contained within the segments. Should this be the case, they must be copied from the SL to the TL segments.

If you use another CAT tool, we recommend that you look at the possibility of introducing similar settings, so that the DocEP tags do not need to be handled by the translator. If such a possibility does not exist, then the standard XLIFF files delivered must contain the relevant tags in the appropriate location in the TL.

2.3 Multilingual documents without two-column amendments

If the document contains text in non-contractual source languages (for a which a Pivot is required), the general rules are as follows:

- Set up the target language document with the document specific DocEP template (if recommended by the SWP) and copy the body of the original multilingual document into the new TL document.
- Split or divide the multilingual TL document by source languages. Leave the splits corresponding to non-contractual source languages aside.
- For the contractual languages: add each SL-TL split document to the corresponding SL-TL Studio project. Perform the following steps for each SL-TL document split:
 1. Pre-translate against the BRTM at 99% in order to get the unchanged text from the basic reference document and/or any relevant standard text from the Normative TMX file.
 2. Open the pre-translated file. Since the target language document was initially set up with DocEP '*Create Document*', the source language column in Studio may contain DocEP text in the target language already. Change the status of such segments to 'locked' and leave the corresponding TL cells empty. If you are using a different CAT tool, copy the source (i.e. DocEP text already in the TL) to the target cells.
 3. Go through the document, checking that the target language segments in the pre-translation come from the appropriate TM and the correct reference document(s). Use the WTM to complete the translation.
 4. Generate the translated TL document splits and a bilingual sdxliff file (or a standard bilingual xliiff file) for each contractual SL-TL combination.
- When the Pivot package arrives: take out from the Pivot document the parts corresponding to the source languages that you do not cover and save them in a separate document.
- Add this split to the PIV project (you can also use the EN project from the initial zip package).

- Perform steps 1 - 3 described above. When the translation has been completed, generate the TL translation and a bilingual PIV-TL sdxliff file (or a standard bilingual xliff file).
- Once you have all the translated TL document splits in MS Word, assemble them to get the final complete TL document.
- Run DocEP Change Document ID.

2.4 Multilingual documents with two-column amendments

In multilingual documents with two-column amendments, it is recommended to:

- either work with the Studio project from the EN - Target Language package which already contains the Coloured complete original multilingual document pre-translated with the Normative Memory (Normative.docx.sdxliff);
- or pre-translate yourself the complete Coloured original multilingual document¹ at 98% using the Normative Memory of the EN - Target Language Studio project.

This will allow you to get the standard text of the **amendment headings**, which are **always in English**. After that, you can split the original document source language-wise using the DocEP *AM Split* macro.

Follow the steps described below, under point 2.5.1.

Once you have all the translated TL document splits, assemble them to get the final complete TL document with the DocEP *AM Plus* macro.

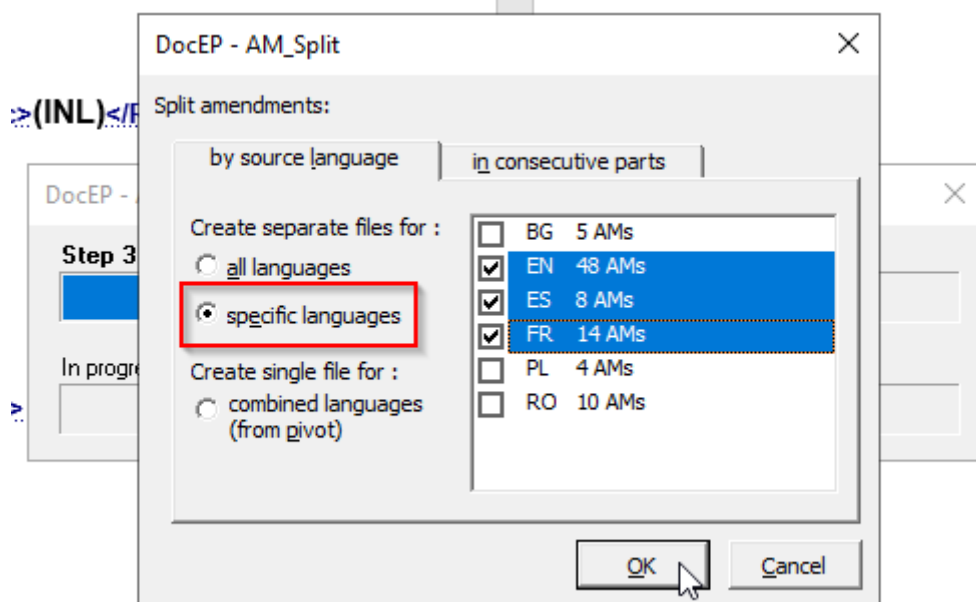
2.5 Handling of documents with two-column amendments which contain non-contractual source languages (for which a Pivot is required)

2.5.1 For the translation using the original zip package and multilingual document

Open the Studio project from the individual EN-TL package. It contains the original multilingual document pre-translated with the Normative TM in order to get the standard segments of each amendment heading and the DocEP text of the cover page. Use this pre-translated sdxliff file (*Normative.docx.sdxliff) for all further steps. Check that all the pre-translated standard segments are correct and complete the translation of the amendment headings using the Normative TM.

¹ Having the amendment headings and left-hand column text of amendment tables in blue will facilitate the identification of the corresponding segments in the Studio Editor and the use of the cascading translation memories provided.

