TRADOS Translator's Workbench for Windows

Workflow Manual

Version 1.2

1999

TRADOS GmbH

Stuttgart, Germany

Table of Contents

Introduction

About This Manual

This manual is a "Workflow" guide for Translator's Workbench. It is meant as an accompaniment to the user manuals for Translator's Workbench, MultiTerm, WinAlign and The S-Tagger (for FrameMaker and for Interleaf). It is assumed throughout this guide that the reader has the user manuals that are relevant to their translation process and is familiar with the fundamentals of these tools.

The Workflow manual is divided into two sections. The first section is intended primarily for first-time users of TRADOS Fine Translation Tools, and, in particular, of Translator's Workbench. The more experienced user of Translator's Workbench will also find some useful tips and information in this section. The second section of the manual is intended primarily for the person responsible for creating translation memories and for preparing files for translation using Translator's Workbench, but it will also provide the project manager with an overview of the tasks involved when using Translator's Workbench and MultiTerm.

More specifically, Section 1 deals with issues such as prerequisites, project-specific requirements, creating translation memories, general workflow scenarios, update management, terminology management and the integration of machine translation tools with Translator's Workbench. Section 2 deals with what file set-up is best for Translator's Workbench and what preparation and handling is required for different file formats. The most common file formats are covered in Section 2.

List of Abbreviations Used

MTW	MultiTerm
TW	Translator's Workbench
TM	Translation Memory
WA	WinAlign
LAN	Local Area Network
CAT	Computer Aided Translation
STF	Sarah's Tagged Format

Diagram Conventions

The conventions used in diagrams throughout the manual are as follows:

- So as to keep diagrams as clear as possible, both *Processes* and *Entities*, are represented by rectangles. An example of a process might be analysis, an example of an entity might be source files. One exception to this is the *Conversion Process*, for example converting source files into a tagged text format. This is represented by a triangle.

- Databases are represented by cylinders.

- Where a step is optional, it is represented by a broken arrow. An optional process is represented by a grey rectangle or triangle.

- Where a step is not optional, it is represented by a continuous arrow.

Trademarks

- TRADOS® Translator's Workbench for Windows, TRADOS® WinAlign, TRADOS® MultiTerm® '95 Plus! are trademarks of TRADOS GmbH. The S-Tagger ® is a trademark of TRADOS Ireland Ltd.

- Microsoft is a registered trademark; Word for Windows, WinWord, and Windows are trademarks of Microsoft Corporation.

- Corel is a registered trademark; Corel Ventura and WordPerfect are registered trademarks of Corel Corporation.

- LOGOS is a registered trademark of LOGOS Computer Integrated Translation GmbH.

- Systran PROfessional for Windows is a registered trademark of Systran Software Inc.

- IBM is a registered trademark, and Translation Manager is a registered trademark of IBM Corporation.

- FrameMaker and PageMaker are trademarks of Adobe Systems Incorporated.

- QuarkXPress is a registered trademark of Quark Inc.

- Interleaf is a registered trademark of Interleaf Incorporated.

- RoboHelp is a trademark of BlueSky Software Corporation.

- CopyFlow is a registered trademark of North Atlantic Publishing Systems.

Other product names are generally trademarks or registered trademarks of their respective owners. They are used without representation as to their release for general use.

Copyright

Copyright TRADOS GmbH 1997-1999.

Section 1: Project Management with Translator's Workbench

Section 1 is a general introduction to project management and workflow using Translator's Workbench and MultiTerm. This section is aimed both at those who are new to Translator's Workbench and those who have some experience with TRADOS Fine Translation Tools. For those who are new, Section 1 highlights many of the issues which arise when first implementing Translator's Workbench in a working environment and aims at reducing the learning curve. For those who have some experience with Translator's Workbench, it is hoped that you will find some new and useful information. Where possible, workflow diagrams have been included to clarify the issues being discussed.

Section 1 is divided into seven sub-sections.

- **When is Translator's Workbench suitable for a translation project?** This is a question often asked by our customers. This section is then divided into two: one which deals with preliminaries and prerequisites for using Translator's Workbench and one which examines project-specific requirements.

- **Creating Translation Memories.** The discussion here centres on when it is possible and desirable to create translation memories using the alignment tool, WinAlign.

- **Preparation of a Translation Project using Translator's Workbench** is an extensive discussion of the preparation required before using the Translator's Workbench and, indeed, the tasks during and after a project where the Workbench has been used. There are some important tips and pointers in this section on how the special functions in Translator's Workbench can be used to your best advantage.

- **Workflow Scenarios** summarises the most common workflow scenarios and specifies what tasks are involved in each.

- **Update Management** provides hints on how to manage updates.

- **Terminology Managment** discusses the advantages of using MultiTerm in conjunction with Translator's Workbench and gives pointers on how to manage terminology updates.

- Finally, in **Translation Memory and Machine Translation**, there is a short description of how Translator's Workbench can be integrated with machine translation tools and of the issues which have to be considered if you are choosing this option.

1.1 When is Translator's Workbench Suitable for a Translation Project?

There are many factors to be considered when deciding whether or not one should use Translator's Workbench for a translation project. These factors will vary depending on the translation environment you work in. For example, if you are a project manager in a large translation company or in the translation department of a production company, you will examine factors which would not be of as much concern to the freelance translator and, of course, vice versa. Having said that, there are many factors that are of equal importance to different types of users. In this section we will outline some of these factors.

It is almost impossible to list factors in order of priority, since something that is at the top of one user's list might not even appear on another user's list. So, factors are listed according to **Preliminaries and Prerequisites** first (Section 1.1.1) and **Project-Specific Requirements** second (Section 1.1.2). Under each topic, the most common questions related to that topic are asked. However, the answers are not provided since they will depend entirely on each translation scenario and can only be provided by you, the user. Several of the topics mentioned here are discussed in more detail later in the manual.

At the end of this section, you will find two diagrams, which summarise the questions and possible responses outlined in the **Preliminaries and Prerequisites** and **Project Specific Requirements** sections.

1.1.1 Preliminaries and Prerequisites

The following are the prerequisite and preliminary points which should be asked by anyone who is considering Translator's Workbench and MultiTerm as translation tools.

Source and Target Languages

Can Translator's Workbench handle the languages in your translation project? Translator's Workbench is based on the Unicode standard and therefore supports all languages available under Windows 95/98/NT (except bidirectional languages). Please refer to the TRADOS website, http://www.trados.com for a complete list of supported languages.

Machine-Readable Files

Are your files in machine-readable (i.e. electronic) format? If not, can you successfully convert them to electronic format within the time-scale of your translation project?

File Format

Is the format of your files supported by Translator's Workbench? Translator's Workbench supports many file formats, including tagged text. Please refer to the TRADOS web site http://www.trados.com and the Translator's Workbench User's Guide for a complete list of the file formats supported.

Text Type

Some text types are more suitable for translation using Translator's Workbench than others. The text types which are most suitable are those where there is a high level of repetition, for instance, technical texts, user manuals, on-line help and texts which are frequently updated. Are the text types you translate suitable for translation using the Translator's Workbench or MultiTerm?

Number of User Licences

How many user licences will you require to complete the translation project in the time allocated? Do you have a sufficient number of licences? If you do not, can you obtain more licences before the start date of your project?

Number of Trained Users

Do you have an adequate number of trained Translator's Workbench users? If the answer is no, can you organise training and practice sessions for them before the start of your translation project? TRADOS offer training which lasts approximately two days. Please contact your local TRADOS vendor for more information. Can your project budget afford the cost of this training and how will your project schedule be affected by this training period? Should the training be seen as a long-term investment, i.e. will these trained translators be used again for another translation project?

Machine Specification

A Pentium with 32 MB of RAM is recommended to run Translator's Workbench. Do your users have the appropriate machine specification?

Support Personnel

If you are planning a large translation project, which will require many users who are perhaps scattered over many locations, you will need support personnel to answer technical questions. Do you have personnel available to support your translators during this project? You might also consider availing of a support contract from TRADOS. Contact your local TRADOS office for information on support options. You will also require one person (or more – depending on the size of the project) to carry out the tasks associated with the use of computer-aided translation technology, e.g. distribution and receipt of translation memories, merging and reorganising translation memories and cleaning and converting files. Do you have these resources at your disposal? Are you, perhaps, going to do these tasks yourself? Does your project schedule allow time for the tasks related to using translation memory tools? For more information on key roles, see Section 1.3, **Preparation of a Translation Project Using Translator's Workbench.**

1.1.2 Project-Specific Requirements

Having examined the preliminaries and prerequisites and having made sure that everything is in order, it is then appropriate to examine the particular project(s) you have in mind and to assess the following points.

Networks and Consistency

Translator's Workbench allows the user to share translation memories over a LAN. Although a LAN is not a prerequisite for using Translator's Workbench, working in this type of scenario helps to guarantee consistency, which is extremely important. In addition to this, if there is a high level of repetition in the files to be translated, sharing memories can speed up translation since one translator may already have translated a passage which appears in a section being translated by another translator. The second translator simply has to accept the translation as a 100% match. Translators will also benefit from similarity throughout files with Translator's Workbench fuzzy matching and concordance capabilities. Will your translators have access to a LAN?

Size and Life-Cycle of the Translation Project

No job is too small or too large for Translator's Workbench. However, the size and potential life-cycle of the project are important factors when deciding if you should use Translator's Workbench. For example, if there is a great number of words to be translated in a very short period of time, and there is a high percentage of repetition throughout the project, then Translator's Workbench might prove to be the only way you can meet the deadlines imposed. Or, as another example, you may have to translate only a small number of words, but you are aware you will receive updates in a number of weeks and you want to have a translation memory so as to facilitate the translation of the update. Even if you are dealing with a project where the life cycle is uncertain, you may still decide to use Translator's Workbench because of the number of words involved and the amount of repetition throughout the files to be translated.

Another reason in favour of using Translator's Workbench is its excellent tag handling feature. If the files for translation have been created in a complex DTP package, with which translators are not familiar, Translators Workbench offers a means of translation which eliminates the necessity of training translators in complex DTP packages and often reduces the amount of DTP "tidy up" to be done after the translation phase of a project.

Updates – Size, Type, Frequency, Timing and File Format

Having considered many of the factors outlined here, it is possible that you might decide that Translator's Workbench is not suitable for your current project. For example, maybe there are no translation memories in existence? Perhaps you have no trained users or no LAN? However, if there are updates expected during the life cycle of the project, you should reconsider. The batch translation feature and **Translate to Fuzzy** command are extremely useful when it comes to

updates. These features automate the tedious processes of finding text that is similar or identical to the original source text and cutting and pasting translations.

There are five factors to take into account when managing updates with Translator's Workbench: size, type, frequency, timing, and file format. Updates are discussed in more detail in Section 1.5, **Update Management**, but we will outline the issues here:

- It is better to receive updates at the early stages of a project rather than at the end of the project, when nearly all translation is complete. If you receive an update late in the project, you will have to translate almost everything again using Translator's Workbench automatic batch processing feature (**Translate**, 'no 100% match' command combinations) or the interactive **Translate to fuzzy** comand approach. While the batch translation process in itself is fast, it requires a quality assurance cycle of the batch translated text.

- You should always take the nature of updates into account. For example if the update consists of a change to the product name, you should consider using the Translator's Workbench automatic substitutions feature or the TM Maintenance feature to process the new files.

- You should always aim to receive any updates before the files have been "cleaned" using the **Clean Up** command. Otherwise, you will have to repeat that step in your process.

- You should encourage the originator of the documents to keep the number of updates to a minimum – large, infrequent updates are better than small, frequent ones.

- Finally, the file format is also something to be considered as some file formats require more preparation work and conversion into suitable formats than other file formats.

Existing Translation Memories

- Are there translation memories in existence which you can use to leverage into the current translation project?

- Are these translation memories compatible with Translator's Workbench?

- If not, can they be converted to Workbench format (Translator's Workbench has an import filter for IBM's Translation Manager translation memories, for example).

- Is the translation in the memories of an acceptable quality? If not, what quality control procedures can you implement?

- If there are no translation memories in existence, do you have access to previous source and target language files?

- If so, are these files suitable for alignment?

There should be a parallel target language file for every source file. They should preferably be named in the same way for ease of identification and the structure of the translated file should reflect the structure of the source file as much as possible. Then you must consider the format of such files.

- Is this format supported by the alignment tool?

- Does a conversion solution exist which will allow you to convert the file format to one which is supported by the alignment tool?

- Will the formatting be preserved during alignment?

For more information, see: Section 2, **How to Deal with Specific File Formats**.

Customer Expectations[1]

- Does your customer expect a translation memory (or memories) as a deliverable at the end of the translation project?

- If so, how are you planning on meeting this requirement? See Section 1.2, **Creating Translation Memories.**

Some customers may already be familiar with translation memory technology. Others may be new to the technology. It is important to inform the customer of the issues involved in the use of this technology and to set realistic expectations. Translator's Workbench is an extremely powerful tool, but it does not do away with the need for professional human translators.

Analysis Figures

If there are previous translation memories in existence, or, alternatively, if you have created translation memories using the alignment tool, you should then run an analysis on your files for translation using the translation memory. These figures will give you an indication of the percentage leveraging. Sometimes these figures might influence your decision on whether or not to use Translator's Workbench. For example, if there is a low percentage of 100% matches, repetitions, and fuzzy matches, you might decide not to use Translator's Workbench. However, other factors such as project life cycle, number of updates expected, percentage repetition or the existence of a MultiTerm database might influence you to use Translator's Workbench. (For more information, see Section 1.3, **Preparation of a Translation Project Using Translator's Workbench.**

Terminology

Translator's Workbench and MultiTerm work very well together to ensure terminological consistency. Terms which are in the MultiTerm database and which appear in a source sentence are highlighted and can be pasted into the translation with the click of a button. It is highly recommended that, when using Translator's Workbench, terminology is stored in a MultiTerm database. You should, therefore, make sure that your terminology is in MultiTerm format before translation and that your translators are familiar with MultiTerm. You may also find that even if Translator's Workbench is not suitable for a translation project, MultiTerm is ideal. For more information, see Section 1.6, **Terminology Management.**

Diagrams

Two diagrams follow, illustrating the points outlined here. The first diagram covers the decision flow for Preliminaries and Prerequisites. It begins with the question "Are both source language and target language supported by Translator's Workbench?" and works its way to a point where all prerequisites are fulfilled.

---

[1] The term "customer" is used throughout this manual to mean both internal and external customers. For example, if you work in the translation department of a large production company, then your "customer" might be the Documentation Department. If you work in a large translation agency, then your "customer" will most likely be external.

Preliminaries and Prerequisites

```
                          ┌─────────┐
                          │  Start  │
                          └────┬────┘
                               │
                          ╱─────────╲
                         ╱ Source and ╲─────────────No──────────────────────────────────┐
                         ╲   target   ╱                                                  │
                          ╲ language ╱                                                   │
                          ╲supported╱                                                    │
                               │                                                         │
                          ╱─────────╲                                                    │
                         ╱   Text    ╲─────────────No──────────────────────────────────┤
                         ╲ type      ╱                                                   │
                          ╲suitable?╱                                                    │
                               │                                                         │
                          ╱─────────╲                           ╱─────────╲             │
                         ╱   File    ╲─────No────────────────►╱  Work     ╲───No───────┤
                         ╲  format   ╱                         ╲  around   ╱             │
                          ╲supported?╱◄────────Yes────────────╲ possible? ╱             │
                               │                                ╲────────╱              │
                          ╱─────────╲                           ╱─────────╲             │
                         ╱  Machine  ╲─────No────────────────►╱           ╲──No─────────┤
                         ╲ readable  ╱                         ╲Conversion ╱             │
                          ╲  files?  ╱◄────────Yes────────────╲possible? ╱              │
                               │                                ╲────────╱              │
                          ╱─────────╲                           ╱─────────╲             │
                         ╱Sufficient ╲─────No────────────────►╱ Purchase  ╲──No─────────┤
                         ╲   user    ╱                         ╲  more?    ╱             │
                          ╲ licences?╱◄────────Yes────────────╲          ╱              │
                               │                                ╲────────╱              │
                          ╱─────────╲                           ╱─────────╲             │
                         ╱  Trained  ╲─────No────────────────►╱ Training  ╲──No─────────┤
                         ╲   users?  ╱                         ╲   or      ╱             │
                          ╲         ╱◄────────Yes─────────────╲alternatives?╱           │
                               │                                ╲────────╱              │
                          ╱─────────╲                           ╱─────────╲             │
                         ╱ Adequate  ╲─────No────────────────►╱           ╲──No─────────┤
                         ╲  machine  ╱                         ╲ Upgrade?  ╱             │
                          ╲spec'n?   ╱◄────────Yes─────────────╲          ╱              │
                               │                                ╲────────╱              │
                          ╱─────────╲                           ╱─────────╲             │
                         ╱  Support  ╲─────No────────────────►╱  TRADOS   ╲──No─────────┤
                         ╲ personnel?╱                         ╲ support   ╱             │
                          ╲         ╱◄────────Yes─────────────╲contract? ╱              │
                               │                                ╲────────╱              │
                     ┌──────────────────┐                                               │
                     │  Proceed with    │                                               │
                     │  project using TW│                                               │
                     └────────┬─────────┘                                               │
                          ┌───┴───┐                                                      │
                          │  End  │◄───────────Alternative solution─────────────────────┘
                          └───────┘
```

The second diagram outlines the topics discussed under **Project-Specific Requirements**. The main requirements are located in the centre of the diagram. They are not ordered according to any priority, as priorities will be different from company to company. The positive elements lead to the right, which favours the use of Translator's Workbench. The negative elements lead to the left, which does not favour the use of the Workbench.

Project-Specific Requirements

Against using ←	Size and life-cycle of project	→ In favour of using
← Low Repetition		High repetition →
← Small number of words		Large number of words →
← No updates expected		Updates expected →

	Existing translation memories	
← No		Yes →
← No alignment possible		Alignment →
← Not compatible with TW		Compatible with TW →
← Low quality translation		Good quality translation →

	Analysis figures	
← Low 100% match rate		High 100% match rate →
← Low fuzzy match rate		High fuzzy match rate →
← Low repetition rate		High repetition match rate →

	LAN	
← No		Yes →
← No alternatives		Alternatives →
← Not stable		Stable network →
← Slow network		Fast network →

	Customer expectations	
← No translation memory		Translation memory →

	Terminology	
← Not compatible with MultiTerm		MultiTerm format →
← No conversion possible		Conversion possible →

Against using Translator's Workbench

In favour of using Translator's Workbench

1.2 Creating Translation Memories

1.2.1 What is required for successful alignment?

If you are a first-time user of Translator's Workbench, you may possess previous source and translated documentation, which you want to convert to a translation memory so that, even as a first-time user, you can gain immediate advantage from Translator's Workbench. TRADOS's visual alignment tool, WinAlign, allows you to do just this – WinAlign converts previous source and target documentation into a translation memory. Certain criteria should be fulfilled before the translation memory can be created. These are outlined below. In this section, you will also find a general description of the alignment workflow.

Formats Supported by WinAlign

WinAlign supports two types of file format: WYSIWYG (What You See is What You Get) and tagged text formats (non-WYSIWYG).

Tagged text formats are produced by converting the original format into tagged text using a conversion or a markup utility. The S-Tagger (for FrameMaker or Interleaf) converts FrameMaker or Interleaf files into a tagged text format (STF). Source and target STF files may be aligned using WinAlign. The ITP Filter Pack marks up exported text from Corel Ventura, PageMaker and QuarkXPress as tagged text which is suitable for alignment. SGML files are marked up with the SGML2RTF conversion tool. DCF Bookmaster, RC and Troff files are marked up using a series of macros from a WinAlign template (`WinAlign.dot`). For more detailed information on the alignment of specific file formats see Section 2, **How to Deal with Specific File Formats**.

WinAlign supports the following WYSIWYG formats:

- Word rich text format.

- RTF based Windows help files.

- HTML.

WinAlign supports the following tagged text formats:

- FrameMaker.

- Interleaf.

- Ventura.

- PageMaker.

- QuarkXPress.

- SGML.

- DCF BookMaster.

- RC.

- Troff.

Comparing old source files with new source files

One way of assessing how much similarity there is between old source files and new source files is by using the Translator's Workbench **Analyze** command. This allows you to compare two versions of source documents. The results will often decide for you whether or not time should be spent on alignment.

Quite often, expectations are that files have only been updated slightly and that there should be 90% leveraging between the old files and the new files. Unfortunately, it is often the case that the

expectation of high leveraging is not met since the original source document has been changed more than the documentation department expected or planned. Comparing old and new source files will provide you with exact details on similarity. To compare old and new source files, follow the instructions below:

- Create a translation memory using the **New** command in the **File** menu of Translator's Workbench. Specify the source language for your files and any target language (it does not matter in this context which target language you specify as long as it is not the same as the source language).

- From the **Tools** menu, select **Analyze** and select all *old source files* to be analysed, click **Analyze**.

- When this analysis is complete, click **Clear** to remove files that have just been analysed.

- Check the **Use TM from Previous Analysis** box, select the *new source files*, click **Analyze**.

Translator's Workbench will now compare the new source files against the old source files and will show you how many 100%, fuzzy, repetition and no matches there are. For more information on match types, see Section 1.3.4, **Analysis Figures and Project Planning**.

Alignment Workflow

You should allow sufficient time for alignment before the commencement of translation. The amount of time required depends on the following:

- How many files are to be aligned?

- How many languages are to be aligned?

- Which file format?

- How much file management is involved if the source and target language files have changed names?

