

# Analyse des Systems zur computergestützten Übersetzung Weidner – Version Französisch-Englisch 2.5

Analyse des Systems zur computergestützten Übersetzung Weidner – Version Französisch-Englisch 2.5 .....		1
Danksagung .....		2
Zusammenfassung .....		3
Abstract.....		4
VORBEMERKUNG.....		5
EINLEITUNG.....		6
1. Geschichte von Weidner .....		7
2. Anwender von Weidner .....		9
2.1 Translation Bureau Canada.....		9
2.2 ITT.....		10
2.3 Weidner .....		11
2.4 Mittel.....		13
3. Weidners Konkurrenten .....		13
3.1 Logos Computer Systems .....		14
3.2 Automatic Language Processing Systems (ALPS) .....		15
3.3 SMART AI .....		17
4. AUFBAU VON WEIDNER.....		18
4.1 Die Hardware.....		18
4.2 Die Programme .....		18
5. Vorbearbeitung (Präedition) .....		40
6. Nachbearbeitung (Postedition).....		43
7. Auswertung.....		45
8. GESAMTBEWERTUNG.....		65
8.1 Häufigkeit bestimmter Fehler .....		65
8.2 Lexik .....		71
8.3 Schlußfolgerung .....		71

## ***Danksagung***

Mein besonderer Dank gilt den EDV-Spezialisten Eckhardt Henne und Markus Goldberg, mit deren tatkräftiger Hilfe es mir gelang, mit den Tücken der Text- und Datenverarbeitung fertigzuwerden.

## **Zusammenfassung**

Kommerzielle Systeme zur maschinellen Übersetzung sind heute in einer wachsenden Zahl von Übersetzungsabteilungen zu finden. Es ist daher für Übersetzer von Bedeutung, über Kenntnisse der Funktionsweise und der Qualität der Übersetzungen solcher Systeme zu verfügen.

Diese Arbeit befaßt sich mit dem System zur maschinell gestützten Übersetzung Weidner, Version Französisch-Englisch 2.5. Sie liefert einen Abriß der Geschichte des Systems Weidner und der seiner Konkurrenten, sowie eine Beschreibung dieser Systeme. Eine eingehende Beschreibung der Weidnerschen Systemkomponenten und ihrer Funktionsweise schließt sich an.

Darüberhinaus wird eine detaillierte Analyse von 18 maschinell übersetzten Texten sowie eine allgemeine Beurteilung des Gesamtsystems auf der Grundlage der Analyseergebnisse geliefert.

## ***Abstract***

Commercial machine translation systems have found their way into a growing number of translation departments. It is therefore important for translators to know about the way these systems work and what their output is like.

This study describes the Weidner Computer-Aided Translation System, Version French-to-English 2.5. A history of Weidner and its competitors as well as an outline of those systems is given, followed by a detailed description of Weidner's system components and the way they work. There is a detailed analysis of 18 machine translated texts and a general assessment of the entire system based on the results of that analysis.

## ***Vorbemerkung***

Folgende Abkürzungen werden in dieser Arbeit verwendet:

NB: Nachbearbeitung

NBA: Nachbearbeitungsaufwand

WB: Wörterbuch

## **Einleitung**

Diese Arbeit befaßt sich mit dem System zur maschinell gestützten Übersetzung Weidner, Version Französisch -> Englisch 2.5. Während meines Aufenthalts am University of Manchester Institute of Science and Technology in England hatte ich Gelegenheit, dieses System zu testen. Im Laufe zweier Monate ließ ich zehn im Schnitt zwei Seiten lange Texte aus den Bereichen Wirtschaft, Technik und Sprachwissenschaft von Weidner ins Englische übersetzen und führte Vokabelsuchläufe sowie Ergänzungen des Wörterbuchs durch. Es kam mir hierbei darauf an, herauszufinden, wie hoch die Qualität der von Weidner gelieferten Rohübersetzung zu bewerten ist, wo System-schwachstellen liegen, wie häufig diese Schwächen zu wie schweren Fehlern führen, bei welchen Textsorten sich welche Systemschwächen besonders bemerkbar machen und wie stark die Verständlichkeit der Übersetzungen dadurch beeinträchtigt wird. Auf dieser Grundlage sollen außerdem Vermutungen darüber angestellt werden, wie hoch sich die Akzeptanz eines Systems wie Weidner, das auch stellvertretend für andere im kommerziellen Einsatz befindliche Systeme steht, die im Rahmen dieser Arbeit ebenfalls angesprochen werden sollen, bei den Übersetzungsabnehmern einerseits und den Übersetzern andererseits voraussichtlich entwickeln wird. Als Anhaltspunkte sollen auch einige wenige in der Literatur zu findende Berichte von Weidner-Anwendern dienen.

Die Auswertung des oben erwähnten kleinen englisch-französischen Übersetzungskorpus habe ich in Form einer „Fehlerdatenbank“ vorgenommen, die ca. 470 aus den Übersetzungen herausgegriffene mehr oder weniger fehlerhafte Passagen enthält, die die Grundlage meiner Bewertung des Systems bilden. Die einzelnen Einträge dieser Datenbank enthalten Einstufungen der Verständlichkeit der falschen Übersetzung, eine linguistische Bewertung, die Einschätzung des Korrekturaufwandes sowie erläuternde Bemerkungen.

Weitere acht Texte aus den Bereichen Technische Chemie, Maschinelle Übersetzung, Datenverarbeitung, Nachrichtentechnik und Europapolitik wurden bei einem zweiten Manchesteraufenthalt vom 1.7. bis zum 9.7. 1987 bearbeitet. Zwei dieser Texte liegen

auf Französisch, Deutsch und Englisch vor, drei auf Französisch und Englisch, drei nur auf Französisch, davon einer von mir selbst, der ausgewählte Beispielsätze enthält. Die letzten beiden Texte wurden wählte ich aus, da sie mir aus Übersetzungslehrveranstaltungen vertraut waren.

Für diese Texte habe ich ein etwas anderes Auswertungsverfahren gewählt. Die Texte wurden satzweise in Tabellen eingetragen, die nach den Kategorien Syntax, Grammatik, Lexik, Stil und Zeichensetzung aufgeteilt sind. Im Anschluß an die tabellarische Aufarbeitung jedes Textes findet sich die Bewertung der Übersetzung hinsichtlich der Verständlichkeit und des Nachbearbeitungsaufwandes insgesamt anhand der Tabellen.

Eine vollgültige Bewertung der Leistungsfähigkeit des Systems Weidner geht über den Rahmen dieser Arbeit hinaus. Sie soll in erster Linie eine phänomenologische Beschreibung der Arbeit mit einem rechnergestützten Übersetzungssystem liefern, die sich an den real auftretenden Bedingungen in der Übersetzungspraxis orientiert, d. h., den Faktor Mensch mitberücksichtigt. Es besteht ein Unterschied zwischen dem, was ein Übersetzungssystem unter idealen Bedingungen leisten kann - beispielsweise mit fehlerfreien und perfekten Wörterbüchern - und den tatsächlich erzielbaren Übersetzungsergebnissen. Die Qualität jeder maschinell erstellten Übersetzung läßt sich durch Verbesserung der Voraussetzungen - beispielsweise durch eine weitergehende Vorbearbeitung und Wörterbuchverbesserung - noch erhöhen.

## **1. *Geschichte von Weidner*<sup>1</sup>**

Die Firma Weidner wurde 1977 von Bruce Weidner gegründet. Es handelte sich um die erste kommerzielle Anwendung des Projektes der Mormonen zur maschinellen Übersetzung an der Brigham Young University in Provo, Utah, USA. Die ersten Systeme wurden 1980 auf den Markt gebracht. Hierbei handelte es sich in der Regel um Versionen, die auf Kleinrechnern der US-Firma Digital Equipment Corp. liefen. Aufgrund des langsamen Fortschritts bei der Entwicklung geriet die Firma Weidner 1982 in

---

<sup>1</sup> Vgl. Tim Johnson: *Natural Language Computing: The commercial applications*, London 1985: 165, 170, 335-336 sowie John Hutchins: *Machine Translation: Past, present, future*, Chichester 1986

finanzielle Schwierigkeiten und wurde in der Folge von einem ihrer japanischen Kunden, der Übersetzungsfirma BRAVIS International, übernommen, die die Aktienmehrheit erwarb und auf der Grundlage der Weidner-Technologie in Zusammenarbeit mit dem führenden japanischen Softwarehaus CSK ein japanisch->englisches Übersetzungssystem zu entwickeln begann. In diesem Jahr verlegte Weidner seinen Hauptsitz von Provo nach Northbrook bei Chicago. Die Forschungs- und Entwicklungsabteilung blieb in Provo.

1984 kündigte Weidner die erste Systemversion an, die auf einem Microcomputer lief, dem um eine 20 Megabyte-Festplatte erweiterten IBM-PC-XT. Dieses System wurde entsprechend „MicroCat“ genannt - für „Micro-Computer Assisted Translation“ - Übersetzung per Mikrocomputer. Das bereits auf dem Markt befindliche System hieß fortan „MacroCat“, da es, wie bereits erwähnt, auf einem größeren Rechner läuft. Die Übersetzungsgeschwindigkeit von „MacroCat“ liegt bei 14 000 Wörtern pro Stunde, die von „MicroCat“ bei 1 600 W/h.

1985 waren folgende Weidner-Versionen auf dem Markt:

Englisch->Spanisch und Spanisch->Englisch

Englisch->Französisch

Englisch->Portugiesisch

Englisch->Italienisch

Englisch->Deutsch

Nur auf „MacroCat“ war zu haben:

Englisch->Arabisch

Folgende Versionen befanden sich in der Entwicklung:

Französisch->Englisch

Deutsch->Englisch



Deutsch->Französisch

Französisch->Spanisch

Japanisch->Englisch

## **2. Anwender von Weidner**

Bedeutende europäische Anwender von Weidner sind die Firmen ITT (International Telephone & Telegraph Co., Harlow, England), ICL (International Computers Ltd.) und DEC (Digital Equipment Corp.) Publierte Erfahrungsberichte liegen bisher vor von der kanadischen Firma Mitel (Telefonbau) <sup>2</sup>, der Firma ITT <sup>3</sup>, dem kanadischen staatlichen Übersetzungsbüro <sup>4</sup> sowie dem eigenen Übersetzungsbüro der Firma Weidner <sup>5</sup>.

### **2.1 Translation Bureau Canada**

Das kanadische „Translation Bureau“ untersteht dem kanadischen

---

2 Michael G. Hundt: Working with the Weidner machine-aided translation system, in: Veronica Lawson (Hg): Translating and the computer 4 - Practical experience with machine translation, London 1982:45-51

3 Ulla Magnusson-Murray: Operational experience of a machine translation service, in Veronica Lawson (Hg): Translating and the computer 5 - Tools for the trade, London 1983, S.171-180; Tim Johnson ebd:283-286

4 Translation Bureau Canada, Project No. 5-5462, 1985, Trial of the Weidner computer-assisted translation system, o.J., o.O.

5 Henrietta Pons: WCC's translation bureau, in: Veronica Lawson 1982, ebd:165-169

Außenministerium und führte vom 11. Mai bis zum 6. September 1985 in Zusammenarbeit mit dem kanadischen Amt für Statistik und den Ministerien für Verteidigung und Landwirtschaft einen Versuch mit der englisch -> französischen Version von Weidner durch, an dem sechs Übersetzer teilnahmen, die vor Versuchsbeginn eine fünftägige Einführung in das System erhielten. Bei diesem Versuch sollte festgestellt werden, ob 1) Weidner die Produktivität der Übersetzer erhöhte, 2) wie hoch die relativen Kosten der computergestützten Übersetzung seien und 3) wie ihr Einfluß auf die Übersetzer und ihre Aufgaben sei. Die bei dem Versuch angewandte Methode bestand darin, Übersetzungen zuerst „von Hand“ und anschließend von Weidner erstellen zu lassen und diese dann auf die Wirtschaftlichkeit der Anfertigung und ihre Qualität zu prüfen. Auf

diese Weise wurden 24 Texte mit 15760 Wörtern übersetzt. Diese Texte wiesen eine verhältnismäßig einfache Struktur auf und entstammten den Bereichen Landwirtschaft, Arbeitswelt, Bildungs- und Militärwesen. Trotz dieser die maschinelle Übersetzung begünstigenden Verfahrensweise stellte sich heraus, daß die „Humanübersetzung“ zwischen 30% und 34% schneller war; diese kam im Schnitt auf 10,77 Wörter/min., während es Weidner mit Nachbearbeitung auf 7,09 bis 7,54 Wörter/min. brachte. Für einen wirtschaftlichen Einsatz von Weidner wäre nach Auffassung der Organisatoren des Versuchs eine Erhöhung der Übersetzungsleistung um 12.1% und 23.6% erforderlich; um ihn aber lohnend zu machen, sei eine Verdoppelung der Leistung notwendig.

Trotz der im Laufe des Versuchs entstandenen umfangreichen Fachwörterbücher war keiner der sechs daran beteiligten Übersetzer daran interessiert, weiterhin mit Weidner zu arbeiten.

## **2.2 ITT**

Aufgrund des Engagements der ehemaligen Leiterin des ITT-Sprachendienstes, Ulla Magnusson Murray, schaffte ITT 1981 ein Weidner-System an, und zwar zunächst für eine sechsmonatige Erprobungsphase, während derer das System geleast wurde. In dieser Phase wurden Fragen der Ergonomie, Kommunikation, Verarbeitungszeit und Kosten geprüft. Es wurden bewußt Übersetzer ausgewählt, die „computerfreundlich“ waren. Sie wurden außerdem fest angestellt, während ITT seinen Übersetzungsbedarf

normalerweise extern erledigen ließ. Dies motivierte die Übersetzer zusätzlich. Es wurden, ähnlich wie bei dem Versuch des kanadischen „Translation Bureau“, 13 Texte parallel maschinell und manuell übersetzt, wobei umfassende Wörterbücher entstanden. Man kam zu dem Ergebnis, daß sich die Übersetzungsgeschwindigkeit mit Weidner verdoppeln ließe, woraufhin das System gekauft wurde. 1985 besaß ITT sechs Weidner-Versionen (Spanisch<->Englisch, Französisch<->Englisch und Deutsch<->Englisch), die auf einem Rechner vom Typ VAX 11/780 von Digital Equipment liefen, an den fünf Übersetzer-Terminals von Hewlett-Packard angeschlossen waren. Mit dem System können Texte durch ITTs internationales Textverarbeitungsnetz sowie über andere ITT-Datenübermittlungssysteme und über das Paket-Datenübertragungsnetz von British Telecom, der privaten Betreiberin des englischen Fernmeldenetzes, übertragen werden. Es existiert ein mehrsprachiger Drucker (mit dem sich auch arabische Schriftzeichen drucken lassen), ein Blattleser und ein Fotosatzsystem. Es war geplant, die Rohübersetzungen, die auf dem Hauptrechner mit Weidner erstellt wurden, auf verteilte Arbeitsplatzrechner zu übertragen, um die Belastung des Hauptrechners zu senken. Diese Terminals sollten dann auch in den Büros freier Übersetzer aufgestellt werden und mit dem ITT-Rechner verbunden sein. In den zweieinhalb Jahren vom Erwerb Weidners durch ITT bis 1985 wurde das mit 5000 Einträgen mitgelieferte Wörterbuch auf 12000 Einträge je Sprache erweitert. Nach Auffassung der Initiatorin Magnusson Murray besteht der eigentliche Vorteil der computergestützten Übersetzung neben der verdoppelten Übersetzungsleistung darin, daß sich die gesamte Übersetzungsbearbeitung auf elektronischem Wege erledigen läßt und lange Eingabezeiten entfallen. Sie betont, daß Kompatibilität und Normung der Hardware von Übersetzungskunden und Übersetzungsdienst sehr wesentliche Faktoren seien.

