

Übersetzungsunterstützende Systeme

Seminar Computerlinguistik

Referentin:

Virginija Siderkeviciute

INHALT

1. Einleitung
2. Begriffe und Grundlagen
 - 2.1 Funktionsweise
3. Vorhandene Programme
4. Evaluationskriterien
5. Fazit
6. Literatur

1. Einleitung

Übersetzungsunterstützende Systeme

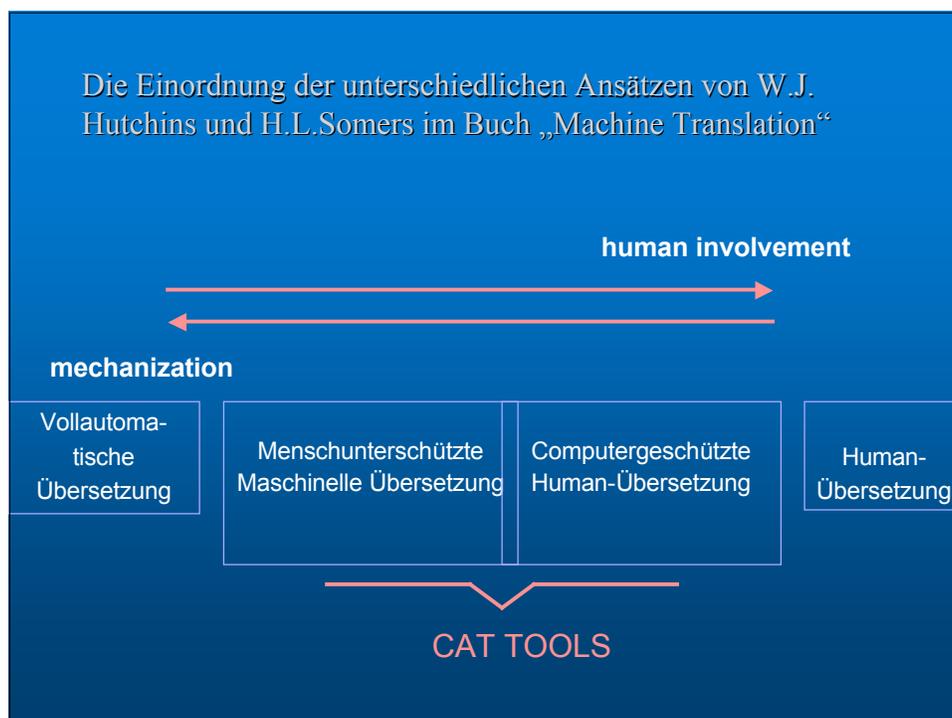
Den Übersetzer unterstützende Systeme (bzw. Begriffe)

1. Maschinelle Hilfen für den Übersetzer
2. Maschinelle Übersetzungshilfen
3. Computergeschützte (Human-) Übersetzung
4. Mensch-unterstützte maschinelle Übersetzung
5. Maschinelle Übersetzung

Die Computergeschützte Übersetzungssystemen (Strategien):

1. Präedition
2. Postedition

Die Einordnung der unterschiedlichen Ansätzen von W.J. Hutchins und H.L.Somers im Buch „Machine Translation“



Translation Memory

Translation Memories (TM's) stellen bereits übersetzte Satz- und Segmentpaare in der Ausgangs- und Zielsprache zur Verfügung.

Ein ähnlicher oder identischer Satz wird nicht mehr übersetzt, sondern kann direkt übernommen oder noch einmal überarbeitet werden.

Nicht bekannte Sätze überträgt der Übersetzer einmal in die Zielsprache. Diese werden in einer Datenbank abgelegt und stehen für weitere Anwendungen zur Verfügung.