- Are there differences in the structure of the source and target files?

- How many human resources can you dedicate to the task of alignment?

- Do you intend checking the alignment of the translation memories before they are used for translation?

Each of these points is discussed in more detail below:

How Many Files?

The greater the number of files, in general, the longer the alignment will take.

How Many Languages?

Again, the greater the number of languages, the more time will be required for the alignment process.

Which File Format?

Some file formats will require more time than others. For example, if you want to create translation memories from FrameMaker or Interleaf files, you will first have to convert those files using The S-Tagger 2.0 for FrameMaker or The S-Tagger 2.0 for Interleaf. This is an additional step that is not required for Windows help files, for instance.

How Much File Management?

When aligning, the user instructs the alignment tool to align file X in the source language with file Y in the target language. It does not matter if these files do not have the same name, as long as you can easily identify which source language file corresponds to which target language file. In general, the task is easier if the files have the same name in the source and target languages.

Differences in File Structure?

If the structure of your target files has changed considerably in comparison to that of your source files, your alignment results might not be acceptable. If only minor changes have occurred, however, it may be worthwhile restructuring your files during the alignment process. WinAlign's Alignment Editor allows you to review alignment both at a structural level and at a segment level.

How Many Resources?

If you have large quantities of text to align, you might want to consider allocating the task to more than one resource in order to speed the process up. If you are working as a freelancer, you will need to consider what time, if any, you can dedicate to the task of alignment.

Checking Translation Memories?

WinAlign allows the user to align in batch mode and to review the alignment result. A linguist opens the aligned source and target files in WinAlign, compares the macro-structure of the documents, makes any changes necessary and then reviews the alignment on a sentence-by-sentence basis. Correctly aligned sentences are accepted and saved into the alignment project. Incorrectly aligned sentences are fixed. The result of this type of alignment is a 100% accurate translation memory, ready for use. However, as you can imagine, this activity can take a significant amount of time.

The alternative is to export the alignment result without reviewing. This process takes significantly less time than reviewing the alignment (a few seconds rather than a few hours). However, since it is not controlled by a linguist, there may be some misaligned units (i.e. a source sentence mapped incorrectly to a target sentence) in the translation memory. The quantity of such units depends on several factors, for example how much the structure of the target file has changed when compared to the source file, how many formatting tags are contained in the files. Misalignment can be reduced significantly if a linguist reviews the alignment. Where alignment has not been reviewed by a linguist, misalignments may be corrected by the translator during the translation phase.

There are two possible alignment workflows. The first workflow includes steps for a linguist to check the alignment. The second alignment workflow consists of automatic alignment without any linguistic review.

Alignment workflow with linguistic/structural review:

- Receipt/retrieval of previous source and target language files.

- Comparison of new source files with old source files.

- Preparation and conversion of tagged text formats.

- Alignment.

- Review Alignment /Re-structuring of source and target language files.

- Export of result.

- Import into Translator's Workbench.

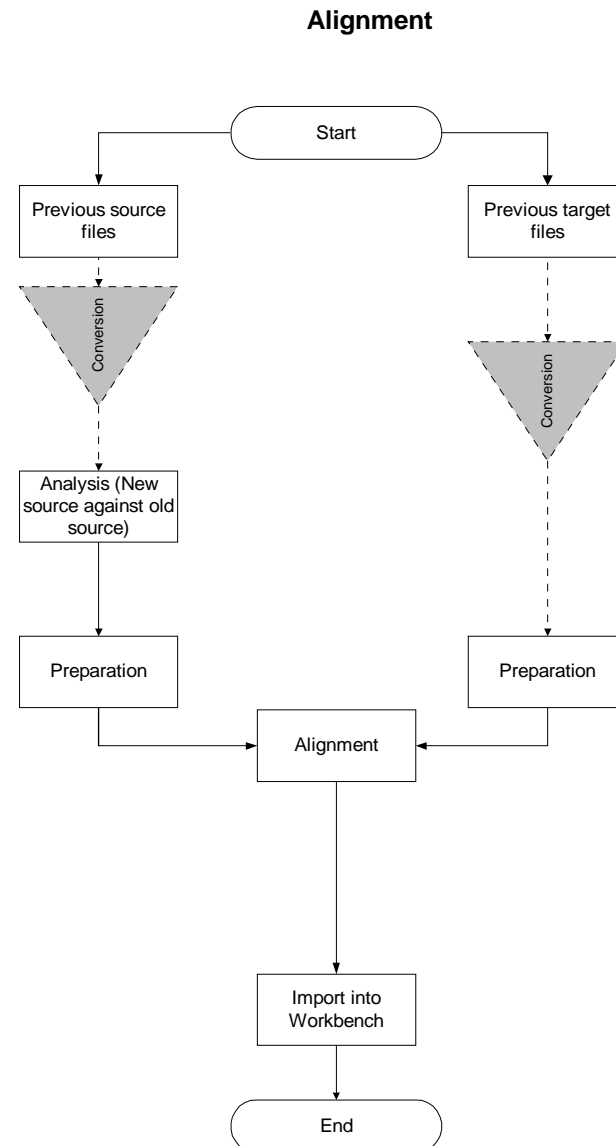Alignment workflow without linguistic/structural review:

- Receipt/retrieval of previous source and target language files.

- Comparison of new source files with old source files.

- Preparation and conversion of tagged text formats.

- Alignment.

- Brief checking of results.

- Export of result.

- Import into Translator's Workbench.

- Fixing of misaligned units during translation.

The following diagram illustrates the generic alignment workflow:

Alignment

```
                              ┌──────────┐
                              │  Start   │
                              └──────────┘
            ┌────────────────────┘    └────────────────────┐
   ┌─────────────────┐                            ┌─────────────────┐
   │ Previous source │                            │ Previous target │
   │     files       │                            │     files       │
   └─────────────────┘                            └─────────────────┘
            ┊                                              ┊
        ▽ Conversion ▽                                ▽ Conversion ▽
            ┊                                              ┊
   ┌─────────────────┐                                     ┊
   │ Analysis (New   │                                     ┊
   │ source against  │                                     ┊
   │ old source)     │                                     ┊
   └─────────────────┘                                     ┊
            │                                              ┊
   ┌─────────────────┐                            ┌─────────────────┐
   │   Preparation   │                            │   Preparation   │
   └─────────────────┘                            └─────────────────┘
            └──────────────┐          ┌──────────────┘
                       ┌─────────────────┐
                       │   Alignment     │
                       └─────────────────┘
                               │
                       ┌─────────────────┐
                       │   Import into   │
                       │   Workbench     │
                       └─────────────────┘
                               │
                          ┌──────────┐
                          │   End    │
                          └──────────┘
```

Analysis Figures

When you have imported your new translation memory into Translator's Workbench, you will, of course, wish to analyse the files you are going to translate using the translation memory. It is always difficult to predict how successful the alignment will be since many factors, previously mentioned, will influence the results. WinAlign is a tool which is meant to give Translator's Workbench users a greater advantage the first time they use the tool. The real advantage comes, however, when translators start using Translator's Workbench *interactively*, adding translations and leveraging from their colleagues' translations and from the MultiTerm database as they work. For more information on Analysis Figures, see Section 1.3.4, **Analysis Figures and Project Planning**.

A Note on Quality

If you want to create translation memories from previously translated material, you should always assess the linguistic quality of the previous translation before commencing alignment. Translators quite rightly find it difficult to work with translation memories where the linguistic quality of 100% and fuzzy matches is poor.

Alternatives to Alignment

What alternatives are there to alignment? If you are not in a position to align files, or if there have been too many changes between the previous source files and the new source files, then you may decide to build a translation memory *during* translation rather than prior to translation. This is done simply by using Translator's Workbench while translating. All segments are saved to the translation memory. The users will get the benefit of any repetition throughout the files and they will also have consistent terminology, thanks to MultiTerm and Translator's Workbench's **Concordance** command. As was mentioned above, WinAlign is there to help users get instant benefit if they are first time users of Translator's Workbench. However, the *real* benefit is to be seen when you are using a translation memory to translate second, third or fourth generation updated files.

You may also be in a position where you do not own sufficient user licences for Translator's Workbench, but you are expected to deliver a translation memory at the end of your project. If this is the case, you have the option of creating translation memories using WinAlign after translation. However, the same issues prevail with regard to misalignment and checking of translation memories. It is unlikely that you will achieve alignment of 100% of your source text. A more realistic figure is 80%. In addition, alignment following translation means that the translators do not gain the benefit of concordance searches, 100%, fuzzy and repetition matches and it is more difficult to guarantee translation consistency.

1.3 Preparation of a Translation Project Using Translator's Workbench

1.3.1 Key Roles

There are several key roles involved in the management of a job using Computer Aided Translation (CAT) technology. Firstly, there is the role of "Project Manager". This is the person who sees the project through from start to finish, who books resources for each of the tasks involved and who draws up a schedule and budget and aims to adhere to both. Throughout the translation project, the project manager is usually responsible for the following tasks:

- Organising the training of translators in Translator's Workbench.

- Organising the receipt or creation, of translation memories and terminology databases.

- Organising the preparation of files.

- Tracking files during translation.

- Liaising between translators and support personnel.

- Delivering translated files and translation memories to the customer.

The Project Manager will often require a project administrator to help out, depending on the size of the project.

The next important role is that of "Translator's Workbench Specialist". This person (there may, of course, be more than one) is responsible for:

- Creating translation memories using the alignment tool.

- Preparing files for translation.

- Writing guidelines for the translators.

The Translator's Workbench Specialist may also be responsible for:

- Setting up translation memories.

- Creating and maintaining terminology databases.

- Cleaning files after translation and merging or updating translation memories.

Finally, support personnel will be required throughout large projects to answer technical questions from the translators. It is sometimes possible to merge the roles of Translator's Workbench Specialist with that of Technical Support Specialist. You might also consider availing of a support contract from TRADOS. Contact your local TRADOS office for information on support options. There may, of course, also be a requirement for other specialists such as software engineers, localisation, product, and DTP specialists.

1.3.2 Translator Preparation

It is recommended before any project begins that you should ensure that all translators have received training in Translator's Workbench and have ample experience with it before commencing their first translation job. Training should be scheduled a few weeks before the start date for translation and should preferably be done with files similar to those that will be translated during the translation project. TRADOS offer training programs which take approximately two days. TRADOS also provide a self-running demo. It is recommended that this be distributed to translators before they receive training so that they can familiarise themselves with the TRADOS family of tools. For more information, contact your local TRADOS vendor (information can be obtained from the TRADOS website, http://www.trados.com).

The time required for training can vary depending on the file formats your translators have to work with. For example, if your translation project involves FrameMaker or Interleaf, training may take

more time as the translators must be familiar with The S-Tagger format and with the process of verifying tags.

1.3.3 File Preparation

The project manager should always allow time for the preparation of the files to be translated. If alignment of files is necessary, then additional time will be required. Some file formats will not require much preparation, for example, Microsoft Word or Windows help files. A representative sample of these files should be tested with Translator's Workbench before translation to make sure that there are no unusual features in the files.

Other file formats, e.g. FrameMaker, require a little more preparation time. It is wise to prepare your source files well, especially if you are translating into several languages, since if your files are prepared well, you will not have to spend as much time doing DTP on the translated files. However, if you make changes to the source documents so that they are more suitable for use with Translator's Workbench, then you should alert the originator of the documents to this fact, especially if you are expecting updates. There is little point in customising the source documentation for translation if the producer of those documents continues to work on the original files. This would mean that every time you receive updates, you would have to customise the source files again.

If you have to make changes to the source documents so that they are more suitable for translation using Translator's Workbench, the originator of the documents should agree either to implement the same changes or to work on a copy of the files that you have customised when updating the files. This guarantees consistency and reduces the amount of preparation required during a cycle of updates.

Information on the preparation required for specific file formats can be found in Section 2: "How to Deal With Specific File Formats", under each sub-heading.

> **Note**
>
> You should always ensure that all translators are working with the same version of Translator's Workbench during a project.

1.3.4 Analysis Figures and Project Planning

Analysing files against translation memories is an important step during preparation. The analysis figures provide information that can be useful when scheduling translation tasks or making decisions about whether or not to use Translator's Workbench. The analysis figures may also have a bearing on your translation budget. Although there are no hard and fast rules about the relation between the analysis figures and translation speed, some general observations can be made.

100% matches

First of all, you should always alert the originator of documents that, when updating source documents, even a small alteration such as changing lowercase to uppercase or bold to italics will result in the loss of a 100% match. This has an impact on translation speed as it will reduce the number of sentences that can be batch translated.

It is difficult to state what level of 100% matches is "good" because, for example, 30% of 1 000 words is low, but 30% of 100 000 words is a significant number. Therefore, you should always take the *total* number of words into account when assessing the leveraging benefit. You should also keep in mind that translation throughput does not normally increase in proportion to the percentage leveraging. Many factors have to be taken into account here. For example:

QUESTION	NOTE
Are the translators first-time users of Translator's Workbench?	If they are, throughput figures may initially be lower than normal.
Is the batch translation feature used or is translation done in interactive mode?	Batch translation will speed up translation, but may require sample checking afterwards.
Is the file format tagged or WYSIWYG?	Tagged formats are more complex than WYSIWYG formats for translators who are not familiar with the formats and will therefore require more time for translation.

As with any new technology, you need to allow time for the users to familiarise themselves with the tools.

Once translators become used to the tool and translation memories grow in size, the translation speed will increase. Experience in the past has shown that the 100% match level has to reach approximately 70% before translation throughput can be doubled. This level of leveraging would, of course, promote the use of the **Translate/'no 100% match'** command combinations or the **Translate to Fuzzy** command. If you feel nervous about using these commands at first, you should carry out some sample tests of your files afterwards in order to reassure yourself that these features are highly dependable.

Fuzzy Matching

The categories of fuzzy matching are important. A fuzzy match in the category of 95-99% generally requires very little editing. A fuzzy match of 55%, on the other hand, might require complete re-translation. Many translators find that fuzzy matches below 60-65% are unusable. It is for that reason that the recommended default percentage for fuzzy matching is 70%. Therefore, when you are assessing the benefits of fuzzy matching, you should consider what proportion of words are contained in the higher fuzzy matching categories and what proportion are in the lower categories.

No Match and Repetition Matches

The "no match" level is, of course, also very important as it gives an indication of the number of words that will have to be fully translated. These words will normally take the same amount of time to translate as they would without a translation tool, but translators can use both MultiTerm and the concordance feature as aids. Again some external factors also need to be considered, like, experience of translators, file format and translation environment (i.e. network or not?).

The "repetition" matches are related to the no match and fuzzy match counts. A sentence that is at first classified as a no match or fuzzy match during analysis is classified as a repetition match if it reoccurs in the same file or in another file. The percentage repetition can often have a bearing on your decision to use or not to use Translator's Workbench. In general, repetition between 20% and 30% is deemed to be quite high. It is unlikely that the level of repetitions will rise above this, unless your text is extremely repetitive. You should remember to analyse all parts of the translation project *at the same time* in order to assess the percentage of repetition across all parts. If there is a substantial number of words in the repetition match category, you should consider exporting "frequent segments" for translation. For more information see Section 1.3.5, **Managing Translation Memories During a Project.**

Budget Implications

If you are a frequent user of external translators, you will have to consider what impact the use of translation tools has on the rates that are paid to and expected by external translators. When translators start using Translator's Workbench, they will soon agree that it is a tool which *aids* rather than hinders the translation process. The Workbench does away with the tedious task of cutting and pasting, for example, and ensures consistency throughout. It is especially useful for managing updates, thanks to the **Translate/'no100% match'** command combinations and **Translate to**

Fuzzy command. There is no doubt that you will have to revise the standard rates that you pay (or are paid). When doing this, you should of course take into account, the cost of the technology, including any hardware upgrades that might be necessary, and the time required for the initial learning period. Even though translation throughput will be increased, time is required for new tasks such as merging translation memories, file preparation and cleaning files.

Analysis Figures and Updates

The "Analyze" function in Translator's Workbench can be used to measure the extent of an update between different sets of source files. To do this, follow these instructions:

- Create a translation memory. Specify the source language for your files and any target language (it does not matter in this context which target language you specify as long as it is not the same as the source language).

- From the **Tools** menu, select **Analyze** and select all *old source files* to be analysed, click **Analyze**.

- When this analysis is complete, click **Clear** to remove files that have just been analysed.

- Check the **Use TM from Previous Analysis** box and select the *new source files*, click **Analyze**.

Translator's Workbench will now compare the new source files against the old source files and will show you how many 100%, fuzzy, repetition and no matches there are. This process is very useful when deciding if WinAlign should be used to create translation memories, for example. See section 1.2, **Creation of Translation Memories** for more details.

Analysis Figures and Different File Formats

It is often the case, especially in the localisation industry, that there are two or more modules with different file formats to be translated. For example, you may have to translate user documentation, on-line help and software. Frequently, on-line help files are generated from the user documentation. Using Translator's Workbench to translate one set of files first, followed by the second set of files can be very rewarding. The **Analyze** command in Translator's Workbench will allow you to assess how much text can be leveraged from one documentation set into another documentation set[2]. Simply follow the instructions outlined above under "Analysis Figures and Updates" to assess the leveraging potential between two sets of documents. In place of the *old source files* and *new source files*, read *document set 1* and *document set 2*.

---

[2]   This process is useful for the preparation stage of a project. Obviously, you can also compare updated files against a translation memory *during* a project in order to assess the exact amount of leveraging.

Tip

If you are leveraging between Word files and RTF help files, which are similar, you should translate the help files first, then write a macro that deletes all hyperlinks from the translation memory. By doing this, you can have sentences in the translation memory which do not have hyperlinks and you will get more 100% matches with the Word files as a result.

You can do this either

- by exporting units from the help translation memory, deleting hyperlinks, creating a new translation memory for documentation and importing the units, or,

- follow the same instructions as above, but instead of importing units into a new translation memory, import them into the original help translation memory, but apply new different attributes to the "new" units.

If you choose the first option, you will have one translation memory per module (i.e. one for help one for documentation). If you choose the second option, you should delete all attributes specific to help and replace them with documentation-specific attributes in the export file. When re-importing, select **Leave Unchanged** under **Existing Translation Units** in the **Import** dialog. This will result in a translation memory that has translation units from both the documentation and help (some with hyperlinks, some without).

It is also possible to adopt such a process when translating tagged formats. The tags can be deleted from the translation memory so that more 100% matches are possible between the documentation and the help.

Another Tip

Frequently in the translation business, document sets that are similar are translated *in parallel*. If you are using translation memory technology to leverage from one document set into another document set, these tasks will have to run sequentially rather than in parallel.

1.3.5 Managing Translation Memories during a project

General

It is worthwhile spending some time planning the management flow for translation memories during a project before that project begins. The management flow depends on the translation scenarios that are in operation. Different scenarios and workflows are discussed in more detail in Section 1.4, **Workflow Scenarios**.

Configuring a Translation Memory

This topic is covered in detail in the Translator's Workbench User's Guide. In general, it is recommended that you select the following system fields when configuring a translation memory: Creation User (especially when working on a LAN), Change Date, Change User and Usage Counter. This type of information allows the translators to see immediately who created a translation unit, for example, or when a translation unit was last changed. This can help enormously in controlling information. The system fields can also be used to filter information being exported from the translation memory and, therefore, are useful tools in the management of translation memories. Be aware, however, that the more fields that are selected, the more processing Translator's Workbench has to do. In addtion to the system fields you may also include descriptive information about the Translation Memory i.e. name, copyright information and a summary of the TM contents. This information is stored with the TM and is often useful at a later stage for example, when there is a significant time lapse between updates to the source material, or where the TM is passed to a new translator.

Specifying Text Fields and Attribute Fields

Again, this topic is covered in detail in the Translator's Workbench User's Guide. There are a few important issues to note here, however. You should decide on your company's general policy for managing translation memories before you begin your first job. The type of questions you ought to be asking are: do you want to create translation memories that contain information relating to more than one customer? If so, are you going to add information on each customer's products into the same translation memory?

The answers to these questions depend on the type of operation you are running. For example, if you are a member of a large translation agency, it is likely that you will have many customers with several different products. On the other hand, if you are working in a translation department of a large company, you will have a small number of customers (i.e. internal departments). In the latter case, it may be appropriate to set up translation memories containing information relating to two or more of your internal customers. In the former case, on the other hand, you should avoid building translation memories containing more than one customer's information, because these translation memories can grow quite large and, consequently, the response time of the translation memory is slower than required, a disadvantage which outweighs the advantage of increased leveraging due to large translation memories. In addition, combining translation memories from different external customers could lead to inconsistencies in terminology.

Using Attributes to Filter Information

Some customers may have different *modules* for translation. As we have already mentioned, there may be a printed documentation set and an on-line documentation set. If this is the case, then it may be useful to create one translation memory with two attributes, one for the printed documentation set and one for the on-line documentation set. When translators are leveraging from the translation memory and adding to it, they can select the appropriate attribute (go to the **Settings** menu, select the **Filter** command). By selecting a penalty for "Not Matching Attributes" (**Penalties** tab under **Options** menu, under **Translation Memory Options**), the translators will be alerted when a translation unit is proposed, which has been assigned an attribute from a different documentation set.

Creating a Smaller Translation Memory from a Master TM using Attributes

It is a good idea to limit the size of translation memories because the bigger a TM is, the slower the response time will be, in general. If you have a translation memory with many attributes, you can export all information with a particular attribute to a Text file. For example, you could export all translation units belonging to the customer "TRADOS" that also have the attribute "on-line documentation" or have been used more than once (i.e. have a Usage Counter greater than 1). That export file can then be re-imported into a new translation memory containing units with those properties only.

