

# *Transit and TRADOS: Converging functions, diverging approaches*

**Tim Hallett**

President of Kanwa Translation, Technical Writing and Consulting, Inc., Illinois, USA

Abraham Lincoln was once reported to have said of a play, "People who like this sort of thing will find this to be the sort of thing they like." Much the same can be said for one's personal preferences with regard to translation memory (TM) software, which will depend greatly on what one expects the software to do.

Recently, I sat down and started to compare two major platforms, Transit XV and SDL/TRADOS 2006. Having extensive experience with both Transit 3.0 and TRADOS, I subjected both to tests specifically designed to address issues that are traditionally associated with the respective software platforms.



## **Transit XV**

Transit XV, developed by STAR Language Technology & Solutions, builds upon the general architecture and approach of Transit 3.0, but represents a marked improvement over its predecessor. Transit 3.0, though in many ways a fine product and my personal choice for a TM system, was, unlike its tank-like predecessor Transit 2.7, plagued by multiple bugs that resulted in crashes, errors, and slow performance, particularly in Excel proj-

ects. The good news is that Transit XV has eliminated these issues and has both good stability and high processing speed.

## **New features**

Among the new features of Transit XV is a new processing mode called "Repetition Mode", which you can use to display, translate, and/or check internal repetitions before starting the "actual" translation. This mode is particularly valuable when dealing with large projects or when there are several translators working on the same project. I tested this function using an Excel spreadsheet with several repetitions intermixed among non-repetitive lines. The only real difficulty was figuring out how to get started, since the cursor did not automatically take me to the first repetition. However, after manually finding a repetitive line and translating it, Transit XV translated all other instances of that line in the document and then moved on to the next repetition. This is a very practical and well thought out tool for managing identical expressions in large projects.

Another interesting feature is "Reference-based Spell-checking", which allows you to include the reference material of your project for spell-checking. Thus, Transit looks up all words of the active language pair in the reference material. With this type of spell-checking, Transit XV can examine the project's context, even for languages for which there is no spell-check dictionary available. The only catch is that you have to have your words spelled correctly in your TM.

Previous versions of Transit always had a good segment filtering functionality, with a fairly easy-to-understand script for regular expressions that allowed users to filter in or out desired strings or search for content. But now Transit XV has added a useful dialog for filtering on the basis of segment attributes, such as for translated or untranslated segments and certain percentages of fuzzy matching, thus greatly expanding an already flexible functionality within Transit.

Already a much-praised feature of SDL/TRADOS, Transit XV now has a "Concordance" function and a corollary "Dynamic Linking" function. The Concordance function allows a user to search the entire corpus of the TM for instances of a word or phrase, and is very useful in ensuring uniformity of expression, even in cases where there is no fuzzy match. The Dynamic Linking function represents a further development of the Concordance function, where a specific word or phrase in the source language can be linked contextually to a specific word or phrase in the target language, not only in the file pair currently being translated, but across all pairs in the TM and the current project.

## "...Ultimately, what I see in both products is a convergence of functionality with an increasing divergence in environment and approach..."

### Transit's TM system

It is here where a short description of the Transit XV TM system is in order. Unlike other TM software that registers segment pairs to a separate TM file for searching, Transit XV creates "virtual databases" out of the source and target file pairs that have been previously translated by the user, and/or those assigned to the current project as reference materials. This has the advantages of minimising the disk space occupied by the TM and keeping the format of the flat data files the same as those operated on within the Transit XV editor. This allows mistakes discovered in the TM to be corrected quickly inside the Transit XV editor, without the need for separate database editing or maintenance applications. Another advantage is that such an arrangement has greater stability, flexibility, and durability than traditional database formats. Transit XV databases do not corrupt and do not require reorganisation. The only real downside of this approach is that you do have to organise your TM in a coherent fashion so that it can be accessed by the program. (Or more accurately, so you can remember where you put it!)

An area of major improvement in Transit XV is the "Report Generation" function. In the past, the Transit report function excelled in giving progress reports and post translation reports that were useful in ascertaining final invoicing, but when clients called and asked for estimates, Transit 3.0 was unable to provide a preliminary analysis of project content against an existing TM. Another weakness of the Transit 3.0 reporting function was that it lacked gradations for fuzzy matching. Fuzzy matches were simply classified as either "Fuzzy Match" (i.e., 100% matches) and "Fuzzy Match" (Edited). Transit XV has included four user configurable gradations for fuzzy matches (e.g., 99-95%, 94-85%, 84-75%, and 75-50%), as well as three types of reports: Import (the above-mentioned pre-translation analysis), Progress, and Translation. The Translation reports are, as always, extremely useful in creating invoices after translation is completed, as they give a comprehensive and accurate tally of the actual final numbers of pre-translations, fuzzy matches (with gradations), and manual translations. Thus, Transit XV now offers a comprehensive tool for issuing reports before, during, and after the translation process, all in the same package, without the need to purchase expensive enterprise-level project management software. This is ideal for small operators such as myself.