## **2.3 Weidner**

Die Firma Weidner betreibt in den USA ein eigenes Übersetzungsbüro, bei dem sie ihr System universell einsetzt, d.h. ohne fachliche Beschränkungen. Die Aussagen von Henrietta B. Pons, der Vizepräsidentin von Weidner, decken sich im wesentlichen mit denen von Ulla Magnusson Murray. Mit Weidner lasse sich die Textdurchsatzzeit steigern, die Kosten senken, die Terminologie innerhalb einer Übersetzung konstant halten, das Endprodukt qualitativ hochwertig gestalten. Sie sieht als das Ziel der

computergestützten Übersetzung die vollautomatische Erstellung von Dokumenten.

Bei Weidner arbeiten die Übersetzer in täglichen Sechsstundenschichten und erzielen dabei eine Übersetzungsleistung von 600 bis 1200 Wörtern/Stunde. Hierbei ist die auf Ergänzung von Wörterbüchern verwendete Zeit nicht berücksichtigt. Die Übersetzungsgeschwindigkeit variiert je nach Erfahrung und Qualifikation der Übersetzer.

## **2.4 Mittel**

Michael G. Hundt, Leiter des Übersetzungsdienstes der kanadischen Mittel Corporation, betrachtet Weidner als ein Werkzeug, mit dem sich die Erstellung einer Übersetzung schneller bewältigen lasse, als von Hand, da die Nachbearbeitung einer maschinell erstellten Rohübersetzung schneller gehe, als eine konventionell angefertigte Übersetzung. Nach einigen Monaten sei die „persönliche Terminologie“ eines Übersetzers in das Systemwörterbuch eingegeben und bräuchte danach nur noch durch gelegentliche Ergänzungen aktualisiert zu werden. In der ersten Phase nach dem Ankauf des Systems durch Mittel habe man 80% der Zeit mit dem Ausbau des Wörterbuchs verbracht, nach Ablauf einiger Monate seien es nur noch 5% gewesen. Hierzu habe das Photosatzsystem wesentlich beigetragen, da man beispielsweise keine Illustrationen arrangieren müsse. Hundt geht davon aus, daß sich bei Mittel eine Übersetzungsleistung von 800 Wörtern pro Stunde erreichen läßt. Mittel erhält (wie vermutlich alle Weidner-Anwender) im Rahmen eines Service-Abkommens alle drei Monate die neueste Version des Systems. Nach Hundts Aussage sei von Weidner stets prompt eine Lösung angefallener linguistischer Probleme geliefert worden. Nach seiner Überzeugung ist die computergestützte Übersetzung der in Zukunft zu beschreitende Weg.

## **3. Weidners Konkurrenten**

Weidner ist nicht das einzige System zur maschinengestützten Übersetzung, das kommerziell erhältlich ist. Schon am Entstehungsort, der mormonischen Brigham

Young University in Provo, Utah, USA, erwuchs Weidner Konkurrenz in Gestalt der Forschungsgruppe, die unter Zugrundelegung eines völlig anderen Ansatzes ALPS entwickelte, eine Art Intimkonkurrenz, die Tim Johnson (1986) als „occasionally bitter, it seems“ beschreibt. Auch die beiden anderen Konkurrenten, Logos und Smart, wurden in den USA entwickelt. Dies dürfte zum einen darauf zurückzuführen sein, daß die Datenverarbeitungstechnik zuerst in den USA so weit gereift zu sein schien, daß man sich an die automatische Übersetzung von Sprachen wagen wollte, nachdem Weaver<sup>6</sup> sein berühmtes Memorandum an Forscherkollegen verschickt hatte; zum anderen auf den Wettlauf der Supermächte, der dazu führte, daß man sich zunächst um die Übersetzung technischer Forschungsberichte aus dem Russischen ins Englische bemühte. Bezeichnenderweise war das erste Projekt von Logos die Übersetzung von Handbüchern für Hubschrauber ins Vietnamesische während der „Vietnamisierung“ des Vietnamkrieges.

### 3.1 Logos Computer Systems<sup>7</sup>

Logos ist das in der Bundesrepublik Deutschland am weitesten verbreitete System zur maschinell gestützten Übersetzung. Die Firma wurde 1969 vom jetzigen Präsidenten und Cheflinguisten Bud Scott in den USA gegründet und befaßte sich zunächst mit der Übersetzung von Wartungshandbüchern für Hubschrauber ins Vietnamesische. Die Entwicklung einer deutsch-englischen Version wurde Ende der siebziger Jahre in Angriff genommen; ein Prototyp wurde 1982 auf der Hannover Messe vorgestellt und die Auslieferung der ersten Systeme begann 1983. Bis Oktober 1984 nutzten 18 Kunden das System, darunter namhafte deutsche Firmen. Logos vertrieb das System auf Leasingbasis, d.h. der Kunde bezahlte eine einmalige Installationsgebühr und anschließend übersetzungsmengenbezogene Gebühren. Die Logos-Anwender konnten sich unbrauchbare Übersetzungen bei Logos gutschreiben lassen. Dieses System fand bei den Kunden jedoch wenig Anklang, weshalb Logos 1985 die Einführung einer neuen Vertriebsstrategie erwog. Logos läuft auf Bürosystemen von Wang mit 80 Megabyte-Festplatte. Es arbeitet mit einer internen Textdarstellung, die eine semantische Abstraktionssprache verwendet, in die jeder zu übersetzende Satz vor

---

<sup>6</sup> Warren Weaver: Translation, in A.D. Booth/W.N. Locke (Hg.) Machine Translation of Languages New York 1955:15-23

<sup>7</sup> Tim Johnson ebd.:160-164, 286-292, 330-332; Logos-Werbung

Ablauf des Parsings übersetzt wird. Zunächst erfolgt eine einfache semantische Analyse eines Satzes. Anschließend wird der Satz anhand von Regeltabellen geparkt, die bei der deutsch-englischen Version 20 000 Regeln enthalten. Dann wird der Transfer mit Hilfe des Zielsprachwörterbuchs durchgeführt. Bei Logos ist Vorbereitung insofern erforderlich, als Silbentrennstriche und Abkürzungen mit Punkt eliminiert, nicht zu übersetzende Wörter oder Abschnitte gekennzeichnet und die Sätze möglichst kürzer als 60 Wörter gehalten werden müssen, da Logos sonst einfach ein Satzende einfügt und den Rest des Satzes für sich übersetzt. Logos empfiehlt, bei Abkürzungen wie „sog.“ einfach den Punkt fortzulassen und die Abkürzung ins Wörterbuch einzutragen.

Jedes Wort wird nach breit angelegten semantischen Kategorien eingeteilt, die in dem Wörterbuch zusammen mit den Regeln, nach denen sich das Wort unter Berücksichtigung verschiedener möglicher Kontexte verhält, eingetragen werden. Der Übersetzer gibt zusätzlich fachgebietsbezogene Codes ein. Logos arbeitet mit einem left-to-right, bottom-up-Parser. Gelten in einem bestimmten Fall keine spezifischen Regeln, liefert das System eine nach den auftretenden Verbal- und Nominalphrasen gegliederte wörtliche Übersetzung.

Die von den Kunden eingegebene Terminologie kann auch von allen anderen Anwendern genutzt werden, da sie an Logos weitergegeben ist. Die meisten Logos-Kunden berichten von Produktivitätssteigerungen in der Größenordnung von 75%; nur ein Logos-Anwender berichtet, daß sich die Steigerung der Übersetzungsleistung auf lediglich 30% bei einfachen Texten belaufe und bei komplizierten Texten gleich Null sei.

Die mit Logos arbeitenden Übersetzer sind in der Regel mit ihrer Arbeit zufrieden.

### **3.2 Automatic Language Processing Systems (ALPS) <sup>8</sup>**

ALPS ist ein direkter Abkömmling des MÜ-Projektes an der Brigham Young University in Provo, Utah, USA. Die mormonische Kirche begann Anfang der siebziger Jahre, sich für die Übersetzung ihrer Schriften in andere Sprachen zu interessieren und

---

<sup>8</sup> Tim Johnson ebd., S.166-167, 327-329

gründete daher an ihrer Universität in Provo ein übersetzungswissenschaftliches Institut. Da man bald feststellen mußte, daß sich die maschinelle Übersetzung nicht für religiöse Schriften eignete, und zu dem Schluß kam, daß vollautomatische Übersetzung ohnehin kein erstrebenswertes Ziel sei, konzentrierte sich das Institut auf die Entwicklung dialogorientierter Übersetzungshilfen. Aus der „weltlicheren“ Zielsetzung des Instituts folgte wenig später seine Umwandlung in ein kommerzielles Unternehmen, die sich im Januar 1980 vollzog. Erster Anwender des Systems wurde die im Sprachlehr- und Übersetzungsbereich tätige Berlitz-Gruppe. Bis Ende 1984 hatte Alps acht Kunden verschiedener Industriezweige in den USA und fünf in Europa, die sämtlich in der Datenverarbeitungsbranche tätig sind.

Von Weidner und Logos unterscheidet sich Alps durch seine Dialogstruktur, die höheren Anschaffungskosten und vier verschiedene Niveaus der Übersetzerunterstützung. Die Systemprogramme sind zur Vermeidung von Problemen der Portabilität zwischen verschiedenen Rechnern statt in Lisp in Algol geschrieben. Die vier verschiedenen Unterstützungsniveaus sind a) ein mehrsprachiges Textverarbeitungssystem, b) selektive Wörterbuchsuche, c) automatische Wörterbuchsuche und d) das Computer-Übersetzungssystem. Letzteres übersetzt immer nur einen Satz auf einmal und stellt dann dem Übersetzer Fragen zur Auflösung auftretender Mehrdeutigkeiten. Die Philosophie von Alps ist es, Aufgaben wie die Auflösung von Mehrdeutigkeiten dem Übersetzer zu überlassen, da es sich nicht lohne, das erforderliche semantische Wissen in den Rechner zu programmieren. Außerdem fühle sich der mit dem System arbeitende Übersetzer so mehr am Übersetzungsprozeß beteiligt, die Rohübersetzung werde besser und der Zeitaufwand sei letzten Endes auch nicht höher, als bei vollautomatischen Systemen.

Natürlich läßt sich mit Alps keine Stapelverarbeitung über Nacht betreiben. Die Frage ist auch, wie lang die Reaktionszeiten des Systems sind, d.h. wie lange der Übersetzer/die Übersetzerin vor dem Bildschirm auf einen Übersetzungsvorschlag warten muß. Dies wäre ein entscheidender Aspekt für die Beurteilung der Benutzerfreundlichkeit von Alps und des dialogorientierten Ansatzes überhaupt.

Alps plante für 1985 die Einführung einer Mikrocomputer-Version, die dann ein direkteres Konkurrenzprodukt zu Logos und Weidner sein würde. Alps verlangt 30 000\$ Installationsgebühr plus 500\$ pro Monat für ein alle vier Unterstützungsniveaus



bie tendes System; die Mikroversion sollte halb soviel kosten. Das System dürfte sich vor allem für komplexe und Mehrdeutigkeits-reiche Texte eignen, die bei einem voll-automatischen System einen hohen Nachbearbeitungsaufwand erforderlich machen.

### 3.3 SMART AI <sup>9</sup>

Die Firma Smart AI installierte 1978 die ersten Systeme ihres „Smart Intelligent Translator“, der auf verschiedenen Rechnern läuft und sich von den anderen beschriebenen MÜ-Produkten vor allem dadurch unterscheidet, daß er ein Expertensystem ist, d.h. er fertigt Übersetzungen mit Hilfe einer „Wissensbasis“ an. Der zu übersetzende Text muß in großem Maßstab vorbearbeitet werden, indem er anhand von PEP (Plain English Program) radikal vereinfacht und vor allem vereinheitlicht wird. PEP basiert auf ca. 1500 Regeln zur Vermeidung linguistischer Fehler und Mehrdeutigkeiten. Beispiel für eine Regel: „The words END and SIDE must be followed by the words OF or FOR.“ Außerdem hat PEP ein auf den Anwender zugeschnittenes Vokabular. Die Übersetzung erfolgt anhand von weiteren 1500 Übersetzungsregeln. Da es sich bei PEP um eine sehr begrenzte Sprache handelt, zumal sie Anaphern und Ellipsen vermeidet, sind in ihr geschriebene Texte anstrengend zu lesen und teils schwerer zu verstehen als gewöhnliches Englisch. Smart eignet sich am ehesten für kurze Mitteilungen und allgemein für Texte, die keinen kontinuierlich fließenden Stil erfordern. Seine bisher größte Anwendung fand Smart bisher im kanadischen „Ministry of Employment and Immigration“, wo Stellenbeschreibungen aus dem Englischen ins Französische übertragen werden. Zugrunde liegt ein Vokabular von 25 000 Wörtern; der Korrektheitsgrad der Übersetzungen liegt bei 97%.

---

<sup>9</sup> Tim Johnson ebd., 167, 169, 292, 293, 333-334

## **4. AUFBAU VON WEIDNER**

### **4.1 Die Hardware**

Weidner MicroCat läuft, wie oben schon erwähnt, auf einem IBM PC-XT mit zehn oder zwanzig Megabyte Festplattenspeicherkapazität und zwei Diskettenlaufwerken mit 360 Kilobyte und 1,2 Megabyte-Format, über die beim Initialisieren des Systems die Übersetzungssoftware eingegeben und auf der Festplatte mit samt dem Betriebssystem MS-DOS installiert wird. Wird versehentlich eine der Systemdateien beschädigt oder ganz gelöscht, kann sie wieder neu eingelesen werden. Außerdem hat man die Möglichkeit, anderswo geschriebene Texte einzulesen oder auf dem PC erstellte Texte oder Übersetzungen auf einen anderen Rechner zu transferieren. Natürlich muß dieser kompatibel sein, also mit dem gleichen Betriebssystem und dem gleichen Diskettenformat arbeiten. Die Tastatur bietet alle spanischen und französischen Sonderzeichen, ohne daß man viele Tasten drücken muß, wie das normalerweise bei PCs üblich ist.

### **4.2 Die Programme**

#### **4.2.1 Beispiel für den Ablauf einer computergestützten Übersetzung**

Nachdem der Benutzer vom Betriebssystem aus durch den Befehl „W“ das Weidner-Menü aktiviert hat, sieht er sich vor folgender Auswahl:

\*\*\*\*\*

\*

\* WCC COMPUTER-ASSISTED LANGUAGE TRANSLATION

\*

\* English-French

\*

\*

\* Version 2.3

\*

\* A. Word Processing (Textverarbeitung)<sup>10</sup>

\* B. Vocabulary search (Vokabelsuchlauf)

\* C. Dictionary update (Wörterbuchaktualisierung)

\* D. Dictionary utilities (Wörterbuchdienstprogramme)

\* E. Translation (Übersetzung)

\* F. Immediate Translation (Sofortübersetzung)

\* G. Deferred translation (Ü. im Stapelbetrieb)

\* H. File utilities (Dateidienstprogramme)

\*

\*\*\*\*\*

#### **4.2.1.1 Textverarbeitung**

Um einen Text einzugeben, wird er/sie Option A wählen. Nun besteht die Wahl zwischen Ganzbildschirmtextbearbeitung und Textbearbeitung auf geteiltem Bildschirm. Letzteres kommt dann in Frage, wenn bereits eine Übersetzung vorhanden ist; dann erscheint in der oberen Bildschirmhälfte der Quelltext, in der unteren die zu revidierende Übersetzung. In unserem Fall ist die Wahl der ersten Option angezeigt. Der Benutzer versieht seine Datei mit einem Namen, z.B. Original. TXT. Nun wird er vom System in den Textverarbeitungsmodus versetzt. Es erscheint das übliche Zeilenlineal, und der Text kann wie bei anderen Textverarbeitungsprogrammen bearbeitet werden. Nach dem Verlassen des Bearbeitungsmodus findet man die Datei im Inhaltsverzeichnis des Betriebssystems vor.

