online Newsletter
news and insights

ACCTI

::::::::::::::::::::::::

MultiCorpora R&D INC.

::::::::::::::::::::::::

MultiTrans PRO

Technology

CAT TOOLS - SORTING THINGS OUT
by Jost Zetzsche

The old German saying claims that all cats are gray in the dark, and not being able to distinguish between them only adds to their mystery and menace. But if we shed a little light on the critters, it may be easier to dispel the shadows and distinguish between them.

By definition, any tool that is specifically designed to aid in the translation process falls under the category of CAT tools (computer-assisted [or -aided] tools). This includes tools for machine translation, project management, glossaries, localization, and translation memory, along with numerous utilities for word counts and invoicing, among others.

For now, let's focus on the applications for translation memory. In a literal sense, these are applications that extend the memory of translators by allowing them to build up databases of translated material and leverage those databases against newly translatable content. In reality, however, they're usually more than that because most CAT tools also:

- allow the user to build terminology databases that complement and extend the functionality of the translation memories, and

- allow translators to work in complicated file formats that they may not understand or otherwise be able to support by hiding or protecting the code and only displaying translatable content.

Furthermore, many of the tools provide methods for analysis, project management, quality assurance, and productivity that go way beyond the typical functions of a mere "translation memory tool."

The fact that new tools are still appearing at a fairly rapid pace (and others are disappearing at an equally fast pace, with Cypresoft's Trans Suite 2000 being the latest casualty) shows that this is still perceived as a relatively fluid market with plenty of room for enhancements and innovation.

Categories of Translation Memory Applications

In an effort to classify the tools into certain categories, I've come up with these main categories:

- tools that use external databases for their translation memories, and
- tools that generate translation memory through file associations of original and translated files.

:::::::::::::::::::::::::::

MultiCorpora
R&D INC.

:::::::::::::::::::::::::::

MultiTrans PRO

The benefit of the association approach is that there is no need for an "additional" database that may tie up computer or server space. The drawback lies in the large amount of translated file pairs that have to be retained to provide for the necessary "reference material."

Memories by Association

There are only three applications in this category—the veteran Star Transit (see www.star-transit.com**)** and the relative newcomers MultiTrans (see www.multicorpora.com) and Translator's Intuition (see www.termseekinc.com). These can be further classified into two sub-categories:

- Star Transit does its matching on the traditional sentence-by-sentence level.

- MultiTrans and Translator's Intuition do their matching on larger segments, such as paragraphs.

External Memories

The category of tools that build external translation memories includes a much larger group of applications.

A useful sub-classification for this group would be the work environment:

- The first group does all or most of its work through macros in Microsoft Word that allow an association with translation memories and terminology databases.

Users who like the Microsoft Word interface point out that, when translating DOC or RTF files, it's helpful to work in a fully formatted file, and they appreciate the general familiarity of the Word interface.
The most well-known application of the Word plug-ins is Trados (see **www.trados.com**— note that Trados is also increasingly using other interfaces as well); others include Wordfast (see **www.champollion.net** and **www.wordfast.org**), MetaTexis (see **www.metatexis.com**), and WordFisher (see **www.wordfisher.com**).

The three smaller tools are primarily focused on Word and RTF files or, in the case of MetaTexis and WordFisher, on some other programs that can be accessed through Word. Trados, on the other hand, covers a very large range of additional applications and formats through its numerous RTF conversion utilities and its relatively new TagEditor and T-Windows interfaces.

The tools of the second sub-category let the translator work in an independent, mostly tabular environment.

This last group that presents all files in a uniform tabular format includes SDLX (see **www.sdlx.com**) and Déjà Vu (see **www.atril.com**). Supporters of this independent environment approach praise the uniformity of the work environment and the independence of other programs (i.e., an upgrade to Word does not have any effect on the CAT tool).

Similarities

In some areas, the major tools (Star Transit, MultiCorpora, Trados, SDLX and Déjà Vu) have become increasingly similar (aside from the inherent similarity that they all have translation memories and display translatable text). They all provide:

- A freely configurable terminology component that plays a fairly major role in the translation workflow.
- Processing of texts in Unicode, thus giving access to all languages that are supported by the Windows operating system.
- A feature that allows "alignment," i.e., the matching up of already translated information.
- Support for TMX, the XML-based translation memory exchange format.
- Concordance searches (i.e., searches for one single word or expression within larger segments) in their translation memories.
- General word processing features, such as spell-checking or search & replace.
- Fairly detailed analysis features.

This is more or less where the similarities end.

Further Differentiators

Aside from the above differentiators (external memory vs. memory by file association and work environment), there are a number of other questions that need to be asked in finding out how the applications differ and how to decide which is best for you. Discussing these and other important questions with each of the CAT vendors will further help you clarify their differences:

- Terminology Handling—How closely is the terminology processing integrated in the translation workflow?
- Code handling—How are handled and stored the formatting codes that occur within segments?
- Supported file formats—Though most tools support a similar range of formats, are "my" formats supported?
- File processing—Are files processed individually or in batch processes?

::::::::::::::::::::::::

MultiCorpora

::::::::::::::::::::::::

MultiTrans PRO

In my opinion, everyone in the translation industry can benefit from some features that translation memory tools offer. But this conviction doesn't necessarily make it any easier to choose the right one. Though all CATs may look gray in the dark, learning to ask the right questions does dispel the murkiness and bring a little clarity to their confusing quirks and idiosyncrasies.

Parts of this article are excerpted from the author's e-book "A Translator's Tool Box" (**www.internationalwriters.com/toolbox**). The author is aware that the list of tools in this article is incomplete.