Unchanged from earlier versions of Transit is the architectural openness of the translation system. This feature, which is much appreciated by more discerning users, allows easy and quick modification, correction, or repair of both project and TM data. The most common example would be, suppose you were translating a project and a

fuzzy match appeared for the current segment. In the fuzzy match offered, you notice a misspelling or mistranslation. In conventional TM systems, you are somewhat wedded to that mistake, unless you either have tools to edit the TM database or undergo a fairly irksome process to access the TM and make the change/correction. In Transit XV, such corrections are as easy as right-clicking the fuzzy match and selecting "Open Reference File". Upon doing so, Transit XV will open the source file pair for that fuzzy match within the Transit XV editing environment at the exact location of the segment in question, where correction of any mistake can be easily and directly made.

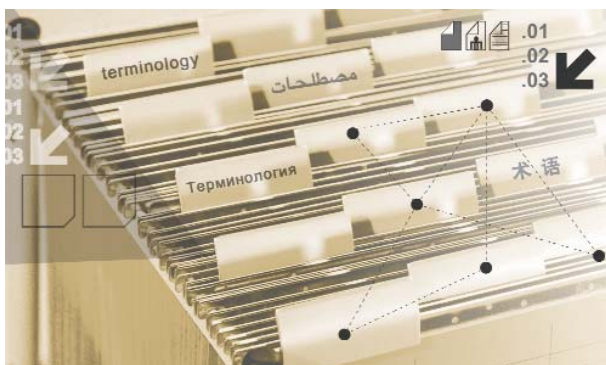
Press Ctrl+S, and the TM is corrected, a process that from beginning to end seldom takes more than five seconds. Tags and other protected items can also be easily moved within the editing environment simply by disabling the tag protection (although this should only be done as needed, and not maintained as a normal translation environment).

Finally, if any damage to the objects or delimiting information makes export impossible, the diagnostic tools give an accurate report of the location of the error, which can be easily repaired by opening the file pairs in a text editor tool and rectifying any discrepancy between the tags and delimiters of the source and target file at that location. All files in Transit XV adhere to standard XML markup conventions, and can be easily understood and repaired by anyone with even a rudimentary knowledge (such as mine) of markup languages (XML, SGML, HTML, etc.). Transit XV also has a "Source Export" function that exports the information from the source file rather than the target file to confirm whether there is an inherent XML form violation in the source material, which is a very useful diagnostic tool.



### Transit's TermStar

Transit XV comes bundled with its own terminology tool, TermStar, which is fully integrated into Transit XV and allows a terminology list to be displayed onscreen along with the translation file pairs. Once again, the greatest advantage of TermStar, like Transit XV, is its flexibility and openness. Other terminology management tools do an excellent job of easily importing glossary information that has been previously compiled into a proper database format used by the program in question, but honestly, when is the last time one of your clients kept their terminology list in XML or SQL format? In my own language field of Japanese, such lists are invariably kept on an Excel spreadsheet. TermStar is the only platform I have used that allows the user to easily import a terminology list from an Excel spreadsheet simply by treating it as tab-delimited text. This one capability has paid off for me personally.



TermStar also has tools for structuring the most complex of multilingual dictionaries, which ideally we would all love to have, but in reality, we all know that we are much more likely to receive a simple two-column word list in Excel from our clients. Transit XV allows you to incorporate this terminology information, unadorned as it is, into a searchable format integrated into the translation environment. Thus, terminology matches along with the translated segment are displayed onscreen in the same environment as the translation, and no clicking around to other programs is required. You can also quickly enter new terminology during the translation process using the "Rapid Entry" feature, which is a real time-saver. (True story: A client once sent me nearly 10,000 glossary items arranged over seven Excel files, so imagine her surprise when I told her that I had incorporated all of them into my TM database and that the entire process took less than half an hour!)

As far as the file types handled by Transit XV, the basic package handles a good assortment, including ANSI/ASCII/Unicode text files, Corel WordPerfect, HTML, XML, SGML, MS Word, Excel, PowerPoint, RTF and RTF

for WinHelp, Windows Resource Files (RC), and SVG. Optional filters for FrameMaker, PageMaker, InDesign CS, QuarkXPress, Visio, AutoCAD, and Interleaf/QuickSilver can be ordered separately.

### SDL/TRADOS 2006

With the purchase of TRADOS by SDL, Inc. in 2005, users of TRADOS around the world wondered what would be the resulting effect on TM. Well, the answer has arrived in SDL/TRADOS 2006. TRADOS still has Translator's Workbench, with all the features and functionality that made it such a hit with translators worldwide; namely its interface with Microsoft Word™, creating a flexible, yet familiar environment for translators to work in. In terms of functionality, there are some important changes made in this version, with the long-desired incorporation of a direct TM editing function similar to that of Transit XV. My only negative point about this feature is that it is accessed through the "Concordance" function, when having a dedicated "Editing" button calling up the segment pair currently in the Translator's Workbench would be more convenient for translators by eliminating an extra step.

