

**An International Communication by
Design, Inc. (ICD) White Paper**



**1726 N. First Street
Milwaukee, WI 53212
414.265.2171
www.icdtranslation.com**

Saving Translation Costs and Improving Consistency with Translation Memory

*By Catherine Deschamps-Potter
Vice President, Sales and Marketing*

May 12, 2008

© 2008, International Communication by Design, Inc. All rights reserved.
All product names used in this paper are the trademarks or registered trademarks of their
respective owners.

Contents

Introduction	2
Problem Statement	2
Other Options	2
Implementation	2
Summary	6
Want to know more?	6

Introduction

This white paper describes the use of translation memory to reduce translation costs and increase consistency of content.

Problem Statement

Companies and organizations in today's marketplace seek to reduce their costs for translation while maintaining and improving quality and consistency.

Other Options

Other methods for saving translation costs include using controlled language, implementing consistent terminology, and rationalizing content. This white paper covers use of translation memory and optional use of a translation memory content management system (CMS) for cost reduction.

Benefit: Cost Savings

Use of translation memory consistently reduces costs for companies. The more content that companies translate, the more savings are realized.

Benefit: Consistency

Use of translation memory increases consistency of translated content because previously translated and reviewed content is reused. It can also assist in improving the quality and consistency of source content. This improvement can be accomplished by reviewing the English content in the translation memory, identifying similar but not identical wording, and choosing standard phrases.

Implementation

Several major software tools exist in the marketplace today for creating and using a translation memory database. Among these tools are such software as SDLX, Trados, Across, Déjà vu, and Star Transit. This paper uses examples from ICD's translation memory tool to illustrate the methodology. However, similar costs savings apply and similar approaches may be used with other translation memory tools.

Word count analysis

When a document or other content is to be translated, companies typically submit their content or documents to a localization service provider (LSP).

The types of files accepted by the LSP depends on the translation memory tool that the LSP uses. For example, the ICD translation memory tool supports these files: OpenTag (OTF), Code File, CopyFlow Filter, Generic Delimited, HTML (Web pages), JAVA, MIF (FrameMaker), RC (Program Resource File), RTF (Documentation and Help), SGML, TXT (Plain Text), Web-based formats, XML, and XPress Tags (Quark), In Design (CS, CS2, CS3), and RC. Other translation memory tools may

support other file types, such as PageMaker with Trados.

After receiving the files, the LSP will typically analyze the files to provide the client with an analysis, including a word count, and a quote. This thorough analysis process involves running the content through the translation memory tool. In situations where graphics are involved, the analysis may need to be performed manually. Other aspects of the initial analysis involve evaluating the design of the document and the software used to create it so that an accurate project cost can be determined.

The application used to create the content is as important as how the document was created. Situations that increase the level of difficulty in the desktop publishing stage include use of Adobe® Illustrator® or Photoshop® files imported into InDesign®, Visio® files in Word, imported tables in Quark® and InDesign, and the lack of style sheets. Use of style sheets and a consistent font saves time in formatting content.

To track cost savings with precision and ensure correct estimates, an accurate word count at this point is essential. The more experience the LSP and the

translation memory software have in handling diverse file types, the more accurate the word count and estimate.

Translation memory analysis

When the files are run through the memory tool, the tool also analyzes the source content to see if any phrases were previously translated. This analysis looks for both identical matches and what are called “fuzzy” matches. The analysis provides customers with an accurate word count and the amount of reuse from the customer’s translation memories.

For phrases that have been translated in their entirety, the analysis will show 100% reuse. For fuzzy matches, the tool uses a threshold to look for similar phrases that have been translated.

For example, if the sentence “I live in a brick house.” was previously translated, and the customer later requested translation of the phrase “I live in a wooden house.”, the translation memory tool would identify the new sentence as a fuzzy match. Below is a sample analysis.

Example of an ICD Translation Memory Tool Analysis

GERMAN

Analyse Total (27 files):

Match Types	Segments	Words	Percent	Placeables
Context TM	0	0	0	0
Repetitions	503	1,760	8	161
100%	1,193	4,232	19	318
95% - 99%	77	631	3	58
85% - 94%	87	653	3	61
75% - 84%	78	451	2	55
50% - 74%	61	368	2	70
No Match	1,849	13,929	63	1,688
Total	3,848	22,024	100	2,411
Chars/Word	5.22			
Chars Total	115,037			

Repetitions are the number of segments of text that are exact matches for another segment in the content. Repeated segments need be translated only once, saving costs.

Your LSP will use data from the analysis, including the amount of reuse, the number of words to be translated, the target language, and the price per word to determine the cost of your translation project.

This value shows the number of segments of text that are a 100% match for content that is already in the translation memory. These segments would not need to be retranslated, saving costs.

The translation phase

With a translation memory tool, a special file type format is created as part of the analysis. Some tools use files that tag the English and the translation or intermediate translation document files. Others go through a server-based application that uses XML files. The translators use these files to translate.

This is the step where the major cost savings are realized. Because translators have to proof the matches instead of translate them, the number of words for which you are charged for translation is reduced. As matches increase, less translation is needed, but more proofing is performed to help ensure accuracy and consistency. Fuzzy matches can also cut down the time needed to translate as the entire phrase may not need to be retranslated. The translator can instead reuse portions of the text that were previously translated.

Benefits of a CMS

Another feature of some translation memory tools is a translation memory content management system (CMS). With a CMS, clients send files to the LSP through the CMS, and the LSP translates the files directly without creating separate files. All content is seamlessly transferred from the client to the LSP to the translator and back again. The analysis and quote can be performed on the server-based CMS to which both the LSP and client have access.

