

# Translator's Workbench II

by  
Sabine Bell

The *Translator's Workbench II* is a translation tool for PC use, comprising a text editor, a translation memory facility and a terminology database, MultiTerm, which is also available as a separate package. It replaces the *Translator's Workbench*, which consisted of three separate programs, one of which could be used as a memory-resident text editor within other common word processors, providing term recognition and substitution facilities. This idea will be resumed in the *Workbench* for Windows.

We have been working with version 1.16 of the *Translator's Workbench II*. In the meantime, this has been superseded by one with additional features such as a conversion program for ANSI files (WINDOWS output format) into ASCII and back again, and a 'text delimiter' definition facility to identify only certain parts of a text (a program, for instance) for translation. These features do not change the actual program described here, but they increase the scope of its application.

The minimum hardware requirements are a 80386 processor, 2 MB of RAM, a hard disk and a colour monitor (as we shall see later on, the use of colour is significant in the program). The installation of the software is very straightforward: a simple copying of the files into the desired directory makes the program instantly available, providing the dongle is in place.

The *Translator's Workbench II* can deal with a variety of file formats, such as ASCII, RTF, INTERLEAF and VENTURA but, as with other programs, the relevant conversion files, i.e. special tag files for the input of text from different word processing systems, are required and form part of the package. For formats not catered for by the *Workbench II*, the text could be input as an ASCII file and would then have to be post-edited to be formatted. A problem I encountered is that the line breaks from the *Workbench II* ASCII document are transformed into hard returns when loading it straight into Microsoft Word, so one has to make sure that the word processor provides the option of loading an ASCII file without returns, as in WordPerfect.

Translation using this software is not explicitly project-based, but to use the facilities available, the user needs to make decisions on the required tag file, the appropriate terminology databases or dictionaries, and the translation memory. So, when these are grouped together, the document is turned into a project and is treated as such by the program: when the translator returns to the work after exiting, s/he is taken straight back to where the document was left,

with all the reference materials in place. The user may also choose to define a project attribute for the translation, which can be useful for matching processes.

The program is menu-driven, although many of the menu functions can also be achieved by pressing brief key combinations. The keys have been assigned meaningfully and are therefore easy to remember. This is one of the numerous aspects which make this package very user-friendly.

The interface languages in the version discussed here are English and German and are instantly interchangeable at any time. The source languages handled by the program are more limited in number than in TRANSIT, due to the matching process described below. This version of the *Workbench II* caters for English, German, French, Spanish and Italian.

The *Translator's Workbench II* in its most basic application is like any other editor or word processor, making use of familiar key assignments and functions. The screen is divided into a menu bar, the edit area, a command line, a status line and a message line.

The command line offers easy access to DOS and the use of a variety of extremely useful and conveniently shortened commands, e.g. open, find, rename. The <Esc> key acts as a toggle between the edit area and the command line, which, incidentally, can be used in normal word processing or translation memory mode.

When the Translation Memory mode is selected, however, the text is processed sentence by sentence or segment by segment. Unlike TRANSIT, the *Workbench II* segments texts according to fixed rules, which, in most cases, makes things easier. The translator is presented with two fields, one of which is an editing field within the edit area, which takes on all the attributes of the latter and supports the same key functions. This field receives the translation, which, once accepted by the translator, will be inserted into the text.

The other field is a display field, containing the source language segment(s) and, if applicable and desired, the translation of words found in the dictionary by ATR, the Active Terminology Recognition facility. ATR automatically looks for known terms, i.e. terms stored in the selected MultiTerm database, and interactively tells the user that these terms are known, by highlighting them in white. This also works in the normal text mode, when a text is translated without referring to the translation memory built up from previous translations.

**As part of our series on machine-assisted translation (Language International 4/1, 4/6), this article focuses on the 'Translator's Workbench II', developed by Jochen Hummel at Trados GmbH in Stuttgart, and tested for the purposes of this article, as with previous items in this series, at the Leeds office of Praetorius Limited.**

```

File Database Terms Functions Options Utilities Help
-----
Bottom of F Xlate Until Next Fuzzy Alt-Z
             Display DB Sentence Alt-R
             Display DB Info Alt-I
             JTerm Recognition Alt-T
             JSIeno Recognition Alt-O
             JProtect Tags Alt-P
             JPhysical Tabs
             Beep

Line 1 of 1 Col 1 Pr St Tr In Wd Ll
F1 Help 2 Save 3 Close 4 Save&Close 5-6 --Trans 8=Next 9=Undo 10 Menu

