

# Erfahrungen im Einsatz mit Translation-Memory-Systemen

*Antje Pesch*

*DELTA International CITS GmbH*

## 1 Einleitung

Im Rahmen der Arbeiten des Arbeitskreises Maschinelle Übersetzung zum Thema Leistungsfähigkeit und Einsatzmöglichkeiten von Translation Memory (TM) Systemen soll der folgende Artikel einen Überblick über die praktischen Erfahrungen eines Übersetzungsunternehmens beim Einsatz von TM-Systemen geben.

Unter einem TM sei im folgenden (entsprechend der am meisten verbreiteten TM-Architektur) eine Datenbank zu verstehen, in welcher Übersetzungseinheiten abgespeichert werden. Eine Übersetzungseinheit setzt sich zusammen aus einem ausgangssprachlichen Satz und seiner zielsprachlichen Entsprechung.

### TM Systeme

- basieren auf dem Prinzip des Recycling: sie garantieren die Wiederverwertbarkeit der Übersetzungseinheiten schon während des Übersetzungsprozesses sowie ihre Konservierung für Folgeprojekte;
- arbeiten in der Regel satzbasiert<sup>1</sup>;
- ermöglichen dem Übersetzer – im Gegensatz zum passiven Charakter der maschinellen Übersetzung – einen interaktiven Arbeitsprozeß: das System bietet Übersetzungseinheiten aus der Datenbank an, die der Übersetzer annehmen, editieren oder ablehnen kann.

TM-Systeme lassen sich bezüglich ihrer Funktionsweise in datenbank- und sprachpaarbezogene Systeme untergliedern. Als Vertreter auf der Basis von Datenbanken organisierter Systeme können beispielsweise die Translator's Workbench von TRADOS, der TranslationManager von IBM und Déjà Vu von ATRIL genannt werden. Das System Transit von STAR hingegen arbeitet sprachpaarbezogen.

Im folgenden sollen einige Eigenschaften der Translator's Workbench von TRADOS – dem derzeitig marktführenden TM-System – und des TranslationManager von IBM – einem TM-System, das die Delta International CITS GmbH oft einsetzt und dessen Vertrieb sie unterstützt – gegenübergestellt werden, um bestimmte Erfahrungen aus dem praktischen Umgang mit den TM-Systemen zu veranschaulichen.

## 2 TM-Systeme in der Praxis

Der Einsatz von TM-Systemen betrifft vor allem Übersetzer, den technischen Support und das Projektmanagement.

### 2.1 Erfahrungen der Übersetzer

Für den Übersetzer, der mit einem TM-System arbeitet, ist zunächst die Benutzerfreundlichkeit des jeweiligen Systems von entscheidender Bedeutung. Weiterhin ist es für ihn von Vorteil, während des Übersetzungsvorgangs auf das TM und die Terminologieerkennung zugreifen zu können.

#### 2.1.1 Benutzerfreundlichkeit

- **Arbeitsumgebung:** Die Schnittstelle der Translator's Workbench zu den Textverarbeitungsprogrammen Word für Windows bzw. WordPerfect ermöglicht dem Übersetzer die Arbeit in einer vertrauten Umgebung. Ferner arbeitet die Workbench nach dem WYSIWYG-Prinzip.

Beim TranslationManager muß der Übersetzer im systemeigenen Editor arbeiten. Dessen Aufbau ist für den Übersetzer zunächst fremd und verlangt eine gewisse Einarbeitung. Zudem arbeitet der TranslationManager nicht nach dem WYSIWYG-Prinzip.

- **Shortcuts:** Die Workbench akzeptiert einerseits die dem Übersetzer aus den Textverarbeitungsprogrammen bekannten Shortcuts und bietet ihm andererseits die Möglichkeit, die Funktionen der Workbench mit spezifischen Shortcuts zu bedienen.

Anders beim TranslationManager: das System akzeptiert die bekannten Shortcuts nicht, sondern verwendet dem Übersetzer völlig fremde Shortcuts. Teilweise geraten die im TranslationManager definierten Shortcuts sogar in Konflikt mit den Shortcuts der Textverarbeitungsprogramme.

- **Bearbeitungsfunktionen** (z. B. Suchen und Ersetzen): Die Translator's Workbench stellt dem Übersetzer zum einen die aus dem Textverarbeitungsprogramm bekannte Funktion zur Verfügung und bietet ihm zum anderen eine spezifische Funktion (*File/Maintenance/Global Changes*), die es ihm erlaubt, Änderungen auf Datei-Ebene vorzunehmen.

Beim TranslationManager ist das Verfahren komplizierter: Möchte der Übersetzer Änderungen während des Übersetzungsvorgangs vornehmen, so ist er gezwungen, in den *Post-Editing-Style* zu wechseln. Globale Änderungen sind hier sowohl auf Datei- als auch auf Folder-Ebene<sup>2</sup> möglich – die Datei bzw. sämtliche Dateien des Folders müssen jedoch vorher vollständig bearbeitet worden sein.