If you would rather store all information in one translation memory, you can of course create separate memories at the start of a project and then merge all translation memories into a "Master Translation Memory" at the end of a project. To create separate translation memories from a master translation memory, either follow the instructions above on exporting all units with a specific attribute (or attributes) or use the **Create Project TM** command (under **Analyze** in the **Tools** menu). For more information, see **Creating Project Translation Memories**.

Creating Project Translation Memories

The **Create Project TM** command allows the user to create a subset of a translation memory from a "Master" translation memory. The subset contains only that data which is relevant to the files you have analysed. It is, therefore, smaller than the Master TM and the response time will generally be quicker. To create a project TM, you should first analyse the files, (under **Analyze** in the **Tools** menu) and then click on the **Create Project TM** button in the left hand corner of the **Analyze Files** dialog box. Translator's Workbench will prompt you for the name and location of the project TM. Once you have specified a name, all relevant translation units are copied to that file.

When the translation cycle is complete, you may want to merge the project TM with the Master TM again. This will add any changed or new units to the Master TM. To do this, see **Merging Translation Memories**.

Translating Frequent Segments

If there is a high degree of repetition in the files you have analysed, it is worth considering the use of the **Export Frequent Segments** command in the **Analyze Files** dialog box. This option will export all segments that are new and occur more than a specific number of times. The number of occurrences is specified by the user. For example, you might decide to export all segments that occur more than five times. The segments are exported to a text file with the following format:

```
<TrU>
<CrU>Brockmann
<ChD>180297
<ChU>O'Brien
<UsC>7
<Seg L=EN_GB>This is a frequently used segment.
<Seg L=DE_DE>n/a
</TrU>
```

The translator inputs his or her translation after the section "<Seg L=DE_DE>" (overwriting the "n/a" in the process). When all frequent segments have been translated, the file is re-imported into the translation memory and the segments are thereby added to the memory. This guarantees a consistent translation for those frequently occurring segments and increases the percentage of 100% matches.

Translating Unknown Segments

It is also possible to translate "unknown segments" and import them into the translation memory before translation begins. Unknown segments can either be segments for which no match was found, or segments below a certain fuzzy match value. This option is useful if you are combining translation memory technology with machine translation technology (see Section 1.7, **Translation Memory and Machine Translation**). To export unknown segments, analyse the files in the usual way (see the Translator's Workbench User's Guide), then click on **Export Unknown Segments** in the **Analyze Files** dialog box and select the highest percentage match value of those segments you wish to export. For example, if you specify 65%, all segments of match value 65% and below will be exported. Then, select the export format which complies with the machine translation system you wish to use i.e. Systran or Logos. When machine translation is complete, re-import the file into the translation memory.

Merging Translation Memories

Merging of translation memories is done via the **Import** command in the **File** menu. You might have to merge translation memories if, for example, you have created a project TM at the start of a translation project and you have to merge that TM with a master translation memory at the end of the project. Alternatively, you may be working in a translation environment where you have multiple users of Translator's Workbench who are not on-site and who are not sharing translation memories on a LAN. In this type of scenario, you may have to merge translation memories at various intervals throughout the project.

Instructions for merging translation memories can be found in the Translator's Workbench User's Guide. You should take some time to examine and understand this section of the Translator's Workbench manual since the choices you make during import affect the information stored in the translation memory.

The main question you should ask is: What will happen to *existing translation units* during import? Translator's Workbench offers five options: **Leave unchanged, Keep most recent, Keep oldest, Merge** or **Overwrite**. Let us call the translation memory *into which* you are importing the "Master Translation Memory". During import, Workbench compares source sentences in the translation memory with source sentences in the master TM. If there are two identical source sentences, the *attributes* assigned to each one will be compared.

If you select **Leave Unchanged** in the **Import** dialog box, the imported translation unit with an identical source sentence will be rejected. This means that there will be no duplicate source sentences in the master TM, but it also means that you might lose translation units that have been translated differently due to different contexts.

If you select **Overwrite**, then the imported translation unit will overwrite the translation unit with the best matching attributes in the master TM. If you make this selection, you should be aware that the target segments from the imported TM might replace target segments in the master TM.

If you select **Merge**, a more comprehensive analysis is done of the *target sentences* and the *attributes*. This option will compare the target sentences of the best matching translation units. If they are identical, then the attributes and text fields of the imported translation unit are *merged* with those of the translation unit in the master translation memory. That is, no new unit is created, but the attributes and text fields are added.

On the other hand, if the target sentences are *different*, then Translator's Workbench analyses the attributes. If the attributes are the same, then the imported translation unit *overwrites* the master translation unit. If the attributes are different, a new translation unit is created, and the old one is kept.

If you select **Keep Most Recent**, Translator's Workbench compares the Creation Date of the existing and incoming translation unit and keeps the translation unit with the most recent Creation Date.

If you select **Keep Oldest**, Translator's Workbench compares the Creation Date of the existing and incoming translation unit and keeps the translation unit with the older Creation Date.

In summary, the safest option is to select **Merge**. When this option is selected, and, if all attributes in project TMs are identical to the master TM, but a source sentence is translated differently to an identical sentence in the master TM, the imported translation unit will overwrite the unit in the master TM. On the other hand, if different attributes are assigned to each project TM, and a source sentence is translated differently to an identical sentence in the master TM, both the new translation and the old translation will be stored in the translation memory. You should assess these options carefully and set up the TMs appropriately at the start of a translation project.

Batch Translation and Segmentation

To conclude this section we should mention the batch translation and segmentation options. Both have a bearing on the management of translation memories and are especially useful if you are dealing with a translation environment where you employ many freelance translators or where you do not have an adequate number of trained users or user licences for a project.

Batch translation allows you to automatically insert translations into the source files. This is done via the **Translate** command from the **Tools** menu. The user can specify what percentage match to translate. For example, by typing 85 in the **% or higher match** box, Translator's Workbench will translate all matches of 85% and above. If a MultiTerm glossary is available, the user has the option of automatically inserting found terms from the glossary into the files.

In addition, the user can also tick the **Segment Unknown Sentences** box. This will make a copy of all sentences which have not been batch translated and will place the segment markers around the original sentence and the copy so that the text will look like this:

{0>This is a segmented sentence.<}0{>This is a segmented sentence.<0}

The translator who receives a batch translated and/or segmented file should check the translated segments (and make any changes necessary) and translate segmented sentences, overwriting the copy of the segmented sentence with a translation.

When the translation is complete and the files have been returned from the freelance translators, the translation units can be imported into a translation memory using the **Update TM** command during the **Clean Up** stage.

Batch translation when combined with the 'no 100% match' commands is also a useful approach to small updates where translators are working in a Translator's Workbench environment. The translator receives files which have been translated using Translator's Workbench batch translation function as outlined above. Using the same TM the translator opens the partially translated file and from the TRADOS menu in Word selects the **Open next no 100% match Get** command. Translator's Workbench skips over all 100% matches and opens the first less than perfect match. The translator translates this segment and continues translating using the **Set Close Open Next**

no **100% match Get** command. This feature allows translators to process minor updates to large files in a more efficient manner.

Tip

If you decide to batch translate and segment sentences, it is useful to enable the **Translated Text Colours** command in the **Options** menu. Here you can specify what colour 100% matches, fuzzy matches and no matches should be. This makes life much easier for the translator. There are some recommendations in Translator's Workbench on-line help on restrictions when using this option. Please consult this beforehand. Also, it is recommended that you carry out some preliminary tests to see how your specific files are suited to the use of this option.

Another Tip

Although the batch translation feature and the **Segmentation** command are very useful, it is generally recommended that they are only used where there is no other alternative. Interactive translation is a better option as the translators gain the benefit of repetition matches in this case and consistency can be guaranteed. Consistency cannot be guaranteed to the same extent when freelancers work either on segmented files or interactively as single users. In addition, working with segmented files using a word processor does not allow the user to take full advantage of Translator's Workbench tag handling features.

Reorganising Translation Memories

Translation Memories and their associated neural network files must be "tidied up" at regular intervals. This process is called "Reorganising" and is initiated by selecting **Reorganize** in the **File** menu. You must have a minimum level of Reorg/Maintenance access rights to carry out this function. At the beginning of a translation project the person responsible for setting up the TM determines the access rights. Translator's Workbench allows you to assign four levels of access rights: SuperUser, Reorg/Maintenance, Read/Write and Read Only. The number of passwords you allocate depends on the number of users, size of the project and responsibilities of each member of the translation team. See the Translator's Workbench User's Guide for more information on TM access rights. It is recommended that translation memories are reorganised at reasonable intervals, for example, every time 500 translation units have been added or at the end of each week.

If translation memories are being shared on a LAN, all translators must close the translation memory before reorganising. One user can then open the translation memory. It is safer to make a copy of the translation memory on your hard drive when reorganising, as network crashes during the process might damage the neural network files. If you have a stable and speedy network this will not be necessary.

The larger the translation memory, and the more users there are, the more frequently the TM will require organising. For example, if you have a TM of 5 MB in size with 5 users accessing it on a daily basis on a slow LAN, you should reorganise the TM once a day.

Network stability is important when several users are accessing translation memories simultaneously. They do not want to be interrupted by network downtime, nor do you want to have corrupt translation memories as a result of a network crash. If your network has heavy traffic on it, it is recommended that the network segment where translation memories are stored and accessed is on a segment of the network where the least traffic is contained. You can set up a "bridge" to divide one segment of the network from the other.

Tip

If you are using twisted-pair cabling, migrating to a port-switched hub may be advisable under extremely heavy loadings.

Cleaning Files and Updating Translation Memories

Remember to allocate time in your schedule for the "Cleaning" of files after translation. This is a simple batch process which strips the hidden source text and all segment markers from your translated files. During the **Clean Up** process, you can also tick the **Update TM** box. This will

update the translation memory with any segments that are new or have been changed. It is a useful feature when the **Translate** and **Segment Unknown Segments** commands have been used or when a file has been edited in any way without opening each translation unit.

Note

By default Translator's Workbench keeps a backup copy of files before you run either the **Translate** or **Clean Up** function. The backup copy of the file has the same name with a *.bak extension and is stored in the same directory as the cleaned or translated file.

If you do not require a backup copy of files, simply disable this function by unchecking the **Keep backup** box in the **General** tab of the **Translation Memory Options** dialog.

It is inadvisable to make significant changes to the translated text once you have cleaned up your documents. These changes will not be contained in the translation memory. If it is necessary to make changes after **Clean Up**, you should make the same changes in the translation memory either by using the TM Database Maintenance function or by manually editing translation units using the Concordance function. See section 1.5, **Update Management** for more information on Translator's Workbench TM Database Maintenance feature.

Tip

It is not necessary to Clean STF files. STF is the tagged text format generated by The S-Tagger. The S-Tagger itself performs this function when converting files back to FrameMaker or Interleaf format. However, if changes have been made to the files without updating the translation memory, you will have to:

- Make a copy of the translated STF files

- Clean Up and Update the TM

- Convert the original translated (segmented) files using The S-Tagger.

Post-Translation

The steps following translation and clean-up of files depends on the file format:

- Microsoft Word files will require minimal DTP and can then be printed and published.

- Windows help files will require building and testing.

- Any files created in DTP packages will require conversion back to the original format (more details in Section 2, **How to Deal With Specific File Formats**).

1.3.6 Summary Checklist

The following table is a summary checklist for the issues mentioned above which you can use as a quick reference when planning translation projects with Translator's Workbench. It is sequentially ordered where possible, but the stages and decisions will depend on each user profile and on the translation environment you are working in.

Stage/Action	Topics for Consideration
Preliminary Preparation Stage	
Check prerequisites	Languages? Machine-readable files? File format? User licences? Trained users? Machine specification? Key personnel? Support personnel?
Check preliminaries	Size and life-cycle of project? Updates? Existing translation memories? Customer expectations? Terminology database? Analysis figures (new source vs. existing TM or new source vs. old source)?
Project Preparation Stage	
Assign key roles	Project manager? Project administrator? Workbench specialist? Technical support?
Train translators	Time required?
Sufficient number of user licences?	Workflow scenario?
Make policy decisions re: translation memories	Number of TMs per customer? Number of TMs per file format? Number of TMs per documentation set? TM Information? Attributes and text fields? (Filtering, merging TMs) TM access rights
Schedule time for alignment	How much preparation?
Schedule time for terminology database creation	MultiTerm format or allow time for conversion from other format?
Make workflow decisions	Interactive translation? Translate to fuzzy? Batch translation? (% match value) Automatic term replacement/ insertion? Segment unknown segments? Use translated text colour options? Merge translation memories (during and/or at end of project)?

Source file preparation	Conversion required? Changes to source files required? Consult with customer on update management.
Analysis	100% level? - Batch translation? Fuzzy level? - Translate unknown segments? No match and repetition level? - Translate frequent segments? Create project translation memories?
Schedule translation	Consult analysis figures. Refer to workflow decisions.
Draw up budget	Pricing issues - investment on the part of external resources?
During Project	
Schedule time for merging TMs	Attributes and text fields - leave unchanged, merge or overwrite?
Allow for support time	Number of personnel required?
Allow for reorganising TMs	Who and how often?
Post Translation	
Schedule time for clean up	No changes to be made after this stage? Update TM necessary?
Schedule time for conversion	Tagged text formats only.
Merge TMs with master TM?	Attributes and text fields - leave unchanged, merge or overwrite?
Archive translation memories	Deliver to customer?

1.4 Workflow Scenarios

Introduction

Firstly, a short recap on what we've already discussed:

- Section 1.1 dealt with the type of questions that should be asked before a project begins. These questions fall into the two categories of (1) preliminaries and prerequisites and (2) project-specific requirements.

- Section 1.2 discussed the process of creating translation memories, the prerequisites for successful alignment and the general workflow.

- Section 1.3 discussed what tasks and key personnel are necessary for preparing a project and what tasks are required during a project. Some of these tasks are optional. Whether or not they are undertaken by you, depends on the translation environment you work in.

Section 1.4 looks more closely at the possible workflow of different translation environments. Firstly, the scenario of "in-house translators only" is examined. This is the type of situation one might encounter in a large production firm with a translation department. Next, a translation scenario is examined where both in-house translators and external resources are used. This type of situation might be encountered in a large translation agency. The third profile is that of external resources who use Workbench or MultiTerm and, finally, we will look at a situation in which external resources (who do not use the Workbench or MultiTerm) are responsible for all translation.

There are numerous combinations of these four workflow scenarios, but it is impossible to cover every possibility. Instead, we will cover the four main scenarios and you can adapt these to your own workflow.

1.4.1 In-house Translators Only with Workbench and MultiTerm

This is probably one of the easiest environments to control and enables you, under the right circumstances, to gain maximum benefit from Translator's Workbench and MultiTerm. In order to do this, all translators should be connected to a LAN and should have an adequate machine specification (see section 1.1, **When is Translator's Workbench Suitable for a Translation Project**, for more details). One of the key roles required here (and in all other scenarios) is that of a "Translator's Workbench Specialist" who will create translation memories, prepare source files, analyse files, create project translation memories and merge and clean translation memories after the translation stage. These tasks can also be shared among the translation staff.

The following tasks may be undertaken in this type of scenario:

- Alignment of previous source files (this is an optional step).

- Creation of a "master translation memory" (again, optional if alignment has taken place).

- Preparation (and conversion/markup[3]) of new source files.

- Preparation (and conversion[4]) of terminology databases.

- Analysis of new files against translation memory.

- Translation of frequent segments (optional).

- Creation of a project TM (optional, but recommended).

- Translation with TM on LAN.

- Reorganisation of TM during translation.

---

[3] Relevant for tagged text formats.

[4] Relevant if terminology exists in a format other than MultiTerm.

- Conversion of tagged files for proof-reading[5].

- S-Tag verification (only relevant for STF file format).

- Clean Up of translated files[6].

- Update of TM[7].

- Conversion of files back to original format (e.g. FrameMaker or Interleaf).

- Merging of project TMs with master TMs (if the project TM option was used at the start).

Note

For all workflows, batch translation is an option.

How the **Translate** feature is implemented (whether it is used in conjunction with **Segmentation** or with the **'no 100% match'** commands) depends on the workflow scenario.

Translate with the **'no 100% match'** commands may be used in all workflow scenarios except External Translators without Workbench and MultiTerm. In this case the **Translate** command is combined with the **Segmentation** command.

The following diagram outlines the workflow described above.

---

[5] When translating FrameMaker or Interleaf files, for example, you will want to convert the files from STF back to their original format for proof-reading at least once during the translation stage. If any changes are required, these should be implemented in the STF files.

[6] This step is not required when converting files using The S-Tagger or when using the ITP Filter Pack (see Section 2.4, **ITP FilterPack Formats and SGML**).

[7] Relevant only if files have been edited without using Translator's Workbench.

In-House Translators Only With Workbench and MultiTerm

```
                              ┌──────────┐
                              │  Start   │
                              └────┬─────┘
         ┌─────────────────────────┼─────────────────────────┐
 ┌───────────────┐       ┌──────────────────┐       ┌──────────────────┐
 │ Terminology in│       │ Old source files │       │ Old target files │
 │ alternative   │       └──────────────────┘       └──────────────────┘
 │ format        │                    Alignment
 └───────┬───────┘
         │
      Conversion                 ┌──────────────────┐      ┌──────────────────┐
         │                       │                  │─────▶│ Translate        │
         ▼                       │   Master TM      │      │ frequent/unknown │
 ┌───────────────┐               └────────┬─────────┘      │ segments         │
 │ MultiTerm     │                        │                └──────────────────┘
 │ database      │                        ▼                    Conversion
 └───────────────┘               ┌──────────────────┐  ◀──  ┌──────────────────┐
                                  │    Analysis      │       │ New source files │
                                  └────────┬─────────┘       └──────────────────┘
                                           ▼
                                  ┌──────────────────┐
                                  │ Create project TM│
                                  └────────┬─────────┘
                                           ▼
                                  ┌──────────────────┐
                                  │ Network          │
                                  │ TM &             │  Reorganise
                                  │ Multiterm        │
                                  │ database         │
                                  └──────────────────┘
                   Translator      Translator      Translator
                                           ▼
                                  ┌──────────────────┐
                                  │ Translated files │
                                  └────────┬─────────┘
                                           ▼
                                  ┌──────────────────┐
                                  │ Clean up         │
                                  │ Update TM        │
                                  └────────┬─────────┘
                                       Conversion
                                           ▼
                                  ┌──────────────────┐
                                  │ Post translation │
                                  │ processing       │
                                  └────────┬─────────┘
                                           ▼
                                      ┌──────────┐
                                      │   End    │
                                      └──────────┘
```

1.4.2 Combination of In-house and External Translators with Workbench and MultiTerm

The second workflow describes a combination of in-house translators and external resources. In this scenario, some translators work with translation memories on a LAN while other translation memories (or copies of translation memories) are sent to external resources who copy the memories to their hard drives. When the translation cycle is complete, the TMs are returned by the external resources and are merged with the other TMs and/or the master TM.

The following tasks are relevant for workflow 2:

- Alignment of previous source files (optional).

- Creation of a "master translation memory" (optional).

- Preparation (and conversion/markup[8]) of new source files.

- Preparation (and conversion[9]) of terminology databases.

- Analysis of new files against translation memory.

- Translation of frequent segments (optional).

- Creation of a project TM (optional, but recommended).

- Delivery of TMs and MultiTerm databases to external resources.

- Translation with TMs on LAN and with TMs on external hard drives.

- Reorganisation of TM during translation.

- Conversion of files for proof-reading[10].

- S-Tag verification (STF files only).

- Clean Up of translated files[11].

- Update of TMs[12].

- Conversion of files back to original format (tagged text format only).

- Merging of project TMs with master TMs.

The following diagram illustrates this workflow.

---

[8]   Relevant for tagged text formats.