### Workbench

One negative aspect of the MS Word interface of previous versions of Workbench was its tendency to slow down or even stall when confronted by large graphical objects in Word documents. After running some tests in newest version with documents that had proved problematic in the past, this issue still seems largely unresolved.

There is still significant slowdown in bypassing large graphical objects, and I eventually had to translate one file I tested manually, since I was unable to progress past a certain point. Given the persistence of this issue, I assume it is a problem inherent to the MS Word interface. Sometimes one must sacrifice performance for luxury, I guess.

### File types

In terms of file types, the usual ones were there along with new compatibility with Linux-based software such as Open Office and Star Office. This is an interesting development and may translate into a competitive advantage in such markets as India and China, where free open-source software and Linux are increasingly popular.

### Reporting functions

As far as the reporting functions go, TRADOS's "Analysis" function is as good an estimation tool as exists on the

## "These advantages, particularly the latter, derive not from any particular programming skill or elegant functionality, but are the by-products of the unitary XML-based approach employed by Transit XV"

market, but a "Post-translation Reporting" function, perhaps integrated into the "Clean-up" function, should be added to give a detailed account of the final number of characters, segments, and words translated by various means and at various fuzzy percentages so as to aid in accurate invoicing.

Although MultiTerm 7.1 is supposed to be bundled with SDL/TRADOS 2006, the version I downloaded was bundled with Multi-Term 5.5, an earlier version than the 6.5 I already had. So my opinions on this software are only those that can be gleaned through reading the documentation. First, I was very gratified to see that a function for rapid entry of new terminology had been incorporated in MultiTerm 7.1, eliminating an ongoing headache for TRADOS users in adding newly-discovered terminology to MultiTerm. However, the importation functions still do not support the easy importation of Excel files or off-the-Web glossary information in .txt or .html format without extensive re-scripting for field assignment in the source file level.

TagEditor has been improved by increasing its functionality with MultiTerm, so that it can now have all the TM functionality available in the word interface available for non-RTF based projects.

The interface is radically different from the MS Word environment of Workbench in that data is displayed as flat text in a graphical structured text/tag-editing environment. The tool is powerful and multi-functional, but because of its largely GUI-type interface, it requires the user to learn certain procedures to perform certain tasks, such as moving and inserting tags (index sorting information for FrameMaker files), or dividing and/or merging segments. It is an adjustment for those used to operating in the MS Word environment only, but is a marked improvement over their previous environments for translating FrameMaker, Excel, and other non-RTF document formats.

### **SDLX**

SDL/TRADOS 2006 also comes bundled with SDLX, SDL's long-time TM system, which was much loved by freelance translators for its high degree of functionality combined with low cost. SDLX uses a traditional paired tabular translation format. An interesting feature is how SDLX simultaneously translates all fuzzy matches in a document as soon as information is entered into the TM. Fuzzy matches in need of editing are marked with a different colour from segments that have been manually translated or 100% matches. SDLX can be used across the entire spectrum of TRADOS file formats, and is a useful tool for those who desire a more traditional editing environment than the MS Word interface of Workbench. SDLX has its own internal termbase management tool,

and can import both MultiTerm and \*.txt files. However, when I attempted to import a simple tab-delimited Japanese glossary, the system managed only to import the first word pair, and then did not display the Japanese text properly, even though I specifically designated the field code page to be Shift-JIS (Japanese). It is likely that this was due to some error on my part, though it is not immediately apparent what that error was.

### **Conclusions**

Ultimately, what I see in both products is a convergence of functionality with an increasing divergence in environment and approach. Transit has always prided itself on the unity of its package: A single environment suitable for all translations in all file types, using universal scripting conventions that allow the user maximum access to the translation text, and ultimately, control over it. Its performance is extremely stable, and the integration of the TM and terminology tools is very good. The terminology tool itself also partakes of this open and flexible approach, and particularly aids the individual translator in tackling the real-world issue of quickly and easily incorporating Excel and Web page glossaries into terminology data that can be used by the TM.

SDL/TRADOS 2006, on the other hand, is going in the opposite direction. It now offers users three markedly different translation environments unified by Multi-term. Workbench maintains its MS Word interface, with all the benefits and drawbacks that entails. TagEditor has become the editing environment for all non-\*.doc, \*.rtf formats. For those used to always working in the MS Word environment, TagEditor does take some getting used to; however, it does have broad functionality, and with the integration of Multi-term, is now a fully-fledged TM system in its own right. SDLX provides an easy-to-use, functional, and traditional TM environment for those who desire such an interface.

Ultimately, people will decide based on look, feel, and comfort level when all functionality is equal. However, I still give the nod to Transit XV for the flexibility of the TermStar tool for importing Excel files, and the ability to easily repair damaged segments that fail to export. These advantages, particularly the latter, derive not from any particular programming skill or elegant functionality, but are the by-products of the unitary XML-based approach employed by Transit XV.