Die Translator's Workbench bietet dem Übersetzer aufgrund ihrer Arbeitsumgebung zunächst einen intuitiveren Zugang, jedoch ist TRADOS gezwungen, bei jeder Veränderung der Textverarbeitungsprogramme, die Workbench entsprechend anpassen zu müssen. Der TranslationManager hat hinsichtlich der Benutzerfreundlichkeit derzeit noch einige Defizite.

### 2.1.2 Zugriff auf das TM

- **Minimum Match Value:** Die Translator's Workbench bietet dem Übersetzer die Möglichkeit, den *Minimum Match Value*<sup>3</sup> sowohl für das TM (Abb. 1, gegenüberliegende Seite) als auch für die Terminologieerkennung selber einzustellen (Abb. 2).

Er muß sich dabei natürlich bewußt sein, daß ein niedrigerer Minimum Match Value in der Regel eine geringere Qualität der Übersetzung impliziert. Jedoch kann es für den Übersetzer dennoch interessant sein, von diesen „schlechteren“ Übersetzungs- bzw. Terminologievorschlägen profitieren zu können.

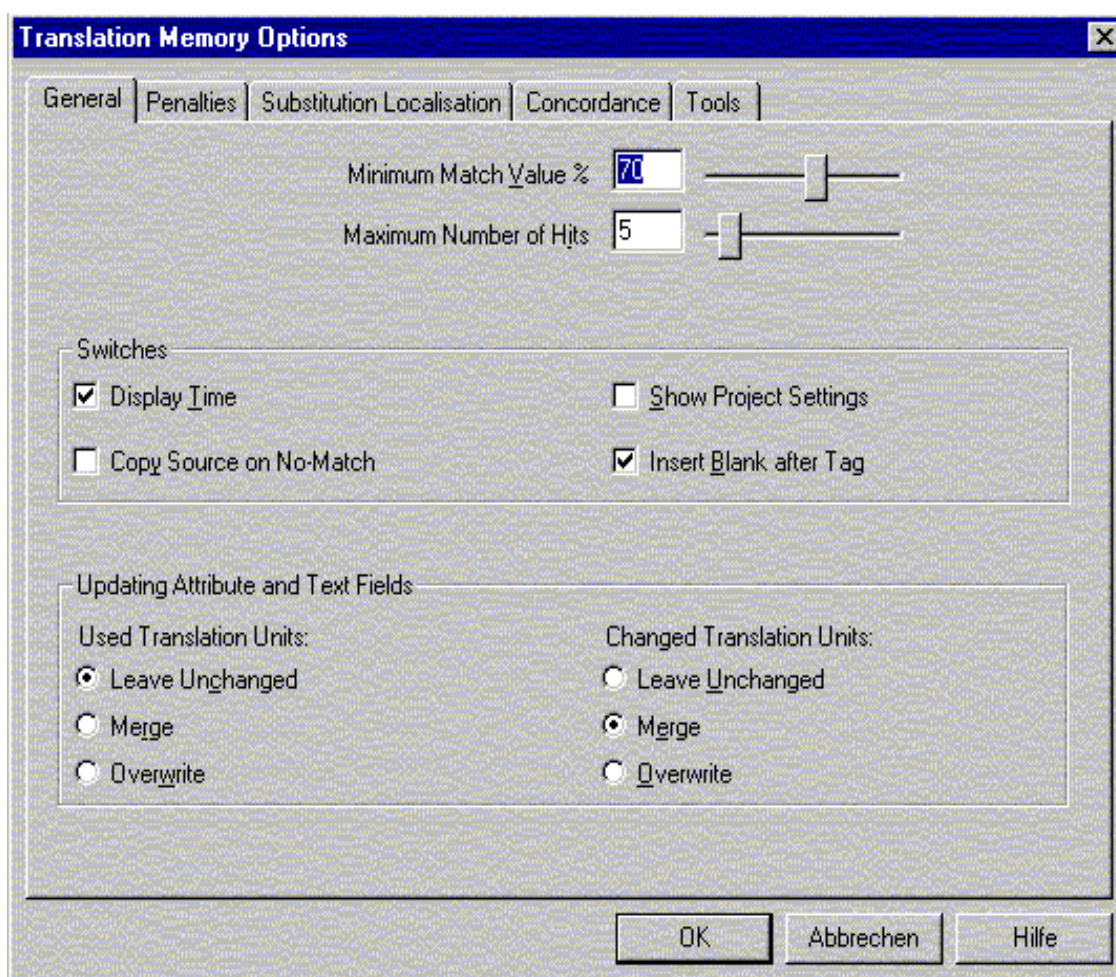


Abb. 1: Angabe des „Minimum Match Value“ für das TM.