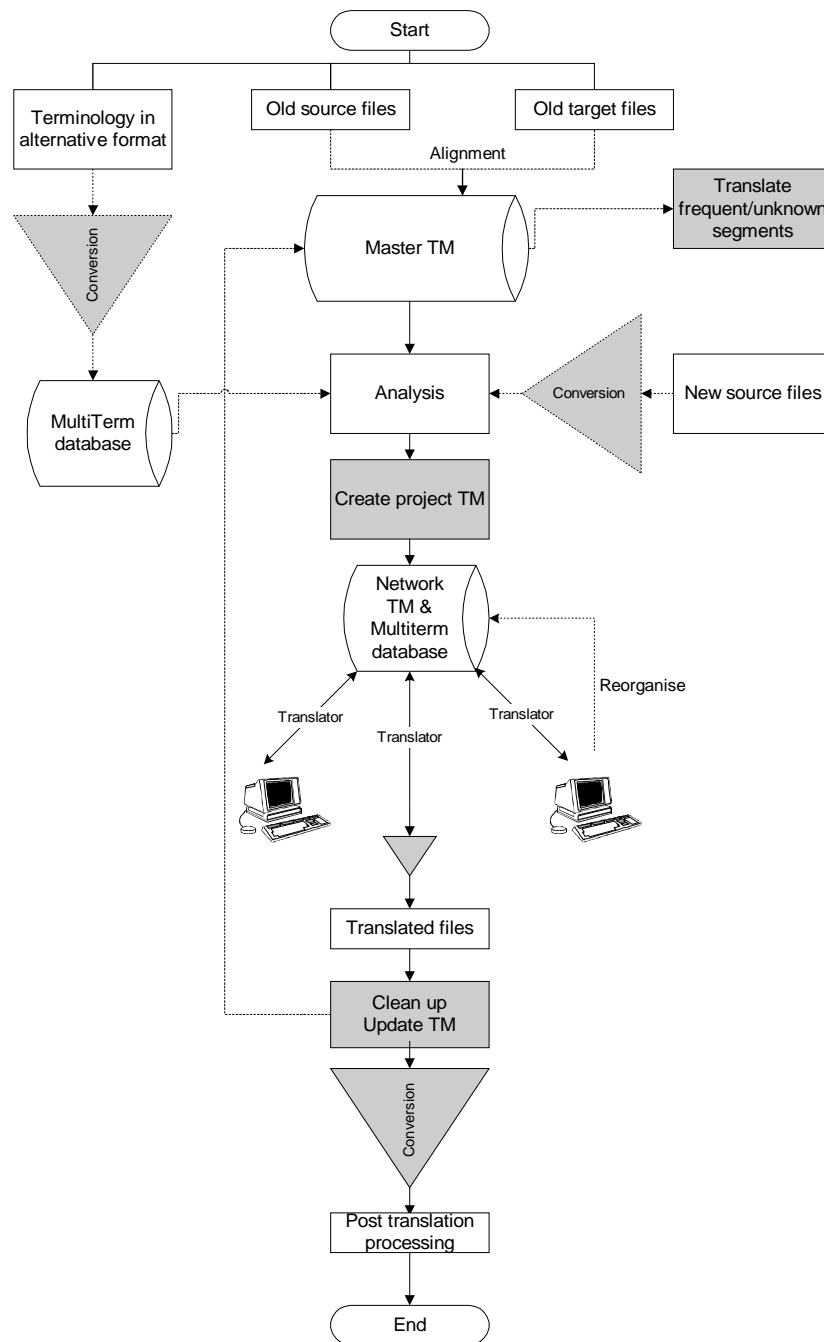[9]   Relevant if terminology exists in a format other than MultiTerm.

[10]   When translating FrameMaker or Interleaf files, for example, you will want to convert the files from STF back to their original format for proof-reading at least once during the translation stage. If any changes are required, these should be implemented in the STF files.

[11]   This step is not required when converting files using The S-Tagger or when using the ITP Filter Pack (see Section 2.4, **ITP FilterPack Formats and SGML**).

[12]   Relevant only if files have been edited without using Translator's Workbench.

In-House and External Translators with Workbench and MultiTerm

Start

Terminology in alternative format

Old source files

Old target files

Alignment

Conversion

Master TM

Translate frequent/unknown segments

MultiTerm database

Analysis

Conversion

New source files

Create project TM

Network TM & Multiterm database

Copies of TMs and MultiTerm database

Reorganise

Reorganise

Translator

Translator

Translator

External translator

External translator

External translator

Translated files

Clean up Update TM

Translated files

Conversion

Post translation processing

End

1.4.3 External Resources Only With Workbench and MultiTerm

In this case, the preparation stages (e.g. alignment, conversion of source files, analysis etc.) can either be done in-house or the external resources might undertake to do all preparation for you as well as translation.

The following tasks are relevant for Workflow 3:

- Alignment of previous source files (optional).

- Creation of a "master translation memory" (optional).

- Preparation (and conversion/markup[13]) of new source files.

- Preparation (and conversion[14]) of terminology databases.

- Analysis of new files against translation memory.

- Translation of frequent segments (optional).

- Creation of a project TM (optional, but recommended).

- Batch translation (optional).

- Delivery of TMs to external resources.

- Translation with TMs on external hard drives.

- Reorganisation of TM during translation.

- Conversion of files for proof-reading[15].

- S-Tag verification (STF files only).

- Clean Up of translated files[16].

- Update of TMs (if necessary).

- Conversion of files back to original format (tagged text format only).

- Merging of project TMs with master TMs (if project TM option used at start).

The following diagram illustrates this workflow.

---

[13]   Relevant for tagged text formats.
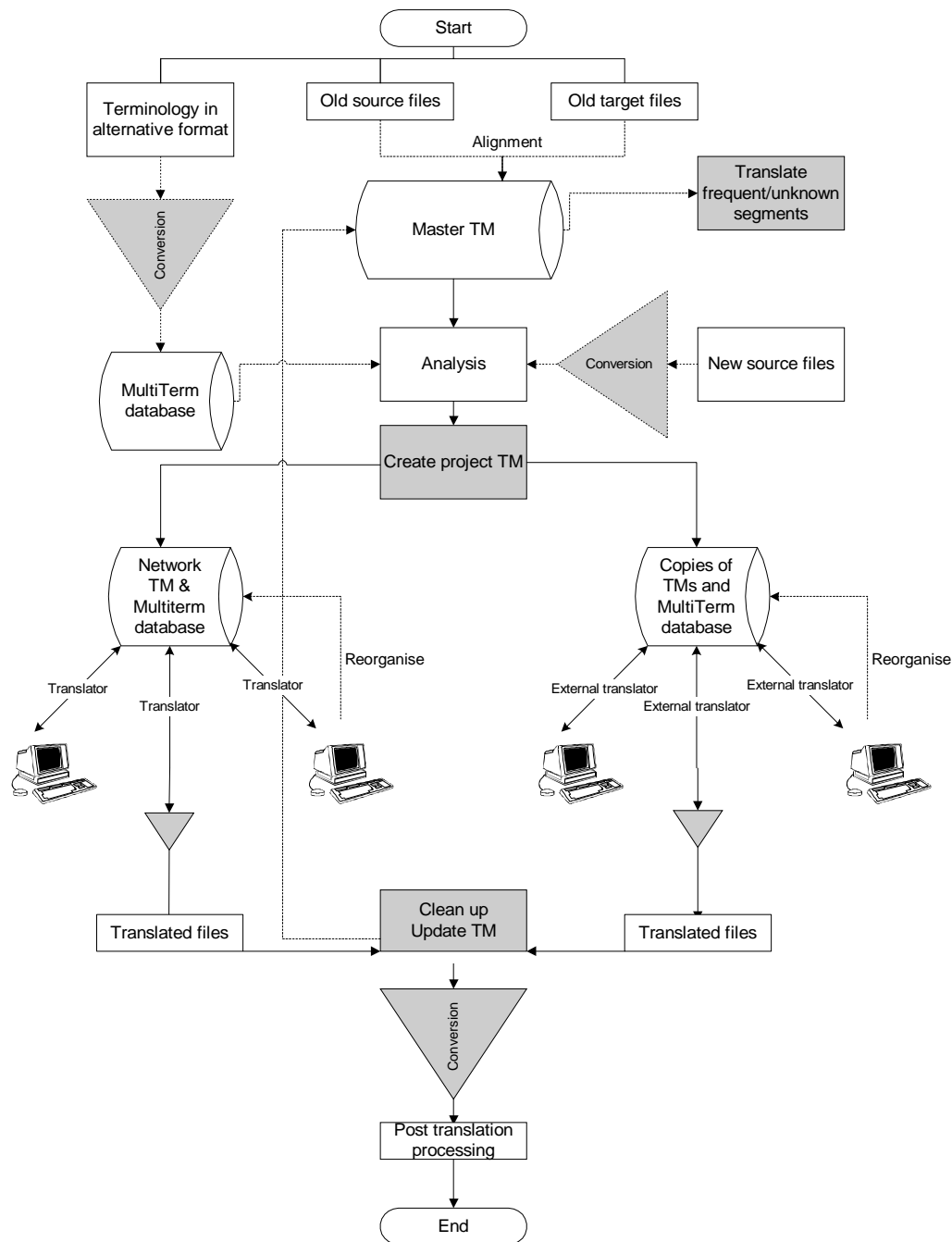
[14]   Relevant if terminology exists in a format other than MultiTerm.

[15]   When translating FrameMaker or Interleaf files, for example, you will want to convert the files from STF back to their original format for proof-reading at least once during the translation stage. If any changes are required, these should be implemented in the STF files.

[16]   This step is not required when converting files using The S-Tagger or when using the ITP Filter Pack (see Section 2.4, **ITP FilterPack Formats and SGML**).

External Translators with Workbench and MultiTerm

```
                              ┌──────────┐
                              │  Start   │
                              └────┬─────┘
        ┌──────────────────┬───────┴────────┬──────────────────┐
┌───────────────┐  ┌──────────────┐   ┌──────────────┐
│ Terminology in │  │ Old source   │   │ Old target   │
│  alternative   │  │    files     │   │    files     │
│    format      │  └──────┬───────┘   └──────┬───────┘
└───────┬───────┘         │  Alignment        │
        │                  └────────┬──────────┘
     Conversion                     │                        ┌──────────────┐
   (triangle)                 ┌─────┴──────┐                 │  Translate   │
        │                     │            │─────────────────▶│frequent/unknown│
        ▼                     │ Master TM  │                 │   segments   │
┌───────────────┐            │            │                  └──────────────┘
│  MultiTerm     │            └─────┬──────┘
│  database      │                  │              ┌──────────┐   ┌──────────────┐
└───────┬───────┘            ┌──────┴─────┐        │Conversion│◀──│ New source   │
        └──────────────────▶│  Analysis  │◀───────│(triangle)│   │    files     │
                            └──────┬─────┘        └──────────┘   └──────────────┘
                          ┌────────┴────────┐
                          │ Create project  │
                          │      TM         │
                          └────────┬────────┘
                            ┌──────┴──────┐
                            │  Network    │
                            │   TM &      │◀···············┐
                            │ Multiterm   │          Reorganise
                            │  database   │
                            └──────┬──────┘
            External translator    │    External translator
                ◀──────────────────┼──────────────────▶
                         External translator
           (computer)          │          (computer)
                          (triangle)
                        ┌───────┴────────┐
                        │ Translated files &│
                        │   translation    │
                        │    memories      │
                        └───────┬────────┘
                        ┌───────┴────────┐
                        │   Clean up     │
                        │  Update TM     │
                        └───────┬────────┘
                          Conversion
                          (triangle)
                        ┌───────┴────────┐
                        │Post translation│
                        │  processing    │
                        └───────┬────────┘
                            ┌───┴────┐
                            │  End   │
                            └────────┘
```
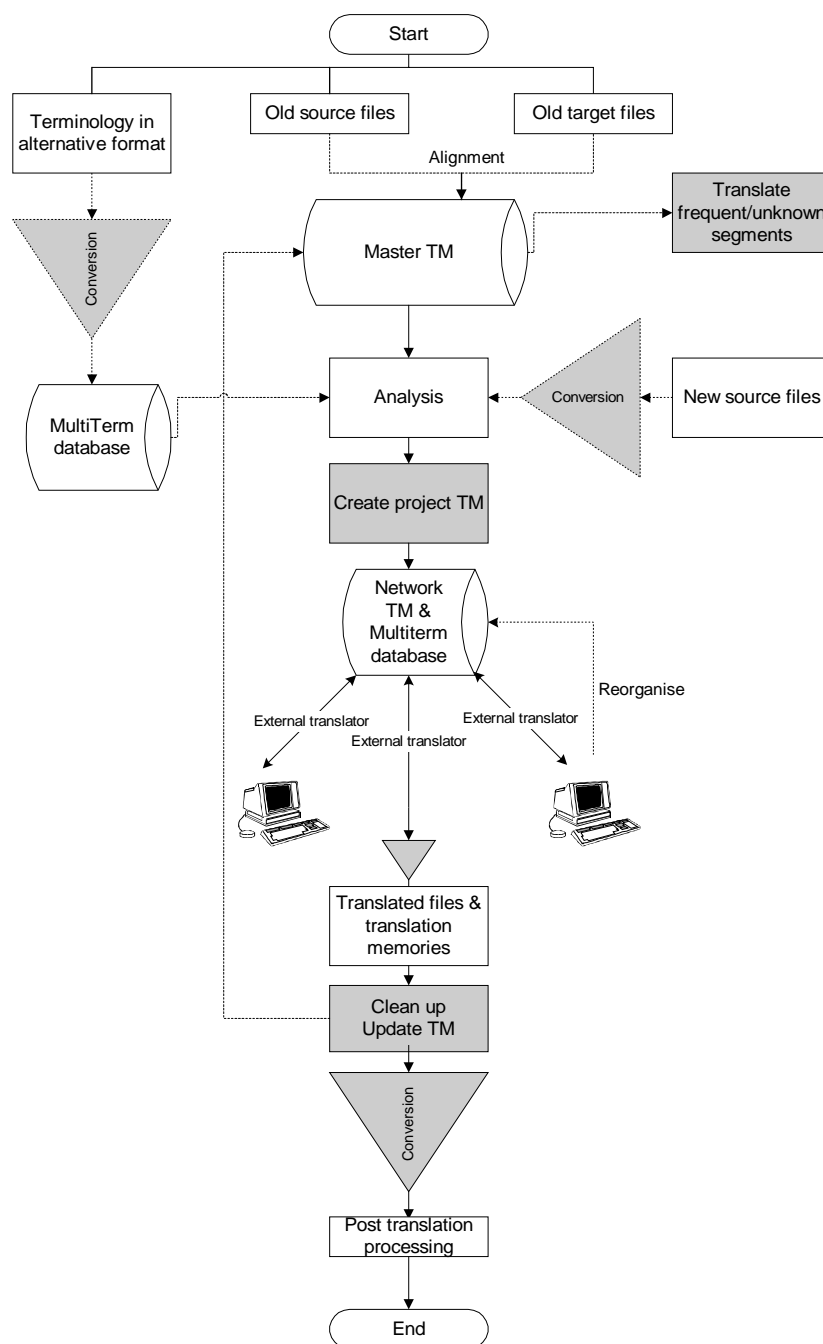
1.4.4 External Translators without Workbench and MultiTerm

The final workflow scenario is one where you have both Translator's Workbench and MultiTerm in-house, but no translation department. All translation is done externally, but your external translators do not have Translator's Workbench or MultiTerm.

The following tasks are relevant for Workflow 4:

- Alignment of previous source files (optional).

- Creation of a "master translation memory" (optional).

- Preparation (and conversion/markup[17]) of new source files.

- Preparation (and conversion/[18]) of terminology databases.

- Analysis of new files against translation memory.

- Translation of frequent segments (recommended).

- Creation of a project TM (optional, but recommended).

- Batch translation, segmentation of unknown sentences and insertion of terms.

- Delivery of semi-translated, segmented files to external resources.

- Translation with word processor.

- Conversion of files for proof-reading (external resources would require The S-Tagger for this)[19].

- S-Tag verification.

- Delivery of translated files back to you.

- Clean up of translated files and update of translation memories[20].

- Conversion of files back to original format (tagged text format only).

- Merging of TMs with master TMs (if required).

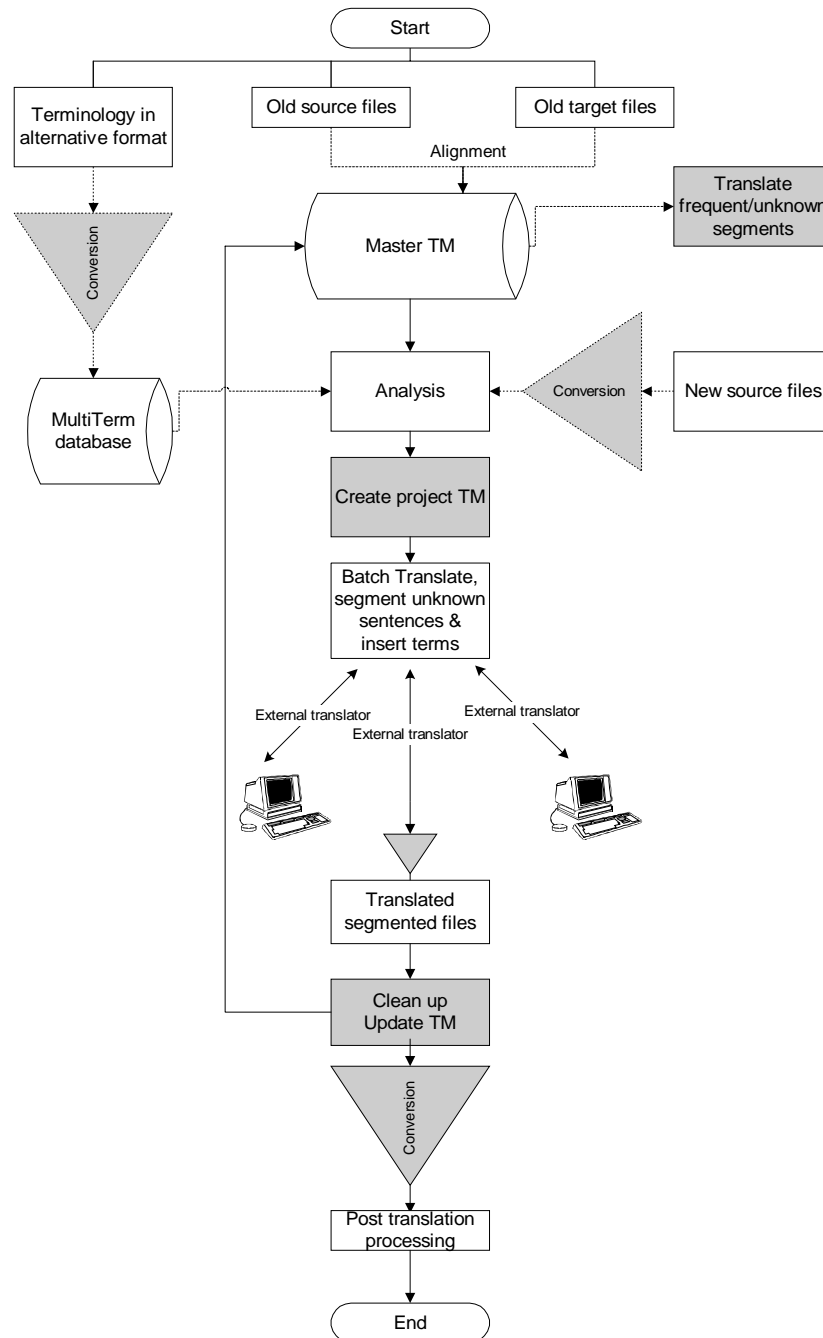The final diagram follows.

---

[17]   Relevant for tagged text formats.

[18]   Relevant if terminology exists in a format other than MultiTerm.

[19]   When translating FrameMaker or Interleaf files, for example, you will want to convert the files from STF back to their original format for proof-reading at least once during the translation stage. If any changes are required, these should be implemented in the STF files.

[20]   This Clean Up step is not required when converting files using The S-Tagger or when using the ITP Filter Pack (see Section 2.4, **ITP FilterPack Formats and SGML**).

External Translators without Workbench and MultiTerm

```
                              ┌─────────┐
                              │  Start  │
                              └─────────┘
                                   │
        ┌──────────────────────────┼──────────────────────────┐
        │                          │                          │
┌───────────────┐         ┌───────────────┐          ┌───────────────┐
│ Terminology in │         │ Old source files│          │ Old target files│
│alternative format│        └───────────────┘          └───────────────┘
└───────────────┘                    Alignment
        │
     Conversion                ┌─────────────┐         ┌────────────────┐
                               │  Master TM  │────────▶│   Translate    │
                               │             │         │frequent/unknown│
┌───────────────┐              └─────────────┘         │   segments     │
│   MultiTerm   │                    │                 └────────────────┘
│   database    │                    ▼
└───────────────┘              ┌─────────────┐    Conversion   ┌───────────────┐
                               │   Analysis  │◀──────────────  │New source files│
                               └─────────────┘                 └───────────────┘
                                     │
                               ┌─────────────┐
                               │Create project TM│
                               └─────────────┘
                                     │
                               ┌─────────────┐
                               │Batch Translate,│
                               │segment unknown │
                               │  sentences &   │
                               │  insert terms  │
                               └─────────────┘
          External translator        │        External translator
                   External translator
                               ┌─────────────┐
                               │  Translated  │
                               │segmented files│
                               └─────────────┘
                                     │
                               ┌─────────────┐
                               │  Clean up    │
                               │  Update TM   │
                               └─────────────┘
                                     │
                                 Conversion
                                     │
                               ┌─────────────┐
                               │Post translation│
                               │  processing   │
                               └─────────────┘
                                     │
                               ┌─────────┐
                               │   End   │
                               └─────────┘
```

1.5 Update Management

Translator's Workbench is a powerful update management tool. It automates the tedious parts of updates, for example, comparing old and new text and cutting and pasting translations. Updates were mentioned briefly in Section 1 under **Project Specific Issues**. Here we will elaborate on the points mentioned.

1.5.1 Assessing the Extent of the Update

The procedure for assessing the extent of an update during a translation project is straightforward. You simply analyse the updated files against the current translation memories or against the original source files.

Analysis Using Translation Memory

If you are using external resources for translation and you receive updates during a project, they will either have to return translation memories to you for analysis or you should request that they do the analysis for you.

If you are using in-house translators only, the analysis can be done in-house.

If you are using external translators who do not have Translator's Workbench, they should return copies of the translated files to you. You will have to "update" a translation memory with the sentences they have translated to date, and then you can perform the analysis. See **Cleaning Files and Updating Translation Memories** in Section 1.3.5, **Managing Translation Memories During a Project**.

Analysis Using Previous Source Files

You can also assess the extent of an update by comparing the original source files against the updated files. This option is useful if you receive an update to files which were translated previously but which have no corresponding translation memory. For instructions on how to do this, see **Comparing Old Source Files With New Source Files** in Section 1.2, **Creating Translation Memories**.

