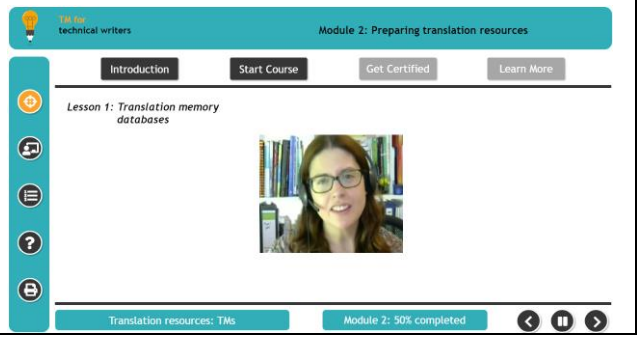
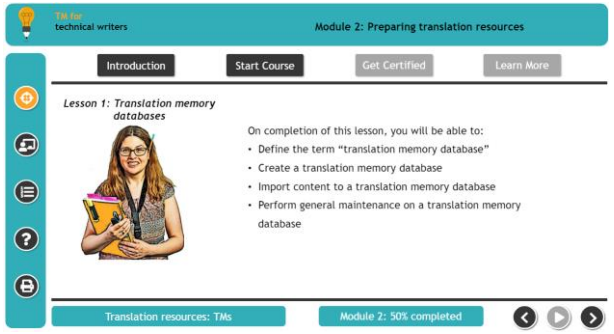
Lesson 1: Translation memory databases

Lesson overview

Screen 1

Welcome	Screen
In Module 1 of this course, we looked at how to create translatable source documents. That's just the start though. If we want clear and consistent translations, we need to provide our translators with the right resources. In this first lesson of Module 2, we will talk about translation memory databases, how to create them, and how to keep them up-to-date.	

Screen 2

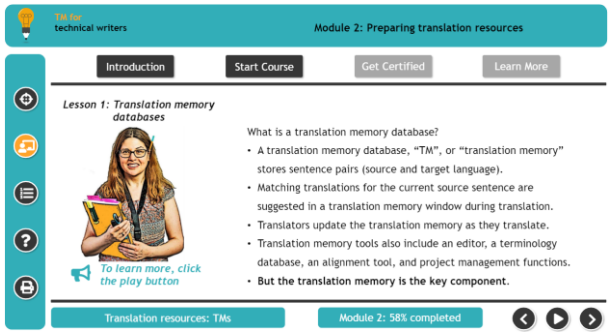
Objectives	Screen
On completion of this lesson, you will be able to: • Define the term "translation memory database" • Create a translation memory database • Import content to a translation memory database • Perform general maintenance on a translation memory database	

Created by: Maria Finnegan

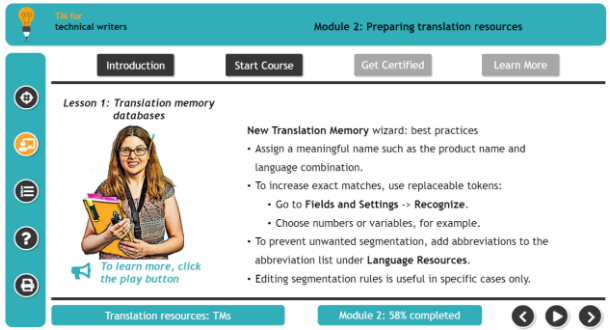
Student no.: 16063651

Lesson Content

Screen 1

What is a translation memory database?	Screen
A translation memory database (“translation memory” or “TM” for short) contains source and target language text saved as matching sentence pairs. When the translator edits content in the editor window of a translation memory tool such as Trados Studio, the content of the current source language sentence is matched against the translation memory. Suggested translations are shown in real time in a translation memory window, and translators update the translation memory with new content as they translate. So, although translation memory tools also include a content editor, a terminology database, document alignment, and translation project management functions, the translation memory is the key component.	