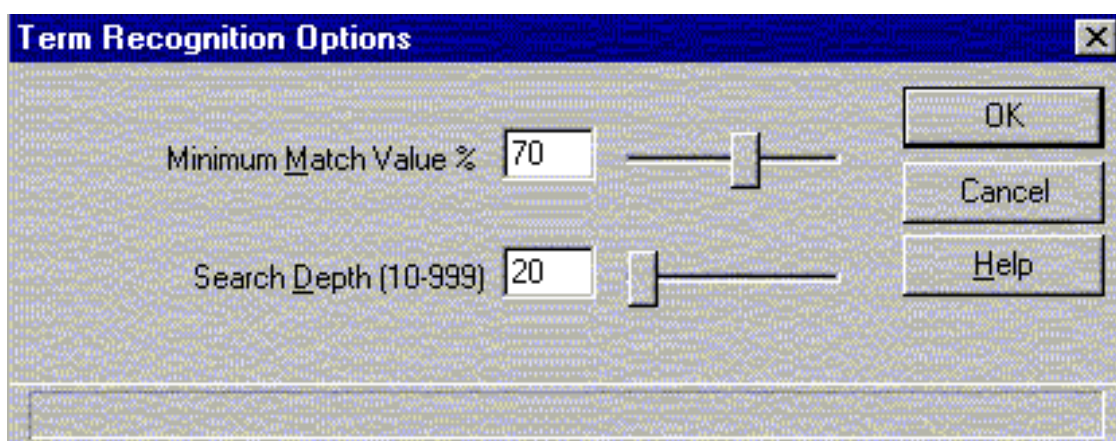


Abb. 2: Angabe des „Minimum Match Value“ für die Terminologieerkennung.



Der TranslationManager bietet diese Möglichkeit nicht. Der Minimum Match Value liegt bei 70 %, kann aber nicht verändert werden.

- **Penalties:** Die Translator's Workbench erlaubt es dem Übersetzer, sogenannte Penalties<sup>4</sup> individuell einzustellen.

