

In: Lebende Sprachen 1/1980 XXV. Jahrgang, 2-5

Im Jahre 1966 erschien in den USA der Bericht „Language and Machines, Computers in Translation and Linguistics“ des Automatic Language Processing Advisory Committee (sog. ALPAC-Bericht). Aufgrund des darin enthaltenen vernichtenden Urteils über die Zukunftschancen der maschinellen Übersetzung kam die einschlägige Forschung durch Einstellung der Förderung - besonders in den USA - weitgehend zum Erliegen. Inzwischen hat sich herausgestellt, dass sie nur scheinbar tot war und zur Zeit eine Renaissance erlebt. Als Schwerpunkt setzen sich die folgenden vier Beiträge unter verschiedenen Aspekten mit dem derzeitigen Stand der Entwicklung auseinander. In einem späteren Heft der „Lebenden Sprachen“ werden wir über die geplante europäische Gemeinschaftsentwicklung des maschinellen Übersetzungsverfahrens EUROTRA berichten. Die Schriftleitung

PROF. HARALD ZIMMERMANN, Regensburg

## **Stand und Probleme der maschinellen Übersetzung**

### *1. Zielsetzung*

Ausgangspunkte und Rahmen der nachfolgenden Bemerkungen zum Stand der maschinellen oder computergestützten Übersetzung seien zwei Zitate: „Alles Übersetzen scheint mir schlechterdings ein Versuch zur Auflösung einer unmöglichen Aufgabe. Jeder Übersetzer muss an einer der beiden Klippen scheitern, sich entweder auf Kosten des Geschmacks und der Sprache seiner Nation zu genau an das Original, oder auf Kosten seines Originals zu sehr an die Eigentümlichkeiten seiner Nation zu halten. Das Mittel hierzwischen ist nicht bloß schwer, sondern geradezu unmöglich.“ (Brief von W. v. HUMBOLDT an A. W. SCHLEGEL vom 23. 7. 1796). In dem im vorliegenden Zusammenhang sehr lesenswerten Artikel „Übersetzung“ in „Handbuch der Linguistik“ - er enthält übrigens auch die vorstehende Bemerkung von W. VON HUMBOLDT - bezeichnet W. WILSS den bei der Humanübersetzung ablaufenden Vorgang zusammenfassend „als einen Prozess der Auswahl, Kombination und Kompensation von interlingual identischen, teilidentischen und nichtidentischen Faktoren (...), die in bestimmtem Umfang präzise angegeben werden können und damit vorhersagbar sind“, und folgert daraus, daß es möglich sein müsse, „durch eine adressaten- und textspezifisch differenzierte syntaktische, syntagmatische und lexikalische Komponentenanalyse die Ursache vieler Übersetzungsschwierigkeiten zu ermitteln und diese mit Hilfe systematisch entwickelter textadäquater Lernstrategien und Lerntechniken zu überwinden“ (Handbuch der Linguistik S. 535).

Festzuhalten ist daher zunächst zweierlei : Eine völlige Identität ausgangssprachlicher und zielsprachlicher Ausdrücke ist nicht immer erreichbar (dies gilt m. E. nicht nur für die Dichtersprache); Übersetzen ist vielfach mehr als die Auswahl zielsprachlicher „Äquivalente“ (vgl. die Kompensationsproblematik, die im übrigen auch in der Kompensation während des Übersetzungsvorganges erkannter ausgangssprachlicher Schwächen oder Fehler bestehen kann).

Man mag sich nun fragen, ob diese Hinweise im folgenden als Argument für oder gegen eine maschinelle Übersetzung (MT) benutzt werden sollen. Sie könnten gleichsam einerseits die derzeitigen Mängel maschineller Verfahren entschuldigen, wobei allenfalls graduelle und nicht prinzipielle Unterschiede zwischen Maschinen- und Humanübersetzung zugestanden

werden; andererseits könnte die menschliche Fähigkeit, „intelligent“ - und vielleicht algorithmisch - zu agieren und zu reagieren und damit gerade im „Kompensationsbereich“ der Übersetzung unentbehrlich zu sein, als Indiz dafür gewertet werden, dass MT letztlich unvergleichbar mit der Humanübersetzung (und damit ein Versuch mit untauglichen Mitteln) sei. Bei den folgenden Überlegungen sollen solche dichotomen Thesen zur Vergleichbarkeit menschlicher und maschineller Übersetzungsvorgänge außer Betracht bleiben - so interessant es wissenschaftlich sein mag.