Screen 2

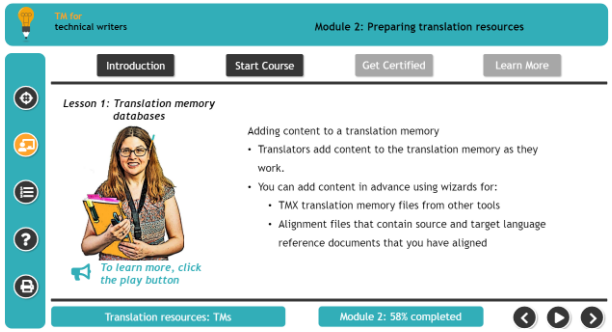
Creating a translation memory database - best practices	Screen
You can quickly create a translation memory using the wizard in Trados Studio. • Choose a location for your new translation memory as well as the source and target language. • Give your translation memory a meaningful name – for example the product name and language combination. • If you select items under **Recognize**, such as, dates, numbers and variables, these items are considered “tokens” if they appear in your source text. This means that if you have multiple sentences where the only difference between them is a token, Trados Studio will automatically insert the correct token from the source text, but otherwise treat these sentences as exact matches. With this function, you can further benefit from consistent source content, increase the leverage from your translation memory tool, and ensure consistent translations. • We just mentioned setting variables as tokens. And we talked earlier about how to handle terms that should not be translated. Here in the **Variable List**, you can define terms that you don’t want translated, such as product names, as variables. This is very helpful if you have similar sentences	

Created by: Maria Finnegan

Student no.: 16063651

Creating a translation memory database – best practices	Screen
for different products where the only difference is the product name. Trados Studio will automatically insert the correct variable from the source text but otherwise treat these sentences as exact matches. • We have also talked about abbreviations and how they are best avoided. However, if you have a specific abbreviation that you must use and that causes sentences to segment incorrectly, then you can add this abbreviation to the **Abbreviation List**. This prevents your text from segmenting after that abbreviation. • You can also directly edit **Segmentation Rules**. Changing the defaults is only recommended in very specific cases though. Once you have made your settings, click **Finish** to create an empty translation memory database.	

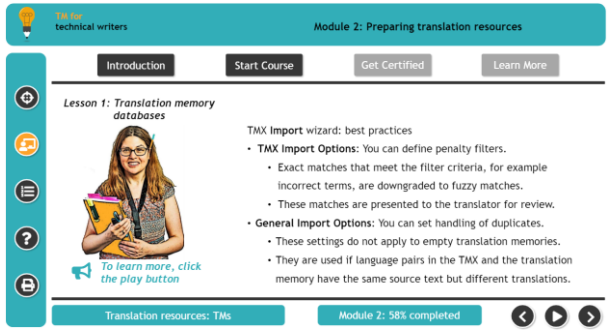
Screen 3

Adding content to a translation memory database	Screen
As I mentioned before, translators update the translation memory with new content as they translate. However, it makes sense to provide translators with some existing translation memory content to work with, if available. • You can import legacy translation memories from other tools if these are saved as TMX (translation memory exchange) files. • Or if you have useful reference documents that are available in the source and target language, but not as a translation memory file, you can align sentence pairs using the semi-automated Trados Studio alignment tool, and add the content to your translation memory. Trados studio has wizards for importing TMX files and for aligning reference documents so it's quite straightforward. There are some things worth noting though because of their impact on translation quality. Let's start with importing a TMX.	

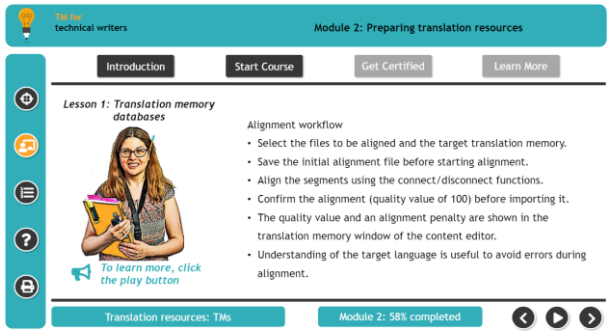
Created by: Maria Finnegan

Student no.: 16063651

Screen 4

TMX import wizard - best practices	Screen
You can import a TMX like this. • Under **TMX Import Options**, you can define penalty filters. This is useful, for example, if you want segments entered by a particular user or that contain a particular term to be reviewed again. With a penalty filter, exact matches that meet the filter criteria are downgraded to fuzzy matches. This makes it clear to the translator that these matches require inspection. • Under **General Import Options**, you can configure how language pairs with the same source text but different translations are handled during import. The thing to remember here is that these settings only apply if you are importing a TMX into a translation memory that already contains content. If your translation memory is empty, and your TMX contains duplicate language pairs or language pairs with the same text in the source but different translations, these will be imported unchanged. That can make for quite a messy translation memory database. We'll look at what to do about this in a few minutes. Once you have made your settings, click **Finish** to complete the import workflow.	