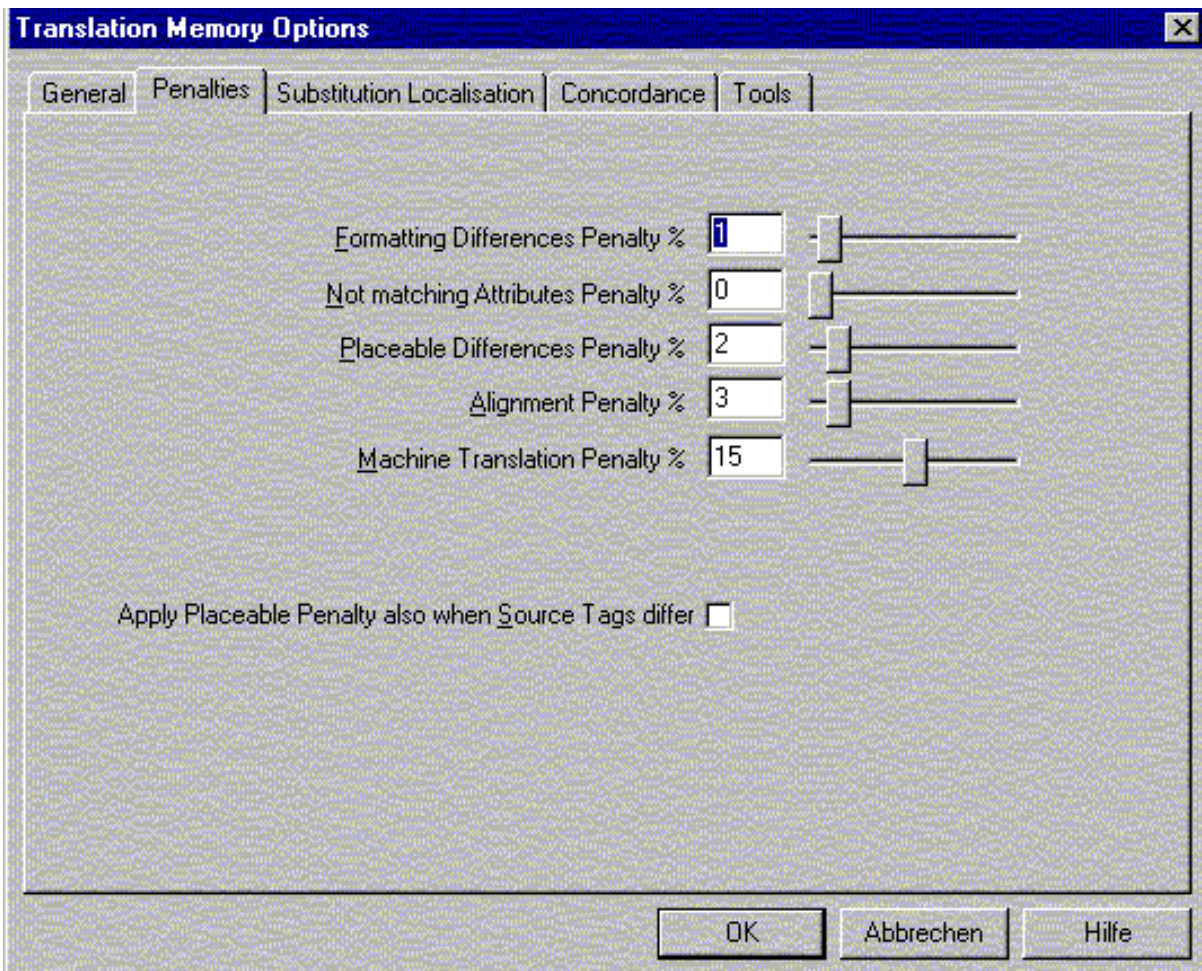


Abb. 3: Individuell spezifizierbare Abzüge für Match-Werte.

Beim TranslationManager ist dies nicht möglich: Weder gibt es eine Differenzierung von Abzügen für Kandidaten aus unterschiedlichen Referenzmaterialien, noch kann man Penalties individuell einstellen.

- **Konkordanzsuche:** Beim Arbeiten mit der Workbench kann der Übersetzer von der Konkordanzsuche, der Suche nach Begriffen oder Phrasen im Kontext, profitieren. Er kann dabei den Minimum Match Value wiederum individuell einstellen.

Der TranslationManager bietet diese Funktion nicht.

### 2.1.3 Zugriff auf Terminologie

Beide Systeme verfügen über eine aktive Terminologieerkennung.

Bei der Translator's Workbench ist es das externe Terminologieverwaltungstool *Multiterm*, welches über eine Schnittstelle in den Übersetzungsprozeß eingebunden werden kann. Die Tatsache, daß es sich bei *Multiterm* noch um eine 16-Bit-Version handelt, was sich durch hohe Suchzeiten bei großen Glossaren bemerkbar macht, bedeutet ein gewisses Defizit, da die Übersetzer dazu tendieren, die Terminologieerkennung zu deaktivieren, was wiederum Mängel in der Konsistenz der Übersetzung zur Folge haben kann.

Der TranslationManager besitzt eine interne Dictionary-Komponente, die für jeden Folder die Auswahl und hierarchische Organisation mehrerer Dictionaries erlaubt.

## 2.2 Erfahrungen des technischen Support

Der technische Support interessiert sich zunächst dafür, welche Dateiformate mit dem jeweiligen TM-System bearbeitet bzw. welche Konvertierungen eventuell notwendig werden können. Weiterhin arbeitet er mit den Analyse- und Vorübersetzungsfunktionen der Systeme und ist zuständig für die Aktualisierung des TM.

### 2.2.1 Dateiformate und Konvertierungen

Mit der Workbench können ausschließlich rtf-Dateien bearbeitet werden. Bei anderen Dateiformaten sind Konvertierungen notwendig (z. B. mit dem S-Tagger bei Interleaf- bzw. FrameMaker-Dateien), die immer ein gewisses Risiko beinhalten.<sup>5</sup>

Der TranslationManager erlaubt ein formatunabhängiges Arbeiten, da er für jedes Dateiformat über ein spezifisches Markup verfügt.

## 2.2.2 Analyse- und Vorübersetzungsfunktion

Die Analysefunktion der Translator's Workbench liefert zum einen die Wort- und Segmentzahl und zum anderen die Anzahl bzw. den prozentualen Anteil an Wiederholungen sowie *No*, *Fuzzy* und *Full Matches*.