Festzuhalten ist jedoch (trivialerweise), dass es gegenwärtig keine der Humanübersetzung der üblichen Qualität vergleichbare vollautomatische Übersetzung für beliebige Texte gibt. Andererseits existiert eine Reihe von Systemen, bei denen man - zumindest in Ansätzen - von Übersetzungsfunktionen reden kann, die über eine reine Wort-für-Wort-Umsetzung hinausgehen. In Entwicklung und Erprobung befindliche Verfahren (z. B. CETA, METEO, CULT und SUSY bzw. das geplante System EUROTRA) streben auf der Grundlage neuer linguistischer Modelle eine Qualitätsverbesserung an, die für textuell breiter angelegte Systeme wie SUSY, EUROTRA usw. zumindest zu brauchbaren Rohübersetzungen bzw. zu fachgebiets-spezifischen Einsätzen (METEO) führen sollen.

Daneben bleibt festzustellen, dass - übrigens entgegen vielen früheren Prognosen - auch Ende der 70er Jahre der eigentliche „Durchbruch“ für die MT nicht erfolgt ist. Dies wird m. E. zu Unrecht dem vielzitierten ALPAC-Report angelastet, obgleich besonders in den USA in seinem Gefolge einige Projekte zur MT (z. B. das GARVIN-Verfahren) nicht mehr weiterentwickelt werden konnten. Dennoch ist auch in Amerika die Entwicklung inzwischen weitergegangen, zumindest im kommerziellen Bereich (vgl. SYSTRAN), bzw. verlagerte sich in verwandte Gebiete (die Sprachverstehenssysteme der „Künstlichen-Intelligenz-Forschung“ von WINOGRAD und WOODS sind hierzu zu rechnen).

In Europa ging nahezu unberührt vom ALPAC-Report die Entwicklung kontinuierlich (wenn auch langsam) weiter: Die Verfahren des Sonderforschungsbereichs (SFB) 100 (SUSY) und der Universität Grenoble (System CETA) sind hierfür Beispiele; in dieser Reihe sind auch zu nennen verwandte Projekte wie PLIDIS/ISLIB am Institut für deutsche Sprache bzw. das (inzwischen eingestellte) LIMAS-Vorhaben: alle diese Systeme sind jedoch eher der Grundlagenforschung oder allenfalls der modellhaften Anwendungsforschung zuzurechnen. Sie stehen insgesamt unter der Belastung, den Nachweis zu führen, dass MT machbar und brauchbar ist. Nur einige wenige Verfahren (z. B. SALAT am SFB 99 in Konstanz/Heidelberg) sind der Konzeption und Erprobung linguistischer Modelle gewidmet und frei von irgendwelcher Anwendungsproblematik.

Dieser Nachweis der Machbarkeit von MT ist bis heute nicht gelungen. Man kommt nicht umhin, nach einer wesentlichen Ursache zu suchen, will man dabei nicht wieder die These der prinzipiellen „Unübersetzbarkeit“ sprachlicher Äußerungen durch Computer verfechten. Ein entscheidender Grund für die Langfristigkeit der Entwicklung von MT-Verfahren (vielleicht der wichtigste) ist m. E. die Komplexität der natürlichen Sprache.

Dies (ver)führte bei vielen Systemen dazu, den Verwendungsbereich (explizit oder meist nur implizit) stark einzuschränken. Einschränkungen können dabei im syntaktischen, morphologischen und semantischen Bereich liegen. Syntaktische Einschränkungen liegen zum Beispiel vor, wenn nur Hauptsätze zugelassen sind oder das Subjekt am Satzanfang stehen muss, semantische Einschränkungen liegen vor, wenn ein Wort nur in einer spezifischen Bedeutung verwendet oder akzeptiert werden kann usf.

Dies führt naturgemäß zu drastischen Beschränkungen für den Benutzer und stellt letztlich die praktische Anwendbarkeit generell in Frage. Daneben sind unter diesen strategischen Voraussetzungen kaum Systemerweiterungen zugelassen, da zur Behandlung von Mehrdeutigkeiten bzw. Alternativlösungen keine Verfahren angeboten werden.

## 2. Komponenten einer maschinellen bzw. computergestützten Übersetzung

Während die Zielsetzung der einzelnen Übersetzungssysteme und -modelle, soweit sie zu „praktischen“ Resultaten (d. h. konkreten Übersetzungen) führen sollen und nicht allein sprach- oder übersetzungstheoretischen Untersuchungen dienen, naturgemäß sehr ähnlich ist, sind die angewendeten Verfahrensweisen und zugrunde gelegten Konzepte recht unterschiedlich. Im folgenden sollen unter Hinweis auf einige der bekannteren Systeme zunächst diese unterschiedlichen Verfahren exemplarisch kurz beschrieben werden.