An important feature of a CMS is that the client can make changes to the content and notify the translator automatically. Clients can also review translated files through the CMS before the content is desktop published. This feature reduces the overall cost of DTP.

Clients can also track the progress of the translation through the CMS.

Returning the translated content

Once the translation is complete, the translator returns it to the LSP, who processes the files back through the translation memory tool. This process converts the content back to the original file format such as Word and InDesign.

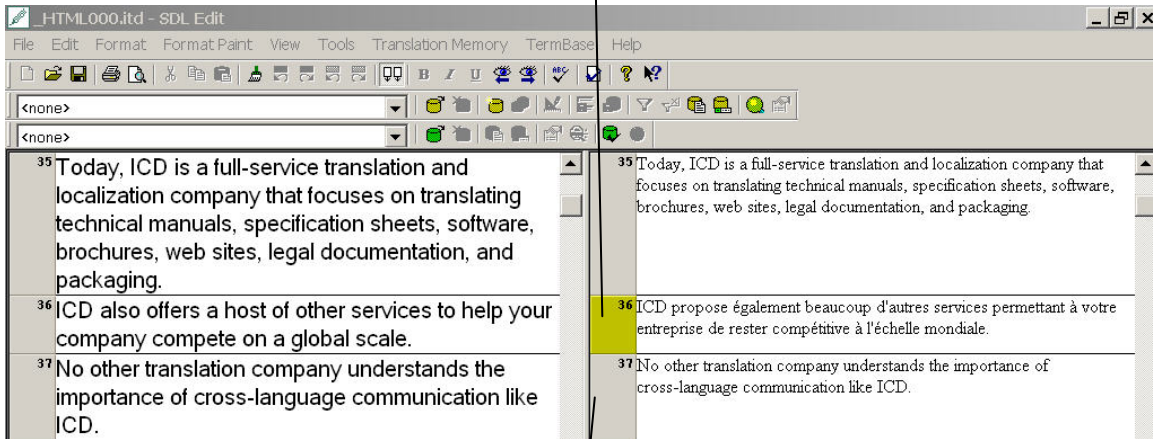
During this process, any newly translated content, along with its English counterpart text, is stored back into the translation memory. The next time that the same English text is used in a document, the corresponding translation can be reused, saving the customer translation costs. Generally, additional desktop publishing is performed on the resulting file to ensure alignment with the English version.

For example, European languages tend to be longer than English. Because English writers often do not write with translation in mind, text can move to different pages. The resulting text must then be adjusted to fit the page formatting and to ensure that the translated document, such as a marketing brochure, agrees with the format of the English document. Use of smaller font sizes can help avoid this problem.

If the company requesting the translation reviews the finished product and has stylistic preferences, then the translator who work for the LSP will review the client's changes and make the necessary changes. When such changes are made, the translation memory is updated so that the company's future projects can take advantage of the changes. An example of the translation file used by the translator is shown below.

Text segments marked as gold have already been translated and the translations are stored in the translation memory. The translator need not translate these segments, but must proofread them.

The more segments that appear in gold, the more cost savings potential.



For text segments marked as gray, the translator reviews the source content in the left column and enters the corresponding translation in the right column.

Sample Savings: ABC Company

Customer ABC Company, who has never translated content, requests an LSP to translate a 25,000 word document. The initial analysis of the customer's file would indicate 0% reuse. At \$.19 per word the translation costs are \$4,750. The source content and corresponding translated content for ABC Company are stored in that company's translation memory.

When ABC Company subsequently submits another source document that reuses content from the source document that was previously translated, the LSP analyzes the document using the memory tool.

In some cases, the customer's translation memory is considered proprietary and the LSP will run the analysis only against the customer's own translation memories. However this is not the case with all translation memory tools. Customers

should check with their LSPs to determine if their memory tool has this feature.

In general, generic memories are not used, but some governmental agencies offer memories for general use. Customers should check with their LSP to determine if they offer such memories.

Additional savings and consistency

For customers who wish to realize further increase cost savings using translation memories, a review of the translation memory can help find ways that the customer can both reduce costs and increase consistency.

Through an analysis of the translation memory, a company may find that its writers use various phrases in the source language to convey the same meaning.

For example, the translation memory can reveal that writers use three different ways to direct the user to back up files:

- Ensure that you back up your files daily.
- Back up your files daily.
- Create a backup of your files on a daily basis.

For this example, by standardizing on the sentence construction to be used to instruct users to back up their files daily, companies can reduce translation costs. They can also increase the consistency and the quality of their content.

To view translation memories, ask your translation vendor to provide you with the tools they use, so you can proof the documents. After determining the standard ways to write certain phrases in the source language, the customer can then create a style sheet for their writer.

Summary

Using and maintaining translation memories has clear benefits for companies doing business in the global marketplace.

Among the benefits are:

- Costs savings through leveraging existing translations
- Reduced review time as reused translations have previously undergone review by the LSP and by customer reviewers
- Reduced processing time through the use of a CMS
- Improved quality of translation through reuse of translated content that has previously undergone customer review
- Improved consistency of translated content through reuse of the same translations
- The potential for consistency in source content through standardizing on source phrases found in the translation memory

Want to know more?

For more information on using translation memories and other ways to improve translation and source quality, contact International Communication by Design, Inc. at info@icdtranslation.com.

When you need complete globalization services, trust ICD for on-time and accurate language translation services.