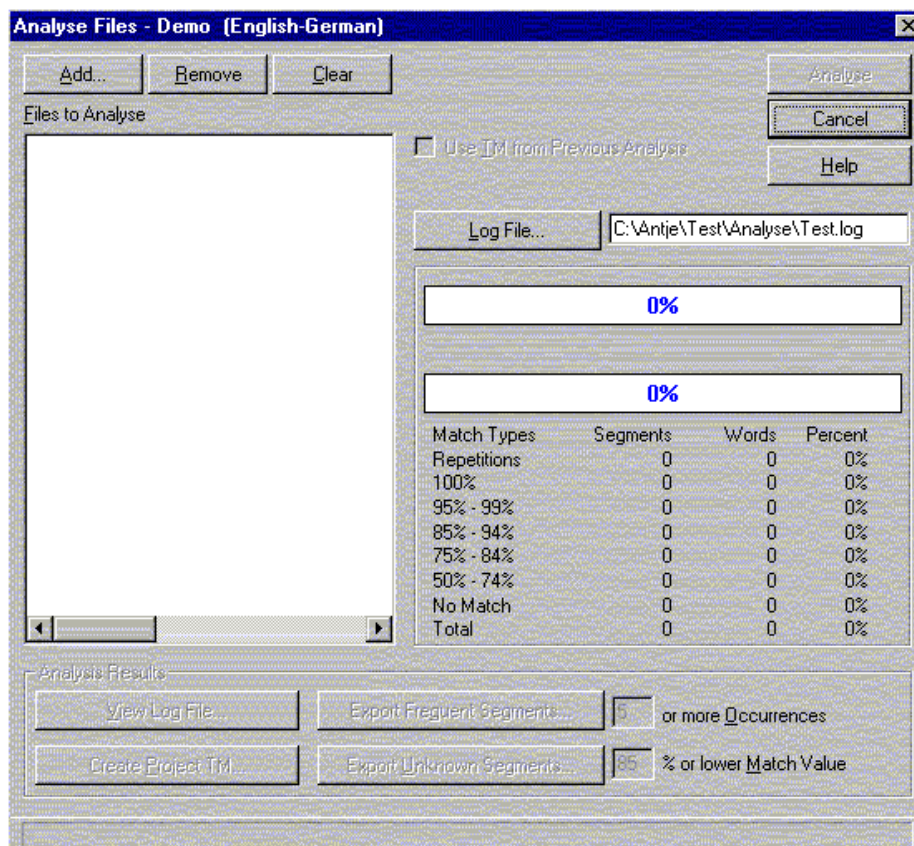


Abb. 4: Dateianalyse mit der Translator's Workbench.

Bei der Vorübersetzung kann der Minimum Match Value individuell eingestellt werden. In der Regel werden allerdings nur Full Matches ersetzt. Es besteht die Möglichkeit, zum einen häufig vorkommende, andererseits aber auch unbekannte Segmente zu exportieren und auf maschinellem Wege übersetzen zu lassen.