2. Begriffe und Grundlagen

Übersetzungseinheit / translation units: abgespeicherte Satz- und Segmentpaare

Translationdatenbanken: Ansammlung von Übersetzungseinheiten

Fuzzy matching: "Unschärfes Suchen" = ermöglicht das Auffinden von Daten, die nur eine gewisse Ähnlichkeit mit dem Suchargument aufweisen

Exact matches : 100% übereinstimmende Daten

Automatische Terminologieunterstützung: Das System sucht in der Terminologiedatenbank nach bereits übersetzten Einzelbegriffen, bzw. neu übersetzte Einzelbegriffe werden dort abgelegt.

Alignment: nachträgliches Segmentieren von bereits vorhandenen nicht in TM erstellten Übersetzungen

Tagger: Modul, das bestimmte Formate wie Framemaker oder Interleaf in das Format des Übersetzungsprogramms konvertiert.

Filter: dienen zur Abgrenzung unterschiedlicher Terminologien (Projektbezogene Arbeit)

2.1 Funktionsweise

- Ausgangspunkt: die vom Auftraggeber der Übersetzung gelieferten Dateien.
- Der Ausgangstext wird in einen Editor übernommen.
- Der Text wird segmentiert.
- Die Software sucht nach gleichen oder ähnlichen Segmenten.
- Ergebnis: Exact Matches oder Fuzzy Matches.
- Dem Übersetzer wird eine automatische Terminologieunterstützung geliefert.

2.1 Funktionsweise

- Die Angebote Lösung \neq dem Ausgangssatz, überträgt der Übersetzer den Satz neu in die Zielsprache.
- Die neue Übersetzung wird in das TM übernommen.
- Am Anfang der Arbeit mit dem TM ist dieses zunächst noch leer und wird im Laufe der Zeit gefüllt.
- Alignment Tool übernimmt bereits angefertigte Übersetzungen in das TM.

2.2. Einsatzbereiche und Vorteile

- ✂ Hat man eine TM-Datenbank erst einmal gefüllt, ergibt sich daraus eine Arbeits- und damit Zeit- und Kostenersparnis.
- ✂ Ein TM garantiert die Qualität und Konsistenz der Übersetzungen.
- ✂ TMs ermöglichen durch die Nutzung bestimmter Filter darüber hinaus das Bearbeiten von Formaten, die andernfalls nicht ohne weiteres importierbar oder exportierbar wären (z.B. Framemaker-Dateien)
- ✂ Der Einsatz von TM bringt die grössten Vorteile, wenn ein hoher Wiederholungsgrad innerhalb eines Textes oder einen ganzen Reihe von Texten (z.B. bei Updates) gegeben ist.

3. Vorhandene Programme

- ✂ TRANSLATORS WORKBENCH, Fa. Trados, Stuttgart
<http://www.trados.com/>
- ✂ TRANSIT, Fa. Star GmbH, Böblingen
<http://www.star-ag.ch/>
- ✂ TRANSLATION MANAGER, Fa. IBM, Heidelberg
<http://www.ibm.com/software/ad/translate/tm/>
- ✂ DEJA-VU, Fa. Atril
<http://www.atril.com/>
- ✂ SDLX, Fa. SDL-International,
<http://www.sdlintl.com/>
- ✂ OPTIMIZER, Fa. Eurolang, München
- ✂ PERSONAL TRANSLATOR 2000, die linguatéc Sprach-
technologien GmbH
<http://www.linguatéc.de>
- ✂ MoBiMem, Fa. MorphoLogic
www.morphologic.hu

3.1. Star Transit

Module

- [Translation Memory](#)
- [Translation Editor](#)
- [TermStar](#)
- [Alignment Tool](#)