Bei der „vollautomatischen“ Übersetzung, wie sie z. B. das System SYSTRAN, aber auch das im Sonderforschungsbereich der Universität des Saarlandes im Aufbau befindliche Verfahren zur maschinellen Übersetzung (SUSY) oder das in Grenoble entwickelte MT-System CETA der G.E.T.A. (*Groupe d'Etudes pour la Traduction Automatique*) praktizieren, ist das entscheidende Merkmal, dass keine - oder nur marginale - problembezogenen Textvorbereitungen (Präkodierungen, Präditionen) vorgenommen werden, dass daneben während einer Text- oder Satzanalyse keine Interaktion zwischen dem System und einem menschlichen Bearbeiter erfolgt. Interaktion wäre z. B. möglich zwischen dem System und einem - sprachpaarbezogenen - Humanübersetzer oder einem Systemspezialisten, der die Probleme der Analyse bzw. auch Synthese kennt und Hilfestellung geben kann, z. B. bei automatisch nicht adäquat auflösbaren Mehrdeutigkeiten.

Dies bedeutet nicht, dass in einem „vollautomatischen“ MT-System kein Feedback zwischen Benutzer und System möglich ist: Diese Rückkoppelung ist jedoch losgelöst von der (Ad-hoc-)Situation einer konkreten Analyse, sie erfolgt im Nachhinein, etwa nach systematischer Auswertung von Fehlern; selbstverständlich gehört dazu eine systematische Erweiterung des Vokabulars. Im allgemeinen begnügen sich solche Verfahren mit einer Art „Rohübersetzung“, denn natürlich weiß man, dass mit Hilfe der vollautomatischen Übersetzung nicht fehlerlos und meist nicht einmal ausreichend gut übersetzt werden kann. Für den (davon entkoppelten) Prozess einer Nachbereitung (Postedition - ähnlich den vielfach in Übersetzungsbüros üblichen „Revisionen“, aber aufgrund der unterschiedlichen Verfahrensweise und Problematik damit kaum vergleichbar -) können durchaus auch Alternativlösungen angeboten werden, z. B. bei SUSY und SYSTRAN nach Bedarf, soweit das System eine erkannte Mehrdeutigkeit (auch im Bereich der Syntax) nicht auflösen konnte.

Die Schwächen solcher „vollautomatischen“ Systeme liegen auf der Hand: Verschiedene Bewertungen von SYSTRAN-Übersetzungen haben etwa deutlich gemacht, dass z. B. eine Postedition - losgelöst vom eigentlichen maschinellen Übersetzungsprozess - einen außergewöhnlich hohen intellektuellen Mehraufwand nach sich zieht.

Man wird also bei den auf die praktische Anwendung ausgerichteten Systemen mittelfristig nicht ohne intellektuelle Interaktionen oder verstärkte Präkodierungen, die dabei in den Übersetzungsprozess integriert sind, auskommen. Ein Verfahren, das diese intellektuelle „Aktion“ schwerpunktmäßig auf den Input-Teil legt, ist das System TITUS (II).

Dieses im Bereich der Textildokumentation zur Übersetzung von Abstracts (multilingual: Französisch - Deutsch - Englisch - Spanisch - Italienisch) eingesetzte System arbeitet mittels

einer streng formalisierten Hilfssprache nach Art einer - allerdings semantisch orientierten - Dependenzgrammatik und legt als Vokabular einen fest gefügten Thesaurus (mit vereinheitlichten Begriffen) zugrunde.

Dieser Thesaurus, der von einem Team von Experten (Terminologen) gepflegt wird, stellt die eigentliche „*Switching language*“ zwischen den Sprachen dar, während eine Anzahl von - mehrsprachig aufeinander abgestimmten - Satzmustern oder -rahmen (nur) mit entsprechenden Vokabeln gefüllt werden muss, die zudem mit formalen Kennzeichen zur Morphologie (z. B. + für Plural) versehen werden. Der Dokumentar muss - neben seiner Muttersprache - (prinzipiell nur) das Vokabular des Thesaurus beherrschen und die alternativen syntaktischen Rahmen kennen. Der erfolgreiche Einsatz von TITUS-II zeigt die prinzipielle Praktikabilität dieser Methode. Zur „Sicherheit“ werden die „Übersetzungen“ (besser sollte man hier von „Umsetzungen“ reden) nochmals kontrolliert und ggf. nachbereitet; trotz einiger Mängel, die in dem doch beträchtlichen Kodieraufwand begründet sind (und durch ein verbessertes System - TITUS-IV - überwunden werden sollen), hat sich diese Vorgehensweise deutlich bewährt.