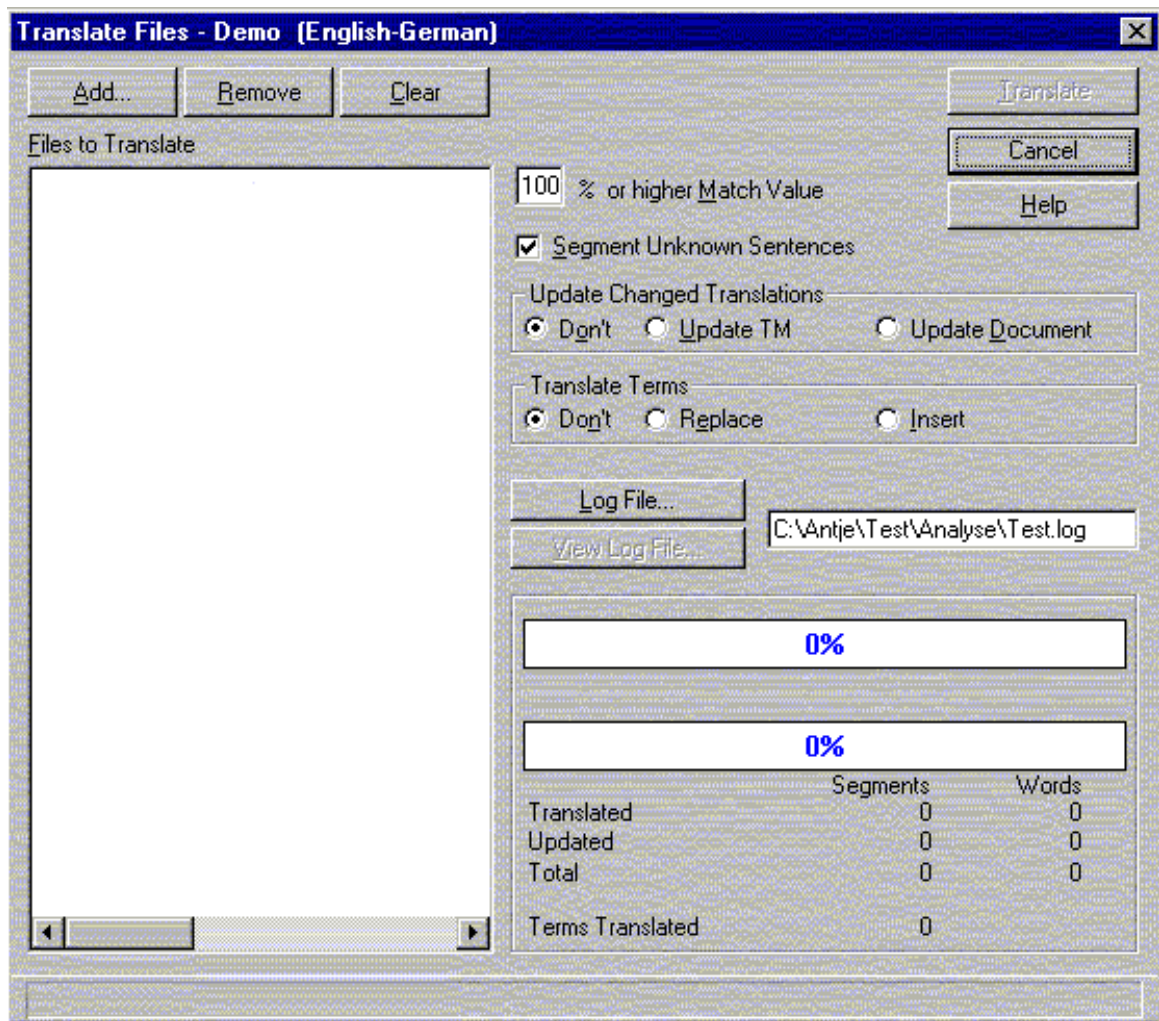


Abb. 5: Einstellungen für die Vorübersetzung.

Beim TranslationManager liefert die WordCount-Funktion die Wortzahl sowie die Anzahl an *No*, *Fuzzy* und *Full Matches*. Die Segmentierung wird erst durch die Analysefunktion durchgeführt, wobei gleichzeitig eine Vorübersetzung durchgeführt werden kann. Es werden ausschließlich Full Matches<sup>6</sup> ersetzt.



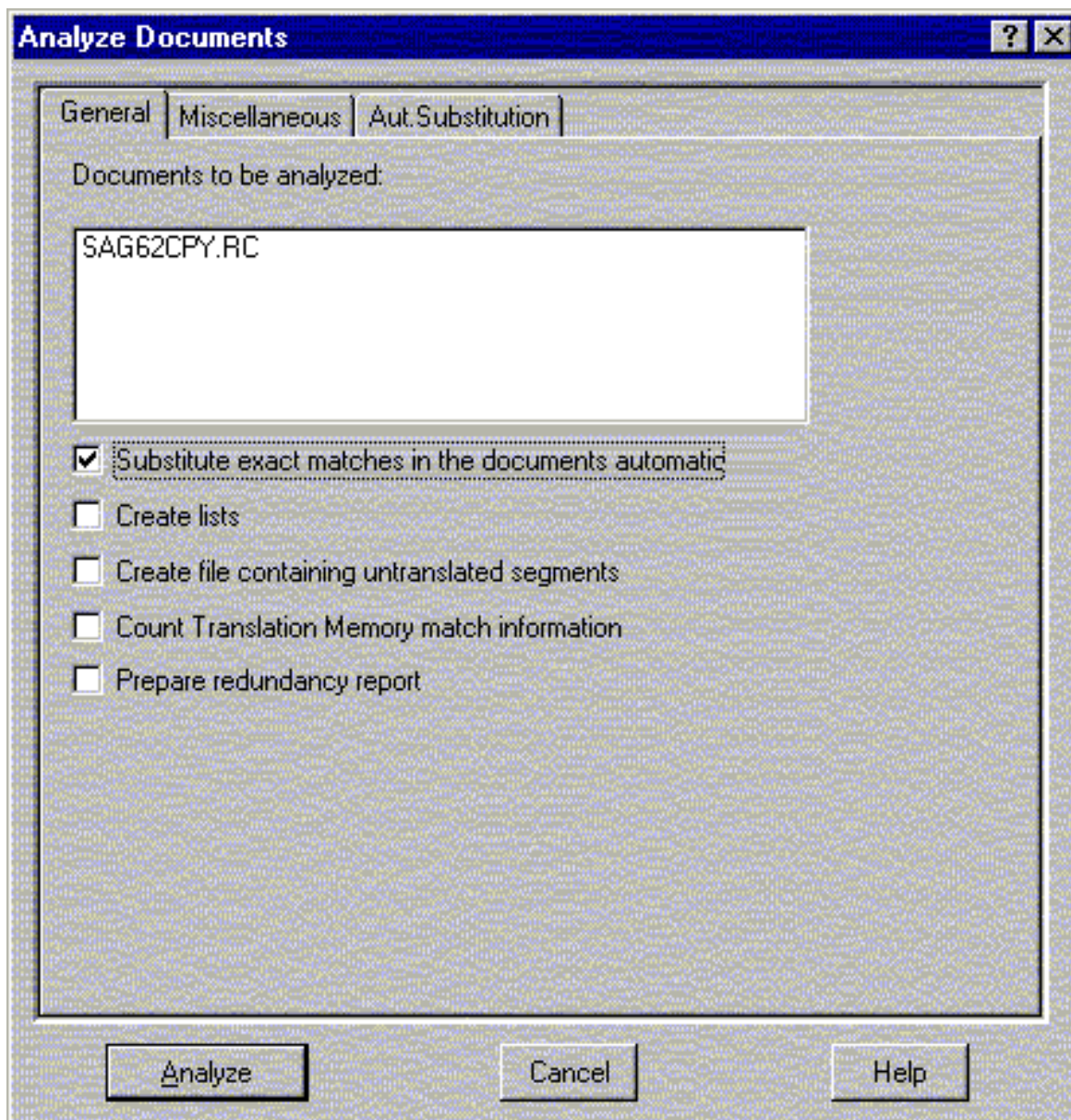


Abb. 6: Dateianalyse mit dem TranslationManager.