---

<sup>10</sup> Deutsche Übersetzungen von mir

#### **4.2.1.2      *Vokabelsuchlauf***

Ehe man nun den Text übersetzen läßt, empfiehlt sich das Starten eines Vokabelsuchlaufs, um festzustellen, welche im Text vorkommenden Wörter nicht im Wörterbuch stehen. Es wird also im Hauptmenü die Option B aufgerufen, die dann folgende Möglichkeiten anbietet: Vokabelliste nach Häufigkeit des Auftretens geordnet, in der Reihenfolge, in der sie im Text vorkommen und mit Kontext, oder alphabetisch. In der Regel ist die zweite Möglichkeit die zweckmäßigste, weshalb sie auch gleich auf dem Bildschirm vorgegeben wird. Soll eine der beiden anderen Möglichkeiten genutzt werden, ändert man den vorgegebenen Buchstaben. Nach Drücken der „Return“-Taste wird die Funktion gestartet. Auf dem Bildschirm erscheint erst die Meldung „Preprocessing“, anschließend „Vocabulary search in progress“. Gleichzeitig wird die Anzahl der bereits mit dem Wörterbuch abgeglichenen Wörter angezeigt. Sobald der Prozeß beendet ist, erscheint wieder das Hauptmenü auf dem Bildschirm.

#### **4.2.1.3      *Wörterbuchergänzung***

Nun ist Option C an der Reihe, Ergänzung des Wörterbuchs. Es ist die Entscheidung zu treffen, ob das Wörterbuch direkt von der soeben erstellten Vokabelliste aktualisiert werden soll, oder direkt über die Tastatur bestimmte Wörter ins Wörterbuch eingegeben werden sollen. Zur Vorbereitung der Übersetzung des neuen Textes wird sinnvollerweise von der Vokabelliste aus ergänzt. Nun wird der Benutzer um die Beantwortung einer Vielzahl von Fragen für jeden einzelnen Eintrag gebeten, die im Kapitel über das Wörterbuch detailliert behandelt werden sollen. Sind alle erforderlichen Vokabeln eingegeben, ist der Text, sofern nicht bereits geschehen, auf nicht zu übersetzende Termini, insbesondere Namen, zu überprüfen. Diese werden dann in eckige Klammern gesetzt. Mit doppelten Schrägstrichen können ganze Sätze oder Absätze von der Übersetzung ausgenommen werden.

#### **4.2.1.4      *Übersetzung***

Nun kann die Datei dem Übersetzungsprozeß unterzogen werden. Es wird Option E gewählt, der Name der zu übersetzenden Datei wird eingegeben, eine

Hierarchie unter den eventuell zahlreichen vorhandenen Wörterbüchern (bis zu 276) wird festgelegt. Außerdem kann ein Begrenzer für nicht gefundene Wörter bestimmt werden, z.B. das „\$“-Zeichen, in das dann nicht im Wörterbuch enthaltene Wörter eingeklammert werden. Nach Drücken der „Return“-Taste erscheint die Meldung „Translation in progress“, und ein Zähler am Bildschirmrand gibt die Anzahl der bereits übersetzten Wörter an. Im Inhaltsverzeichnis des Betriebssystems befinden sich nun vier neue Dateien: Original.txt, Original.voc, Original.out und Original.ref. Die letztgenannte Datei ist eine formatierte Version des Quelltextes, die so gestreckt bzw. gestaucht ist, daß sie die gleiche Länge wie die Übersetzung hat, und bei der Nachbearbeitung parallel mit ihr am (waagrecht geteilten) Bildschirm verfolgt werden kann.

#### **4.2.1.5      *Nachbearbeitung***

Zur Nachbearbeitung des Textes wird wieder die Option Textverarbeitung gewählt, diesmal der geteilte Bildschirm. Nun kann die in der unteren Bildschirmhälfte erscheinende Rohübersetzung nachbearbeitet und anschließend über den angeschlossenen Drucker ausgedruckt werden. Bei der Nachbearbeitung sind natürlich Funktionen wie das Suchen und Ersetzen eines bestimmten Terminus im ganzen Text sehr hilfreich. In einem Sondermodus können auch kleine Programme zum Suchen und Ersetzen eines Wortes in einem bestimmten Textbereich oder zum Löschen und Übertragen ganzer Abschnitte geschrieben werden.

#### **4.2.1.6      *Stapelbetrieb***

Die Funktionen „Vokabelsuchlauf“ und „Übersetzung“ lassen sich auch im sogenannten Stapelbetrieb über Nacht oder in einer Arbeitspause durchführen. Hierzu drückt man den Buchstaben G, wird dann dazu aufgefordert, den Namen der zu behandelnden Textdatei und den gewünschten Vorgang einzugeben, z.B. „Original.txt“, Vokabelsuchlauf. Der durchzuführende Prozeß wird wie derum durch einen Buchstaben aus einem Menü eingegeben. Bis zu 14 Vorgänge lassen sich in den „Stapel“ einfügen, dessen Abarbeitung durch Drücken der „Return“-Taste ausgelöst wird, wann der Benutzer es wünscht. Weitere Optionen sind das Löschen einzelner in der Warteschlange befindlicher Dateien, sei es, weil ihre Bearbeitung bereits erfolgt ist, sei es, weil man sich umentschieden hat. Es läßt sich auch die Reihenfolge der Liste ändern

oder die ganze Warteschlange auf ein mal löschen, wobei der Benutzer aber um Bestätigung gebeten wird, wie bei allen Vorgängen, bei denen ein gesamtes Inhaltsverzeichnis gelöscht wird.

#### **4.2.1.7      *Sofortübersetzung***

Außerdem gibt es die Funktion „Immediate Translation“, bei der eine sofort zu übersetzende Zeile eingegeben werden kann. Wird über die Zeile geschrieben, beginnt der Übersetzungsprozess, und kurz darauf erscheint die Übersetzung auf dem Bildschirm. Diese Funktion ist für eine schnelle Überprüfung der Übersetzung kritischer Passagen im vorhinein nützlich.

#### **4.2.2    *Aufbau des Wörterbuchsystems***

Die Weidner-Wörterbücher werden in einer gemeinsamen Datei mit dem Namen MSDICT.FE (Französisch-Englisch) verwaltet. Ein unberührtes System verfügt über zwei Wörterbücher, das „Kernwörterbuch“ und das „Allgemeine Wörterbuch“. Im Kernwörterbuch befinden sich ca. 10 000 Einträge, das Allgemeine Wörterbuch ist leer. In dieses können nun durch den Benutzer neue Einträge vorgenommen werden. Handelt es sich um fachgebietsspezifische Terminologie, so empfiehlt es sich, ein eigenes Spezialwörterbuch anzulegen. Bis zu 274 zusätzliche Wörterbücher können von Weidner verwaltet werden. Die im Kernwörterbuch vorhandenen Einträge können durch den Benutzer in keiner Weise angetastet werden. Allerdings können Einträge aus ihm in das Allgemeine Wörterbuch übertragen und dort verändert werden, wenn dies erforderlich ist. Wie lassen sich nun neue Unterwörterbücher erzeugen? Vom Hauptmenü aus wird Option D gewählt, „Dictionary Utilities“. Nun besteht die Wahl zwischen „Dictionary list“ und „Dictionary Maintenance“. Erstes bietet die Möglichkeit, ein vorher spezifiziertes Wörterbuch insgesamt ausdrucken zu lassen, wobei angegeben werden kann, von wo bis wo, falls nicht das gesamte Wörterbuch ausgedruckt werden soll. Zur Schaffung eines neuen Wörterbuches muß Option C gewählt werden. Nun erscheint folgendes Menü:

\*\*\*\*\* \*\*

```

*           Dictionary Name Maintenance
*
*
* 1)GENERAL   2)WCC CORE
*
*
*-----*
* (A)dd a dictionary (D)elete a dictionary (R)ename a
* dictionary (S)et a lookup sequence (P)rint a lookup
* sequence
*
*           Use Control Z to exit.
*
*           Selection:
*
***** *

```

Da es um die Neuschaffung eines Wörterbuches geht, wird A gewählt. Nun wird der Name des neuen Wörterbuchs eingegeben und anschließend die neue Wörterbuchliste in alphabetische Ordnung gebracht. Mit D lassen sich Wörterbücher löschen, mit R umbenennen, mit S in eine Hierarchie bringen, die dann bei einem bestimmten Text verlangt werden kann. Dabei handelt es sich um die Reihenfolge, in der bei der Vokabelsuche des Systems nachgeschlagen wird. Befindet sich ein Terminus in mehreren Wörterbüchern, so wird die im ersten angegebenen Wörterbuch der Hierarchie gefundene Bedeutung gewählt. Befindet er sich in keinem, landet er in der Liste der nicht gefundenen Wörter bzw. wird nach der Übersetzung in „\$“ eingeschlossen. Es können bis zu neun Wörterbücher in einer Sequenz angegeben werden; die Sequenzen werden mit Buchstaben benannt, weshalb ihre Zahl durch die Anzahl der Buchstaben des verwendeten Alphabets begrenzt ist. Mit „P“ lassen sich die vorhandenen Sequenzen ausdrücken. So kann geprüft werden, welche Hierarchie für den jeweiligen zu übersetzenden Text adäquat ist.

### **4.2.3 Wörterbuchaktualisierung**

Wie läßt sich nun ein Wörterbuch erweitern? Nachdem man aus dem Hauptmenü in das Wörterbuchprogramm gelangt ist, entscheidet man sich entweder dafür, neue Einträge direkt über die Tastatur in das Wörterbuch einzugeben oder sie von einer vorher erstellten Vokabelliste zu übertragen. Soll nicht das Allgemeinwörterbuch verwendet werden, ist dies vorher anzugeben. Dies geschieht durch Tippen von „\*“ statt eines Quellwortes. Anschließend wird der Name des zu verwendenden Wörterbuchs eingegeben. Nun gibt man das neu einzutragende Wort ein, betätigt die Zeilenschaltung und wird nun gefragt, ob dieses Wort in das WB aufgenommen werden soll, sofern es noch nicht in ihm enthalten ist. Ist letzteres der Fall, kann der vorhandene Eintrag betrachtet werden. Befindet sich das Wort im Kernwörterbuch, jedoch mit einer ungeeigneten Bedeutung, kann auch dieser Eintrag in Augenschein genommen werden. (Sollen „Idioms“ eingetragen werden, d.h. bei Weidner Zusammenstellungen von mehreren Wörtern, bei denen es sich nicht unbedingt um im linguistischen Sinne idiomatische Wendungen handeln muß, müssen die Einträge direkt über die Tastatur eingegeben werden. Eventuell noch nicht im Wörterbuch vorhandene Wörter sind zuerst einzugeben, ehe das Wörterbuchmenü den Mehrwortausdruck akzeptiert). Nehmen wir an, der Benutzer hat sich für die Möglichkeit entschieden, eine bereits existierende Vokabelliste einzugeben.

#### **4.2.3.1 Verben**

Nun erscheint folgendes Menü (Beispiel aus dem Text „Comment vivent les femmes en



France“):

```

***** *
*
*   ***GENERAL DICTIONARY***
*
*   les énormes changes SURVENUS
*
*       SOURCE:
*       SURVENUS
*
*           A) USE SOURCE AS IS
*       B) MODIFY SOURCE
*       C) GET NEXT SOURCE
*           D) SEARCH FOR NEXT SOURCE
*           E) CHANGE CURRENT DICTIONARY
*
*       ENTER OPTION: A
*
*****

```

Im dargestellten Fall muß die finite Verbform verändert werden, da Verbformen nur im Infinitiv ins Wörterbuch gelangen dürfen; sonst wird eine bestimmte finite Verbform immer gleich übersetzt. Ist sie gleichzeitig Partizip und 3. Person Singular Imperfekt, wie z.B. „used“, wird „He used“ als „il usé“ ins Französische übertragen, falls vorher „used“ als „usé“ in das WB eingetragen wurde. Ausnahmen sind Formen wie „unused“, zu denen gar kein Infinitiv existiert.<sup>11</sup> In unserem Beispiel muß also „survenir“ ins Wörterbuch eingetragen werden. Nun erscheint die Frage, ob es sich bei dem eingegebenen Wort um einen Homographen handelt. Je nachdem, ob es zu dem eingegebenen Wort noch davon abgeleitete Adjektive, Substantive oder Adverbien gibt, ist diese Frage zu bejahen oder zu verneinen. (Die Partizipien von Verben werden grundsätzlich als Homographen eingetragen, hier also „survenant“ und „survenu“). Nun muß die Übersetzung für das Verb eingegeben werden. Natürlich aus den oben genannten Gründen nur eine. Es erscheint dann eine Zeile von Nullen, in die grammatische Informationen zu kodieren sind. Die Schreibmarke steht auf der ersten Ziffer in dieser Reihe. Es ist nun, wie die Hilfsinformation in der Fußzeile dem Benutzer mitteilt, die Wortart einzugeben. Diese Hilfsinformation gibt auch gleich an, welche Wortarten kodiert werden können: Verben, Substantive, Adjektive, Adverbien, Präpositionen, Konjunktionen, Interjektionen, Zahlwörter, Artikel... Nachdem das Programm die Information „Verb“ erhalten hat, springt es nur noch an die Stellen in der Ziffernreihe, die für Verben relevant sind. Nun ist anzugeben, ob es sich um ein Verb

---

<sup>11</sup> Handbuch-Beispiel

handelt, das nur im Singular oder Plural auftritt, oder um ein „normales“, das in beiden Varianten auftreten kann. Ist letzteres der Fall, bleibt die Null stehen. Ebenso verfährt man mit der nächsten Frage, in welcher Person das Verb stehe. Beide Fragen sind nur dann nicht mit „0“ zu beantworten, wenn eine konjugierte Verbform eingegeben wurde. Gleiches gilt für die Frage nach dem Tempus. Nun muß eingegeben werden, ob das Verb transitiv ist, ob es allein stehen kann oder nicht, ob es grundsätzlich intransitiv ist (z.B. „tomber“), welche Ergänzungen es nach sich ziehen kann (ein indirektes Objekt, den Subjonctif, ein Partizip Präsens, es handelt sich um ein Verb der Bewegung). Wird das Verb mit être konjugiert? Ist es reflexiv im Französischen, aber nicht im Englischen (Je me rase -> I shave)? Des Weiteren ist festzulegen, wie ein nachfolgendes à oder de zu übersetzen ist. Standard ist bei „de“ „of“, bei „à“ „to“. Das Programm bietet für „à“ „nothing“ (es bleibt unübersetzt), „at“, „in“, „with“ und „for“ und für „de“ „nothing“(s.o.), „from“, „by“, „with“, „for“ und „on“. Nun ist ein englisches Verb aus einer im Handbuch nachzuschlagenden Liste von 84 (!) Beispielsverben herauszusuchen, dessen Konjugation der Übersetzung des neu eingegebenen französischen Verbs entspricht. Dies ist äußerst zeitaufwendig und oft schwer zu entscheiden.

