AUTOMATE, EMIGRATE OR EVAPORATE!
COREL AND TRANSLATION AUTOMATION

Orlagh Neary

Corel Corporation Ltd., 3rd Floor Europa House, Harcourt Street, Dublin 2, Ireland.

Corel Corporation is a Canadian software company which has, in the space of 10 years, grown to be a world leader in the development of graphics, desktop publishing and multimedia software. Corel distributes its software products in 15 languages to more than 70 countries. The Irish office co-ordinates manufacturing and advertising activities outside of Canada, oversees technical support for Europe and Asia and is also responsible for the localisation and quality assurance for product translations. In March 1995, Corel Corporation invested in various translation tools with the goal of reducing translation costs and time to market. In July 1995, these tools were used in the localisation of CorelDRAW!6. This required a considerable rethinking of the traditional approaches to localisation projects. This paper attempts to give an insight into the reasons behind such an investment, the steps involved in integrating translation tools into the Corel localisation process and the lessons learned.

Localisation is the adaptation of software products to suit the requirements of another culture, language and market. There are many tasks involved in software localisation, such as the translation of the software user interface, on-line help and documentation. Other tasks include the engineering and quality assurance of the software and on-line help, printing manuals and manufacturing the finished products.

Before the introduction of Translation Automation, Corel delivered the source files to the localisation vendors who were then responsible for the translation and in some cases the engineering of the products. In recent years, Corel has rapidly expanded to produce more and larger products that it still attempts to make available on international markets as soon as possible. This expansion has driven the cost of localising products upwards leaving us with no alternatives but to automate, emigrate or evaporate!

Corel had to accelerate the localisation process by automating tasks and reduce localisation costs by recruiting in-country vendors. Our other alternative was to let some of our markets evaporate by localising into fewer languages. By moving away from our traditional approach, towards a combination of automation and emigration we have been able to ship the localised products in a more timely and cost-effective manner.

Translation is one of the main localisation tasks, so it was an obvious candidate for automation. After consulting with users and developers of Machine Translation and Translation Memory systems, Corel felt that both of these systems could enhance the localisation process. In March 1995, Corel invested in the translation memory system XL8 and the machine translation system LOGOS. XL8 is a DOS-based system that runs on Windows 3.x and later. The main advantages XL8 had over competitors

at the time, was that it supported two of the three file formats used by Corel. Although it did not support Ventura 5 it was quite easy to develop a filter. XL8 also provided an interface to the LOGOS machine translation system. LOGOS is a UNIX-based system that can be accessed through Windows 3.x and later. Corel purchased four language pairs, English to French, English to Italian, English to German and English to Spanish.

Corel recruited a team of four language specialists to support the two systems and also supplied them with new equipment - a UNIX workstation for the LOGOS software and a software package which allows PC users to gain access to the functionality of the UNIX system. Corel also recruited a UNIX administrator for the LOGOS system. The team's first task was to create a translation memory based on DRAW!5 translations - a process commonly known as alignment. One million words were aligned in French, Italian, German, Spanish, Dutch, Swedish, Brazilian Portuguese and Japanese over a two month period.

In order to maximise the potential of machine translation, the user must invest time, money and effort. Dictionaries are pivotal - the better the dictionaries, the better the translation quality will be. But entering a term into the dictionary is not as simple as entering the source term and its equivalent target term. The user has to provide information about the term that will help the system to think more like a human translator. By July 1995, the Corel Language Specialists had entered more than 10,000 terms in the various dictionaries.

In August 1995, the DRAW!6 project commenced. Software aside, the localisation department was faced with the task of localising more than 900,000 words of documentation and on-line help into seven languages within a period of two months. The average amount of text that we were able to reuse from Translation Memory was 15% - less than we had hoped for. On the Machine Translation side, feedback from the agencies suggested that the average post-editing speed of translators was 3,000 words per day. This represented a 33% increase in translator productivity, which was encouraging. Even though machine translated text helped reduce the amount of time needed for translation, in some areas, the quality was quite poor.

Corel has not used the LOGOS system in any major translation project since DRAW!6. This is because we required time to evaluate and improve the output. To this end, the language specialists have devised an evaluation method whereby texts are rated every month on terminological, syntactic and semantic accuracy. There has been a marked improvement in the translation quality over the last six months and experiments suggest that we could increase translator productivity by 50%. In addition, we have not localised any project that would justify the use of machine translation. The amount of effort required in coding new terminology, submitting files for translation and managing the process could not be applied profitably unless the project had more than 300,000 words. We are now exploring the possibility of using LOGOS to translate the contents of the Corel Internet Web site into French, Italian, German and Spanish - a task which will require a HTML filter for LOGOS.