### 2.2.3 Aktualisierung des TM

Der Prozeß der Aktualisierung des TM wird bei der Workbench *CleanUp* genannt. Beim *CleanUp* werden die ausgangssprachlichen Segmente aus dem vollständig übersetzten Text entfernt und zusammen mit ihren zielsprachlichen Entsprechungen als Übersetzungseinheit im TM abgespeichert. Dabei werden – dif-

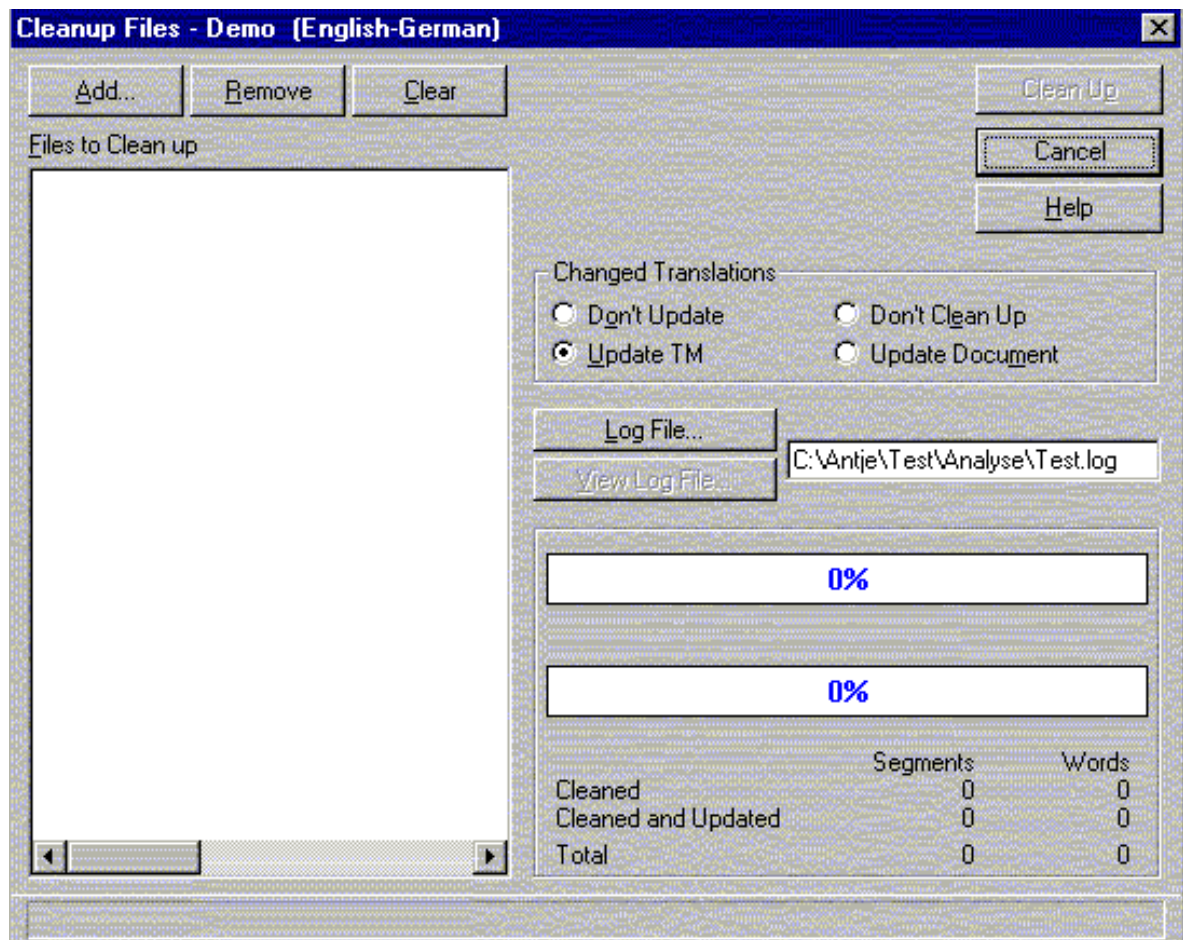


Abb. 7: Aktualisierung des TM bei der Translator's Workbench.

ferenziert nach den den Übersetzungseinheiten zugeordneten projektspezifischen Attributen – schon vorhandene Übersetzungseinheiten überschrieben.