1.5.2 How to Update

Once you have assessed the *extent* of the update, it will then be necessary to decide how the updated files should be translated. The analysis results will help in making this decision (see Section 1.3.4, **Analysis Figures and Project Planning**). If you have a high level of 100% matches, you should consider using the **Translate/'no100% match'** command combinations, or the **Translate to Fuzzy** command. Alternatively, you may decide to translate the updated files interactively. Again, if there is a high percentage of 100% matches (or high fuzzy matches), it is reasonable to expect that translation throughput will be faster than usual, but this option will not be as fast as the **Translate/'no100% match'** command combinations or **Translate to Fuzzy** command. Or, after you have assessed the content of the update, you may choose the translation memory **Maintenance** feature or the translation memory **Substitutions** option.

1.5.3 The Five Update Factors

Previously it was stated that there are five factors to be taken into account when managing updates. They are: Timing, Frequency, Size, Type and File Format.

Timing

While all five factors are important, the most important factor where Translator's Workbench is concerned is the timing of updates. It is better to receive updates at an early stage in a project rather than at a later stage since, at the later stages, more text will have to be re-translated. Also, you should avoid receiving updates *after* the translated files have been cleaned and passed on to post-translation production (e.g. DTP or Engineering). Any changes made to the files after **Clean Up** will not be contained in the translation memories. The options here are either:

- Not to implement the changes in the translation memory

- To implement the changes in *both* the cleaned-up files and the translation memory using the Translation Memory **Maintenance** feature, or,

- To take one step backwards in the process and implement the changes in a copy of the uncleaned translated files and clean the files a second time and use the "Update TM" option.

> **Note**
>
> The translation memory **Maintenance** feature allows you to make global changes to your TM. You can search and replace text within translation units and you can also make changes to information stored in system, attribute and text fields. Refer to the Translator's Workbench User's Guide and on-line help for more information.

Frequency

The more updates you receive, the more work there is to do, especially in terms of preparation of source files. It is therefore important to keep the number of updates to a minimum. It is usually better to have large, infrequent updates, than small, frequent ones.

Size of Update

The size of the update will often determine the translation process. If the update is large, i.e. there is a low percentage of 100% matches and fuzzy matches, interactive translation is probably the best solution. If there is a high level of 100% and fuzzy matches, on the other hand, it is recommended that the **Translate/'no100% match'** command combinations or **Translate to Fuzzy** command are used. Sometimes the update might be so small that it is not worth the effort of retranslating entire files (even using the batch translation mode). In this case, it might be better to cut the source sentences from the updated files and paste them into the translated files. When you re-open this segment, Translator's Workbench will recognise that it is a new source sentence. By clicking on the **SetClose** button, this new segment will be added to the translation memory. In order to keep the translation memory tidy, you should delete old units from the translation memory (see **Editing and Deleting Translation Units** in Translator's Workbench User's Guide).

Type of Update

Where an update consists of a change in a variable element (for example, numeric data has been updated) you should consider using the Translator's Workbench **substitution** option. Translator's Workbench allows the user to determine whether numbers, acronyms and names are treated as variable elements i.e. placeables or as normal text. Variables do not usually change during translation. If you enable these options Translator's Workbench will automatically identify and substitute acronyms and numbers. In addition, you may also supply a list of customised variables containing elements which are specific to your own translation requirements. Where an update consists of updates to variables then you can reconfigure the **substitution** option and process the updated source files using the batch translation feature. Refer to the Translator's Workbench User's Guide and on-line help for more information on configuring the **substitution** option in your TM.

File Format

As more preparation work is required for some file formats than for others, the file format should be considered when managing updates. Files which have been created in a Desktop Publishing package, such as Interleaf or FrameMaker, have to be prepared *and* converted each time before they can be translated using Translator's Workbench. In addition, tag verification and conversion

for proof-reading represent additional steps for these types of files formats. The fewer updates during a project with these formats, the better. If updates are necessary, then they should happen preferably at the beginning of the translation stage, when no conversion or tag verification of the translated files has yet been done, rather than at the end. Word documents or RTF help files, on the other hand, do not have to be converted.

1.5.4 Updates and Preparation of Files

Some effort must be put into the preparation of files before they can be translated using Translator's Workbench. This is especially the case if the files have been created in a sophisticated Desktop Publishing package such as Interleaf or FrameMaker and have to be converted using The S-Tagger. More details about the preparation involved can be obtained in the relevant manuals for The S-Tagger and in Section 2 of this manual.

If you are expecting updates during a project, then each set of files will have to be prepared for translation. Therefore, it is of the utmost importance that you:

- Limit the number of updates.

- Obtain agreement from the file author that they will update the files you have prepared, not the files they originally sent to you.

The reason for the first point is obvious: the more updates there are, the more time must be spent assessing and preparing files for translation. Point number 2 highlights the fact that when files are prepared for translation some changes might be necessary to make those files more suitable for Translator's Workbench. It therefore makes sense for your customer to update the files you have prepared. This will speed up the translation process for them and save you a lot of time. If it is not practical for the customer to wait for you to deliver the prepared files back to them, then you can communicate the changes you have made and ask them to implement those changes themselves.

Tip

Preparation work can be reduced substantially if customers are aware of how they should (and should not) set up their files. Tips with regard to file set-up can be obtained in Section 2 below and in more detail in the relevant manuals for The S-Tagger. It is strongly recommended that you read these and communicate any relevant information to your customers.

1.6 Terminology Management

The handling of terminology by MultiTerm is covered extensively in the MultiTerm User's Guide and will therefore not be repeated here. However, terminology management is an important feature of any translation project and should be considered in as much detail as the management of translation with Translator's Workbench. Indeed, it might be the case that MultiTerm is more suitable for the translation of particular text types than Translator's Workbench.

In this section, some of the issues concerning the interaction of MultiTerm and Translator's Workbench are mentioned. Again, it is impossible to cover all possible scenarios. Instead the issues that you should consider as a manager of a translation project using computer-aided translation technology are highlighted.

1.6.1 MultiTerm and Translator's Workbench

Translator's Workbench User's Guide describes in detail the interaction between Translator's Workbench and MultiTerm. This terminology management feature greatly enhances the translator's working environment, allowing them to paste terms into their translation automatically and to look up more detailed information in the MultiTerm database, thereby ensuring terminological consistency. It is also possible for translators to edit or add terms as they work. Obviously it is important, especially when databases are being shared on a LAN, that terms are either flagged as "proposals", for example, or as "client-approved terms". A database administrator can at any stage during a project change the status of a term or add/delete terms.

1.6.2 Configuring a MultiTerm Database

One of the first issues to consider is whether or not the terminology for a translation project is available in MultiTerm format. If the answer is "no", then there are ways of converting other formats into MultiTerm format. For more detailed information see the MultiTerm User's Guide.

Another consideration is the configuration of the database. In the same way that the configuration of translation memories requires thought, the configuration of terminology databases also requires some thought. Some of the issues that need to be addressed here are:

- Are the databases to be bilingual or multi-lingual?

- How will the terminology be stored, e.g. will the terminology be stored centrally and exported to MultiTerm format when required?

- What text and attribute fields should be available in each database?

- Will there be a "standard" database definition or will each customer have a different one?

- Will translators be allowed to edit, add and delete entries or will the databases be "read only"?

- If the databases are read only, how will you control requests from translators for the update of terminology?

1.6.3 Automatic Insertion of Terms during Batch Translation

It was mentioned earlier that it is possible to automatically insert terms when batch translating files. This option only affects unknown segments and might be useful if you are using external resources for translation who have neither Translator's Workbench nor MultiTerm experience. The choices here are:

- "Don't (Replace Terms)"

- "Replace" or

- "Insert".

The first option is ticked when you do not want to replace terms. The "Replace" option means that the program will replace known terms by their translation. These terms will be formatted with the

style "tw4winTerm". If there are multiple translations for the term, the first one is selected. It is recommended that if this option is being used, the "Segment Unknown Sentences" option is activated. This ensures that the terms are only inserted into the *target* sentence, leaving the source sentence untouched. "Insert" means that known terms and their attributes are inserted as "Annotations" (Word). If multiple translations are present, they are separated by a comma.

1.6.4 Updates to Terminology

Changes in terminology during a translation project are undesirable, but sometimes necessary. The reason terminology changes are undesirable is because they are labour intensive and difficult to control.

If you are working with a central terminology database with translators on a LAN, you can simply implement the changes in the central database as an administrator. If you are using external translation resources, you will also have to regenerate local databases for each resource.

Not only will the terminology databases have to be updated, but also the translation memories and the translated files. It is undesirable to make changes to the files only but not to the translation memory as this will lead to errors and inconsistency in terminology. There are several options for updating both files and translation memories:

- Either you search through the files for terminology to be changed and replace it without opening the translation segment, or

- You open each segment where the term occurs, update the segment and click on **SetClose.**

The first option updates the files only, not the translation memory. The latter option updates the translation memory at the same time as the files.

If you choose the first option, the translation memory can be updated using the **Update TM** command as described earlier. If the number of changes to be made is relatively small, then simply opening each translation segment where the term occurs, updating and closing the segment again is probably the best option to choose. If, on the other hand, there are extensive changes to be made, then the first option will be faster.

It is also possible that terminology changes have to be made at the end of a project (after the **Clean Up** stage). The same options exist here as long as you still have a copy of the translated (segmented) files (which is recommended). If the files have already been cleaned when the terminology changes are requested, you will be faced with the choice of:

- Either taking one step back in the process and implementing the changes in the segmented files and running the **Clean Up** process again, or

- Implementing the changes in the cleaned-up files and in the translation memory separately.

The advantage of option 2 is that it will not delay the critical path of your project - you can implement corrections in the translation memory at the end of a project by using the TM **Maintenance** feature and globally changing all affected translation units in the translation memory.

1.7 Translation Memory and Machine Translation Technology

In this section we will discuss the combination of translation memory technology with machine translation technology. A discussion of the advantages and disadvantages of such an approach is beyond the scope of this manual. If you are considering this option, it is recommended that an extensive evaluation of the chosen machine translation system be carried out beforehand. Some of the issues which should be considered during such an evaluation are:

- Available Language Pairs.

- Supported File Formats.

- Level of Translation Quality Required.

- Hardware and Software Investment Costs.

- Personnel Costs.

- Available Dictionaries.

- Workflow Implications.

- Amount of time required before payback from investment.

1.7.1 Translator's Workbench and Machine Translation Tools

Translator's Workbench can be integrated with two different machine translation tools, *Logos* and *Systran PROfessional for Windows*. Unknown segments (or segments below a certain match value) are exported to the Systran or Logos file format after the **Analyze** stage. The text file is sent to the machine translation system for translation. When translation is complete, the machine translated segments may be edited (optional step) and then re-imported into the translation memory. The user will be presented with a "machine translation match" in place of a no match. The machine translation match is presented in the user's word processor in a grey window as opposed to the normal green or yellow window. In addition, it is possible (and recommended) that a machine translation penalty be set (default 15%) to alert all users that a proposal has come from machine translation and will probably require extensive editing.

1.7.2 Terminology Considerations

For the quality of machine translation to be in any way acceptable to an end user, it is necessary that the correct terminology be used. This means that customer-specific dictionaries must be available to the machine translation system. Systran and Logos require either "coding" of customer-specific dictionaries (which is a time-consuming task) or the user can convert existing MultiTerm databases into a format which can be used by the machine translation engine. This format is usually composed of the source term, the target term and the part of speech along with a specific file header. Although this solution is less time consuming than coding customer-specific dictionaries, you should be aware that the information contained in such converted dictionaries is minimal and perhaps not sufficient to prevent the machine translation system from making some grammatical mistakes.

1.7.3 Diagram

A diagram is presented below illustrating the relationship between Translator's Workbench, MultiTerm and non-interactive MT (Logos/Systran).

Translator's Workbench, MultiTerm and Machine Translation
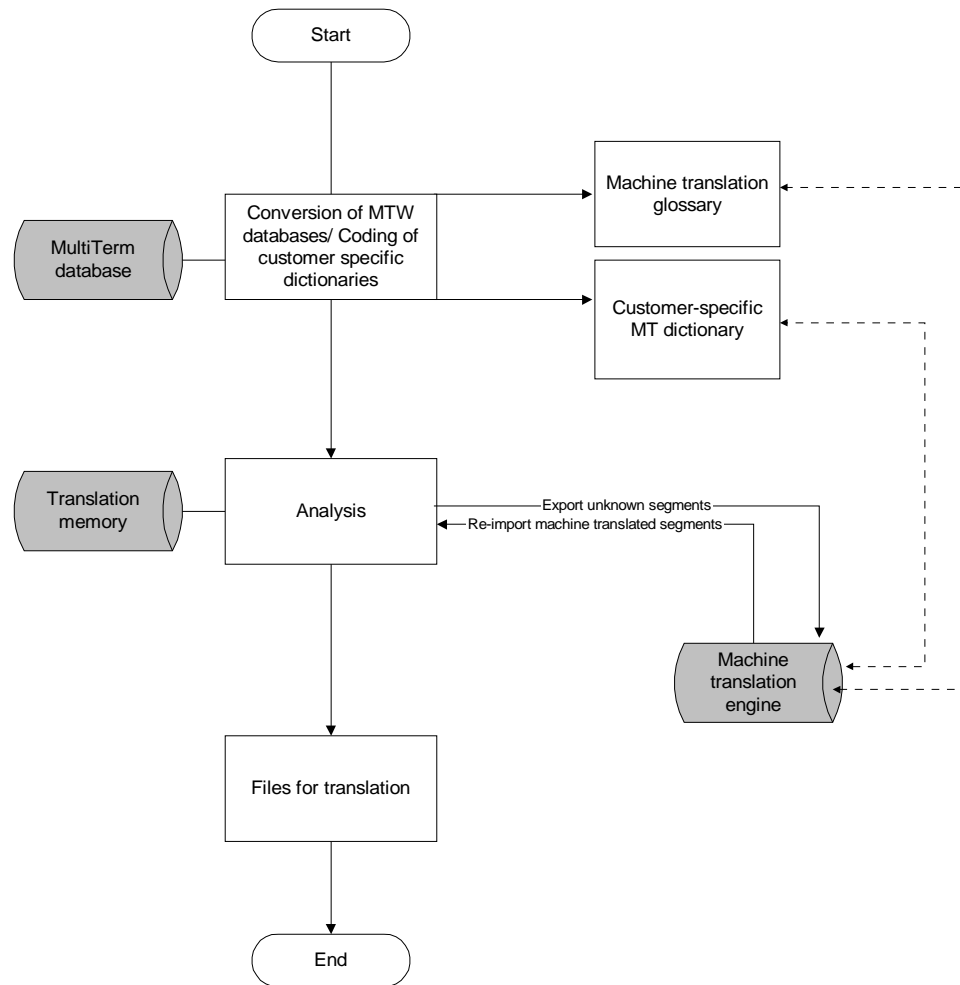
Conclusion

This concludes Section 1. We will now move on to Section 2 where file set-up, preparation and alignment issues will be addressed.

Section 2: How To Deal With Specific File Formats

Introduction to Section 2

Section 2 discusses two categories of files: WYSIWYG and tagged. The term WYSIWYG refers to "What You See Is What You Get". WYSIWYG formats are Word files and RTF based Windows help files.

A non-WYSIWYG file format would be a tagged text format. The original file format is either converted to, or marked up as a tagged format, translated and then rendered as the original format again. Tagged text formats are currently translated in Word. However, at time of press TRADOS was developing the TagEditor, a specialised application designed to facilitate the translation of tagged text formats. The tags represent information on fonts, character styles, formatting and the position of text, for example, whether it is in the "Body" of a document or in a "Table".

Translator's Workbench supports the following file types in a "tagged" format: FrameMaker, Interleaf, HTML, SGML, PageMaker, QuarkXPress, DCF BookMaster, RC, Ventura and Troff.

The S-Tagger is a TRADOS application which converts FrameMaker (MIF) and Interleaf (IASCII) files into a tagged format called STF. STF files can then be translated in Word. The ITP Filter Pack allows HTML, PageMaker, QuarkXPress and Ventura, to be "marked-up" so that they too can be translated in Word (more about this later). The SGML/RTF conversion tool marks up SGML files. Both the ITP Filter Pack and the SGML/RTF conversion tool are freely available from the TRADOS website http://www.trados.com

RC, DCF and Troff files are all marked up by means of macros which are included in the TRADOS template (tw4win.dot or tw4win97.dot). When this template is attached in Word, you can run any one of the macros by going to the **Tools** menu, selecting **Macro** and clicking on, for example, **tw4winPrepareRC** if you are preparing RC files for translation.

Translator's Workbench recognises two different types of tags: "tw4winInternal" and "tw4winExternal". The external tags remain fixed in position during translation and are ignored by Translator's Workbench. The internal tags contain information on character formatting and styles and are treated as placeables by Translator's Workbench, i.e. they can be moved around the sentence to suit the translation. When translating in Word 97 external tags are write protected by default. To deactivate this protection, choose the **Toggle Tag Protection** option from the TRADOS menu in Word 97.

If you need to know more about Workbench's tag handling feature, refer to the Translator's Workbench User's Guide as well as the on-line help and the relevant manual for The S-Tagger.

Word files and Windows RTF Help files are dealt with in some detail in this section. File preparation for FrameMaker and Interleaf files is covered extensively in the relevant manuals. For that reason, only an outline of the tasks involved are included here along with some tips for translation and alignment. The workflow for all other file formats i.e. those supported by the ITP Filter Pack and the SGML/RTF utility is also described here briefly.

Under the Heading **File Set-Up**, tips are given on what to include in your files and what not to include. It is acknowledged that project managers and translators frequently do not have much control over the set-up of files, but this information could be communicated to technical writers. It is, after all, in everybody's interest that files be optimised for translation using Translator's Workbench.

File Preparation and the Alignment Cycle gives some pointers on the preparation required before aligning certain file types. Refer to the WinAlign User's Guide for information on alignment workflows.

Under the heading **File Preparation and the Translation Cycle**, the steps required to prepare files before translation are outlined.

Finally, some tips for the translation process are given under the heading **During Translation** and each section is followed by a workflow diagram outlining the translation cycle.

Note

Please note that the information on file preparation and file set-up are recommendations only. You may find that some recommendations do not apply to your file set-up.

Support for WordPerfect 6.1

Although Translator's Workbench 2.0 (or higher) does not support WordPerfect 6.1, Translator's Workbench 1.16 continues to support this program. As both versions 1.16 and 2.0 (or higher) run on the same dongle, translators who wish to continue using WordPerfect 6.1 can translate WordPerfect projects using version 1.16 and translate Word projects using version 2.0 (or higher). Information stored in Version 1.16 and Version 2.0 TMs may be interchanged using Translator's Workbench import-export feature. Refer to the Translator's Workbench User's Guide for more information.

2.1 Microsoft Word

2.1.1 File Set-Up

The following should be taken into consideration when creating Word files destined for translation:

Size of Files

File size should always be a consideration. Remember that, when using Translator's Workbench, files grow to approximately twice their size because the original text is stored in the file until the **Clean Up** stage. If a translator's machine specification is not adequate and/or the translation memory he or she is working with is large, then the response time of the translation memory will be affected and they will be hampered instead of helped during translation. A general rule to follow would be to keep files to a reasonable size.

Graphics

Graphics should always be linked. They should never be embedded. If graphics are embedded, the file size will be greater than if they are linked. The original art files should be supplied by the customer.

Styles

Translator's Workbench automatically applies the paragraph style of the source sentence to the target sentence. However, if the base paragraph style has been changed manually (i.e. there are "local overrides"), this change may not be applied to the target sentence. Each time this occurs the translators will have to apply the style manually.

For example, if you have set up a paragraph style called "Heading 1", which has the characteristics: Arial Font, 16 point, Space before = 12 pt, Space after = 3 pt, and then you decide to apply "bold" formatting manually to a heading with this style, Translator's Workbench will apply the base style (as described above), but may not apply the bold formatting in the target sentence. If bold formatting is required, the translator will have to apply it manually. By applying styles manually, therefore, you create more work for the translator.

As a rule, styles should only be altered using Word's **Style** command, and should not be altered manually.

Revision Marks

If there are revision marks (Word 7) or tracked changes (Word 97) in your Word files, you should remove them (by accepting them) and deactivate revision mode (Word 7) or the track changes feature (Word 97) before commencing translation using Translator's Workbench. Always consult with the originator of the document before accepting or rejecting revision marks.

Index Markers

Translator's Workbench treats index markers embedded in sentences as sub-segments, i.e. they are opened separately within a translation unit and are stored separately in the translation memory. If the index marker is a leading index marker, on the other hand, it will be treated as an independent segment. If you embed several index markers in a sentence, each one will have to be opened, translated and placed individually by the translator. If you use leading index markers, they can be opened automatically using the Translator's Workbench **SetCloseOpenGet** command. The latter option is speedier and more convenient for the translator.

Field Codes

Translator's Workbench treats field codes as placeables, i.e. items which are not to be translated, but which have to be *placed* in the target sentence nonetheless. When Translator's Workbench is in use, it is recommended that the user works with field codes on view (select **Options** from Word's **Tools** menu, Click on the **View** tab, check **Field Codes**). However, since field codes are treated as placeables, you should not place any text for translation within field codes.

Coloured Text

In Section 1, the **Translated Text Colours** command was mentioned. This option allows you to specify different text colours for 100% matches, fuzzy matches etc. and is very useful for the translator, especially if the **Translate** and the **Segment Unknown Sentences** commands are used. However, colours are sometimes used in source files to communicate information (e.g. that a term or phrase only applies to a specific context). Therefore, if you intend to use the **Translated Text Colours** command, you should first decide if this will interfere with the use of colour in your source files.