Natürlich ist nicht zu verkennen, dass ein Einsatz solcher Verfahren nur dann sinnvoll oder möglich ist, wenn auf die Struktur des „Inputs“ (wie in der Dokumentation z. T. möglich) noch Einfluss genommen werden kann und wenn - auch dies ist zu bedenken - keine großen „stilistischen“ Ansprüche an die zielsprachlichen „Äquivalente“ gestellt werden.

Ein anderer, in diesem Zusammenhang erwähnenswerter Weg wird im LOGOS-System (I bis III) verfolgt. Hier handelt es sich um ein reines „computergestütztes“ System, bei dem der Mensch aktiv in den maschinellen Übersetzungsprozess eingreifen kann. In der „Pre-Editing-Phase“ wird dies vor allem unter lexikalischen Aspekten gesehen: Fehlende Wörter zu einem zu übersetzenden Text können unmittelbar ergänzt werden, ein (zweisprachiges) Ad-hoc-Wörterbuch kann auf die im Text notwendigen Übersetzungen „eingestimmt“ werden u.a.m. In einer Nachbereitungsphase (Post-Editing) kann die probeweise angegebene Übersetzung systematisch (maschinenunterstützt) überprüft und abgeändert werden; zugleich erfolgt dabei ein Feedback zur Wörterbuchergänzung bzw. -korrektur. Ein ähnlicher Weg wird in dem Verfahren CULT besprochen. Hier bietet das System bei Feststellen einer Mehrdeutigkeit vereinheitlichte natürlingsprachige Paraphrasen an, die der Informationsmittler (z. B. Humanübersetzer) dann „entscheidet“, womit in einer frühen Analysephase - ohne viel Aufwand - die „korrekte“ Übersetzung gesteuert werden kann.

All diese Verfahren scheinen einen Hauch von Resignation zu tragen. Wie kann man aber - abgesehen von dieser „Humanunterstützung“ - die Qualität der Übersetzung entscheidend verbessern? Da hier kein historischer Überblick zur Entwicklung der MT (und ihrer ungelösten Probleme) gegeben werden kann, seien die wohl wesentlichen Erkenntnisse der bisherigen MT-Entwicklung thesenartig zusammengefasst:

- die Wort-für-Wort-Übersetzung war eine Sackgasse;
- Syntaxanalyse allein (auch auf der Grundlage der generativen Transformationsgrammatik) reicht nicht aus;
- Semantik (Integration in das Verfahren bzw. als Basis) ist notwendig (aber welche?);
- Satzbezogene Sprachübersetzung - wie nahezu in allen gegenwärtigen Systemen praktiziert - ist unzureichend  
(Problem der anaphorischen Verwendung, der pronominalen Referenz ...)

Natürlich bietet es sich an - und dies wird exemplarisch immer wieder erfolgreich demonstriert - MT-Verfahren auf der Basis von Erkenntnissen der Forschungen im Bereich der künstlichen Intelligenz zu entwickeln (LENDERS 1975), d. h. das sog. „Weltwissen“ in den Übersetzungsprozess zu integrieren. Abgesehen davon, dass der Aufwand für den Aufbau derartiger „Miniwelten“ erheblich ist („Miniwelt“ deshalb, weil das „Weltmodell“ - wegen der Zahl der Verknüpfungen und auch des Problems der [noch] notwendigen Konsistenz/Folgerichtigkeit dieser „Welt“ - nur auf einen sehr, sehr kleinen Ausschnitt der „realen“ Welt beschränkt ist und auch kaum übertragbar erscheint) : Es gibt kaum eine „reale“ Mini-Welt, bei der für ein Übersetzungssystem Bedarf besteht. Allerdings sind inzwischen einige Bereiche „entdeckt“ worden, die relativ geschlossene Fach- oder Sachgebiete enthalten (sei es im Bereich der Syntax oder/und des Vokabulars), die man zwar nicht so eingeschränkt handhaben kann wie die sog. „Mondgesteinwelt“ von WOODS oder die „Bauklötzchenwelt“ von WINOGRAD, die aber abgegrenzt genug erscheinen, um den Einschränkungseffekt auszunutzen, der mit der Auswahl eines engeren Fachgebiets einhergeht. Zumindest im Bereich der Wortsemantik tritt üblicherweise eine starke Reduktion der Mehrdeutigkeiten ein; noch vorhandene Mehrdeutigkeiten lassen sich auch mit (einfachen) heuristischen Methoden recht gut auflösen (z. B. „wenn Wort A mehrdeutig und Wort B im Satz/Absatz, dann ist bei A die Bedeutung i gemeint“); selbst im Bereich der Syntax lassen sich hier leichter Regularitäten bzw. Kriterien entdecken, die mit hoher Wahrscheinlichkeit eine automatische Vereindeutigung erlauben.