### **4.2.3.2 Substantive**

Die erste über ein neu eingegebenes französisches Substantiv zu liefernde Information ist das Geschlecht: „0“ bedeutet „undefiniert“ (z.B. „professeur“), „1“ ist „maskulin“, „2“ „feminin“. Als nächstes ist anzugeben, ob es sich um ein Substantiv handelt, das menschlich, ein Körperteil, belebt, abstrakt (socialisme, amour, éternité; bei solchen Substantiven wird in der englischen Übersetzung der Artikel weggelassen), eine Mengenbezeichnung oder nichts davon ist. Dann ist der Numerus anzugeben: „0“ heißt, daß sowohl Plural als auch Singular vorkommen. Nun ist aus 30 Beispielsubstantiven dasjenige auszuwählen, dessen Pluralbildung der der englischen Übersetzung am nächsten kommt (z.B. house-houses, brother-brethren, quiz-quizzes...).

### **4.2.3.3 Adjektive**

Zuerst ist einzugeben, ob das Adjektiv im Englischen vor- oder nachgestellt wird. Nun ist wie bei den Verben die Übersetzung eines dem Adjektiv folgenden „à“ und „de“ zu spezifizieren. Es muß eingegeben werden, ob es sich um ein „adnominales“ Adjektiv handelt, ob es also auch in einer Reihung von Adjektiven immer direkt beim Substantiv stehen muß. Beispiel: „une jeune femme italienne“ -> „a young Italian woman“. Gegenbeispiel: „une jeune femme riche“ „a rich young woman“. Nun werden Informationen über die Steigerung des Adjektivs verlangt, die aus zehn Beispielen herauszufiltern sind. Wird das Adjektiv nicht gesteigert, bleibt die „0“ stehen. Zum Schluß ist noch die Adverbialform anzugeben - entweder „ly“, „ally“, „ily“, „y“ ...

#### **4.2.3.4 Adverben**

Die meisten Adverben brauchen laut Weidner-Handbuch nicht eingegeben zu werden, da sie sich von bereits vorhandenen Adjektiven ableiten lassen. Es werden also nur Adverben ergänzt, deren zugehörige Adjektive noch nicht im Wörterbuch stehen (wobei es sicher sinnvoller ist, dann gleich das Adjektiv einzugeben), oder weil sie nur als Adverben vorkommen.

Zuerst ist zu entscheiden, ob das jeweilige Adverb stets nach- oder vorgestellt werden muß, oder ob beide Positionen möglich sind. Z.B. vorgestellt: „vraiment“ - „I really like...“; nachgestellt: „efficacement“ - „That works efficiently“. <sup>12</sup> Steigerung: siehe Adjektive.

#### **4.2.3.5 Pronomina**

Das Geschlecht ist anzugeben, die „Wesensart“ (belebt, abstrakt, etc...), Numerus, Person, Kasus (Nominativ, z.B. „il“, „je“ oder Akkusativ, z.B. „la“, „les“), und es ist ein zugeben, ob es sich um ein Reflexivpronomen handelt („se“).

#### **4.2.3.6 Präpositionen**

Es genügt die Angabe, daß es sich bei dem eingegebenen Wort um eine Präposition handelt.

#### **4.2.3.7 Konjunktionen**

Handelt es sich um eine unterordnende Konjunktion, ist „7“ einzugeben, bei einer beordnenden Konjunktion eine „8“.

#### **4.2.3.8 Interjektionen**

---

<sup>12</sup> Handbuch-Beispiel

Es ist nur die Wortart anzugeben.

#### **4.2.3.9      *Zahlwörter***

Es ist anzugeben, ob es sich um eine Zahl handelt, und ob das Zahlwort im Singular oder Plural steht.

#### **4.2.3.10     *Artikel***

Es ist nur die Wortart anzugeben.

#### **4.2.3.11     *Homographen***

Ein Homograph ist gemäß der Weidner-Definition ein Wort, das in mehr als einer Wortart auftreten kann. Verben werden von Weidner grundsätzlich als Homographen betrachtet, da ihre konjugierten Formen in der Regel auch Substantive und Adjektive enthalten. Also gibt man zunächst die Übersetzung des Verbs ein. Als nächstes die Übersetzung für die zweite mögliche Homographenvariante, dann für die dritte und eventuell eine vierte. Bis zu neun Einträge sind möglich, trotz der Empfehlung, nur eine Bedeutung je Wortart einzugeben. Diese Empfehlung ist auch insofern berechtigt, als auf der QS-Seite der Wortgleichung nur der Infinitiv steht, man also auch völlig unterschiedlich geschriebene Substantive, die von dem Verb abgeleitet sind, nicht voneinander unterscheiden kann, da sie schlechterdings nicht dastehen. Zur Reduzierung möglicher Ambiguitäten empfiehlt es sich also, nicht zu viele Varianten einzugeben und nach vermutlicher Vorkommenshäufigkeit zu selektieren. Die Autoren des Handbuchs nehmen einfach an, man wolle verschiedene Bedeutungen sicher auch in verschiedenen Wörterbüchern unterbringen. Als Beispiel wird „terre“ gegeben: eine Übersetzung für den „allgemeinen Sprachgebrauch“ - „the earth“, eine für den Gartenbau - „loam“ oder „clay“ und eine für die Politik - „territory“. Nun ist entweder jede dieser Übersetzungen in ein entsprechendes Gartenbau- oder Politikwörterbuch einzugeben, oder es ist zu entscheiden, welche Übersetzung voraussichtlich die am häufigsten benötigte sein wird. Der Haken hierbei ist, daß nur eine begrenzte Zahl von Unterwörterbüchern für eine Übersetzung verwendet

werden darf und bei einer solchen Vielzahl von Wörterbüchern, wie sie dieser Ansatz verlangt, schnell der Überblick verloren gehen kann. Überdies ist das Verfahren ausgesprochen mühselig und entsprechend zeitraubend. Laut Handbuch kommt es selten vor, daß ein Homograph mit mehr als zwei Wortarten in das Wörterbuch eingetragen werden muß, da sich solche Vokabeln entweder bereits im Kernwörterbuch befänden oder zu selten seien, um eingetragen zu werden. Dies kann ich allerdings nicht bestätigen, da mir durchaus auch gemeinsprachliche Vokabeln vorgekommen sind, die in einem einzelnen Wörterbuch stehen sollten, da sie gleichberechtigt im gleichen Text vorkommen können.

Zusätzlich empfehlen die Systementwickler, nach Durchsicht der bereits vorhandenen Rohübersetzungen auf bestimmte Übersetzungsmöglichkeiten einfach zu verzichten, die voraussichtlich selten vorkommen. Dadurch bleibe auch das System effizient, da es nicht so viele Einträge zu überprüfen brauche. Ein Beispiel für den Wörterbucheintrag eines Homographen:

1) SOURCE=ÉCORCER

2) HOMOGRAPH: Y

3) TRANSLATION A = PEEL (verb)

4) TRANSLATION B = BARK (noun)

Als Faustregel gilt: Alle Wörter, die identisch mit einem Partizip oder der 3. Person Singular eines Verbs sind, sollten grundsätzlich mit dem entsprechenden Verb als seine Homographen eingegeben werden. Beispiel:<sup>13</sup> Das Substantiv „montre“ sollte bei Eingabe des Verbs „montrer“ gleich miteingegeben werden.

Nun ist die „Homograph Rule“ festzulegen. Diese bezieht sich auf die Wortarten, die als homograph eingegeben worden sind. Es ist eine Nummer anzugeben:

---

13 Handbuch-Beispiel

word types	Homograph	Example <sup>14</sup>
noun/determiner	6	son
noun/adverb	11	bien
noun/adjective/adverb	14	fort
noun/adjective	15	directeur
adverb/conjunction	16	si
verb/preposition	17	entre
verb/adjective/(noun)	18	parler
noun/preposition	31	contre
noun/conjunction	32	car
noun/quantifier	33	cent

### „ING“-Regel

Hier wird das Genus des Substantivs eingegeben, das sich von dem eingegeben Verb ableiten läßt. „ING“-Regel deshalb, weil hier in den Systemen mit Englisch als Quellsprache Informationen über entsprechende Endungen stehen sollen. Ist kein Substantiv vorhanden, erfolgt keine Eingabe.

### „ED“-Regel

Hier wird die Verbform angegeben, von der das homographe Adjektiv abgeleitet ist. Ist kein solches vorhanden, bleibt „0“ stehen. Handelt es sich um das Partizip Perfekt, ist eine „1“ einzugeben, beim Partizip Präsens eine „2“, bei der 1.Person Singular Präsens eine „3“.

### „STAT“-Regel

---

14 Handbuch-Beispiel



Hier ist anzugeben, von welcher Verbform sich das französische Substantiv ableitet. „0“ = kein Substantiv vorhanden, „1“ = Partizip Perfekt, „2“ = Partizip Präsens, „3“ = 3.Pers. Sgl. Präs.

### **Querverweis**

Hat die Übersetzung eines Homographen mehrere Wörter, muß angegeben werden, welches zielsprachliche Wort sich an das quellsprachliche Wort anpaßt, sei es in Kasus, Genus, Numerus etc.

### **Homograph Type Flag**

Bei jeder Übersetzung für eine einzelne Wortart muß angegeben werden, um welche Wortart es sich gerade handelt, da dies aus dem Eintrag als solchem ja nicht hervorgeht.

Es wird empfohlen, zuerst die vermutlich am häufigsten vorkommende Übersetzung einzutragen, denn „wenn das System Schwierigkeiten bei der Entscheidung hat, welche Bedeutung zu nehmen sei, nimmt es die erste“. (Handbuch, FE 20)

#### **4.2.3.12 Mehrworteinträge („Idiom“-Einträge)**

Ein Mehrworteintrag kann zwei bis acht Wörter in der Quellsprache umfassen. Als Mehrworteintrag (idiom) gilt je der Eintrag, in dem Leertasten auftreten. Unterscheidet sich die Übersetzung einer Reihung von Wörtern von ihrer wörtlichen Übersetzung, empfiehlt es sich, sie als „idiom“ einzutragen. Zunächst will das System jedoch eine Übersetzung für je des einzelne Wort, vorausgesetzt, es befindet sich noch nicht in einem der Wörterbücher. Jedes Wort muß in seiner Grundform eingetragen werden, z.B. nicht „Seeing is believing“, sondern „See be believe“<sup>15</sup> oder statt „Personne n'est parfaite“ „personne ne être parfait“. Nach Homographie wird bei Mehrworteinträgen nicht gefragt, es kann nur eine Übersetzung eingegeben werden. Nun müssen für jedes einzelne Wort die gleichen Fragen beantwortet werden wie bei einem Einzelworteintrag. Zusätzlich wird hier die Querverweiskomponente wichtig. Jedes

---

15 Handbuch-Beispiel

Wort in der Übersetzung, daß in irgendeiner Form analog zu einem Quellsprachwort gebeugt werden kann, muß auf dieses verweisen. Dies geschieht durch einfaches Anhängen der Nummer des zugehörigen QS-Worts.

Beispiel:<sup>16</sup>

SOURCE	TARGET	FLAGS	PRI	IR	CR	REF
1) ADRESSER	20)SPEAK	35)10000000000000000000	50)	22	80)	1
2) LA	21)TO	36)60000000000000000000	51)	0	81)	0
3) PAROLE						
4) à						

Das Verb „speak“ verweist auf das französische Verb zur Anpassung von Tempus, Person und Numerus. „to“ erhält keinen Querverweis, da es unverändert bleibt. Gleiches gilt natürlich auch für die nicht übersetzten Wörter.

Als nächstes ist ein „Schlüsselwort“ zu benennen, unter dem das „idiom abgespeichert wird. Dies führt dazu, daß das System bei jedem Vorkommen des entsprechenden Schlüsselwortes prüft, ob die im Text gefundene Wortkombination auf die Struktur des im Wörterbuch befindlichen „idiom“ paßt. Hierbei kann es natürlich zu Irrtümern kommen, wenn eine zufällige Wortreihung als „idiom“ erkannt wird, obwohl die gefundenen Wörter gar keine Übersetzungseinheit bilden. Es empfiehlt sich nun, das am seltensten vorkommende Wort des Eintrags als Schlüsselwort einzutragen. Dies beschleunigt letztlich auch die Nachschlagegeschwindigkeit. Es ist weiter anzugeben, welches Wort des „idioms“ die meiste für die Übersetzung relevante grammatische Information trägt. Beispiel aus dem Handbuch: „faire promener le chien“ -> „take the dog for a walk“. Hier soll „faire“ als Haupteintragungswort gewählt werden, da es den meisten Einfluß auf die Übersetzung hat. Der Widerspruch zwischen den beiden Anforderungen an ein Schlüsselwort wird hier klar - „faire“ ist sicher nicht eben eine seltene Vokabel. Ein weiteres Beispiel aus dem Handbuch: „chef de chantier“ -> „foreman“. Hier ist „chef“ zu wählen, weil „chantier“ in der Regel unverändert bleibt. Bleibt eine eingetragene Wendung stets unverändert, bleibt „0“ stehen, sonst ist die Nummer des gewählten Wortes einzutragen. Nun ist dem Ausdruck als Ganzem eine Wortart zuzu-

---

16 Handbuch-Beispiel

ordnen. Hat beispielsweise der Ausdruck insgesamt verbale Funktion wie z.B. „casser la croûte“<sup>17</sup>, wird er als „verbal“ eingetragen und außerdem als intransitiv, ungeachtet der Tatsache, daß ein Teil von ihm, das Verb „casser“, transitiv ist.

Enthält ein Ausdruck Zeichen wie „-“, „#“, „/“, „()“, „,“, „:“, „;“ oder „:“, so können diese als separate Einheit eingegeben werden. Es müssen aber Leerstellen zwischen allen für sich zu betrachtenden Einheiten gelassen werden, sonst gilt der gesamte Eintrag als ein Wort. Hierauf zielt die Frage „Idiom the same without <Zeichen>? [N/Y] N, also: Bleibt die Bedeutung des eingegebenen Ausdrucks gleich, wenn das fragliche Zeichen fortgelassen wird? (Als normale Antwort wird „Nein“ angenommen). Außerdem ist anzugeben, an welcher Stelle im Eintrag das zu berücksichtigende Zeichen steht. Sinn der separaten Betrachtung ist die Vereinfachung der Deklination/ Konjugation. Handbuch-Beispiel: „assurance - automobile“. Ohne Leertasten zwischen den beiden Wörtern und dem Bindestrich könnte nur das letzte Wort in den Plural gesetzt werden, hier ist es aber bei beiden Wörtern erforderlich. Außerdem kann festgelegt werden, ob das „idiom“ auch als solches erkannt werden soll, wenn der Bindestrich fehlt, also in unserem Fall die Kombination „assurance automobile“. Dies kann natürlich zu Irrtümern führen, wenn die beiden Wörter rein zufällig nebeneinanderstehen.

In ein „idiom“ lassen sich auch „Löcher“ eingeben, d.h., daß in den Ausdruck auch noch andere Wörter integriert werden können oder bei Mehrworteinträgen, die nicht allein stehen können, sogar müssen. Muß ein Wort integriert werden, wird an seine Stelle ein „Platzhalter“ in Form eines Prozentzeichens gesetzt, z.B. „à % égard“, „aussi % que“, da diese Formen nicht für sich allein stehen können.