Unused Styles and Fonts

Every Word document has a corresponding style sheet which contains all the paragraph and character styles defined in the Word document. Often style sheets contain many styles that are not used. Translator's Workbench allows you to strip out any unused styles from a document or a series of documents. First, enable the style removal by checking the **Strip Unused Fonts** box in the **General** tab of the **Translation Memory Options** dialog. Then run the **Clean Up** function. Stripping unused styles from your documents will help optimise the response time between Translator's Workbench and Word. Similarly, many Word documents contain unused fonts in the RTF header. You can remove unnecessary fonts by running the **Clean Up** function. By default the **Strip Unused Fonts** option is checked in the **General** tab of the **Translation Memory Options** dialog. For more information on stripping unused fonts and styles please refer to the Translator's Workbench on-line help.

Tip

You can run the **Clean Up** function to strip unused styles and fonts from Word documents before you start translating in Translator's Workbench. Always consult with the originator of the document before altering the style sheet.

2.1.2 File Preparation and the Alignment Cycle

The WinAlign User's Guide provides a detailed description of the preparation required for the alignment of Word files. The main steps are summarised below:

1. Open source and target files and check them visually for any unusual features.

2. Delete generated Table of Contents and Indices.

3. Re-sort sorted lists so that the target list matches the sequence of the source list.

4. Accept all revisions/changes (if any) in the files. Turn off revision mode (Word 7)/track changes (Word 97).

5. Save the files in RTF format.

6. Create an alignment project and link the file pairs.

7. Identify paragraph styles that can be used for WinAlign's structure recongition. If no styles are used try Word's AutoFormat feature to automatically apply styles to documents.

8. Adjust the alignment weightings and set the structure recognition levels.

9. Align the files.

10. Review the alignment. Make changes where required.

11. Export the alignment results. Import the alignment results into a new or existing translation memory.

12. Use the **Concordance** command in Translator's Workbench to search for known phrases and check that characters look OK. If you are aligning languages which use different fonts, make sure that you have installed the correct fonts on your system and have mapped these fonts in Translator's Workbench using the **Font Translations** command in the **Translation Memory Setup** dialog box.

13. Run an analysis using the translation memory you have created and the original source documentation. This will give you an indication of how successful the alignment has been.

14. Open a selection of files in Word and open some segments to make sure that the fonts look OK when units are proposed by the TM.

2.1.3 File Preparation and the Translation Cycle

When Word files are received for translation, you should check and implement the following:

1. Open all files and check them visually for any unusual features.

2. Accept all revisions, if this has not already been done. Turn off revision mode (Word 7)/track changes (Word 97).

3. If there are any files which are unusually large, you should consider dividing them up before translation and merging them again after translation. Before doing this, make sure you have any custom templates attached. Do not paste sections into new files (you will lose custom settings by doing this). Instead, make several copies of the file, and cut out text piece by piece leaving only that section you want translated in the file.

4. Ensure that no styles have manual overrides. If they do, consult with the document designer and make changes to the styles using Word's **Style** command.

5. Check the number of styles and fonts in the document. If necessary run **Clean Up** and remove all unused styles and fonts before starting translation.

6. Link and remove graphics, if this has not already been done. The customer should supply the original art files to you.

7. Turn all Field Codes on (select **Options** from Word's **Tools** menu, click on the **View** tab, check **Field Codes**).

8. If there are too many colours in use in the documents, and you want to use the **Translated Text Colours** feature in Translator's Workbench, you should test the use of the **Translated Text Colours** command and, if this interferes with the original text colours, consult with the technical writer to see if any colours can be removed from the original source files.

9. Open a selection of files and "mock" translate sections using a test translation memory. This will highlight any potential problems with file set-up.

10. Write instructions for the translators, highlighting any unusual features in the files and how they should be dealt with.

2.1.4 During Translation

Apart from the obvious advantages of using Translator's Workbench and MultiTerm, there is little difference between translating files in Word only and translating files in Word using Translator's Workbench and MultiTerm. Below are a few hints for the translation stage:

- Always work with hidden text turned on.

- Always use "Normal View" in Word.

- When you want to spell-check, you will only want Word to check the translated segments. You should therefore turn off the hidden text for the spell-checking stage.

- Similarly, if you want to proof-read on hard copy, you will only want to print your translated sentences, not the source sentences. You should make sure that hidden text is NOT specified as one of the printing characters (select **Options** from Word's **Tools** menu, click on the **Print** tab and, under **Include with Document**, make sure that the **Hidden Text** box is unchecked). You will now be able to print a document which displays only translated segments. Please note, however, that if there are index markers in your files, these will not appear in the print-out because they are formatted as "hidden". This also applies to any other "hidden" text in the translated segments.

- When implementing corrections, you can either choose to open each segment, edit it and use **SetClose** to update both the segment and the TM, or you can implement the correction without opening the segment or using **SetClose**. If you choose to do the latter, then you must update the TM afterwards. To do this:

  1. Make a copy of the file(s).

  2. Open the translation memory.

  3. Go to the **Tools** menu and select **Clean Up**.

  4. Add the file(s) to be cleaned.

  5. Check **Update TM**.

  6. Click on **Clean Up**. Your changes will now be implemented in the TM.

- The final step before files are handed over for printing or DTP tidy up, is to "Clean" the files. The instructions above illustrate how to clean the files. If all changes have been implemented in the TM already, then there is no need to check the **Update TM** box.

Tip

By default Translator's Workbench keeps a backup copy of files before you run either the **Translate** or **Clean Up** function. The backup copy of the file has the same name with a *.bak extension and is stored in the same directory as the cleaned or translated file.

If you do not require a backup copy of files, simply disable this function by unchecking the keep backup box in the **General** tab of the **Translation Memory Options** dialog.

Please refer to the following Workflow diagram, which outlines the translation workflow for MS Word and the workflow diagram in Section 2.2 outlining the alignment workflow for both Microsoft Word and RTF based Windows help files.

Preparation and Translation of Microsoft Word Files

```
                            ┌──────────────┐
                           (     Start      )
                            └──────────────┘
                            ┌──────────────┐
                            │ Check files  │
                            │  visually    │
                            └──────────────┘
                            ┌──────────────┐
                            │ Accept all   │
                            │ revisions/   │
                            │changes (if any)│
                            └──────────────┘
                            ┌──────────────┐
                            │ Assess size  │
                            │   of files   │
                            └──────────────┘
                            ┌──────────────┐
                            │Remove manual │
                            │  overrides   │
                            └──────────────┘
                            ┌──────────────┐
                            │Clean up unused│
                            │styles and fonts│
                            └──────────────┘
                            ┌──────────────┐
                            │ Link graphics│
                            └──────────────┘
                            ┌──────────────┐
                            │Assess use of │
                            │colour in text│
                            └──────────────┘
                            ┌──────────────┐
                            │Turn field    │
                            │  codes on    │
                            └──────────────┘
                            ┌──────────────┐
                            │   Analysis   │
                            └──────────────┘
                            ┌──────────────┐
                            │Mock translate│
                            │& write       │
                            │instructions  │
                            │for translators│
                            └──────────────┘
                            ┌──────────────┐
                            │ Translation  │
                            └──────────────┘
                            ┌──────────────┐
                            │Clean up and  │
                            │   update     │
                            └──────────────┘
                            ┌──────────────┐
                           (      End       )
                            └──────────────┘
```
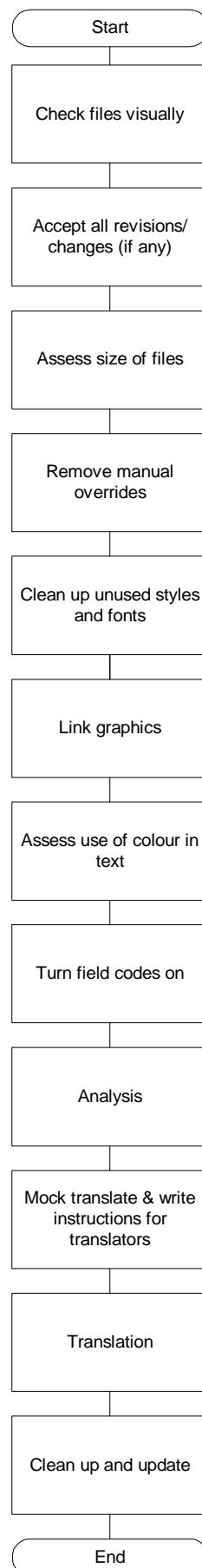
2.2 RTF Based Windows Help Files

2.2.1 File Set-Up

Many of the issues pertaining to the set-up of Word files are also relevant for RTF help files. Please refer to the following headings in section 2.1, **Microsoft Word**:

- Size of files

- Revision marks/tracked changes

- Coloured text

- Unused styles and fonts

In addition to the above, the following points should be noted for RTF based Windows help files file set-up:

Jumps and Popups

You should never apply a paragraph style containing the property "double underline" where there is also hidden text in the paragraph. This leads to problems with recognition of the context ID as a placeable by Translator's Workbench. The context ID should be formatted as "hidden" only.

Note

When translating RTF based Windows help files using Translator's Workbench, the styles "tw4win Jump" and "tw4winPopup" are applied to jumps and popups. This does NOT affect the functionality of the help file in any way.

Footnotes

Footnotes are one of the key characteristics of RTF based Windows help files. Translator's Workbench has two methods of handling footnotes during the translation process depending on whether you are translating in Word 6/Word 7 or Word 97.

Word 6/Word 7 Footnotes

It is preferable not to have long strings of text attached to keyword footnotes. Translator's Workbench will treat these as single segments. For example, if you have a footnote such as the following:

^{K} Trados Translator's Workbench; Workbench; Translator's Workbench

Workbench will open this as *one* segment. No 100% match will be found unless *exactly* the same segment occurs in the translation memory. Even if the phrases "Trados Translator's Workbench", "Workbench" and "Translator's Workbench" occur individually in the translation memory, there will be no 100% match.

If, on the other hand, there is a separate Keyword footnote for *each* of these phrases, then any time the phrase occurs, it will be suggested as a 100% match. Please note that having separate Keyword footnotes does not affect the functionality of the Help files if they are joined again at the final stage.

Word 97 Footnotes

If you are translating using **Word 97** Translator's Workbench creates a temporary document (TW4WinFootnote.doc) in which you translate all footnotes. In addition, for keyword footnotes, if you check the **Multi-segment Footnotes** box in the **General** tab of the **Translation Memory Options** dialog, Translator's Workbench treats each phrase occurring between semi colons as a

separate segment. Footnotes are automatically integrated into the translated RTF file during the translation phase. To activate this option you must reconfigure your segmentation rules. Please refer to the Translator's Workbench User's Guide for more information.

2.2.2 File Preparation and the Alignment Cycle

The steps for alignment of RTF based Windows help files are as follows:

1. Open source and target files and check them visually for any unusual features.

2. Re-sort sorted lists so that the target list matches the sequence of the source list.

3. Accept all revisions (if any) in the files.

4. Make sure the files have been saved in RTF format (sometimes files might have the extension .rtf, but they are saved as Word .doc files).

5. Create an alignment project and link the file pairs.

6. Identify paragraph styles that can be used for WinAlign's structure recongition. If no styles are used try Word's **AutoFormat** feature to automatically apply styles to documents.

7. Adjust the alignment weightings and set the structure recognition levels.

8. Align the files.

9. Review the alignment. Make changes where required.

10. Export the alignment results. Import the alignment results into a new or existing translation memory.

11. Use the **Concordance** command in Translator's Workbench to search for known phrases and check that characters look OK. If you are aligning languages which use different fonts, make sure that you have installed the correct fonts on your system and have mapped these fonts in Translator's Workbench using the **Font Translations** command in the **Translation Memory Setup** dialog box.

12. Run an analysis using the translation memory you have created and the original source documentation. This will give you an indication of how successful the alignment has been.

13. Open a selection of files in Word and open some segments to make sure that the fonts look OK when units are proposed by the TM.

2.2.3 File Preparation and the Translation Cycle

Again, some stages in the preparation stage for RTF based Windows help files are similar to the preparation required for Word documents. There are a few additional steps too:

1.   Open all files and check them visually for any unusual features.

2.   Accept all revisions, if this has not already been done. You should check with the originator of the document before you do this.

3.   If there are any files which are unusually large, you should consider dividing them up before translation and merging them again after translation. Before doing this, make sure you have any custom templates attached. Do not paste sections into new files (you will lose custom settings by doing this). Instead, make several copies of the file, and cut out text piece by piece leaving only that section you want translated in the file.

4.   Ensure that no styles have been changed manually. If they have, consult with the technical writer and make changes to the styles using Word's **Style** command. Again, this is especially important for jumps and popups.

5.   Check the number of styles and fonts in the document. If necessary run **Clean Up** and remove all unused styles and fonts before starting translation.

6.   Mark all bitmaps (e.g. {bmc bitmap.bmp}) as "tw4winInternal" style. This will prevent them from being opened as translation units by Translator's Workbench. For instructions on how to do this, refer to Translator's Workbench User's Guide.

7.   If there are colours in use in the source documents, and you want to use the **Translated Text Colours** command in Translator's Workbench, you should test the use of the **Translated Text Colours** command and, if this interferes with the original text colour, consult with the technical writer to see if any colours can be removed from the original source files.

8.   Open a selection of files and "mock" translate sections using a test translation memory. This will highlight any of the potential problems with file set-up etc. You should pay special attention to footnotes and the formatting of jumps and popups.

9.   Write detailed instructions for the translators, highlighting any unusual features in the files and how they should be dealt with.

2.2.4 During Translation

The tips given previously about the translation of Word files, also apply to RTF based Windows help files (see below). In addition, the translator should pay particular attention to jumps and popups. Hyperlinks should always be recognised as a placeable by Translator's Workbench.

Translator's Workbench should always treat keyword and title footnotes as separate segments. When the cursor is placed beside a footnote symbol, e.g. $, and **OpenGet** is clicked, Translator's Workbench will open a footnote window. The translator will translate the text and then click on the **SetCloseOpenGet** button to move back into the main body of text or on to the next footnote symbol.

In addition, the following tips apply:

- Always work with hidden text turned on.

- Always use "Normal View" in Word.

- When you want to spell-check, you will only want Word to check the translated segments. You should therefore turn off the hidden text for the spell-checking stage.

- Similarly, if you want to proof-read on hard copy, you will only want to print your translated sentences, not the source sentences. You should make sure that hidden text is NOT specified as one of the printing characters (Go to the **Tools** menu, Select the **Options** command, click on the **Print** tab and, under **Include with Document**, make sure the **Hidden Text** box is

unchecked). You will now be able to print a document which includes only translated segments.

- When implementing corrections, you can either choose to open each segment, edit it and use **SetClose** to update both the segment and the TM, or you can implement the correction without opening the segment or using **SetClose**. If you choose to do the latter, then you must update the TM afterwards. To do this:

  1. Make a copy of the file(s).

  2. Open the translation memory.

  3. Select **Clean Up** from the **Tools** menu.

  4. Add the file(s) to be cleaned.

  5. Check **Update TM**.

  6. Click on **Clean Up**. Your changes will now be implemented in the TM.

- The final step before files are handed over for printing or DTP tidy up, is to clean the files. The instructions above illustrate how to clean the files. If all changes have been implemented in the TM already, then there is no need to check the **Update TM** box.
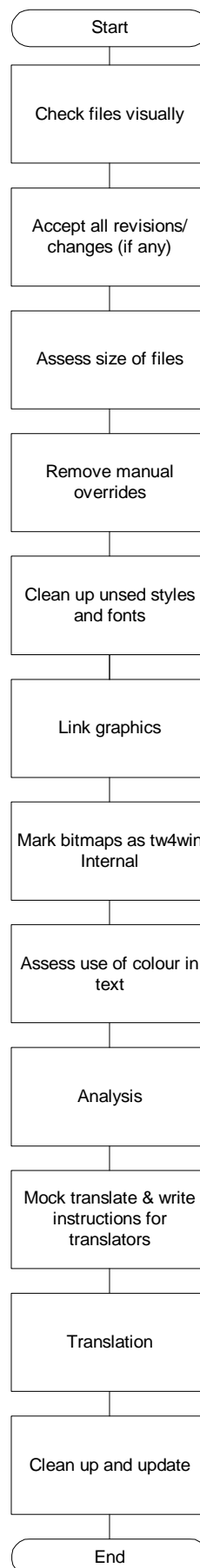
2.2.4.1 Translation of RoboHelp Files

If you have RoboHelp installed, and you are translating a RoboHelp DOC or RTF file using Translator's Workbench while the RoboHelp toolbar is active, there may be some conflicts between the RoboHelp and Workbench environments. The following workflow is therefore recommended for RoboHelp files:

1. Open the file for translation in Word.

2. Detach the RoboHelp template (robohelp.dot). For Word 7, select **Templates** from the **File** menu. For Word 97, select **Templates and Add-ins** from the **Tools** menu. If you do not specify a new template, Word automatically applies the normal.dot.

3. Translate the file using Translator's Workbench.

4. Run the clean up process checking **Update TM** if required.

5. Re-apply the RoboHelp template.

6. Compile the help system in RoboHelp.

7. Implement any changes required in the translated files **before** the clean up process.

8. Repeat steps 4-7 as necessary.

The following diagram illustrates the tasks involved in the translation of RTF based Windows help files.

Preparation and Translation of RTF Based Windows Help Files

```
                    ╭──────────────────╮
                    │       Start        │
                    ╰──────────────────╯
                             │
                    ┌──────────────────┐
                    │ Check files visually │
                    └──────────────────┘
                             │
                    ┌──────────────────┐
                    │  Accept all revisions/ │
                    │   changes (if any)   │
                    └──────────────────┘
                             │
                    ┌──────────────────┐
                    │  Assess size of files │
                    └──────────────────┘
                             │
                    ┌──────────────────┐
                    │   Remove manual    │
                    │     overrides      │
                    └──────────────────┘
                             │
                    ┌──────────────────┐
                    │ Clean up unsed styles │
                    │     and fonts      │
                    └──────────────────┘
                             │
                    ┌──────────────────┐
                    │    Link graphics    │
                    └──────────────────┘
                             │
                    ┌──────────────────┐
                    │ Mark bitmaps as tw4win │
                    │      Internal      │
                    └──────────────────┘
                             │
                    ┌──────────────────┐
                    │ Assess use of colour in │
                    │        text        │
                    └──────────────────┘
                             │
                    ┌──────────────────┐
                    │      Analysis      │
                    └──────────────────┘
                             │
                    ┌──────────────────┐
                    │ Mock translate & write │
                    │   instructions for   │
                    │     translators     │
                    └──────────────────┘
                             │
                    ┌──────────────────┐
                    │    Translation     │
                    └──────────────────┘
                             │
                    ┌──────────────────┐
                    │ Clean up and update │
                    └──────────────────┘
                             │
                    ╭──────────────────╮
                    │        End         │
                    ╰──────────────────╯
```

2.3 S-Tagger Files (STF)

The S-Tagger is a conversion solution which allows FrameMaker MIF files and Interleaf ASCII (IASCII) files to be converted into a tagged text format, STF, which can be translated in Word using Translator's Workbench. At time of press TRADOS was developing the TagEditor, a specialised application designed to facilitate the translation of tagged text formats. Although The S-Tagger 2.0 for FrameMaker is used to convert FrameMaker MIF files and The S-Tagger 2.0 for Interleaf converts Interleaf ASCII files, the principles are the same. For this reason we examine both FrameMaker and Interleaf formats in this section.

The S-Tagger 2.0 for FrameMaker converts FrameMaker MIF files into STF. The following versions of FrameMaker are supported: FrameMaker 4.x, 5.x (including 5.5), FrameBuilder 1.x and FrameMaker + SGML 5.x.

The S-Tagger 2.0 for Interleaf converts Interleaf ASCII (IASCII) files into STF. Versions of Interleaf 5.2 or higher are supported.

Translator's Workbench treats STF tags as either internal or external tags with the style "tw4winInternal" or "tw4winExternal". The external tags are left in position during translation. They do not appear in Workbench's Edit Window nor in the translation memory. The internal tags are treated as placeables and can be moved around in the sentences to fit the translation. If you are translating in Word 97, external tags and segment delimiters are write protected by default.

File set-up and the translation preparation process are covered extensively in The S-Tagger User's Guides and will, therefore, not be covered in detail here. The main steps involved are summarised below. If you need to know more about these steps, please refer to the manuals for The S-Tagger.

2.3.1 File Set-Up and Preparation

FrameMaker

1. File Preparation: Please use the following as a check list and refer to The S-Tagger 2.0 for FrameMaker User's Guide for more information:

   ✓ Turn off all change bars.

   ✓ Delete all hard returns which have been inserted to improve the appearance of a paragraph.

   ✓ Ensure that all anchored frames which contain text boxes, or text strings are positioned at the beginning or end of a paragraph of text.

   ✓ Set any conditional text styles (or formats) which should be translated to "Show".

   ✓ Update cross-references.

   ✓ Break any markers which contain translatable text of more than 200 characters into two markers.

   ✓ Ensure all markers are collapsed if using IXgen to index your documents.

   ✓ If your document contains text insets for translation, you should copy them into the document rather than referencing them. Where possible reference a document by inserting an OLE object rather than a text inset. If your text inset needs to remain referenced to other external documents which will be updated or translated separately you will need to give your translators special instructions.