Translation Editor

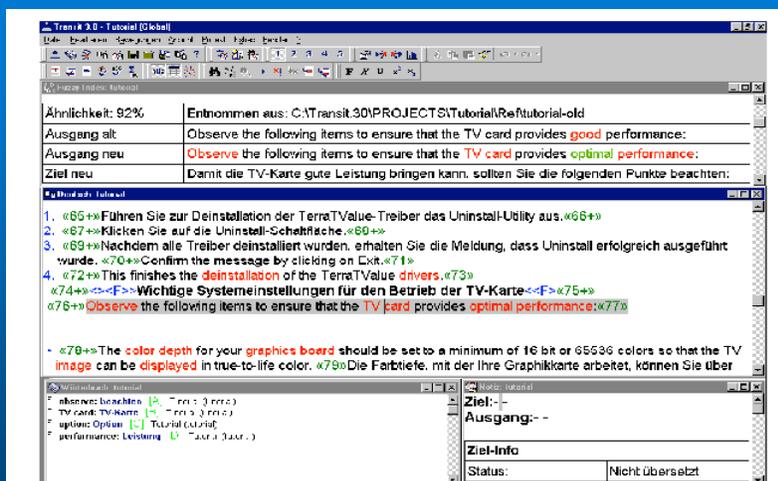


Abb. Transit-Editor 3.0 mit Zieltext-Fenster, Fuzzy-Index Fenster, Wörterbuchfenster und Notizfenster. Das Ausgangstext-Fenster liegt im Hintergrund.

TermStar

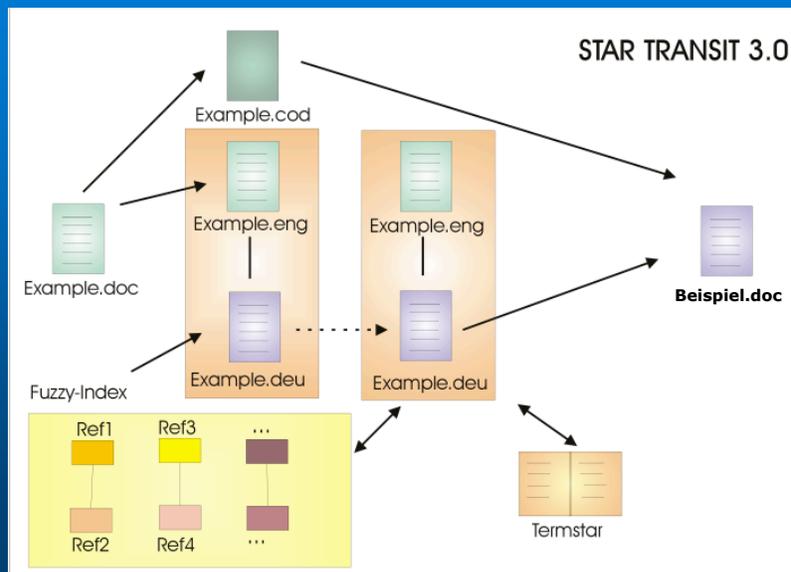
- **TermStar** - das eigenständige Terminologieverwaltungssystem.
- Es arbeitet im Hintergrund und sucht nach bereits in der Datenbank vorhandenen Terminologie.
- Die Sprachrichtung kann jederzeit geändert werden.
- Noch nicht erfaßte Termini können jederzeit in eine Datenbank aufgenommen werden.

Transit

Transit ist eine Arbeitsoberfläche mit drei integrierten Komponenten. Der Übersetzungsprozess findet im integrierten Editor statt.

Das **Alignment Tool**, das bereits angefertigte Übersetzungen formatiert, in Segmente zerlegt, durchnummeriert und als Referenzdateipaar nutzt.

3.2. Arbeitsweise von Transit 3.0



3.3. Translator`s Workbench

Module:

Translation Memory

MultiTerm

Schnittstelle zur Textverarbeitung

Transcend

WinAlign

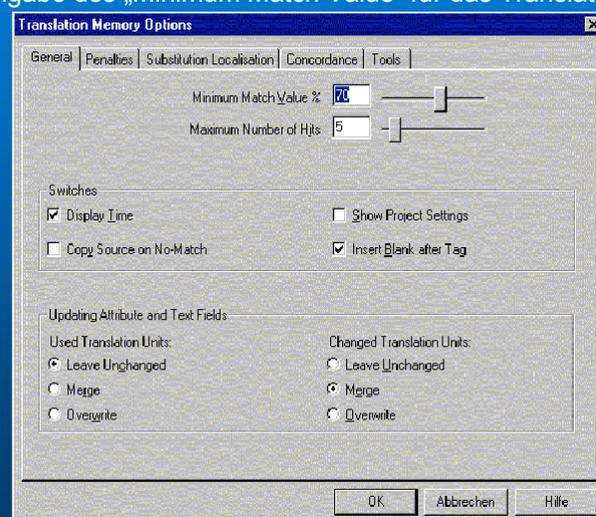
Transcend

- ist eine PC-Software zur vollautomatischen Übersetzung der Texte ins Englische und aus dem Englischen.
- liefert keine Wort-für-Wort - Übersetzung, sondern analysiert morphologische und syntaktische Strukturen sowie semantische Merkmale.
- erstellt Rohübersetzungen, die überarbeitet werden müssen.
- ist "interagiert" mit den Modulen Translation Memory und MultiTerm.
- arbeitet mit eigenen, komplex aufgebauten Wörterbüchern, und zwar einem nicht zugänglichen Hauptwörterbuch sowie mit erweiterbaren Benutzerwörterbüchern.

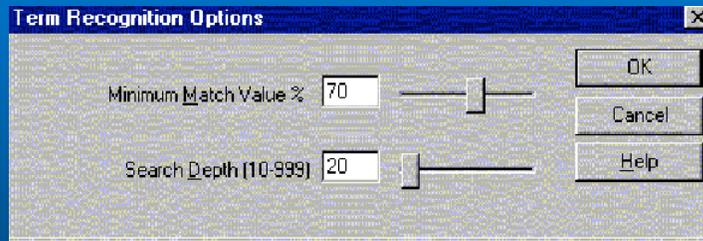
3.3.1. Arbeitsweise der Translator's Workbench

Zugriff auf das TM:

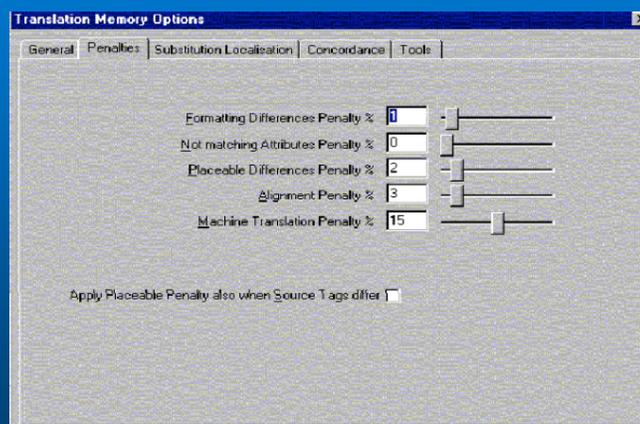
Angabe des „Minimum Match Value“ für das Translation Memory



Angabe des "Minimum Match Value" für Terminologieerkennung



Penalties- individuell spezifizierbare Abzüge für Match-Werte.



4. Evaluationskriterien

Die Systeme werden nach folgenden Kriterien begutachtet:

Funktionalität (Functionality): Fähigkeit der Software, Methoden und Techniken zu gegebenen Anforderungen des Benutzers bereitzustellen; Fähigkeit der Software, korrekte bzw. genaue Ergebnisse zu liefern

Zuverlässigkeit (Reliability): Ist das System fehlerfrei? Wie werden Störungen behandelt?

Benutzbarkeit (Usability): Ist das System einfach/verständlich zu bedienen? Wie groß ist die Einarbeitungszeit bzw. Umgewöhnungszeit bei Systemwechsel?

Effizienz (Efficiency): Ausnutzung von Ressourcen; Reaktionszeit des Systems

Wartbarkeit (Maintainability): Welche Handlungen sind zur Pflege des Systems erforderlich? Welchen Support bietet der Hersteller?

Portierbarkeit (Portability): Welche Möglichkeiten zur Zusammenarbeit mit anderen Systemen bieten sich? Welche Anforderungen an vorhandene Hard- bzw. Software werden gestellt?