Kann ein Eintrag sowohl für sich allein als auch mit einer Ergänzung vorkommen, wird die entsprechende Stelle mit einem „\$“ markiert. Beispiel: „avoir \$ faim“. Natürlich muß die zielsprachliche Variable auf die quellsprachliche Variable querverweisen, es sei denn, die Übersetzung enthalte keine entsprechende Variable. In diesem Fall erscheint das Platzhaltersymbol in der Übersetzung. Hat man einen Platzhalter eingegeben, ist eine weitere Frage zu beantworten. Des weiteren ist anzugeben, ob sich der Platzhalter in der Zielsprache außer bezüglich der Wortart noch in anderer Hinsicht nach dem QS-Platzhalter richtet. Die Frage ist durch „Y“ oder „N“ zu beantworten. Anschließend ist dies näher zu spezifizieren. Bei der nächsten Frage geht es darum, die

für den Platzhalter in Frage kommenden Wortarten einzuschränken, falls erforderlich.  
Das sieht dann so aus:

121) WT # 0: UNKNOWN

0000000000000000

Jede Null entspricht einer Wortart und wird durch eine Eins ersetzt, wenn die entsprechende Wortart in Frage kommt. Allerdings sind mehr Nullen in der Tabelle, als es Wortarten gibt (vielleicht in der Annahme, daß noch weitere Wortarten existieren). Ein Platzhalter darf nicht als Schlüsselwort für den eingegebenen Mehrworteintrag verwendet werden. Allerdings darf er die Wortart des Eintrages bestimmen.

#### **4.2.3.13      *Zur Praxis der Wörterbuchwartung***

Es genügt nicht, sich bei der Ergänzung des Wörterbuchs auf die in der Vokabelliste angegebenen Wörter zu beschränken, denn bei der Lektüre der Rohübersetzung fallen dennoch zahlreiche falsch übersetzte Formen auf. Die Vokabelliste enthält Wörter, die überhaupt nicht im Systemwörterbuch vorhanden sind, sie gibt aber keinen Aufschluß darüber, ob die vorhandenen Äquivalente für die übrigen Wörter auch die im vorliegenden Text in Frage kommenden sind. Die wichtigste Phase der Wörterbuch-Aufarbeitung ist daher der Abgleich anhand der Rohübersetzung. Hier kommt es noch mehr darauf an, darauf zu achten, daß die neuen Einträge auch in das richtige Unterwörterbuch gelangen, und daß die WB-Sequenz für die eventuelle Neuübersetzung des Textes richtig gewählt wird. Hat nämlich ein Fachterminus einen oder mehrere fachliche oder gemeinsprachliche Einträge, kann bei einer falschen Nachschlagerangfolge unter den Wörterbüchern und eventuell unsystematisch auf die verschiedenen WB verteilten Einträgen häufig vom System eine falsche Wahl getroffen werden, die sich dann mehr oder minder negativ auf den Korrekturaufwand auswirkt. Nicht wenige Einträge sind nämlich in mehreren WB gleichermaßen sinnvoll aufgehoben, und wenn hier nicht systematisch vorgegangen wird, droht im technischen Bereich mit den vielen denkbaren (und sinnvollen) Unterwörterbüchern der Überblick verloren zu gehen, was zu terminologisch bunt gesprenkelten Übersetzungen führen kann. Es ist also eine sorgfältig durchorganisierte WB-Hierarchie erforderlich, bei der sich die Verwendung einer Klassifikation lohnen kann.

#### **4.2.3.14 Zur Rolle des Wörterbuchs in einem System zur rechnergestützten Übersetzung**

Ein Maschinenwörterbuch wie das von Weidner unterscheidet sich in wesentlichen Punkten von einem konventionellen Wörterbuch oder einer Terminologiedatenbank.

1) Aus Gründen der Zugriffsgeschwindigkeit, und weil es offenbar technisch nicht machbar ist, dem Rechner eine Reihe detaillierter Informationen einzugeben, wie z.B. eine große Zahl verschiedener kontextgebundener Übersetzungsäquivalente, angereichert durch entsprechende Kontextangaben und/ oder Definitionen, die er dann während des Übersetzungsprozesses auswertet, läßt das Weidner-Wörterbuch nur ein Übersetzungsäquivalent je Wortart und Unterwörterbuch zu.

2) Bedingt durch die bei einem Mikrocomputer gegebene begrenzte Speicherkapazität (zur Zeit maximal 30 Megabyte; das Weidner-Kern-WB benötigt bei ca. 10 000 Einträgen etwa 1,2 MB, was 6% der installierten 20 MB-Festplatte entspricht) läßt sich auf einem solchen System keine anspruchsvolle Terminologiedatenbank installieren. Es ist zu berücksichtigen, daß der vorhandene Speicherplatz auch noch für weitere Dateien benötigt wird, die Systemprogramme verbrauchen weitere 2,5 MB, hinzu kommen die beim Systembetrieb anfallenden Textdateien.

Natürlich wäre es ideal, ein Übersetzungsprogramm auf eine Terminologiedatenbank zugreifen zu lassen, etwa METAL auf TEAM oder SYSTRAN auf EURODICAUTOM.

Angesichts der genannten technischen Grenzen liegt es auf der Hand, daß ein auf einem Microcomputer implementiertes MÜ-Wörterbuch sehr selektiv sein muß und seine linguistischen Informationen auf möglichst engem Raum zu kodieren sind.



## **5. Vorbereitung (Präedition)**

Wie oben bereits erwähnt, kann ein Text nicht ohne eine gewisse Vorbehandlung dem automatischen Übersetzungsprozeß unterzogen werden, da sonst Fehler auftreten, die sich mit mehr oder weniger großem Aufwand von vornherein vermeiden lassen. Dieser Aufwand spielt natürlich auch eine Rolle für die Wirtschaftlichkeit eines maschinellen Übersetzungssystems, da ja die Arbeitszeit des Übersetzers/ der Übersetzerin der wesentlichste Kostenfaktor bei der maschinellen Übersetzung ist, zumal die reine Maschinenzeit bei Weidner nur indirekt über die Systemanschaffungskosten zum Tragen kommt, also bei hoher Auslastung des Systems sehr gering wird. Man kann davon ausgehen, daß der Nachbearbeitungsaufwand den Vorbereitungsaufwand bei weitem übersteigt. Um die erforderlichen menschlichen Eingriffe also so gering wie möglich zu halten, müssen bei der Vorbereitung möglichst alle Elemente ausgeschaltet werden, mit denen das System voraussichtlich nicht fertig wird. Es liegt hierbei im Ermessen des Anwenders, in wie weit der Ausgangstext vor der Übersetzung manipuliert wird, denn natürlich wäre es auch möglich, in diesem Stadium erkennbare Mehrdeutigkeiten syntaktischer wie semantischer Art aufzulösen. Hierbei kommt die vorhandene Erfahrung mit dem System zum Tragen, d.h. die Kenntnis seiner Schwächen bei der richtigen Interpretation grundsätzlich (d.h. eventuell für einen menschlichen Leser nicht) mehrdeutiger Passagen oder Wörter. Hier gilt es, abzuwägen, wieviel Aufwand die Vereindeutigung erfordert und wieviel sie nützt. Läßt sich ein mehrdeutiger Satz durch eine einfach durchzuführende Umstellung so vereinfachen, daß seine Übersetzung Weidner keine Schwierigkeiten macht, lohnt sich dies sicherlich. Besonders problematisch sind natürlich Texte, die von wenig sprachgewandten Autoren verfaßt wurden und möglicherweise Grammatik- und Rechtschreibfehler enthalten, was bei Übersetzungsaufträgen in der Praxis an der Tagesordnung ist.

Notwendige Vorbereitungsmaßnahmen sind das Markieren nicht zu übersetzender Wörter und Passagen - dies empfiehlt sich auch für Tabellen, da bei der zu erwartenden unterschiedlichen Länge der in den Tabellen möglicherweise vorhandenen zu übersetzenden Wörter in der Zielsprache die Tabelle durcheinandergeworfen werden kann. Es ist daher sinnvoll, diese Wörter bei der Nachbearbeitung zu überschreiben. Die



Rechtschreibung ist genau zu überprüfen, sofern nicht ein entsprechendes Programm schon beim Eingeben des Textes Rechtschreibfehler verhindert. Bei der hier besprochenen Version war kein solches Programm implementiert. Da schon ein falscher Akzent zur Nichterkennung eines im Wörterbuch vorhandenen Terminus führt, ist diese Phase sehr wichtig. Dies gilt auch für die Zeichensetzung und für alle im Text vorkommenden Sonderzeichen. Zeichen, die für die Übersetzungsphase eine besondere Bedeutung haben, wie zum Beispiel die Klammern, mit denen nicht zu übersetzende Wörter eingeschlossen werden, dürfen nicht an anderer Stelle im Text vorkommen und müssen da, wo sie in ihrer korrekten Funktion verwendet werden, auch wieder geschlossen werden, da sonst der gesamte einer geöffneten Klammer folgende Text unübersetzt bleibt. Ebenfalls zur Vorbearbeitungsphase möchte ich die Aktualisierung des Wörterbuchs aufgrund der beim Vokabelsuchlauf erstellten Vokabelliste rechnen. Hierbei gilt es, abzuwägen, wann die Eintragung eines bestimmten textspezifischen Terminus sinnvoll ist oder womöglich mehr schadet als nutzt, wenn eine andere Bedeutung häufiger ist.

Nach P.A.Bennett (Hg) (1986) besteht der Sinn der Vorbearbeitung darin, eine höhere Übersetzungsqualität zu gewährleisten, sowie den „mechanischen Teil“ der Übersetzung effizienter und zuverlässiger zu machen:

„Of course, the main aim of pre-editing is to ensure a higher quality of translation, and/or to render the mechanised part of the general task more efficient and reliable. There is an inevitable trade-off between the amount of pre-editing required and the sophistication of the system, though it is generally recognised that there is a limit to pre-editing (perhaps corresponding to the point at which it would have been more efficient to have the text translated by a human) beyond which the machine's role can no longer be described as MT as such.“<sup>18</sup>

---

18 Paul Bennett (HG) 1986: Multilingual Aspects of Information technology, Gower, Aldershot 1986:87



## **6. Nachbearbeitung (Postedition)**

Die Nachbearbeitungsphase ist die entscheidende Leistung des Übersetzers. Hierbei spielt eine Vielzahl von Parametern eine Rolle, am maßgeblichsten die Anforderungen des Übersetzungsnutzers. Besteht der Verwendungszweck der Übersetzung lediglich darin, einen Fachmann grob über den Inhalt eines Artikels zu informieren, genügt die Beseitigung der größten Unverständlichkeiten im Schnellverfahren. Muß das Übersetzungsprodukt höheren Ansprüchen genügen, steigt der Grad der erforderlichen übersetzerischen Sorgfalt. Es muß nicht nur auf die reine Verständlichkeit geachtet werden, sondern auch auf angemessenen Stil und korrekte, idiomatische Wortwahl, und zwar sowohl fachlich wie gemeinsprachlich.

Der Leser eines maschinell übersetzten Textes ist bereit, in Kenntnis der Problematik einige Abstriche an der stilistischen Qualität hinzunehmen, solange der vorliegende Text einigermaßen verständlich ist. Umso krasser fällt der hingenommene Stilverlust auf, wenn eine hochwertige Übersetzung des gleichen Textes ebenfalls zur Verfügung steht. In deren Licht wird zweierlei klar: Erstens kann die Erzielung einer Idealübersetzung auf maschinellm Wege als völlig illusorisch bezeichnet werden, da hierzu der Maschine nicht nur die Fähigkeit zur inhaltlich, sondern auch auf der Sprachebene korrekten Übersetzung, mithin zum Stilempfinden, eingeimpft werden müßte. Zweitens zeigt ein solches Experiment, auf welches Niveau sich der Leser von Fach- und Informationsliteratur herabdrücken läßt, wenn ihm zum Zweck der schnellen Erlangung von Übersetzungen aktueller Literatur keine Alternative zur maschinell erstellten Rohübersetzung zur Verfügung steht. Hieran wird aber auch das Dilemma der Nachbearbeitung maschineller Rohübersetzungen deutlich. Da dem Übersetzer/der Übersetzerin ein vorstrukturierter Text vorliegt, der sich in jeder Hinsicht stark an das fremdsprachige Original anlehnt, nun aber so schnell wie möglich in lesbare Form gebracht werden soll, muß er/sie sich mit einem Mindestmaß an Änderungen begnügen, die sich auf die Herstellung der Verständlichkeit und die eventuelle Änderung technischer Spezifika beschränken, wenn sich das Verfahren wirtschaftlich lohnen soll. So wurde aus einem technischen Zwang, dem der wirtschaftlich durchzuführenden MÜ, die Informationsübersetzung geboren - denn natürlich ist kein Käufer eines MÜ-Systems vom ersten Tag an in der Lage, Einsparungen bei den Übersetzungsausgaben zu erzielen, da vorerst hausspezifische Terminologie eingegeben werden muß, Kinderkrankheiten zu beseitigen sind, die Akzeptanz des Systems bei den Übersetzern sichergestellt werden

muß (und damit auch eine leistungsfördernde Motivation, vgl. hierzu GREEN 1982).<sup>19</sup> Im Prinzip wären auch menschliche Übersetzer in der Lage, eine oberflächliche und stilistisch mangelhafte Übersetzung anzufertigen. Zwei Faktoren stehen dieser Möglichkeit aber im Wege. Erstens das Stilempfinden der betroffenen Übersetzer und zweitens der Langeweile-Faktor, d.h. Inkonsequenz bei der Verwendung von Terminologie. Die Informationsübersetzung kann also mit Fug und Recht als neues Produkt angesehen werden, das offenbar einen schon lange latent vorhandenen Bedarf befriedigt. Ein interessantes Verfahren in diesem Zusammenhang ist die „Schnelle Nachbearbeitung“, wie sie in den EG praktiziert wird (Loffler-Laurian 1986). Hierbei geht es darum, für den EG-internen Bedarf schnell kostengünstige Übersetzungen zu erstellen, die nicht die gleichen Qualitätskriterien erfüllen müssen, die an für die Öffentlichkeit bestimmte Texte gestellt werden. Also werden nur die größten Fehler beseitigt, und stilistische Fehler bleiben im Text. Es liegt auf der Hand, daß sich so der Übersetzungsdurchsatz enorm steigern läßt. Ein Beispiel, zitiert nach Loffler-Laurian 1986:

Original                      Although both governments and the  
Industry train personnel, it is fre  
quently done on a piecemeal basis.

Rohübersetzung    Bien que les gouvernements et l'industrie  
forment le personnel, il est souvent fait  
sur une base fragmentaire.

Schnelle NB        Bien que les gouvernements et l'industrie  
forment les personnels, ils le font  
souvent sur une base fragmentaire.

Normale NB        Tant les pouvoirs publics que l'industrie  
                          donnent une formation au personnel, mais  
                          ils le font souvent par bribes.