2. Save the files as MIF.

3. Open the MIF files again in FrameMaker and verify that no MIF errors are reported.

4. Convert the files to STF (RTF format) using The S-Tagger 2.0 for FrameMaker.

5.   Create the ancillary file when prompted to do so at the end of the conversion process.

6.   Save any errors reported in the Results Window to a log file, and assess their implications for the MIF files.

7.   If necessary, make any changes in the source FrameMaker files and reconvert.

8.   Using a copy of the STF files generated, make a small change to the text (but not to the tags) and run the tag verification process and then, convert the files back to MIF, using The S-Tagger 2.0 for FrameMaker.

9.   If any errors are reported during conversion, isolate them, and make the necessary changes to the FrameMaker files. Repeat steps 2 through 8.

10. Open the newly created MIF files in FrameMaker and examine the results. This is similar to what your files will look like when they have been converted after translation. Make any changes to the formatting that you deem necessary, save the files and repeat steps 2 through 9.

11. You are now ready to commence translation of the STF files.

Interleaf

1.   File Preparation: Please use the following as a checklist and refer to The S-Tagger 2.0 for Interleaf User's Guide for more information:

   ✓   Open each Interleaf file. Make sure that the correct fonts are installed on your system and that they are recognised by Interleaf.

   ✓   Ensure that any relevant control expressions are set.

   ✓   Turn off revision tracking and ensure that there are no revision markers in the files.

   ✓   Check for "NOTAG" references and update each one.

   ✓   Check that all graphics and art files are contained within anchored frames.

   ✓   Turn off automatic hyphenation.

   ✓   Look at each text string in the document. If a sentence is broken into two or more text strings, consider replacing the text strings with a single microdocument.

   ✓   Check the microdocuments. If any microdocument is of a fixed width, decide if you should change it or not.

   ✓   Scan through the document for any instance of a cosmetic hard return. Delete these where appropriate (see The S-Tagger 2.0 for Interleaf User's Guide).

2.   Save the files as IASCII.

3.   Open each IASCII file again in Interleaf and verify that no errors are reported in the Opening/Saving Messages Window

4.   Convert the files to STF (RTF) using The S-Tagger 2.0 for Interleaf.

5.   Create the ancillary file when prompted to do so at the end of the conversion process.

6.   Using a copy of the STF files generated, make a small change to the text (but not to the tags) and then convert the files back to IASCII, using The S-Tagger 2.0 for Interleaf.

7.   If any errors are reported, isolate them, and make the necessary changes to the source Interleaf files. Repeat steps 2 through 8.

8.   Place the new IASCII files in the same directory/book structure as the original files, open them in Interleaf, and examine the results. This is similar to what your files will look like when they have been converted after translation. Make any changes you deem necessary to the formatting of your source files, save the files and repeat steps 2 through 10.

9.   You are now ready to commence translation of the STF files.

Note

Catalogs which contain text in shared components, shared frames, or in numbering formats should be treated in the same way as Interleaf documents. Any text in catalogs which is not used will not be presented for translation.

2.3.2 File Preparation and the Alignment Cycle

The steps for preparing FrameMaker and Interleaf files for alignment are the same as those for translation preparation. Follow the steps as outlined above in **File Set-up and Preparation**. Do not forget that these steps apply both to the source and target language files.

Enabling Structure Recognition for STF Files

STF is a tagged text format. As a result STF files do not contain Word styles such as "Heading 1" that can be used for structure recognition in WinAlign. This means that WinAlign's structure recognition feature cannot be used as a parameter for alignment.

However, in almost all cases, STF files contain external tags that do carry structure information in the form of paragraph style ("**<ps...>**") tags. See the following example:

Example

<ps "Head 1" 1>TRADOS**<:fc 1>**®**<:/fc>** Translator's Workbench for Windows

<ps "Body" 3>The tool for **<:cs "Emphasis" 1>**professional translators**<:/cs>** who expect that intelligent, innovative technology offered in an ergonomic environment will bring both greater success and enjoyment to their work

<ps "Head 2" 4>The Translator's Second Memory

The two tags **<ps "Head 1" 1>** and **<ps "Head 2" 4>** are placeholders for two heading levels in the original FrameMaker or Interleaf file. If you can map the placeholders to Word styles, then you can harness WinAlign's structure recognition feature in STF alignment. There are two ways of converting the external tags to Word styles:

* either by using a Word macro. A sample macro, **WinAlignPrepareSTF** may be found in the document template WinAlign.dot, or,

* by manually applying Word Styles using Word's search and replace function.

Refer to the WinAlign User's Guide for more information on enabling structure recognition for STF files.

Alignment

Below are some points to remember when aligning STF files. Refer to the WinAlign User's Guide for more information on STF alignment.

* In the **General** tab of the **WinAlign Project** dialog, make sure to choose "Word Documents" from the **File Type** drop-down list. In the current version of WinAlign, the program treats STF files in the same way as Word documents.

* In the **Alignment** tab, define the weighting of the various tuning options according to the specific document features of your alignment pairs (formatting, numbers, etc.).

* In the **Structure Recognition** tab, set the structure recognition depth and add the styles you have applied under **Enabling Structure Recognition for STF Files** above. If you did not follow the procedure described in that section, set the structure recognition depth to **Ignore**.

2.3.3 During Translation

General Tips

All translators should be familiar with STF format before commencing a translation project (refer to The S-Tagger User's Guides).

There are specific rules for the addition and deletion of STF tags during translation. Translators should be familiar with these (refer to the relevant chapters in The S-Tagger User's Guides).

Translators should be vigilant about the sequence of tags in the target segment. Sometimes it is necessary to move tags around within a sentence. This can be done either by using the **Copy Source** command (recommended for a no match) or by using the **Get Placeable** command on the toolbar (recommended for a 100% or Fuzzy Match). When tags are being placed, the translator should take care that the spacing between the tag's bracket and the first and last letters of the word to which it applies, are identical to the source segment.

Translator's Workbench automatically inserts tags in the same sequence in the target text as in the source text, independently of how they might appear in the translation memory. The translators should be aware of this and should always check the sequence of tags when translating and make sure that they are placed in the correct position for the translated segment.

If there are more tags in the translation memory proposal than in the source segment, Translator's Workbench will append tags from the TM to the last tag positions in the target sentence. The translator should be vigilant about the sequence of tags, especially when the translation memory proposal has more tags than the current source segment.

It is advisable to translate the ancillary file first and to give a copy of the translated file to all translators. The ancillary file will contain text which is represented in the STF files by tags, such as variables, and text which will not appear in the STF files at all, such as headers and footers. It is helpful to the translators to see how text in the ancillary file has been translated so that they can translate the rest of the STF file appropriately.

The following are some general tips which relate to the translation of all file formats using Translator's Workbench:

- Always work with hidden text turned on.

- Always use "Normal View" in Word.

- When you want to spell-check, you will only want Word to check the translated segments. You should therefore turn off the hidden text for the spell-checking stage.

- Similarly, if you want to proof-read on hard copy, you will only want to print your translated sentences, not the source sentences. You should make sure that hidden text is NOT specified as one of the printing characters (select **Options** from Word's **Tools** menu, click on the **Print** tab and, under **Include with Document**, make sure that the **Hidden Text** box is unchecked). You will now be able to print a document which displays only translated segments.

- When implementing corrections, you can either choose to open each segment, edit it and use **SetClose** to update both the segment and the TM, or you can implement the correction without opening the segment or using **SetClose**. If you choose to do the latter, then you must update the TM afterwards. To do this:

  1. Make a copy of the file(s).

  2. Open the translation memory.

  3. Go to the **Tools** menu and select **Clean Up**.

  4. Add the file(s) to be cleaned.

  5. Check **Update TM**.

  6. Click on **Clean Up**.

Your changes will now be implemented in the TM.

- The final step before files are handed over for printing or DTP tidy up, is to "Clean" the files. The instructions above illustrate how to clean the files. If all changes have been implemented in the TM already, then there is no need to check the **Update TM** box.

Proof-Reading

FrameMaker

To proof-read, it is advisable to convert the STF files back to MIF, open them in FrameMaker, print and proof-read them and then implement the changes in the translated RTF files in Word. Before the files are converted, the translator should run The S-Tag Verifier on the translated files and implement any changes where necessary. You will need at least one licence of The S-Tagger 2.0 for FrameMaker in order to carry out the STF to MIF conversion.

Interleaf

To proof-read, you may want to convert the files back to Interleaf IASCII, open them in Interleaf, print and proof-read them and then implement the changes in the translated STF files in Word. However, conversion for proof-reading is not as straightforward as for FrameMaker. Most translators will have FrameMaker licences, but not all will have Interleaf licences (cost implications if you are using external translators). Also, there must be at least one person available who has a good knowledge of Interleaf. If both of these conditions are fulfilled, then before the files are converted, the translator should run The S-Tag Verifier on the translated files and implement any changes where necessary. During proof-reading, any changes required should be implemented in the translated STF files. Do not forget to update the translation memory (either interactively or in batch mode using the **Update TM** command). You will need at least one licence of The S-Tagger 2.0 for Interleaf in order to carry out the STF to IASCII conversion.

Tag Verification

The S-Tag Verifier is a stand-alone utility which can be freely downloaded from the TRADOS website, http://www.trados.com.

Again, the tag verification process is described in detail in The S-Tagger User's Guides, so, an outline of the steps is all that is required here. The instructions for tag verification are also contained in The S-Tag Verifier's on-line help system.

1. Save your file as RTF and close it.

2. Launch The S-Tag Verifier.

3. Make sure RTF is selected as the file format.

4. Select the translated file(s) you wish to verify.

5. Click OK.

6. Confirm that the path to the corresponding source STF files is correct.

7. Click OK.

The verification process now runs. Any errors are reported immediately. Details of the errors are saved to a CMP file. The translator should open the CMP file, assess the errors and make any changes to the translated STF file and then verify the tags again.

The CMP file will also list any changes made to the tags in the file which are simple text formatting changes, and which are allowed. These are referred to as Alerts and Warnings. When the translator is satisfied that the files contain no tag errors, then the files can be converted to MIF or IASCII and opened in FrameMaker or Interleaf for proof-reading. Note that the translator will require at least one licence for The S-Tagger 2.0 for FrameMaker or for Interleaf in order to convert the files back to MIF or IASCII format. Alternatively, ITP Kitten Tools can be downloaded from ITP's website, http://www.itp.ie). The ITP Kitten Tools can convert translated STF files to text only for proof-

reading purposes. No dongle for The S-Tagger is required for this. Please consult the on-line help for instructions on how to convert the files.

The conversion process from STF to MIF or IASCII is described step-by-step in The S-Tagger User's Guides. During proof-reading, any changes required should be implemented in the translated STF file. Do not forget to update the translation memory (either interactively or in batch mode using the **Update TM** command).

Post-Translation

When the proof-reading changes have been implemented in the translated STF files, the files should be run through The S-Tag Verifier again to ensure that no errors were introduced during proof-reading. It is advisable that the translator writes comments in the final CMP file confirming that any formatting changes outlined in the file were deliberate.

When any errors are fixed, the translated files, the CMP files generated during the final S-Tag Verification and the translated ancillary files should be delivered to the person who will do the final conversion from STF to MIF or IASCII. It is preferable that this job is carried out by a DTP specialist who has a good knowledge of FrameMaker or Interleaf. If the translator has made comments in the CMP file, this will help the person responsible for final conversion.

Once the files have been converted, cross-references should be updated and the table of contents and index should be generated.

TRADOS Font Mapper for Interleaf

TRADOS Font Mapper for Interleaf is a companion application to The S-Tagger 2.0 for Interleaf which has been developed to make the process of changing the fonts in Interleaf documents translated into or from Eastern European languages or Japanese, using TRADOS software, more efficient. TRADOS Font Mapper replaces the fonts in the original document with the fonts that you specify. Other attributes such as font size and some styles may also be changed. The Font Mapper fits into the Interleaf translation workflow as follows:
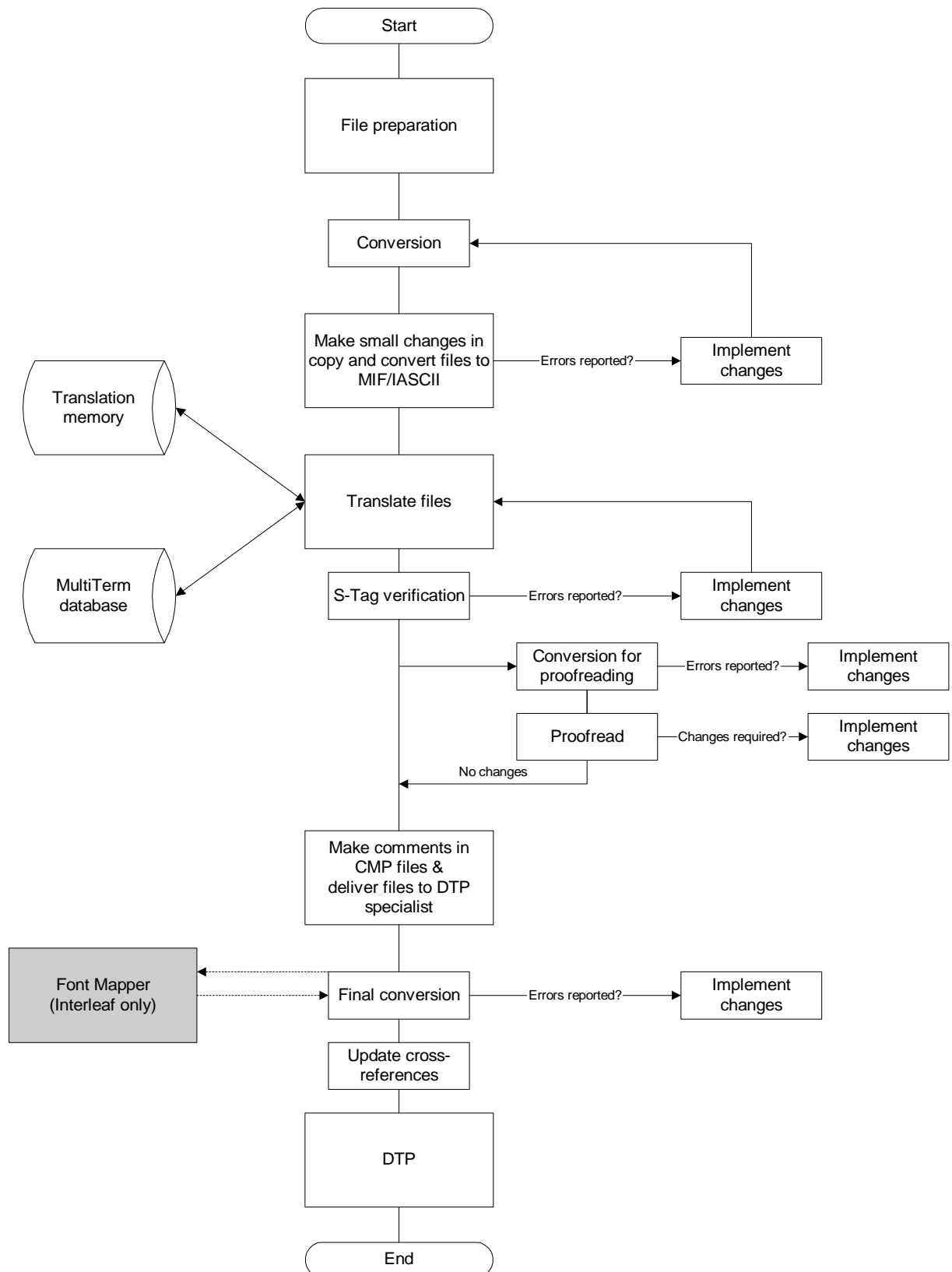
1. Prepare IASCII files for translation (as instructed above).

2. Convert IASCII files to STF using The S-Tagger 2.0 for Interleaf.

3. Translate STF files.

4. Verify Tags.

5. Convert STF files to IASCII using The S-Tagger 2.0 for Interleaf.

6. Run the Font Mapper Utility.

7. Open the IASCII files in Interleaf.

The Font Mapper may be downloaded from the TRADOS website http://www.trados.com and runs on the same dongle as The S-Tagger 2.0 for Interleaf. Refer to the TRADOS Font Mapper's on-line help for more detailed instructions on how to map fonts.

The STF translation workflow is presented below in diagram format.

Preparation/Translation Workflow with STF Files
and Translator's Workbench

```
                          ╭─────────────╮
                          │    Start    │
                          ╰──────┬──────╯
                                 │
                   ┌─────────────────────────┐
                   │     File preparation     │
                   └────────────┬────────────┘
                                │
                       ┌──────────────┐                                    ┌──────────────┐
                       │  Conversion  │◄───────────────────────────────────│              │
                       └──────┬───────┘                                    │              │
                              │                                            │              │
              ┌────────────────────────────┐                    ┌──────────────────┐
              │  Make small changes in     │  Errors reported?  │   Implement      │
              │  copy and convert files to │───────────────────►│    changes       │
              │       MIF/IASCII           │                    └──────────────────┘
     ╭────────────╮       └──────┬─────────┘
     │ Translation│              │
     │   memory   │◄────┐        │
     ╰────────────╯     │        │
                        └──┌──────────────┐                     ┌──────────────────┐
                           │ Translate files │◄─────────────────│   Implement      │
                        ┌──└──────┬───────┘                     │    changes       │
     ╭────────────╮     │         │                             └──────────────────┘
     │ MultiTerm  │◄────┘  ┌──────────────┐  Errors reported?   ┌──────────────────┐
     │  database  │        │ S-Tag verification │──────────────►│   Implement      │
     ╰────────────╯        └──────┬───────┘                     │    changes       │
                                  │         ┌──────────────┐    ┌──────────────────┐
                                  ├────────►│ Conversion for │ Errors reported? │  Implement │
                                  │         │  proofreading  │─────────────────►│  changes   │
                                  │         └──────┬───────┘    └──────────────────┘
                                  │         ┌──────────────┐ Changes required?┌──────────────────┐
                                  │         │  Proofread   │─────────────────►│   Implement      │
                                  │         └──────┬───────┘                  │    changes       │
                                  │   No changes   │                          └──────────────────┘
                                  ◄────────────────┘
                   ┌─────────────────────────┐
                   │  Make comments in        │
                   │  CMP files &             │
                   │  deliver files to DTP    │
                   │  specialist              │
                   └────────────┬────────────┘
    ┌──────────────────┐        │
    │  Font Mapper     │◄┄┄┄┄ ┌──────────────┐  Errors reported?  ┌──────────────────┐
    │ (Interleaf only) │┄┄┄┄►│ Final conversion │───────────────►│   Implement      │
    └──────────────────┘     └──────┬───────┘                    │    changes       │
                             ┌──────────────┐                    └──────────────────┘
                             │ Update cross- │
                             │  references   │
                             └──────┬───────┘
                   ┌─────────────────────────┐
                   │           DTP            │
                   └────────────┬────────────┘
                          ╭─────────────╮
                          │     End     │
                          ╰─────────────╯
```

2.4 ITP Filter Pack Formats and SGML

The ITP Filter Pack contains a series of utilities which mark up text formats: (either HTML or the text exports from certain DTP packages) as tagged RTF which is suitable for translation using Translator's Workbench. Tagged text formats are currently translated in Word. However, at time of press TRADOS was developing the TagEditor, a specialised application designed to facilitate the translation of tagged text formats. This section examines the workflow for translating file formats marked up by the ITP Filter Pack utilities: HTML, Ventura, QuarkXPress and PageMaker. SGML files which are marked up for translation using the SGML2RTF conversion tool are also included here. The workflow is structured as follows:

- **File Set-up and Preparation** makes recommendations on how to set up the original file formats prior to running the ITP Filter Pack or the SGML2RTF conversion tool.

- **File Markup** explains how the markup utilities work.

- **During Translation** provides tips for translators when translating the marked up tagged text format.

- **Post-Translation** outlines the process for converting marked up RTF files back to their original file format.

- **Alignment** looks at the alignment process for each file format.

2.4.1 File Set-Up and Preparation

Recommendations for file set-up are made for each file format separately. Please note that these are only recommendations.

HTML

- **File Management:** A project with files in HTML format can often consist of hundreds of files distributed over several sub-directories. When it comes to translating these files, there is a lot of file management involved. Technical writers of source HTML files, which are destined for translation, should be alerted to this fact and should try to keep the number of files and sub-directories to a minimum.

- **Tag Definition:** The ITP Filter Pack marks up a number of pre-defined HTML tags. In addition, you can create an HTML tag definition file which contains information on how specific HTML tags should be marked up. See ITP Filter Pack on-line help for more information on setting up a tag definition file.

- **Bitmaps:** Text inside bitmaps cannot be translated without the use of an image editor. If you have to use an image editor, you will not gain the benefits of a translation memory. It is, therefore, better not to put text which will require translation inside bitmaps.

- **Text Inside Code Segments:** Avoid putting text which will require translation inside code segments (i.e. "<!--" to "-->"). These tags are marked up as external tags by the HTML conversion utility (described below), but external tags cannot be translated using Translator's Workbench.

- **Comment Tags Inside Sentences:** Avoid putting comment tags inside sentences. Comment tags will be marked as external tags by the HTML conversion utility. If an external tag occurs in the middle of a segment, the segment will be broken.

SGML

- **Internal Tags:** The SGML2RTF conversion tool marks up all tags as external tags unless otherwise specified. Before running the utility, examine your SGML files and create a tag file, `internal.tag,` containing those tags which should be marked up as internal tags. Place one tag on each line.

- **Non Translatable Text Tags:** Some SGML tags may contain text which should not be translated e.g. a serial number. If you specify these tags in your tag file SGML2RTF will mark

up the tags and non-translatable text as one item. Precede non-translatable tags with an asterisk (*) in the tag file.