Die Evaluation wird auf die Funktionen konzentriert

Ist es notwendig, ein TM zu benutzen, oder eine andere Methode erfolgversprechender ist?

(Ausgangstext weist Wiederholungen auf und ist reich an Terminologie. Ist Konsistenz innerhalb des Dokuments nötig?)

Welches TM-System zu benutzen ist? Ein Kriterium ist die benutzte Sprache. Obwohl TMs KEIN linguistisches Wissen benutzen, müssen beispielsweise notwendige Zeichensätze unterstützt werden. Dafür gibt es noch keinen Standard. Man muß untersuchen, welche Formate das jeweilige System unterstützt.

Welche Software und Hardware ist nötig? Womit hat der Benutzer Erfahrung?

Arbeitet ein einzelner Übersetzer oder ein Team? (Im Fall des Teams ist von Bedeutung, ob alle lokal in der Nähe und greifbar oder verstreut an verschiedenen Orten sind. Müssen Arbeiten übers Netzwerk möglich sein? Arbeiten Leute u.U. gleichzeitig am selben Dokument? Kann man in dem System individuelle Lese-/Schreibrechte vergeben? Ist Exportieren von TMs möglich, um sie beispielsweise auszuliefern? Projektspezifische TMs möglich?)

5. Fazit

- Die starke Verbreitung der TM-Werkzeuge
- Die eingesetzte TM-Technik der Programme unterscheidet sich nicht wesentlich und ist von der Grundidee her einfach.
- Die Leistungen der Produkte liegen in der Integration in die Arbeitsumgebung des Übersetzers, der Verknüpfung mit Datenbanken und anderen Übersetzungswerkzeugen und der Projektverwaltung.
- Die individuelle Evaluation des Marktangebots ist zu aufwändig. Man muss sich auf Werbeversprechen verlassen.

6. Literatur

1. *R. Nübel, U. Seewald-Heeg: Translation-Memory-Module automatischer Übersetzungssysteme, in LDV-Forum – Forum der Gesellschaft für Linguistische Datenverarbeitung Band 16 • Nr. 1/2 Dezember 1999, S.16-36*
2. *H. Bohn: Translator's Workbench: Funktionalität – Entwicklungen – Kundenanforderungen, in LDV-Forum – Forum der Gesellschaft für Linguistische Datenverarbeitung Band 16 • Nr. 1/2 Dezember 1999, S.36-41*
3. *J.Klein: Professionelles Übersetzen mit STAR Transit, in LDV-Forum – Forum der Gesellschaft für Linguistische Datenverarbeitung Band 16 • Nr. 1/2 Dezember 1999, S.41-52*
4. *A. Pesch: Erfahrungen im Einsatz mit Translation-Memory-Systemen, in LDV-Forum – Forum der Gesellschaft für Linguistische Datenverarbeitung Band 16 • Nr. 1/2 Dezember 1999, S.52-64*
5. *U. Reinke: Evaluierung der linguistischen Leistungsfähigkeit von Translation Memory-Systemen – ein Erfahrungsbericht, in LDV-Forum – Forum der Gesellschaft für Linguistische Datenverarbeitung Band 16 • Nr. 1/2 Dezember 1999 S. 100-118*
6. *A. Blatt, K.-H. Freigang u.a.: Computer und Übersetzen. Eine Einführung. G.Olms Verlag: Hildesheim.Zürich.New York 1985, S.71-79*
7. *Gerd Willée, B. Schröder: Computerlinguistik. Was geht, was kommt? Gardez Verlag: Sankt Augustin 2002, S. 263-268*
8. *H.D. Luckhardt, H.H. Zimmermann: Computergeschützte und maschinelle Übersetzung. A.Q-Verlag, Augsburg 1991, S.9-16*
9. *Martina Schwanke: Maschinelle Übersetzung, Springer Verlag: Berlin Heidelberg New York 1991, S.47-67*