---

<sup>19</sup> Roy Green: The MT errors which cause most trouble to post-editors, in Lawson (Hg) 1982:101-105

Ein solches Verfahren stellt der Kreativität des Übersetzers zusätzlich zum üblichen Zeitdruck ein weiteres Hindernis in den Weg, nämlich eine neue Qualität von Beeinflussung durch die hypnotische Kraft der schwarz auf weiß vorhandenen Rohübersetzung. Das System hat bereits eine Vokabularauswahl getroffen und gibt eine Syntax vor. Zwei Möglichkeiten sind nun denkbar:

1) Syntax und/oder Wortwahl sind vollkommen falsch. Nun wird der Nachbearbeiter entweder versuchen, die Fehler mit möglichst geringem Aufwand geradezubiegen, oder eine Radikalkur im Sinne einer völligen Ersetzung der fraglichen Passage vorzunehmen.

2) Syntax und Wortwahl sind zwar nicht korrekt, aber auch nicht völlig falsch, d.h. unidiomatisch bzw. terminologisch unpassend. Dies ist der kritische Punkt, an dem sich das übersetzerische Sprachgefühl beweisen (und behaupten) muß, insbesondere hinsichtlich des oben genannten Dauerhypnose-Effekts. Denn die Versuchung ist groß, eine zwar verständliche, aber stilistisch mangelhafte oder unschöne Passage stehenzulassen. Außerdem besteht noch die Möglichkeit, zwar den unzutreffenden Terminus durch einen korrekten zu ersetzen, aber die untypische Syntax stehenzulassen (oder, weniger wahrscheinlich, umgekehrt).

Der Faktor der eigenen Unsicherheit des Übersetzers ist eine nicht zu vernachlässigende Größe. Was man genügend oft gelesen und/oder gehört hat, geht letzten Endes in den eigenen, zumindest passiven, Wortschatz über. Dies zu erforschen, wäre ein lohnendes Thema für eine Langzeitstudie an mit der Nachbearbeitung maschineller Übersetzungen beschäftigten Übersetzern.

## **7. Auswertung**

Die Auswertung der in den Rohübersetzungen aufgetretenen Fehler hat mehrere Ziele. Zunächst soll festgestellt werden, welche Arten von Fehlern auftreten, wie häufig sie auftreten, und in welchen Textsorten mit welchen Fehlertypen verstärkt zu rechnen ist. Es sollen Schlußfolgerungen darüber gezogen werden, worin die Fehlerursache besteht. Allerdings muß ich mich mangels Detailkenntnissen über die von Weidner verwendete Grammatik und das Regelsystem auf Vermutungen beschränken; hinsichtlich des Wörterbuches liegen die Verhältnisse etwas klarer, da der Systemanwender

zwangsläufig über Detailkenntnisse des WB verfügen muß, um mit ihm arbeiten zu können. Über die Programmabläufe gibt das Benutzerhandbuch aus naheliegenden Gründen keinen Aufschluß.

Es soll weiterhin eine Einschätzung der Schwere der auftretenden Fehler gegeben werden, und zwar auf dem Umweg über die Verständlichkeit der fehlerhaft übersetzten Passage. Diese Einschätzung gilt natürlich mit der Einschränkung, daß der Autor selbst kein „native speaker“ der englischen Sprache ist.

Ferner wird eine Einschätzung des zu erwartenden Nachbearbeitungsaufwandes geliefert, die sich nach der für die Korrektur einer unzureichend bzw. völlig falsch übersetzten Passage erforderlichen Zeit richtet. Hierbei spielt der Grad der Verständlichkeit keine Rolle, da auch die Korrektur eines stilistisch mangelhaft übersetzten Satzes, der völlig verständlich sein kann, in der Regel mehr Aufwand erforderlich macht, als die Verbesserung eines Satzes, der durch ein völlig falsch übersetztes Wort vollkommen unverständlich geworden ist. Ebenso kann natürlich auch die falsche Analyse der Wortart eines Wortes zu einer weitgehend unbrauchbaren Übersetzung eines ganzen Satzes führen.

Nach Maßgabe dieser Kriterien habe ich aus den ersten zehn mit Weidner angefertigten Übersetzungen, die den Bereichen Sprachwissenschaft, Wirtschaft und Technik entstammen, etwa 470 fehlerhaft übersetzte Passagen herausgegriffen und diese in eine Datenbank eingetragen, deren einzelne Einträge aus dem französischen Original, der englischen Übersetzung, einer Einteilung des/der jeweiligen Fehler 1.) in die linguistischen Kategorien „grammatikalisch“, „lexikalisch“, „orthographisch“ und „syntaktisch“, 2.) in die Kategorien „verständlich“ (trotz Fehlers), „eingeschränkt verständlich“, „unklar“ und „unverständlich“ sowie „sinnentstellt“. Als lexikalischer Fehler gilt ein Fehler, der auf eine falsche Wörterbuchauswahl bzw. dessen Versagen zurückzuführen ist. Als eingeschränkt verständlich habe ich eine Passage dann eingestuft, wenn der Sinn einem Leser mit etwas Hintergrund hinsichtlich der maschinellen Übersetzung oder Kenntnis des Französischen erschließbar ist. Bemerkungen über Stil enthalten nur Einträge, die als verständlich eingestuft worden sind, da sich bei unverständlichen Passagen nicht mehr von Stil sprechen läßt.

Die Bewertung weiterer acht Texte erfolgt in Tabellenform. Die Tabellen enthalten je

einen Satz des übersetzten Textes in der französischen Originalversion, der autorisierten englischen Übersetzung (bei zwei Texten nicht vorhanden) und der besten Weidner-Rohübersetzung, gefolgt von einer linguistischen Bewertung, die ähnlich wie bei den übrigen Texten nach den Kategorien Syntax, Grammatik, Lexik, Stil und Zeichensetzung erfolgt. Die fehlerbehafteten Passagen werden mit einer korrekten Übersetzung versehen in die jeweils relevanten Felder eingetragen, eventuell in mehrere. Die korrigierte Version steht durch einen Doppelpunkt getrennt hinter der falschen Rohübersetzung. Bei einigen Fällen wird eine Korrektur in zwei Stufen angegeben - die erste geht auf den primär vorliegenden Fehler ein, die zweite liefert einen besseren Übersetzungsvorschlag. Den Tabellen schließen sich eine Bewertung der Verständlichkeit und des Nachbearbeitungsaufwandes sowie Angaben über zur Verbesserung der Rohübersetzung möglichen Maßnahmen an. Die Angaben über den NBA richten sich nach den Anforderungen einer stilistisch einwandfreien Übersetzung. Einer der Texte wird durch den Vergleich von drei vorhandenen Rohübersetzungen bei ausgewählten aussagekräftigen Sätzen ergänzt. Der tabellarischen Auswertung folgt ihre ebenfalls satzweise Bewertung hinsichtlich Verständlichkeit, erforderlicher Nachbearbeitung und Maßnahmen zur Verbesserung des Übersetzungsergebnisses.

### **Beschreibung des Textes „Kayak ou les bases des services informatiques“**

Bei diesem Text handelt es sich um eine Projektbeschreibung, die dem Bulletin der französischen „Agence de l'Informatique“ entnommen ist. Er beschreibt ein Forschungsvorhaben zur Entwicklung eines kommerziell verwertbaren Büroinformationssystems. Die Terminologie dichte ist hoch.

### **Satzweise Bewertung der Rohübersetzung**

SATZ 1: UNVERSTÄNDLICH wegen des Fehlens eines Schlüsselwortes. Dieser Fehler ist absolut unerklärlich, da die korrekte Übersetzung „basis“ als substantivischer Homograph des Verbs „baser“ im Kernwörterbuch steht, wogegen bei der Wurzel „bas“ kein Wort zu finden war, dem im Kern-WB „bar“ als Übersetzungsäquivalent zugeordnet ist.

NACHBEARBEITUNGS-AUFWAND (NBA): GERING

MAßNAHMEN: Separater Eintrag von „base“ mit der korrekten Übersetzung „basis“

sowie von „services bureautiques“.

SATZ 2: TEILWEISE VERSTÄNDLICH; für einen Kenner der französischen Sprache eingeschränkt verständlich.

NBA: MITTEL

MAßNAHMEN: Eintrag von „maîtrise“ ins WB als Homograph von „maîtriser“.

SATZ 3: TEILWEISE VERSTÄNDLICH; die unverständlichen Satzteile enthalten syntaktische Fehler wie z.B. „the-know how to do research“, wo der Bindestrich an der falschen Stelle steht und „how“ falsch bezogen ist.

NBA: MITTEL

MAßNAHMEN: Eintrag von „épanouissement“, „programme de recherche“, „savoir-faire“, und „en cours“ ins WB.

SATZ 4: UNVERSTÄNDLICH wegen massiver lexikalisch-syntaktischer Fehler: „the clean harshness to the research“ enthält zwei lexikalische und einen syntaktischen Fehler. „Propre“ gehört zu „à la recherche“ und nicht zu „rigueur“.

NBA: UMFANGREICH

MAßNAHMEN: Eintrag von „propre“ und „rigueur“ ins WB.

SATZ 5: GRÖßTENTEILS VERSTÄNDLICH abgesehen von zwei lexikalischen Fehlern.

NBA: GERING

MAßNAHMEN: Eintrag von „composant“ (Homograph von „composer“) ins WB.

SATZ 6: GRÖßTENTEILS UNVERSTÄNDLICH, vor allem aufgrund der falschen syntaktischen Einbindung und Übersetzung von „contrôle“. Ein Eintrag von „à contrôle %“, wobei % als Platzhalter für ein Adjektiv spezifiziert wird, würde die sen Fehler beseitigen. DAS \ blieb wegen des versehentlich davor angeordneten Anführungszeichens stehen.

NBA: UMFANGREICH

MAßNAHMEN: Eintrag von „bande passant“ ins WB („passante“ muß in der Grundform eingetragen werden).

SATZ 7: VERSTÄNDLICH



NBA: GERING

MAßNAHMEN: Evtl. Eintrag von „être ouvert sur“ in ein SpezialWB.

SATZ 8: UNVERSTÄNDLICH aufgrund komplexer lexikalisch-syntaktischer Fehler.

NBA: UMFANGREICH

MAßNAHMEN: Eintrag von „poste de travail“, „unique“, „ensemble“ und „commande“ ins Wörterbuch.

SATZ 9: VERSTÄNDLICH

NBA: MITTEL; lexikalisch-stilistische Korrekturen

MAßNAHMEN: Eintrag von „autonome“ ins WB

SATZ 10: GRÖßTENTEILS UNVERSTÄNDLICH aufgrund syntaktischer und lexikalischer Fehler.

NBA: UMFANGREICH

MAßNAHMEN: Eintrag von „plan“, „modulaire“, „extensible“, „entrée“ und „sortie“ ins Wörterbuch.

SATZ 11: EINGESCHRÄNKT VERSTÄNDLICH; enthält syntaktische Unklarheiten wie „developed system“ und „oriented program network“.

NBA: MITTEL

MAßNAHMEN: eventuell Eintrag von „orienter réseau“ als „idiom“.

SATZ 12: EINGESCHRÄNKT VERSTÄNDLICH wegen „local ones“.

NBA: GERING

MAßNAHMEN: Eintrag von „local“ als Substantiv-Homograph des Adjektivs.

SATZ 13: EINGESCHRÄNKT VERSTÄNDLICH WEGEN DER Übersetzung von „cours“ durch „courtyards“.

NBA: GERING

MAßNAHME: Eintrag von „en cours de“ ins WB; eventuell mit der Übersetzung „under“.

**Kommentar der Rohübersetzung insgesamt**

Dieser Text wurde nur einmal nach Eingabe eines Teils der in der Vokabelliste angegebenen Wörter übersetzt, weshalb auch einige mit \$\$ markierte Wörter in der Übersetzung auftauchen. Durch massive Erweiterung des Wörterbuches lie ße sich die Lesbarkeit und lexikalische Korrektheit der Übersetzung noch deutlich steigern. Die vorliegende Rohübersetzung erfordert bei vier größtenteils oder völlig unverständlichen, vier eingeschränkt bzw. teilweise verständlichen und nur zwei verständlichen Sätzen einen mittleren Nachbearbeitungsaufwand, der bei einer von mir durchgeführten Nachbearbeitung eine halbe Stunde betrug.

### **Beschreibung des Textes „Les fibres optiques“**

Bei diesem Text handelt es sich um ein Kapitel aus einem Lehrbuch über Nachrichtentechnik. Die Terminologiedichte ist entsprechend hoch und sorgte für eine lange Vokabelliste bei der ersten Auswertung durch Weidner.

### **Satzweise Bewertung der Rohübersetzung**

SATZ 1: VERSTÄNDLICH trotz lexikalischer Fehler.

NBA: MITTEL

MAßNAHMEN: Eintrag von „optique“ und „guide lumière“ ins WB.

SATZ 2: EINGESCHRÄNKT VERSTÄNDLICH, aber die falsche Übersetzung von „coeur“ nimmt dem Satz den Informationsgehalt.

NBA: MITTEL

MAßNAHMEN: Eintrag von „coeur“ ins WB.

SATZ 3: EINGESCHRÄNKT VERSTÄNDLICH wegen mangelhafter Syntax und falscher Übersetzung von „gaines“.

NBA: MITTEL

MAßNAHMEN: Eintrag von „gaine“ in ein fernmeldetechnisches WB.

SATZ 4: UNVERSTÄNDLICH wegen zahlreicher unübersetzter Wörter, falsch übersetzter Wörter und wirrer Syntax.

NBA: UMFANGREICH

MAßNAHMEN: Eintrag von „bande passante“, „MHZ“, „GHZ“ und „micromètre“ und „fixer les idées“ („idée“ muß in der Grundform eingetragen werden, daher nicht „fixer

les idées“).

SATZ 5: VERSTÄNDLICH

NBA: MITTEL

MAßNAHMEN: s. voriger Satz

SATZ 6: VERSTÄNDLICH

NBA: GERING

MAßNAHMEN: s.o.

SATZ 7: VERSTÄNDLICH trotz syntaktischer und lexikalischer Mängel.

NBA: MITTEL

MAßNAHMEN: Eintrag von „effectuer“ (als reflexiv, damit bei der Übersetzung das Reflexivpronomen verschwindet), und „mode“ mit der Übersetzung „mode“.

SATZ 8: EINGESCHRÄNKT VERSTÄNDLICH; hinter „the empties“ kann ein Fachmann eventuell „vacuum“ vermuten.

NBA: MITTEL

MAßNAHMEN: Eintrag von „vide“ (Substantiv) ins WB, eventuell auch die wissenschaftlichen Einheiten und Bezeichnungen.

SATZ 9: TEILWEISE VERSTÄNDLICH, da die unübersetzten Wörter „physique-ment“ und „différencie“ syntaktisch unaufgelöst bleiben.

NBA: UMFANGREICH

MAßNAHMEN: Eintrag von „physique“ und „différencier“ ins WB

SATZ 10: VERSTÄNDLICH

NBA: MITTEL

MAßNAHMEN: s.o.

SATZ 11: UNVERSTÄNDLICH wegen lexikalisch bedingter syntaktischer Fehler

NBA: UMFANGREICH

MAßNAHMEN: Eintrag von „illustrer“ und „mode de fonctionnement“ ins WB

SATZ 12: TEILWEISE VERSTÄNDLICH wegen erheblicher syntaktischer Mängel

NBA: MITTEL

MABNAHMEN: Eintrag von „fibre à saut d'indice“ und „assurer“ ins WB.

SATZ 13: VERSTÄNDLICH

NBA: MITTEL

MABNAHMEN: Eintrag von „réfraction“ und „généralement“ ins WB.

SATZ 14: Nur TEILWEISE VERSTÄNDLICH wegen falscher Übersetzung von „il en est de même“.