Beim TranslationManager ist der Prozeß um einiges aufwendiger. Die einzelnen Translation Memories müssen zunächst mit dem Befehl *Build archive TM* auf die Aktualisierung vorbereitet werden. Sie werden dann in einem neuen TM zusammengeführt. Sämtliche Einträge dieses TM müssen schließlich – zusätzlich differenziert nach den unterschiedlichen Markups – durchgesehen und unifiziert werden.

## 2.3 Erfahrungen des Projektmanagements

Die praktischen Erfahrungen des Projektmanagements beim Einsatz von TM-Systemen sind eher indirekter Art. Im Projektmanagement treffen die unterschiedlichen Motivationen der Kunden einerseits und der Übersetzer andererseits aufeinander. Es ist die Aufgabe des Projektmanagers, zwischen dem Interesse der Kunden, die Zeit- und Geldaufwand der Übersetzung möglichst gering halten wollen und dem Interesse der Übersetzer, die eine angemessene Bezahlung und die Anerkennung des kreativen Aspektes ihrer Tätigkeit verlangen, zu vermitteln.

Der Konflikt äußert sich vor allem in der Frage der Bezahlung von 100%-Matches. Die Kunden sind oft nicht bereit, für diese zu bezahlen, da sie meinen, der Übersetzer brauche sie nicht zu bearbeiten. Die Praxis zeigt allerdings, daß es keine 100%-Matches gibt. Es gibt in der Übersetzung keine mathematischen Äquivalenzen, sondern es handelt sich bei jeder Übersetzungseinheit um eine semantische Einheit, die im Kontext überprüft werden muß. Logische Konsequenz für den Übersetzer ist es jedoch, diese notwendige Überprüfung nicht vorzunehmen, da sie ihn Zeit kostet, die ihm nicht bezahlt wird. Dies geht zu Lasten der Qualität der Übersetzung. Außerdem kann bei den Übersetzern Demotivation hervorgerufen werden, denn ihre vorher kreative Übersetzeraufgabe wird reduziert auf ein simples Bestätigen der 100%-Matches.<sup>7</sup>

## 3 Fazit

Viele Faktoren beeinflussen die Entscheidung für oder gegen den Einsatz von TM-Systemen bzw. die Auswahl eines bestimmten TM-Systems.

Von Bedeutung ist zunächst die Akzeptanz des TM-Systems bei den Übersetzern, die vor allem von der Benutzerfreundlichkeit des Systems beeinflußt wird. Der technische Support interessiert sich für die Vor- und Nachteile im Vergleich der technischen Leistungen, primär bezüglich der unterstützten Dateiformate und hinsichtlich der Möglichkeiten der Pflege und Aktualisierung des TM.

In der Praxis steht der Benutzerfreundlichkeit und dem technischen Leistungsvermögen eines TM-Systems jedoch immer die wirtschaftliche Betrachtung gegenüber.



---

**ANMERKUNGEN**

- <sup>1</sup> Die Segmentierung des zu übersetzenden Textes nach Sätzen und damit die Bildung der Übersetzungseinheiten aus in sich schlüssigen Satzeinheiten hat sich in der Praxis bewährt. Dennoch besteht die Möglichkeit, bei asiatischen Sprachen sogar die Notwendigkeit, den zu übersetzenden Text nach Absätzen zu segmentieren.
- <sup>2</sup> Bei TranslationManager werden die Dateien in Foldern (Ordnern) organisiert.
- <sup>3</sup> Der *Minimum Match Value* ist ein prozentualer Schwellenwert, ab welchem Übersetzungseinheiten bzw. Terminologie dem Übersetzer als Vorschlag angezeigt werden. Im allgemeinen wird ein Wert von 70 % empfohlen.
- <sup>4</sup> Penalties sind prozentuale Abzüge vom Match-Wert auf Grund von Unterschieden in der Formatierung, den Attributen oder den *Placeables* (nicht zu übersetzende Texteinheiten wie Zahlen oder Datumsangaben) bzw. auf Grund von Alignment oder maschineller Übersetzung.
- <sup>5</sup> Für die Bearbeitung von HTML-Dateien hat TRADOS inzwischen den Tag-Editor entwickelt.
- <sup>6</sup> Der TranslationManager unterscheidet bei den Full Matches nach dem Kontext in *exact context match* und *last exact context match* bzw. *Full Matches* aus sogenannten *joined segments*.
- <sup>7</sup> In verschärfter Form existiert diese Problematik im Falle von maschineller Übersetzung, wo den Übersetzern lediglich die Aufgabe des Post-Editing bleibt und es für sie außerdem meist weit schwieriger ist, die falsch übersetzten Sätze zu korrigieren als diese selber zu übersetzen.