```

```

----- Top of File -----
If you require pure desktop publishing power, boundless creativity,
galanted quality in other words, everything you could possibly wish
for - choose this machine.

With this printer, you can compose complex text pages and intricate
illustrations for perfect presentation materials, bulletins and reports.

There's a full range of models to meet your needs, from 5 to 14 pages
per minute. Whether you wish to scale, relate, move or otherwise change your
drawing, the printer will put your ideas to paper.

New: The compact, affordable machine described in this booklet is our new
page:
GER: Drucker

----- No Match -----
C:\TWI\INCADB ASC.TXT Line 1 of 15 Col 1 Pr St Tr In Wd Ll
F1 Help 2 Save 3 Close 4 Save&Close 5-6 --Trans 8=Next 9=Undo 10 Menu

```

```

----- Top of File -----
Das in dieser Broschüre beschriebene kompakte, preisgünstige Gerät ist unser
neuer Seitendrucker.

Wenn Sie reine DTP-Power, grenzenlose Kreativität und eine auf Lebensdauer
garantierte Qualität suchen - kurz, alles, was Sie sich überhaupt nur
wünschen könnten, wählen Sie dieses Gerät.

Mit unseren Drucker können Sie komplizierte Graphiken für perfekte Vorträge,
Bandschreiben und Jahresberichte in komplexe Textseiten einarbeiten.

Es gibt eine umfassende Auswahl von Modellen, um Ihren Bedürfnissen gerecht
zu werden, mit Modellen von 4 bis 15 Seiten pro Minute.
- Enquire today!!

New We offer a wide range of models to meet your needs, from 4 to 15
pages per minute.
Old: There's a full range of models to meet your needs, from 4 to 15 pages
per minute.
-----Date: 92/04/19 Head: 0 Alt: User: RC- Gls. Match-----
C:\TWI\INCADBZASC.TXT Line 11 of 13 Col 1 Pr St Tr In Wd Ll
F1 Help 2 Save 3 Close 4 Save&Close 5-6 --Trans 8=Next 9=Undo 10 Menu

```

The translation/s for a highlighted term is/are obtained by pressing a function key. If there is more than one translation, they are separated by commas and highlighted one after the other by pressing the key repeatedly. If there is more than one highlighted term, a toggle function is used to scroll through them. Once highlighted, each of these can be pasted directly into the translation by pressing <Alt A>.

The dictionary lookup provides an interesting difference from the programs discussed so far: both

the IBM Translation Tool and TRANSIT have a separate dictionary window listing all the terms and translations from the segment which appear in the selected database(s). It may seem that the advantage of this would be that the translator cannot help but look at the correct term, instead of relying on his/her memory and confidence, but from my own experience, I know that firstly, the fact that the term is displayed does not necessarily mean it is being noticed, and secondly, long lists (segments containing several known terms) require scrolling through the dictionary window, which is time-consuming. By incorporating the dictionary window information into the display field, the *Translator's Workbench II* gives the translator less visual detail to cope with and more space on the screen.

The program offers a so-called 'Steno' function ('Macro' in TRANSIT) which allows the user to define stenos or replacement characters for frequently appearing phrases (this can be very useful when a long name containing unexpected capital letters appears frequently in a text!). The phrases are automatically inserted when the steno is typed followed by a space. The stenos are stored in an ASCII file which can be created and edited directly from within the *Workbench II*, taking effect instantly. Consequently, it is very easy to add new stenos in the course of a translation.

But to come to the most important aspect of *Translator's Workbench II*: the biggest achievement of this sort of program is the comparison of new source texts with previously translated ones and their translations, namely, the translation memory function. In TRANSIT, the translation memory consists of a number of comparable source documents stored together with the target language documents, with which the new text is compared in the process of importing it into TRANSIT. Previously translated segments are automatically replaced with the original translation if desired. Basically, the translation memory is a freely chosen group of previously translated, thematically related documents and their translations.

In Trados's program, however (like the IBM Translation Tool), the translation memory is a separate database which is updated after each segment. It stores only one exact translation for a segment, so if a segment appears a second time under the same project attribute and the translation is changed, the first version will be overwritten.