NBA: UMFANGREICH

MABNAHMEN: Eintrag von „élargissement“ und „atténuation“ ins WB.

SATZ 15: EINGESCHRÄNKT VERSTÄNDLICH

NBA: MITTEL

MABNAHMEN: Eintrag von „maximal“ und „fonctionnement“ ins WB.

SATZ 16: UNVERSTÄNDLICH wegen massiver lexikalisch-syntaktischer Mängel

NBA: UMFANGREICH

MABNAHMEN: Eintrag von „fibre à gradient d'indice“, „décroître“, „continu“ (Homograph von „continuer“), und „précédemment“ ins WB.

SATZ 17: TEILWEISE VERSTÄNDLICH

NBA: MITTEL

MAßNAHMEN: Eintrag von „réfocaliser“ ins WB.

SATZ 18: Nur TEILWEISE VERSTÄNDLICH wegen idiomatischer Mängel.

NBA: UMFANGREICH

MAßNAHMEN: Eintrag von „monomode“ ins WB.

SATZ 19: EINGESCHRÄNKT VERSTÄNDLICH wegen Überetzung eines Superlativs als Komparativ und falscher Übersetzung von „nul“.

NBA: MITTEL

MAßNAHMEN: Eintrag von „performant“ ins WB

SATZ 20: VERSTÄNDLICH

NBA: GERING

MAßNAHMEN: s.o.

SATZ 21: UNVERSTÄNDLICH wegen falscher Übersetzung von „donne“ als „data“.

NBA: GERING

MAßNAHMEN: /

SATZ 22: Nur EINGESCHRÄNKT VERSTÄNDLICH wegen falscher Übersetzung von „pratique“.

NBA: MITTEL

MAßNAHMEN: /

SATZ 23: VERSTÄNDLICH

NBA: GERING

MAßNAHMEN: /

SATZ 24: EINGESCHRÄNKT VERSTÄNDLICH

NBA: UMFANGREICH

MAßNAHMEN: Eintrag von „matière première“ „silice“ und „largement“.

SATZ 25: EINGESCHRÄNKT VERSTÄNDLICH

NBA: MITTEL

MAßNAHMEN: /

SATZ 26: Nur EINGESCHRÄNKT VERSTÄNDLICH wegen falscher Übersetzung von „plan“.

NBA: UMFANGREICH wegen stilistischer Mängel

MAßnahmen: Eintrag von „plan“ in ein Spezialwörterbuch

SATZ 27: VERSTÄNDLICH

NBA: MITTEL

MMAßNAHMEN: Eintrag von „capacité“ ins WB

SATZ 28: EINGESCHRÄNKT VERSTÄNDLICH

NBA: MITTEL

MAßNAHMEN: Eintrag von „insensibilité“, „parasite“ und „isolant“ (Substantiv) ins WB.

SATZ 29: TEILWEISE VERSTÄNDLICH

NBA: MITTEL

MAßNAHMEN: Eintrag von „poste“ und „liaison“ ins WB.

Satz 30: TEILWEISE VERSTÄNDLICH

NBA: MITTEL

MAßNAHMEN: Eintrag von „encombrement“ und „encombrant“ (Homograph von „encombrer“)

SATZ 31: EINGESCHRÄNKT VERSTÄNDLICH

NBA: MITTEL

MAßNAHMEN: Eintrag von „télécommunications“ und „fiabilité“ ins WB.

SATZ 32: UNVERSTÄNDLICH

NBA: UMFANGREICH

MAßNAHMEN: Eintrag von „répéteur-régénérateur“ und „régénération“ ins WB.

SATZ 33: TEILWEISE VERSTÄNDLICH

NBA: UMFANGREICH

MAßNAHMEN: Eintrag von „micron“, „voisin“(Adj.) und „respectivement“ ins WB:

SATZ 34: VERSTÄNDLICH

NBA: GERING

MAßNAHMEN: /

SATZ 35: TEILWEISE VERSTÄNDLICH

NBA: UMFANGREICH

MAßNAHMEN: Eintrag von „répéteur“, „centrale téléphonique“ ins WB.  
Erstaunlicherweise wird hier „capacité“ richtig wie dergegeben!

SATZ 36: EINGESCHRÄNKT VERSTÄNDLICH wegen falscher Übersetzung von „débit“; ein Techniker könnte aufgrund der Einheit vermutlich die Bedeutung erschließen.

NBA: GERING

MAßNAHMEN: Eingabe von „débit“ ins Wörterbuch

SATZ 37: VERSTÄNDLICH

NBA: GERING

MAßNAHMEN: s.o.

SATZ 38: UNVERSTÄNDLICH

NBA:UMFANGREICH

MAßNAHMEN: Eintrag von „coaxial“ ins WB; s.o.

SATZ 39: TEILWEISE EINGESCHRÄNKT VERSTÄNDLICH

NBA: MITTEL

MAßNAHMEN: Eintrag von „marché“ als Homograph von „marcher“.

### **Kommentar der Rohübersetzung insgesamt**

Bei sechs unverständlichen, 12 eingeschränkt und 10 teilweise verständlichen sowie 11

verständlichen Sätzen, bei denen es sich meist um die kürzeren handelt, bereitet die Lektüre dieser Rohübersetzung einige Mühe, die sich durch eine entsprechende Aufarbeitung des Wörterbuchs erheblich reduzieren ließe. Die erforderliche Nachbearbeitung ist bei sieben Sätzen gering, bei 20 Sätzen mittel und bei immerhin 12 Sätzen umfangreich.

### **Beschreibung des Textes „Peroxyde d'hydrogène et traitement des surfaces métalliques“**

Bei diesem Text handelt es sich um die französische Version der englisch/französisch/deutschen Hochglanzbroschüre einer international tätigen Chemiefirma, die für den Einsatz von Wasserstoffperoxid bei der Metallveredelung wirbt. Dieser Text ist, wie auch der vorstehende Text, sehr terminologieintensiv. Da es sich um einen Werbetext handelt, sind die Unterschiede zwischen der maschinell erstellten Rohübersetzung und der englischen Version des Prospektes natürlich besonders ausgeprägt. Die erforderliche NBA richtet sich natürlich nach der Zielgruppe des Textes. Soll die Übersetzung die Werbefunktion erfüllen, müssen an sie strenge Anforderungen gestellt werden, d.h. sie muß ebenso gut sein wie die im Prospekt enthaltene. Handelt es sich nur um eine interne Informationsübersetzung, genügt eine sprachlich korrekte Wiedergabe des Textes. Daher gebe ich bei meiner Bewertung des NBA erst den für eine funktionserhaltende Übersetzung, dann den für eine Informationsübersetzung erforderlichen NBA an.

### **Satzweise Bewertung der Rohübersetzung**

SATZ 1: VERSTÄNDLICH

NBA: GERING/GERING

MAßNAHMEN: /

SATZ 2: TEILWEISE VERSTÄNDLICH

NBA: UMFANGREICH/MITTEL

MAßNAHMEN: /

SATZ 3: VERSTÄNDLICH

NBA: MITTEL/MITTEL



MAßNAHMEN: /

SATZ 4: EINGESCHRÄNKT VERSTÄNDLICH

NBA: MITTEL/MITTEL

MAßNAHMEN: /

SATZ 5: UNVERSTÄNDLICH

NBA: UMFANGREICH/UMFANGREICH

MAßNAHMEN: /

SATZ 6: TEILWEISE VERSTÄNDLICH

NBA: UMFANGREICH/MITTEL

MAßNAHMEN: Eintrag von „frappe de monnaie“ und eventuell „construction automobile“ ins WB.

SATZ 7: TEILWEISE VERSTÄNDLICH

NBA: UMFANGREICH/UMFANGREICH

MAßNAHMEN: Eintrag von „fonte“, „sûr“ und „échelle“ ins WB.

SATZ 8: EINGESCHRÄNKT VERSTÄNDLICH

NBA: MITTEL/GERING

MAßNAHMEN: /

SATZ 9: EINGESCHRÄNKT VERSTÄNDLICH

NBA: MITTEL/MITTEL

MAßNAHMEN: /

SATZ 10: VERSTÄNDLICH

NBA: MITTEL/MITTEL

MAßNAHMEN: /

SATZ 11: EINGESCHRÄNKT VERSTÄNDLICH

NBA: UMFANGREICH/MITTEL

MAßNAHMEN:

/

### **Kommentar der Rohübersetzung insgesamt:**

Bei einem unverständlichen, zwei teilweise und fünf ein geschränkt verständlichen und zwei verständlichen Sätzen, von denen bei Informationsübersetzungs-Ansprüchen zwei geringen, sieben mittleren und zwei hohen Nachbearbeitungsaufwand verlangen, liest sich diese Rohübersetzung über weite Strecken trotz ergänzten Wörterbuchs mühsam. Hier liegen die meisten Probleme bei der Übersetzungsauswahl auf der Wortebene und der Syntax, die häufig quellsprachlichen Mehrdeutigkeiten und der Homographenproblematik zum Opfer fällt.

### **Beschreibung des Textes „Langues étrangères et communication internationale“**

Bei diesem Text handelt es sich um die Einleitung zu einer Resolution, die sich mit den Voraussetzungen für die Vertiefung des innereuropäischen Gedanken- und Kulturaustausches beschäftigt.

### **Satzweise Bewertung der Rohübersetzung**

SATZ 1 (Überschrift): VERSTÄNDLICH (KORREKT)

NBA: /

MAßNAHMEN: /

SATZ 2: UNVERSTÄNDLICH wegen zwei wesentlichen lexikalischen Fehlern.

NBA: MITTEL

MAßNAHMEN: Eintrag von „note préparatoire“ in ein Unter-WB.

SATZ 3: VERSTÄNDLICH trotz stilistischer, syntaktischer und lexikalischer Mängel.

NBA: MITTEL

MAßNAHMEN: Eintrag von „à l'initiative“ und der Eigennamen in die entsprechenden WB.

SATZ 4: TEILWEISE VERSTÄNDLICH; zahlreiche lexikalische Fehler, einer durch einen Tippfehler im Quelltext bedingt. Erhebliche stilistische Mängel.

NBA: UMFANGREICH

MABNAHMEN: /

SATZ 5: VERSTÄNDLICH; im wesentlichen stilistische Mängel.

NBA: UMFANGREICH

MABNAHMEN: /

SATZ 6: VERSTÄNDLICH, trotz syntaktischer und lexikalischer Mängel

NBA: MITTEL

MABNAHMEN: Eintrag von „notamment“ mit der Bedeutung „particularly“.

SATZ 7: Nur TEILWEISE VERSTÄNDLICH wegen zwei wesentlichen lexikalischen Fehlern.

NBA: UMFANGREICH

MABNAHMEN: Eintrag von „maitre“ mit der Bedeutung „school teacher“ und „continu“ als „continuous“.

SATZ 8: Nur TEILWEISE VERSTÄNDLICH wegen eines wesentlichen lexikalischen Fehlers.

NBA: MITTEL

MABNAHMEN: Eintrag von „manifeste“ in der Bedeutung „manifesto“ in ein Unter-WB.

SATZ 9: EINGESCHRÄNKT VERSTÄNDLICH wegen syntaktischer Unklarheiten .

NBA: UMFANGREICH

MABNAHMEN: Eintrag von „mettre \$ en oeuvre“ mit der Bedeutung „implement \$“; evtl. außerdem „mise en oeuvre“ als „implementation“.

### **Kommentar des Textes insgesamt**

Bei vier verständlichen, einem eingeschränkt verständlichen, drei teilweise verständlichen und einem unverständlichen Texten hat der Text insgesamt einen relativ hohen Verständlichkeitsgrad, er erfordert aber einen recht hohen Nachbearbeitungsaufwand zur Bereinigung stilistischer Mängel, die sich auf die Syntax auswirken. Durch Erweiterung der Wörterbücher ließe sich die Lesbarkeit dieses Textes

nur in geringem Umfang steigern, zumal teilweise auch Homographie im gemeinsprachlichen Bereich vorliegt.

### **Beschreibung des Textes „Erasmus“**

Bei diesem Text handelt es sich um einen Fragenkatalog bezüglich der Fremdsprachenkenntnisse, die zukünftige Stipendiaten der europäischen Austauschprogramme ERASMUS und COMETT aufzuweisen haben und der Möglichkeiten, die ihnen an den jeweiligen Gastinstitutionen geboten werden. Die Terminologiedichte ist gegenüber den technischen Texten vergleichsweise gering, das vorhandene Fachvokabular entstammt dem Bildungsbereich.

### **Satzweise Bewertung der Rohübersetzung**

SATZ 1: VERSTÄNDLICH

NBA: MITTEL

MAßNAHMEN: /

SATZ 2: VERSTÄNDLICH

NBA: UMFANGREICH

MAßNAHMEN: /

SATZ 3: UNVERSTÄNDLICH wegen syntaktischer Unklarheiten.

NBA: MITTEL

MAßNAHMEN: Eintrag von „étranger“ (Adj.) ins WB.

SATZ 4: VERSTÄNDLICH

NBA: /

MAßNAHMEN: /

SATZ 5: VERSTÄNDLICH

NBA: MITTEL

MAßNAHMEN: /

SATZ 6: VERSTÄNDLICH

NBA: UMFANGREICH

MABNAHMEN: /

SATZ 7: VERSTÄNDLICH

NBA: UMFANGREICH

MABNAHMEN: Eintrag von „pays d'origine“ ins WB.

SATZ 8: UNVERSTÄNDLICH

NBA: UMFANGREICH

MABNAHMEN: Eintrag von „organisme“, „durée“ (Homograph von durer), und „supporter le coût“ ins WB.

SATZ 9: EINGESCHRÄNKT VERRSTÄNDLICH

NBA: UMFANGREICH

MABNAHMEN: s.o.

SATZ 10: VERSTÄNDLICH

NBA: UMFANGREICH

MABNAHMEN: /

SATZ 11: VERSTÄNDLICH

NBA: MITTEL

MABNAHMEN: /

SATZ 12: VERSTÄNDLICH

NBA: GERING

MABNAHMEN: s.o.

SATZ 13: EINGESCHRÄNKT VERSTÄNDLICH wegen syntaktischer Mängel

NBA: UMFANGREICH

MABNAHMEN: Evtl. Eintrag von „socio-culturel“ ins WB

SATZ 14: EINGESCHRÄNKT VERSTÄNDLICH wegen syntaktischer Mängel

NBA: MITTEL

MABNAHMEN: Eintrag von „étudiant“ (Subst. Homograph von „étudier“) ins WB.

## **Kommentar der Rohübersetzung insgesamt**

Bei drei unverständlichen, drei eingeschränkt verständlichen und immerhin neun verständlichen Sätzen liest sich die se Rohübersetzung relativ flüssig. Einen Strich durch die Rechnung macht die für eine öffentlich zu verwendende Übersetzung erforderliche sehr umfangreiche Nachbearbeitung, da die stilistischen Abweichungen von der autorisierten Übersetzung erheblich sind - sieben Sätze verlangen umfangreiche, fünf mittlere und nur einer geringe Nachbearbeitung.

## **Beschreibung des Textes „Résumé 1“**

Bei diesem Text handelt es sich um die französische Version einer mehrsprachigen Zusammenfassung eines Fachartikels über die Auswertung maschineller Übersetzungen (sic), der der einschlägigen Zeitschrift „Multilingua“ entnommen ist. Er enthält entsprechend linguistisches Fachvokabular, das sich aber aufgrund der Kürze des Textes quantitativ im Rahmen hält.