Convert the *Normative.docx.sdlxliff back to Word and split the document source language-wise (with the DocEP *AM Split* macro) in order to get the amendment batches from the contractual source languages (EN, FR, DE, ES, IT). Use the option *Create separate files for specific languages*.



The steps below have to be repeated for each contractual SL-TL split.

Add each split document to the corresponding SL-TL Studio project. At this stage, the source language column in the Studio Editor already contains the DocEP text and the amendment headings in the target language. Lock these segments.

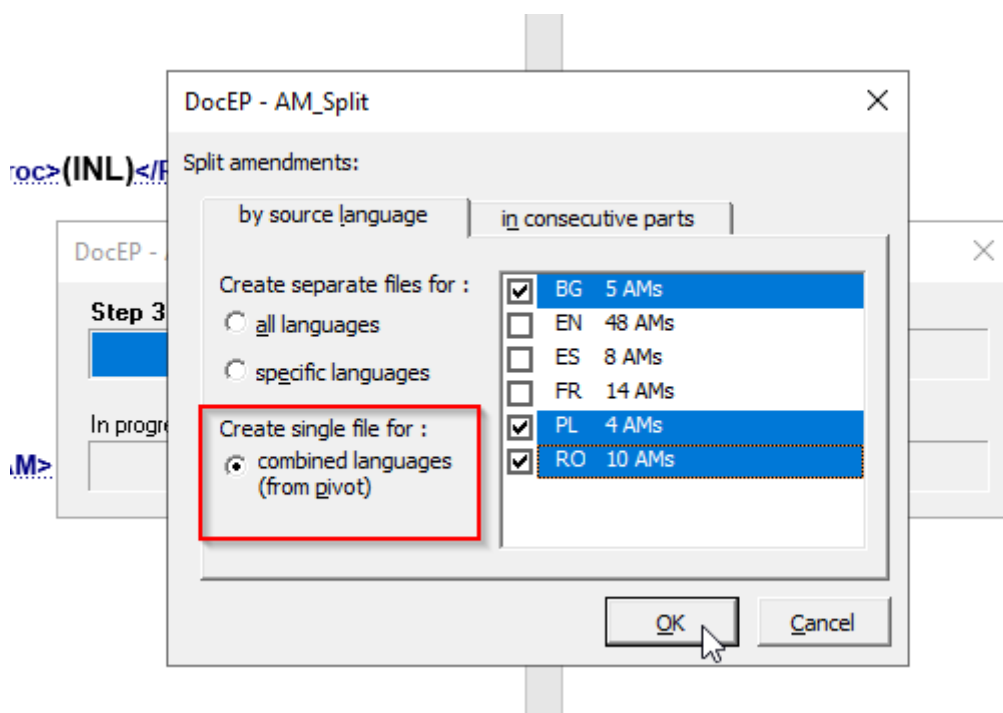
Pre-translate the TL document using the BRTM at 100% in order to get the reference document text into the left-hand column. Check that all the pre-translated segments (including those marked as 100%) are correct and complete the translation of the left-hand column using the BRTM.

Enable the WTM and use it to complete the translation of the document. Save the file frequently. While interactively translating in Studio, mark the differences between the left-hand column and the right-hand column in Bold and Italic.

When the translation has been completed, generate a bilingual sdlxliff file (or a standard bilingual xliff file) for each contractual SL-TL combination.

2.5.2 For the translation using the PIVOT document and package

Split the pivot document source language-wise (with the DocEP *AM Split* macro) in order to get the amendment batches from the source languages you do not cover. Use the option *Create single file for combined languages (from pivot)*. You will obtain a single document named *FdRnumberpivot2TL*.



Add the TL document to the PIV project (you can also use the EN project from the initial package).

Pre-translate the TL document using the Normative TM at 98% in order to get the standard segments of each amendment heading. Check that all the pre-translated standard segments are correct and complete the translation of the amendment headings using the Normative TM.

Pre-translate the TL document using the BRTM at 100% in order to get the reference document text into the left-hand column. Check that all the pre-translated segments (including those marked as 100%) are correct and complete the translation of the left-hand column using the BRTM.

Enable the WTM and use it to complete the translation of the document. Save the file frequently. While interactively translating in Studio, mark the differences between the left-hand column and the right-hand column in Bold and Italic.

When the translation has been completed, generate a Pivot-TL bilingual sdlxiff file (or a standard bilingual xdiff file).

2.5.3 Generation of complete TL document

Generate all split source language-specific TL documents treated according to points 2.3.1 and 2.3.2 above.

Assemble the final complete TL document with the DocEP *AM Plus* macro.

Run DocEP Change Document ID.

2.6 Delivery of the TL document

Deliver

1) a monolingual MS Word document corresponding to your translation of the complete multilingual document.

If the multilingual file has been split into monolingual parts during the translation process, reassemble it using the DocEP function *AM Plus (for document type AM)* or do it manually (in case of all other types of documents) before the delivery.

All tagging and formatting from the original document must be kept or replicated in the MS Word file that you will deliver. Check that the DocEP tags have the correct *HideTWBExt* and *HideTWBInt* styles. If they take the *Normal* style instead, use the *FormatAllTags* macro on the DocEP Macro Lists menu.;

AND

2) a bilingual SDLXLIFF file (or a standard bilingual XLIFF file) for each SL-TL combination, for the contractual languages and, if applicable, PIV-TL for the other languages. All tagging and formatting from original document must be kept or replicated in the bilingual (SDL)XLIFF files you will deliver.