Although, the Translation Memory XL8 is very easy to use, it has limited word processing functionality. During the DRAW!6 project, Corel used a Fuzzy Matching level of 40%, which was too low. Since then, we have discovered that translators are more productive working with a Fuzzy Matching level of above 65%. XL8 is also restricted in that the translator can only see one sentence at a time. This means that post-editing out of context introduces some problems. To ensure that translations were accurate, files were converted back to their native format and then proofed. Any changes which were implemented in these files also had to be implemented in the XL8 files. At the time, XL8 did not support the new code in Windows 95 Help files and this was problematic. Corel is now evaluating other Translation Memory systems based on criteria formulated during our DRAW!6 XL8 experience.

In brief, Corel now views Machine Translation and Translation Memory as useful tools that can be effectively applied to certain projects only. Therefore, we now evaluate each project independently and then decide which tools will yield the best results. For DRAW!6, we automated the translation of Help and Documentation but had no tools for accelerating the translation and engineering of the Software. The software is needed at an early stage in the localisation process to ensure terminology consistency with the Help and Documentation. In addition, most manuals contain many screen shots that must be taken from the localised software. The Corel Localisation Tools Group was set up in November 1995 with the task of developing a suite of tools which would speed the delivery of the localised software.

The Tools Group have developed tools to improve Terminology Management in Corel - GlossMan and WordGate. GlossMan compares an English language DLL/EXE with its localised equivalent. The tool matches the resource in each file and outputs the English and language terms to a tab delimited report file. This report file can then be imported into the WordGate Terminology Management system. WordGate allows interactive and batch entry of terms. The tool supports 20 languages including French, Italian, German and Spanish. WordGate can be also used to generate partially translated glossaries and improve terminology consistency within and across products. Both GlossMan and WordGate have enabled Corel to produce glossaries for most of its products and improve terminology consistency across products.

The Tools Group have also developed Corel Trinity to help translators, engineers, quality assurance specialists to localise the software. Trinity is more than a translation memory system for software. This tool allows the translator to reuse text from previous versions of the software in a visual translation environment. Trinity provides translators with a suite of editors that can be used to translate and update 32-bit application files. Each editor provides a near WYSIWYG environment for the translation of Strings, Menus, Dialog boxes, RC Data Tables and various other custom controls that Windows 95 and Windows NT support. The Menu Editor enables the translators to select a menu and then translate it in a near - WYSIWYG manner. The Short Menu Prompt, Long Prompt and Tool Tip text is displayed in the same editing

environment, making it easier to ensure that the translation of these three separate text entities is consistent. The Menu Editor also checks each menu for duplicated hot keys with each popup and ensures that the translator is consistent in the use of Hot Keys. The Dialog editor enables Translators to view dialogues as they appear in Corel Applications. The controls on the Dialog frame can be scaled and positions using the mouse and edited by a simple double-click. A useful toolbar at the bottom of the screen lists all the controls within the Dialog box. Additionally, the Dialog Editor checks for Duplicate Hotkeys and clipped text during the translation process. Maintaining the context of information as well as giving the translators WYSIWYG translation environments have helped increase the quality of translation and reduce the number of problems detected by internal QA and Engineering teams.[1]

Although Trinity is still in development, it has already been used in a number of projects. It has proved such a success among translators that Corel is now exploring the possibility of developing a tool similar to Trinity for use on the Help and Documentation. Many tools are already on the market for Help and Documentation but none of them seem to offer a total, integrated localisation solution for Help, Documentation and Software. This was confirmed by a translation tool survey that we conducted among members of the Irish Software Localisation Industry Group. We were surprised to find that Translation Memory systems are not used as extensively as one would imagine, in that 31.25% of those surveyed use translation Memory systems, but only 6.25% use them on a regular basis. We found that more people have used and evaluated Machine Translation systems than any other translation tool. We also discovered that the most frequently used translation tools in the localisation industry are Spell Checkers and Terminology Management systems.

With the recent acquisition of the WordPerfect family of products, Corel has re-thought its translation automation strategy. We have decided not to impose one Translation Memory system on our vendors but instead we have let them use whatever Translation Memory system they are most comfortable with. In this way, there is no learning curve involved for the translators and so their productivity is not inhibited. We do however require that the vendors deliver the memories to us in a standard format for easy import into a system.

The Localisation Services Department in Corel has definitely not evaporated but instead has gone from strength to strength. By automating and emigrating our process we have been able to deliver quality products to our international markets in a timely fashion.