Note

Non translatable text must be surrounded by opening and closing tags if it is to be marked up as one item by the SGML2RTF utility.

For example:

The `*ISBN` tag in the `internal.tag` file must correspond to the following tags in the SGML file:

`<ISBN>0-201-48345-9</ISBN>`

Absence of either the opening or closing tag will result in a markup error.

PageMaker

- **Text Flow:** All stories must be contained in a single flow. Any threaded stories must be joined, to avoid multiple tag files per PageMaker file.

- **Tagged Paragraphs:** All paragraphs must be tagged. Apply a defined paragraph tag to all paragraph tags with the custom 'no style' tag.

- **Artwork:** All artwork must be linked. In-line graphics (anchored graphics) are not exported. Although they are lost, they can be easily cut and pasted from the source material into the translated material and thus re-linked. Generate a list of graphics using the PageMaker plug-in "Save for Service Provider". This will help when recreating the graphics links.

- **Export:**
  - ✓ With the file open, select **Edit Story** from the **Edit** menu.
  - ✓ Place the cursor in the text stream, and select **Export Text** from the **File** menu.
  - ✓ In the **Export Document** dialog, select **tagged text** as the type, make sure that **entire story** radio button is selected, and that **export tags** box is not checked. (Not selecting **export tags** may appear confusing. However, if the **export tags** option is checked the files will not be marked up correctly by the ITP Filter Pack.)

- **Ancillary File:** Create an ancillary file for text which is not exported to the tagged text format (master page text, master text items and callouts). Translate the ancillary file and incorporate the text into the translated PageMaker file.

TRADOS Story Collector

At time of press TRADOS was developing TRADOS Story Collector, a tool which will facilitate the preparation of PageMaker files for translation in Translator's Workbench. TRADOS Story Collector will allow the export of multiple stories from PageMaker.

QuarkXPress

There are two possible methods of exporting tagged text from QuarkXPress, either by manual export or by using CopyFlow a QuarkXPress plug-in. Both are examined here.

- **Manual Export:** QuarkXPress documents often consist of many stories. To export text manually from a QuarkXPress document:
  - ✓ select the story and choose the **Save Text...** command from the file menu.
  - ✓ The **Save Text** dialog box displays. Save the output file as Xpress tags (*.xtg). Select the **Entire Story** radio button. Using this method, one export file is created for each story. Try and assign meaningful names to the export files so you can easily import them back into the right locations after translation.

- **CopyFlow Export:** A QuarkXPress plug-in, CopyFlow provides additional menu selections in QuarkXPress which allow you to automatically import and export text and artwork. CopyFlow is produced by North Atlantic Publishing Systems, (Inc.) http://www.napsys.com. You need to carry out the following preparatory steps before running the ITP Filter Pack markup utility:

  ✓ **Naming Boxes:** Before exporting QuarkXPress document to tagged text, you must assign names to all text and picture boxes. Use the **Name Box** command from the **CopyFlow** menu in QuarkXPress.

  ✓ Select **CopyFlow Preferences** from the **CopyFlow** menu in QuarkXPress. Select the **Use QuarkXPress Style Tags** radio button. Select **Style Tags (Single File)** as the export format. Specify the folder to which CopyFlow will export the files.

  ✓ Set the **Batch Export Folder** (from the **CopyFlow** menu in **QuarkXPress**).

  ✓ Select **Batch Export** form the **CopyFlow** menu. The **Export Text** dialog displays. The named text boxes appear on the left hand side. Click the **Select All** button. The text boxes export to a single file with a `.TTG` extension in the location specified above.

  ✓ Rename the `.TTG` file to a `.XTG` file (for marking up in the ITP Filter Pack).

CopyFlow

CopyFlow is a third party tool produced by North Atlantic Publishing Systems (Inc.). CopyFlow is **QuarkXPress** version and platform dependent. It does not ship with **QuarkXPress.** Contact North Atlantic Publishing Systems, (Inc.) for more information. Website: http://www.napsys.com.

Ventura

- **Text Frames:** Tag all text frames.

- **Callouts:** Place callout text in frames and tag all frames.

- **Graphics:** Anchor all graphics or free form frames.

- **Automatic Generation:** Automatically generate variables, auto-numbered lists, cross-references, headers and footers.

- **Page Breaking Tags:** Open the files in Word. If there are page breaking tags, delete them using Word's search and replace function.

- **Ancillary File:** Create an ancillary file for text which is not exported to the tagged text format (callout text, text in headers and footers and frame text). After translation incorporate the ancillary file into the translated Ventura file.

- **Export:**

  ✓ Select the frame or page with text for e xport.

  ✓ Select **Export** from the **File** menu.

  ✓ Save the export file as ANSI text with a `.TXT` extension.

2.4.2 File Markup

The ITP Filter Pack

The ITP Filter Pack contains a series of utilities which mark up text formats: (either HTML or the text exports from certain DTP packages) as tagged RTF which is suitable for translation using Translator's Workbench. At the time of press TRADOS was developing the TagEditor, a specialised application designed to facilitate the translation of tagged text formats.

Once the marked up RTF files have been translated, the ITP Filter Pack provides a backwards conversion utility which converts the translated RTF files back into a tagged text format which can be displayed in or imported into their original file format.

Mark Up Tags

In addition to the standard menus, the ITP Filter Pack has one menu per file format: Ventura, PageMaker, **QuarkXPress** and HTML. Each menu has a **Mark Up Tags** option which marks up tags in the text files as internal and external tags, using the character styles tw4winInternal and tw4WinExternal and saves the output as an RTF file.

For HTML files, you can supplement the tags that the ITP Filter Pack marks up with additional tags which are specific to the HTML files you wish to translate. Refer to the ITP Filter Pack on-line help for more information on setting up a tag definition file.

For Ventura files, there are two separate markup utilities: one for marking up files which originated in Ventura 4, the other for files which originated in Ventura 5 or later.

SGML2RTF Conversion Tool

The SGML2RTF tool is a 32-bit command line tool which runs from a DOS box under Windows. Like the ITP Filter Pack, it marks up SGML files as tagged RTF so that they can be translated using Translator's Workbench for Windows. Tags are marked up as internal and external tags, using the character styles tw4WinInternal and tw4winExternal and saved out as RTF. Non translatable text within tags may also be marked up as one internal tag (See **SGML File Set-Up and Preparation**). The SGML2RTF conversion tool is freely available from the TRADOS website: http://www.trados.com. Refer to the readme.doc which ships with the tool for more information on marking up SGML files as RTF.

2.4.3 During Translation

After the files have been marked up as RTF, either using the ITP Filter Pack or the SGML2RTF conversion tool, they contain external and internal tags. Translator's Workbench ignores external tags and treats internal tags as placeables.

The following are some general tips which relate to the translation of all file formats using Translator's Workbench:

- Always work with hidden text turned on.

- Always use "Normal View" in Word.

- When you want to spell-check, you will only want Word to check the translated segments. You should therefore turn off the hidden text for the spell-checking stage.

- Similarly, if you want to proof-read on hard copy, you will only want to print your translated sentences, not the source sentences. You should make sure that hidden text is NOT specified as one of the printing characters (select **Options** from Word's **Tools** menu, click on the **Print** tab and, under **Include with Document**, make sure that the **Hidden Text** box is unchecked). You will now be able to print a document which displays only translated segments.

- When implementing corrections, you can either choose to open each segment, edit it and use **SetClose** to update both the segment and the TM, or you can implement the correction without opening the segment or using **SetClose**. If you choose to do the latter, then you must update the TM afterwards. To do this:

  1. Make a copy of the file(s).

  2. Open the translation memory.

  3. Go to the **Tools** menu and select **Clean Up**.

  4. Add the file(s) to be cleaned.

  5. Check **Update TM**.

  6. Click on **Clean Up**.

Your changes will now be implemented in the TM.

Proofreading HTML files

During the translation process you can view your partially translated RTF file in a web browser to ensure that it appears correctly. To do this:

1. Close the segment you are working on. Save the file as RTF.

2. Run the ITP Filter Pack. From the **HTML** menu select the **RTF to Text (Multilingual)** option. This converts the RTF file to HTML.

3. Open the HTML file in a web browser.

2.4.4 Post Translation

Clean Up

When translating with Translator's Workbench the **Clean Up** feature removes the source text and segment delimiters stored in the translated document. For all formats (except SGML) the ITP Filter Pack performs the clean up operation as part of the **RTF to Text** conversion process.

RTF to Text

Each **Mark Up Tags** command has a corresponding **RTF to Text** command which converts the translated RTF file back to its original text format. This text format may then be imported into the original DTP package, or, in the case of HTML, opened in a web browser.

For HTML files there are three backwards conversion options:

- **RTF to Text (Multilingual),** which converts the RTF to a language specific HTML file. Hidden text and segment markers are removed and a META tag is inserted which instructs the browser as to which code page it should use to display the HTML file. When you select the **RTF to Text (Multilingual)** option, the **Select Language** dialog appears prompting you to choose a language. ITP Filter Pack automatically associates a character set with the language you select. If you wish to choose a different character set, check the custom character set box and specify the code page you wish to associate with the selected language.

- **RTF to Text (Compatible),** which also converts the RTF to a HTML file. Hidden text and segment markers are removed. However, no META tag is inserted. Since not all special characters are supported by all web browsers, this tool converts special characters to an equivalent character which is supported by all browsers. For example, curly quotes are not supported by all browsers so the **RTF to Text (Compatible)** converts curly quotes to straight quotes. Refer to the ITP Filter Pack on-line help for a list of special characters which are converted to equivalent characters.

- **RTF to Text**, which is similar to RTF to Text (Compatible) except that it does not convert any special characters.

Note

In most cases the RTF to Text (Multilingual) option will yield the best conversion results.

SGML Clean Up

Converting SGML files which have been marked up as RTF back to SGML is a slightly different process.

- After translation, run Translator's Workbench **Clean Up** Utility.

- Open the cleaned RTF files in Word and save out as text only with an `.SGM` or `.SGML` extension.

> **Tip**
>
> Word will prompt you to save the file out with a `.TXT` extension. To save the file as text only with an `.SGM` or `.SGML` extension simply enclose the file name in quotation marks in the **Save As** dialog. E.g. To name a file as `intro.sgm`, type `"intro.sgm"` in the **File Name** box of Word's **Save As** dialog.

Import

Once you have converted the translated RTF files back to text format you need to import the files back into their original program. Although the principles are the same the steps required vary according to the program in question.

HTML

- Open the HTML files in a browser.

- Check that all links are functional.

SGML

- Open the SGML files in an SGML editor.

- Visually check the files.

PageMaker

- Open the original file in PageMaker.

- Select **Place** from the **File** menu. In the **Place** dialog, select the file to import and specify the following settings: check the **As New Story** radio button and check both the **Retain Format**, and **Convert Quotes** boxes. Do not select any other options.

- Enable **Autoflow** by selecting the **Autoflow** option from the **Layout** menu.

- Place the translated text.

- Re-link the art files.

- Cut and paste the ancillary text into the translated PageMaker file.

> **Tip**
>
> Re-linking PageMaker art is a simple process. Open the original story and the translated story in two windows side by side. Simply cut and paste the graphic anchor from the original story into the corresponding place in the translated story. This will restore all art references.
>
> **Warning**
>
> There is a PageMaker 6.5 or lower bug which results in the truncation of index markers longer than 50 characters during import. This has been fixed for PageMaker 6.5.2.

QuarkXPress

The two methods for exporting QuarkXPress text to tagged text have corresponding import approaches.

Manual Import

- The ITP Filter Pack converts the translated marked up RTF file to an `*.XTG` file.

- Open the source file in QuarkXPress.

- Place the cursor where you wish to insert the text. Select **Get Text...** from the **File** menu. The **Get Text** dialog box displays. Ensure the **Include style sheets** box is checked. Select the `*.XTG` file for import.

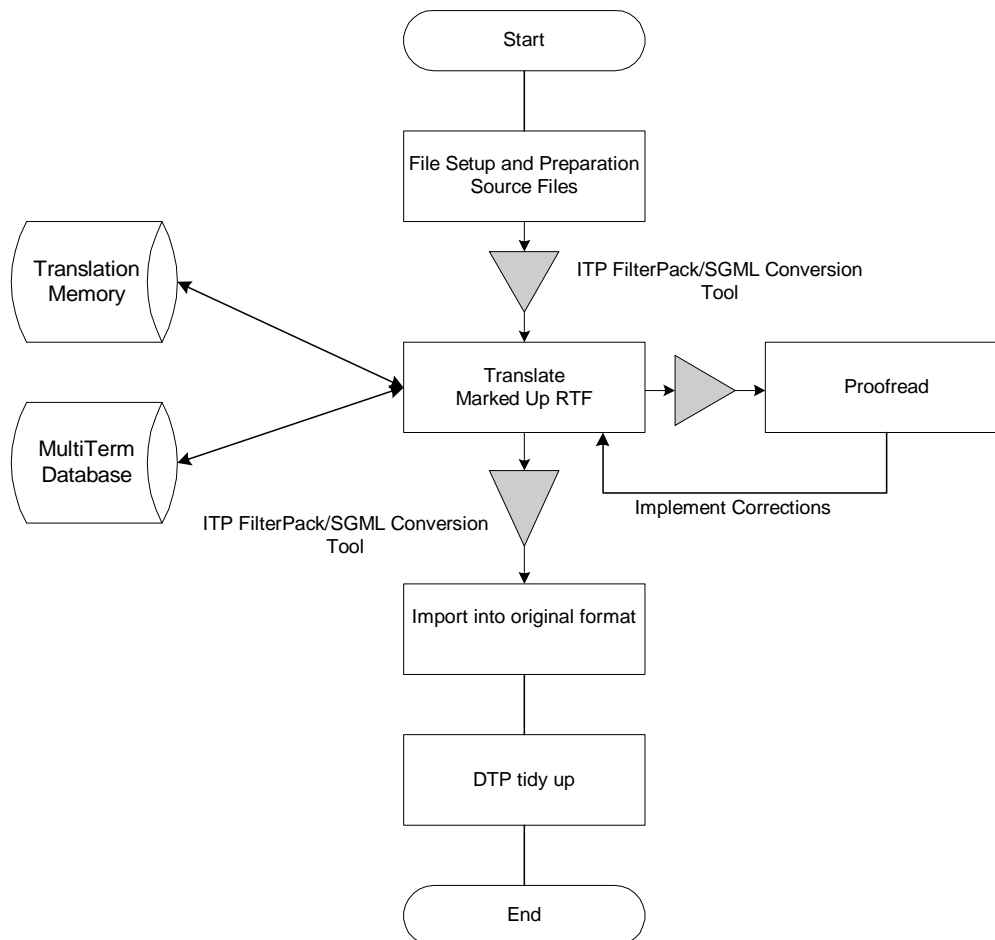- Click on **Open**.

Copy Flow Import

- The ITP Filter Pack converts the translated marked up RTF file to an `*.XTG` file. Rename this file with a `*.TTG` extension before importing it into QuarkXPress.

- Open the source file in QuarkXPress.

- Set the CopyFlow Preferences as per the export process.

- Set the Batch Import Folder to the location of the translated files.

- Select **Batch Import** from the Copyflow menu. Check the text box radio button. The named boxes will appear. Click the **Select All** button. Click on **OK**. QuarkXPress imports the tagged translated files.

- Paste the translated text from the ancillary file into the QuarkXPress document.

- Perform DTP tidy up: adjust callout text, pagination, tables, update cross-references and regenerate the TOC and index.

Ventura

- Open the files in Corel Ventura.

- Paste translated text from the ancillary file into the Ventura document..

- Perform DTP tidy up: adjust callout text, pagination, tables, update cross-references and regenerate the TOC and index.

The following workflow diagram illustrates the tagged text translation process.

Translating Tagged Text Files

```
                          ┌─────────────────────────┐
                          │          Start          │
                          └─────────────────────────┘
                                       │
                                       ▼
                          ┌─────────────────────────┐
                          │  File Setup and Preparation │
                          │      Source Files       │
                          └─────────────────────────┘
                                       │
                                       ▼
                                       ▽        ITP FilterPack/SGML Conversion
                                                           Tool

┌─────────────┐
│ Translation │ ◄────────┐          ┌─────────────┐        ┌─────────────┐
│   Memory    │          │          │  Translate  │───▷───  │  Proofread  │
└─────────────┘          │          │ Marked Up RTF│       └─────────────┘
                         │          └─────────────┘
┌─────────────┐          │                 │
│  MultiTerm  │ ◄────────┘                 ▽          Implement Corrections
│  Database   │                                                   
└─────────────┘    ITP FilterPack/SGML Conversion
                              Tool
                                       │
                                       ▼
                          ┌─────────────────────────┐
                          │  Import into original format │
                          └─────────────────────────┘
                                       │
                                       ▼
                          ┌─────────────────────────┐
                          │       DTP tidy up       │
                          └─────────────────────────┘
                                       │
                                       ▼
                          ┌─────────────────────────┐
                          │           End           │
                          └─────────────────────────┘
```

2.4.5 Tagged Text Format Alignment

With the exception of HTML, the SGML, PageMaker, QuarkXPress and Ventura alignment processes are very similar to the STF alignment process. You must first convert the original file format into a format supported by WinAlign and then align the files. Preparing SGML, PageMaker, QuarkXPress and Ventura files for alignment is identical to preparing files for translation.

Once you have exported the original file format to tagged text format use the ITP Filter Pack or the SGML2RTF conversion tool to mark up the files as RTF. Like STF files, marked up tagged text files do not contain Word styles such as "Heading 1" that can be used for structure recognition in WinAlign. This means that WinAlign's structure recognition feature cannot be used as a parameter for alignment. There are two possible options: either set WinAlign's structure recognition option to **Ignore**, or, assign Word styles to tags which represent styles in the original DTP package.

Note

Assigning Word styles to PageMaker, QuarkXPress or Ventura tagged text files requires extensive knowledge of the DTP package and the files.

If in doubt, disable WinAlign's structure recognition option.

Align the files using the standard alignment workflow.

In summary the SGML, PageMaker, QuarkXPress and Ventura alignment workflow is as follows:

1. Carry out any file preparation necessary in the original file format.

2. Export to tagged text format (except HTML).

3. Mark up the text with external and internal tags (using either The ITP Filter Pack or the SGML2RTF conversion tool).

4. If possible, assign Word styles to the marked up RTF files. Otherwise, disable WinAlign's structure recognition option.

5. Align the filenames.

6. Align the files.

7. Check alignment (optional). Review the alignment. Make changes where required.

8. Export the alignment results. Import the alignment results into a new or existing translation memory.

9. Use the **Concordance** command in Translator's Workbench to search for known phrases and check that characters etc. look OK. (If you are aligning languages which use different fonts, make sure that you have installed the correct fonts on your system and have mapped these fonts in Translator's Workbench using the **Font Translations** command in the **Translation Memory Setup** dialog box.)

HTML Alignment

WinAlign supports WYSIWYG alignment of HTML files which means that your files appear in WinAlign just as they would appear in your browser. The only exception is tags that occur mid-segment but do not contain any formatting information e.g. the anchor tag `<A HREF>`. These tags appear in red in the alignment editor. Some tips for aligning HTML files:

- Re sort any alphabetical lists so that the source text corresponds to its translation.

- If your HTML files are spread over several folders, use WinAlign's drag and drop feature to load the files into a WinAlign project.

- Before aligning, assess the formatting in the HTML files and set your structure recognition options accordingly.

> **Note**
>
> Remember, although WinAlign supports alignment in a WYSIWYG fashion, once you export the alignment project and set up a Translator's Workbench TM, the formatting information is exported as tags i.e. tw4winInternal. This allows you to translate your new HTML source files using Translator's Workbench.
>
> At time of press TRADOS was developing a TagEditor which will support WYSIWYG translation of HTML files.

The following workflow diagram illustrates the alignment process for tagged text files.

Aligning Tagged Text Files

```
                                    ┌─────────────┐
                                    │    Start    │
                                    └─────────────┘

┌─────────────────────┐                                    ┌─────────────────────┐
│ File Setup and      │                                    │ File Setup and      │
│ Preparation         │                                    │ Preparation         │
│ Source Files        │                                    │ Target Files        │
└─────────────────────┘                                    └─────────────────────┘

         ▽             ITP FilterPack/SGML Conversion              ▽
                                    Tool

┌─────────────────────┐                                    ┌─────────────────────┐
│   Marked Up RTF     │                                    │   Marked Up RTF     │
└─────────────────────┘                                    └─────────────────────┘

┌─────────────────────┐                                    ┌─────────────────────┐
│ Apply Heading styles│                                    │ Apply Heading styles│
│ for Structure       │                                    │ for Structure       │
│ Recognition, re-    │                                    │ Recognition, re-    │
│ save as RTF         │                                    │ save as RTF         │
│ or                  │                                    │ or                  │
│ Set WinAlign        │                                    │ Set WinAlign        │
│ Structure           │                                    │ Structure           │
│ Recongition to      │                                    │ Recongition to      │
│ 'Ignore'            │                                    │ 'Ignore'            │
└─────────────────────┘                                    └─────────────────────┘

                     ┌─────────────────────┐
                     │    Alignment        │
                     └─────────────────────┘

                                    ┌─────────────────────┐
                                    │  Review Alignment   │
                                    └─────────────────────┘

                     ┌─────────────────────┐
                     │   Export results    │
                     └─────────────────────┘

                     ┌─────────────────────┐
                     │ Import into Workbench│
                     └─────────────────────┘

                           ┌─────────────┐
                           │     End     │
                           └─────────────┘
```

Index