Translations created outside the *Workbench II* can be used for reference by segmenting them and merging the source and target language segments into an ASCII document, which subsequently requires structuring according to clearly defined rules. Due to the nature of translation, this process cannot be fully automated and can be very time-consuming,

but it is effective. Going in the other direction, the translation memory can also be exported as an ASCII file, either for reference or for editing purposes.

The basis on which matching takes place in this program is not pattern-matching on a character-by-character basis as in TRANSIT, but morphological: nouns and adjectives are stored in their root form and consequently recognised in their different forms. This results in increased terminology recognition (or decreased dictionary input requirements) and added consistency. The program also recognises sections within segments and offers 'fuzzy matches' in addition to 100% agreements. These processes require a thorough knowledge and analysis of the source language, which is the reason for the limited number of languages covered by the *Workbench II*. Needless to say, the advantage of an increased number of matches outweighs this disadvantage (especially for those of us who work with the languages offered!).

Although 100% matches are inserted automatically into the editing field, the user has the option of editing them, and they have to be formally accepted by pressing <Ctrl><Enter>, which also leads to the next segment. The display field shows not only the source text segment, but also the project attribute and user identity of the previous translation. These can be used as a filter, for instance if a translator wishes to use only his/her own work for reference.

The more interesting aspect of matching, however, is 'fuzziness'. The user can determine the degree of acceptable deviation by setting the percentage of agreement between the segments. The lower the percentage, the more matches there will be. For the matching to be effective, a percentage of between 65% and 75% is recommended, although any value between 1% and 99% can be set.

In cases where the only difference between segments is a term which can be found in the dictionary, the program, in spite of regarding the match as fuzzy, automatically implements the change by inserting the new term instead of the old one, so in most of these cases, the translator just needs to confirm the translation. Numbers, by the way, do not constitute a difference: the system automatically applies a number change, even following punctuation conventions.

If words (names for instance) have been replaced, this is highlighted in blue. Green signifies words added to the source sentence, whereas red identifies text which no longer appears in the source segment. And brown indicates that a phrase has moved. If, as is often the case with 'brown differences', the translator does not change the found (fuzzy) match, s/he is asked to confirm the translation before the system allows her/him to move on to the next segment.

This colour coding considerably speeds up the process of editing the fuzzy match (i.e. the automati-

cally appearing target language segment) and helps to grasp the changes more quickly. It is the most striking feature of the program and works convincingly. Trados generally place great emphasis on the appearance of their programs, which was one of the reasons for being nominated in 1992 for the 'Best German Software' at the European Software Excellence Award in Cannes, the annual event staged by the SPA (Software Publishers' Association) and Ziff-Davis.

The one problem a translator may have with the program is one for which there is generally no easy answer: the tag protection in files other than ASCII. External tags, not appearing within a segment, are just stored in the translation memory with the text, but tags within a segment may have to be specially inserted or moved by the translator: in the case of a match being found, the original tags will appear as suggested options, requiring the translator's approval. Normally, the tags as such are protected to prevent destroying the format, but if they do not match the requirements, the tag protection has to be switched off to make the necessary amendments. If no match is found, the tags have to be inserted in the appropriate places. As far as the export of these files back into the word processors is concerned, however, there are no problems.

The terminology database, MultiTerm, is fully accessible from within the program, i.e. in addition to being used by the automatic process and the display facilities, it can be called up and added to, edited or browsed. Any change takes effect immediately. As this database can also be operated independently of the *Workbench*, the preparation of dictionaries is very convenient. In fact, MultiTerm, although being a major aspect of the *Workbench*, is described in a separate manual and is a product which deserves a review in its own right.

As with the translation memory, import of otherwise collected data is possible in a similar way as described for the memory database. Having familiarised him/herself with MultiTerm, the translator will find it an invaluable tool.

The User's Guide to the *Translator's Workbench II* is a helpful reference and contains detailed but concise information about the program, a trouble shooting section, a comprehensive list of key assignments and a glossary; in other words, it is user-friendly. The same text also constitutes on-line Help, accessible either as the entire file or via an index. In this context, it seems less helpful.

My overall impression of the package is that it is very responsive and easy to use. The decisions a translator has to make are all relevant to the process of translation itself. The *Translator's Workbench II* is attractive in its appearance and does make one's work easier.

The assistance of *Language International* software editor, Bob Clark, in the compilation of this article is acknowledged.