## **Satzweise Bewertung der Rohübersetzung**

SATZ 1: TEILWEISE VERSTÄNDLICH

NBA: UMFANGREICH

MABNAHMEN: /

SATZ 2: EINGESCHRÄNKT VERSTÄNDLICH wegen massiver syntaktischer Mängel

NBA: UMFANGREICH

MABNAHMEN: /

SATZ 3: TEILWEISE VERSTÄNDLICH

NBA: MITTEL

MABNAHMEN: /

## **Kommentar der Rohübersetzung insgesamt**

Bei einem eingeschränkt und zwei teilweise verständlichen Sätzen, die umfangreichen bzw. mittleren Nachbearbeitungsaufwand erforderlich machen, ist die Qualität dieser Rohübersetzung als mäßig zu bewerten, zumal sich über Maßnahmen der Wörterbuchpflege keine wesentliche Qualitätssteigerung erzielen läßt.

### **Beschreibung des Textes „Résumé 2“**

Ähnlich wie bei „Résumé 1“ handelt es sich um die französische Version einer mehrsprachigen Zusammenfassung eines Fachartikels aus dem Bereich der MÜ; hier geht es um die Kategorisierung der in maschinellen Rohübersetzungen auftretenden Fehlertypen. Der Text stammt ebenfalls aus „Multilingua“.

### **Satzweise Bewertung der Rohübersetzung**

SATZ 1: EINGESCHRÄNKT VERSTÄNDLICH

NBA: MITTEL

MAßNAHMEN: Eintrag von „brut“ in ein Spezial-WB.

SATZ 2: EINGESCHRÄNKT VERSTÄNDLICH

NBA: MITTEL

MAßNAHMEN: /

SATZ 3: TEILWEISE VERSTÄNDLICH

NBA: UMFANGREICH

MAßNAHMEN: Überprüfen des Eintrags für „viser“ im Kern-WB!

SATZ 4: VERSTÄNDLICH

NBA: MITTEL

MAßNAHMEN: /

SATZ 5: EINGESCHRÄNKT VERSTÄNDLICH wegen möglicherweise durch den Einschub verursachter syntaktischer Mängel bei der Auflösung von „corrigés“.

NBA: UMFANGREICH

MAßNAHMEN: /

SATZ 6: UNVERSTÄNDLICH wegen eines wesentlichen lexikalischen und eines grammatikalischen Fehlers

NBA: MITTEL

MAßNAHMEN: Eintrag von „grille“, „correction“ und „dresser“ ins WB. Da „correction“ kein Homograph von „correct“ ist, ist die Übersetzung mit „correct ones“ erstaunlich.

SATZ 7: UNVERSTÄNDLICH

NBA: GERING

MAßNAHMEN: Eintrag von „spectre“ ins WB.

SATZ 8: TEILWEISE VERSTÄNDLICH

NBA: UMFANGREICH

MAßNAHMEN: Eintrag von „grille“ und „traditionnel“ ins WB.

### **Kommentar der Rohübersetzung insgesamt**

Bei zwei unverständlichen, zwei teilweise und drei eingeschränkt verständlichen Sätzen sowie einem verständlichen Satz, die in zwei Fällen umfangreichen, in sechs Fällen mittleren und in einem Fall geringen Nachbearbeitungsaufwand notwendig machen, ist die Qualität dieser Rohübersetzung ähnlich mangelhaft zu bewerten wie die von „Résumé 2“, was aufgrund der Ähnlichkeit der Textsorte und der Wörterbucheinträge, auf die zugegriffen wurde, zu erwarten war.

### **Beschreibung des Textes „Sätze“**

Dieser Text besteht aus einer kleinen Sammlung von Beispielsätzen, in denen bestimmte linguistische Probleme für MÜ-Systeme stecken, wie Anapher und Ellipse. Die Sätze stammen von mir. Die Übersetzungen sprechen dafür, daß Weidner Probleme mit der Auflösung von Inversionen und Est-ce que- Konstruktionen des Französischen hat.



## **8. GESAMTBEWERTUNG**

### **8.1 Häufigkeit bestimmter Fehler**

#### **8.1.1. Syntax**

Satzbaufehler machen das umfangreichste Fehlerkorpus aus. Sie lassen sich in folgende Typen unterteilen:

##### **8.1.1.1 Bezugsfehler**

Satzteile werden falschen Partnern im Satz zugeordnet, was zu entsprechenden Umstellungen in der Übersetzung führt. Beispiel: „...le transfert du savoir-faire de la recherche à l'industrie“ -> „... the transfer the-know how to do research to the industry“. Trotz des Bindestriches zwischen „savoir“ und „faire“ wird die adverbiale Bestimmung „de la recherche“ als Objekt von „faire“ interpretiert, wodurch der Sinn des Satzes völlig verändert wird.

### **8.1.1.2 Wortartfehler**

Dieser Fehlertyp gehört in den Bereich der Homographieprobleme, da er meist auf der Wahl des falschen Homographen oder auf dem Fehlen des richtigen Homographen beruht. Beispiele: „Certains ont pour but ...“->„Certain airplanes for goal...“ Hier wird die Form „ont“ auf „avoir“ zurückverfolgt. Unter dem Eintrag von „avoir“ findet sich als Übersetzung eben auch „airplanes“ als Übersetzung der homographen Imperfektform „avons“. Es ist anzunehmen, daß Weidner über eine Regel verfügt, die besagt, daß a) am Satzanfang gewöhnlich das Subjekt steht und b) ein Adjektiv („certains“) meist von einem Nomen gefolgt ist. Ein weiteres Beispiel: „Les composants du système“->„the composing system“. In diesem Fall fehlt der adjektivische Homograph des Partizip Präsens von „composer“ im Wörterbuch, und Weidner hat daher keine Alternative zur partizipialen Übersetzung und der damit verbundenen Umstellung des Wortes vor das Substantiv „system“.

### **8.1.1.3 Reihenfolge Nomen-Attribut**

Bei Reihungen von Attributen bekommt Weidner Schwierigkeiten mit der Auflösung und richtigen Reihenfolge in der Übersetzung: „Un programme multitâches et orienté réseau“ ->„A multitask and oriented program network“. Gerade bei im Französischen durch „et“ verbundenen nachgestellten Adjektiven wird es problematisch.

### **8.1.1.4 Segmentierung komplexer Attribute**

Wird ein Nomen von einem Attributsatz gefolgt, wird die zielsprachliche Zuordnung von dessen einzelnen Komponenten problematisch: „un ensemble de fonctions accessibles par des commandes...“->„a body of accessible functions by orders“.

„Accessibles“ steht direkt hinter dem Substantiv und wird daher vor dieses gestellt. Der Rest des Attributsatzes wird separat behandelt. Hieraus ist zu schließen, daß die syntaktische Analyse auf sehr engem Raum innerhalb des Satzes erfolgt; es ist deshalb logisch, daß die Rohübersetzungen mit zunehmender Länge und Verschachtelung der Sätze schlechter ausfallen. Ein weiteres Beispiel: „La rigueur propre à la recherche“ ->

„The clean harshness to the research“. Hier wird ebenfalls „propre“ „rigueur“ zugeordnet (und beide falsch übersetzt), wodurch diese Phrase völlig unverständlich wird, da von der eigentlichen Struktur nichts mehr bleibt.

#### **8.1.1.5 Fehlende Artikel in der Rohübersetzung:**

„Le traitement de surface du cuivre“->„the surface copper treatment“. Richtig wäre:

„The surface treatment **of** copper“.

„Les produits de réaction du peroxyde“->„The reaction products peroxide. Richtig:

„The reaction products **from** peroxide.“

### 8.1.1.6 Falsche Relativpronomina

„...ce qui évite ...“ -> „... what avoids ...“. Besser: „which avoids“; richtig: „thus avoiding“.

### 8.1.2. Homographieprobleme

Ein immer wieder auftretende Fehlertyp war die Verwechslung von Wortarten und dadurch hervorgerufene syntaktische Fehlbezüge. Diese Fehler sind durch das Weidnersche Homographensystem bedingt. Da alle Homographen unter einem Lemma eingetragen werden, also Ableitungen, die mit den Partizipien und Personalformen eines Verbs identisch sind, auch unter die sem Verb gespeichert werden, entstehen häufig Ambiguitäten, bei denen das System in vielen Fällen eine falsche Entscheidung bezüglich der Wortart trifft, also ein Substantiv als Partizip (Übersetzung von „marché“ mit „walked“ statt „market“), ein Adjektiv als Substantiv, ein Verb als Adjektiv oder umgekehrt übersetzt wird. Es kam aber auch vor, daß ein Infinitiv als Substantiv übersetzt wurde: „On a tendance à les juger suffisants“ -> „One has \$tendance\$ to the sufficient judge“. Diese Fehlinterpretationen führen meist zu einer Umstellung von Satzteilen, hier zur Positionierung des Adjektivs vor das vermeintliche Substantiv. Da hier keine buchstäbliche Homographie vorliegt - es hätte ja „juge“ heißen müssen - sind Zweifel am korrekten Funktionieren des Homographensystems anzumelden. Zudem traten einige Fälle auf, bei denen sich keine lexikalische Erklärung für die angebotene Übersetzung finden ließ - z.B. „L'étude des erreurs vise d'une part à ...“ wurde zu „The study of sold one out on one errors hand to ...“. Nach dem Herausfiltern des syntaktisch versprengten „on one hand“ bleibt „The study of sold one out errors to“, nach Entfernung der Übersetzung von „L'étude des erreurs“ steht noch „sold one out to“ als Übersetzung von „vise à“, also „sell one out“ für „viser“.

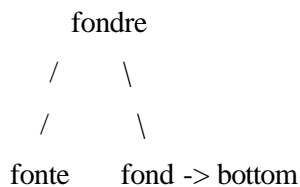
Sehr häufig war die Übersetzung von mit Partizipien homographen Substantiven in der Form „Adjektiv + one(s)“, z.B.: „les locaux“->„the local ones“, „les coopératives“ -> „the cooperative ones“. Diese Art von Homographieproblemen dürfte sich durch entsprechende Wörterbucheinträge reduzieren lassen.

Es entstehen aber auch Probleme durch vorhandene Homographeneinträge: „La figure

4.9 donne...“ -> „The figure 4.9 data ...“

Komplexere und schwer nachzuvollziehende Homographieprobleme sind diejenigen, bei denen ein Wort auf ein Verb zurückverfolgt wird, von dem aus dann ein falscher Homograph in der Übersetzung gewählt wird (siehe auch unter Syntax).

Beispiel: „Le traitement de fonte“ -> „the bottom treatment“. Man kann sich den Vorgang der Wörterbuchsuche in Baumform vorstellen:



Richtig wäre natürlich die direkte Zuordnung „fonte“ -> „cast“ gewesen.

### 8.1.3 Präpositionen

Präpositionen werden sehr häufig falsch übersetzt; das System hat seine Standardvorgaben für jede einzelne Präposition, die sich bei „de“ und „à“ gleich bei der Eingabe eines Verbs ins WB ändern läßt; handelt es sich aber um andere Präpositionen, muß die entsprechende Wortkombination als „idiom“ eingetragen werden. Beispiel: „dotée d'un langage“ -> „endowed a language“. Vermutlich steht das Verb „to endow“ als transitiv in WB, ohne daß die korrekte Übersetzung „endowed with“ für ein nachstehendes „de“ eingetragen wurde.

### 8.1.4 Pronomina

Pronomina werden ebenfalls sehr häufig falsch wie dergegeben, da die Analyse nicht über den Satz hinausgeht. Beispiel: „La maison est belle. Elle est verte.“ würde zu der Übersetzung „The house is beautiful. She is green“ führen und nicht „It is green“.

### 8.1.5 Tempus von Verbformen

Bei zusammengesetzten Verbformen des französischen Perfekts kommt es vor, daß die

Formen separat übersetzt werden und so ein Präsens konstruiert wird: „La femme que j'ai vue est morte“ -> „The woman that I saw is dead“. Hier liegt insofern ein Sonderfall vor, als entweder eine adjektivische Übersetzung mit „dead“ oder eine verbale Übersetzung wie „The woman I saw died“ möglich wäre.

Ein anderer Fehlertyp ist die Beibehaltung des französischen Tempus in Fällen, die im Englischen ein anderes Tempus verlangen würden, z.B.: „plusieurs systèmes sont déjà installés“ -> „several systems already are installed“. „Several systems have already been installed“ wäre richtig.

### **8.1.6 Zeichensetzung**

Alle Zeichen werden so übernommen, wie sie im Quelltext stehen. Bei Zahlen werden französische Kommata nicht in Punkte, Punkte nicht in Kommata umgewandelt. Ebenso werden Doppelpunkt-Konstruktionen übernommen, die im Englischen stilistisch weniger eingesetzt werden.

### **8.1.7 Lautbindung**

Vor mit einem Vokal beginnenden englischen Wörtern wird „a“ nicht in „an“ umgewandelt.

## **8.2 Lexik**

Abgesehen von falschen Übersetzungen, die durch im Wörterbuch fehlende Einträge bzw. durch Homographie verursacht werden, treten immer wieder Fälle auf, in denen gemeinsprachliche Verben, Adjektive oder Substantive verschiedene Bedeutungen haben, die in ein und dasselbe Wörterbuch gehören, da sie unabhängig voneinander in einem Text auftreten können. Beispiel: „cours“ kann mit „course“ oder „courtyard“ übersetzt werden, „entrainer“ mit „cause“, „bring about“ oder „induce“. Im Wörterbuch steht immer nur eine dieser Varianten. Häufig handelt es sich nur um stilistische Alternativen wie bei „cause“ und „bring about“; bei „course“ und „courtyard“ ist dies anders. Probleme mit Eigennamen und Abkürzungen, Akronymen und Siglen lassen sich durch einen einmaligen Eintrag ins WB aus der Welt schaffen, kommen sie nur selten vor, lassen sie sich im Quelltext markieren. Vergißt man dies, entstehen eventuell massive Probleme bei der Nachbearbeitung, insbesondere bei übersetzbaren Eigennamen; es sei hier nur an die Übersetzung des holländischen „van“, das in einem von Systran zu übersetzenden englischen Text vorkam, mit „camion anglais“ erinnert.

## **8.3 Schlußfolgerung**

Aus dem Gesagten geht klar hervor, daß dieses System sich nicht für Übersetzungen eignet, die hohe Ansprüche in punkto Stil stellen, die einen hohen Grad von Ambiguität enthalten oder stark verschachtelt sind. Weidner eignet sich für die Übersetzung von sich stark wiederholenden Texten, bei denen seine Stärke, nämlich die der Terminologiekonsistenz, voll zum Tragen kommen kann. Durch umfangreichen Ausbau des Wörterbuchsystems läßt sich die Übersetzungsqualität auf ein Niveau bringen, das die Nachbearbeitung auf ein wirtschaftliches Maß reduziert.

Zweitens eignet sich ein computergestütztes Übersetzungssystem, wie bereits im Kapitel über Nachbearbeitung beschrieben, für die Herstellung von zum schnellen Gebrauch bestimmten Informationsübersetzungen, die nach der beschriebenen EG-Methode der schnellen Nachbearbeitung in verwendbare Form gebracht werden. In diesem Bereich tut sich ein großer Markt auf, der den rechnergestützten Übersetzungssystemen zu rasch wachsender Verbreitung verhelfen dürfte.



Eidesstattliche Erklärung:

Hiermit versichere ich an Eides statt, keine anderen als die angegebenen Hilfsmittel verwendet zu haben.

Hildesheim, den 30.07.1987

.....