Screen 5

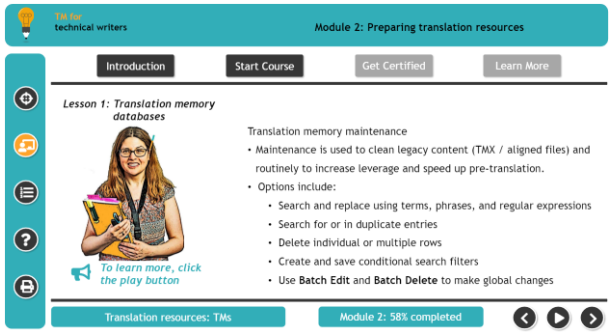
Alignment workflow	Screen
Alignment is also quite a straightforward procedure: • Just select the source and target file and the translation memory where you want to import the alignment. • An alignment file is created. Save this before continuing. Now you can start aligning the segments using the connect/disconnect functions. • Something to remember here: if the alignment that you have completed contains reliable content, confirm the segment pairs before importing the alignment to the translation memory. Confirmation assigns a quality value of 100 to the alignment, which the translator can see and use as a guide during translation. • Import the alignment file into the translation memory. • When you open the content editor, you can see the quality value that I mentioned here in the translation memory window. Also, content imported from an alignment file is automatically	

Created by: Maria Finnegan

Student no.: 16063651

Alignment workflow	Screen
assigned a 1% penalty. So, sentences that would normally be exact matches are presented as fuzzy matches to the translator for review. • One last thing: if you are completing an alignment, it makes sense to have some understanding of the target language to avoid errors.	

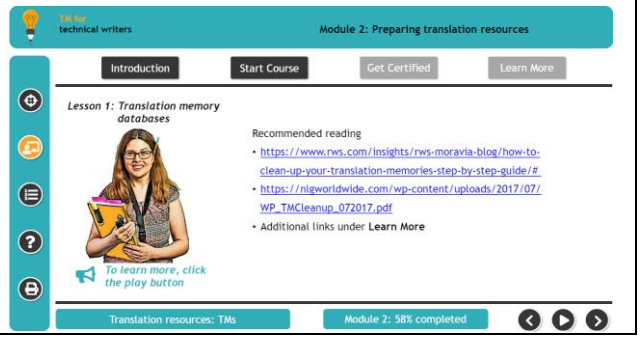
Screen 6

Translation memory maintenance	Screen
So now you have a translation memory with some content. But, perhaps the TMX file contained a lot of data from different translation memories that were never thoroughly reviewed. Or, the reference files you aligned contained out-of-date terminology. This is where translation memory maintenance or “cleanup” comes in. When you open a translation memory, you can carry out the following tasks: • Run general searches to find and replace terms or phrases • Search for or in duplicate entries • Run pattern searches using regular expressions (for example, to find dates of a particular format) • Manually delete individual or multiple rows • Create and save conditional filters to search through the translation memory more quickly • Use **Batch Edit** and **Batch Delete** to make global changes automatically in your translation memory based on the conditional filters that you have created. Of course, these functions are not just useful for cleaning up legacy content. It generally makes sense to review your translation memories regularly and to make sure that they are clean. Regular maintenance increases the leverage from your translation memory during project creation, and speeds up pre-translation, saving you time and translation costs. Match recognition during translation is also more efficient and less ambiguous, which saves a lot of questions, and a lot of time.	

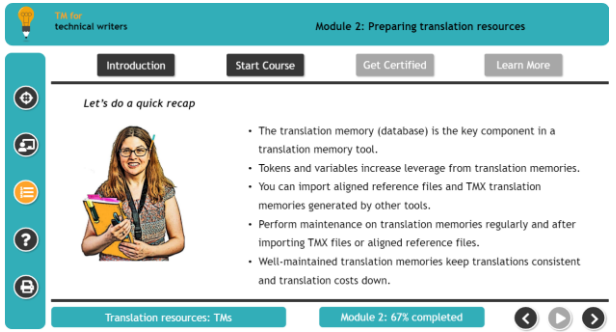
Created by: Maria Finnegan

Student no.: 16063651

Screen 7

Closing	Screen
The translation vendors RWS and NLG have some excellent advice on keeping translation memories clean. I've also included some resources on alignment and on translation memory tools in the **Learn More** section, which you can take a look at once you complete this module.	

Lesson Summary

Summary	Screen
• The translation memory (database) is the key component in a translation memory tool. • Tokens and variables increase leverage from translation memories. • You can import aligned reference files and TMX translation memories generated by other tools. • Perform maintenance on translation memories regularly and after importing TMX or alignment files. • Well-maintained translation memories keep translations consistent and translation costs down.	

Created by: Maria Finnegan

Student no.: 16063651