Ein derartiges, auf die Übersetzung spezieller Texte ausgerichtetes Verfahren stellt das in Montreal entwickelte System METEO dar. METEO soll der durchaus (für Kanada) sinnvollen Übersetzung von Wettervorhersagen dienen (böse Zungen könnten behaupten, dass bei der bekannten Vorhersageschwäche von Wetterämtern - abgesehen von Sonnen- und Mondaufgang - eine falsche Übersetzung hier vielleicht sogar die richtige Wettermeldung bringen würde) und wird auch praktisch eingesetzt. Wetterberichte stellen zwar keineswegs einfache, aber relativ stark schematisierte Ausschnitte zu „natürlichen“ Sprachen dar, sie sind also nicht unbedingt typisch für die Allgemeinsprache oder auch die wissenschaftlichen Fachsprachen.

Das zweite Verfahren aus Montreal, AVIATION, soll der Übersetzung von Handbüchern (Manuals) zur Flugzeugwartung dienen. Trotz dieser Fachgebiets- und Textsortenbeschränkung ist - wie (sogar) bei METEO - auch eine menschliche Interaktion (vor allem Prä- und Postedition) eingeplant.

Inzwischen sind wohl alle Entwickler und Administratoren von produktionsorientierten Übersetzungssystemen aus Erfahrung klug geworden: Eine „Fully Automatic High Quality Translation“ (sog. „FAHQT“) ist nirgendwo mehr explizit das Ziel. Es beginnt sich mehr und mehr die Erkenntnis durchzusetzen, dass die Beschränkung auf Fachgebiete nicht nur aus ökonomischen Gründen sinnvoll ist.

### *3. Linguistische Modelle in der MT*

Seither sind seit dem ALPAC-Report von 1966, der nach Meinung vieler Experten ein vernichtendes Urteil über die möglichen Anwendungen maschineller Übersetzungsverfahren gefällt hat, schon mehr als ein Dutzend Jahre mit durchaus angestiegener und bewegter theoretischer und praktischer Sprachforschung vergangen. Es ist zu fragen, ob von dieser Seite aus inzwischen neue Konzepte vorliegen oder in MT-Verfahren integriert sind, die evtl. vorhandene neuere Erkenntnisse umsetzen und damit eine „neue Epoche“ der MT einleiten können. In der Tat sind einige m. E. viel versprechende linguistische Modelle entwickelt worden, die auch - abgesehen von der schon erwähnten Miniwelt-Konzeption - in der „Praxis“ einer fach-

gebietsorientierten, humanunterstützten Computerübersetzung (um einmal den allgemeinen Trend der MT verbal zu charakterisieren) von Relevanz sind oder sein könnten.

Das in der Generativen Transformationsgrammatik von N. CHOMSKY entwickelte, syntaxorientierte Sprachmodell - angereichert um die FODOR-KATZ-Semantik (sog. interpretative Semantik, in der MT meist benutzt zur kontextsensitiven Restrangierung von syntaktischen Mehrdeutigkeiten) - ist etwa mit Modifikationen die semantische Grundlage des Systems SUSY; generative Transformationsregeln zur Ermittlung sog. Tiefenstrukturen oder kanonischer Strukturen, wobei eine Oberflächenstruktur auf verschiedene Tiefenstrukturen bzw. verschiedene Oberflächenstrukturen auf eine gemeinsame Tiefenstruktur abgebildet werden (z. B. Passiv-Aktiv-Transformation im Deutschen), werden auch verwendet in dem in Austin/Texas 1972 begonnenen und inzwischen wieder fortgesetzten Projekt METALS (*Mechanical Translation and Analysis of Languages System*).

Ein stärker dependenzielles Konzept - jedoch ebenfalls im Blick auf „Normierung“ und Kanonisierung - wurde in Grenoble im MT-Projekt CETA verfolgt, wobei auch Konzepte des russischen Linguisten MEL'CHUK zugrunde gelegt wurden, der eine Zeitlang ebenfalls MT-Entwicklung betrieben hatte: Eine Tiefenstrukturnotation ist dabei der baumartigen Oberflächenbeschreibung zugeordnet, ebenso Informationen zur semantischen Interpretations-/Repräsentationsfunktion der syntaktischen Kategorien.

Das Ende der 60er Jahre von FILLMORE entwickelte Modell der Tiefenkasus ist - obwohl es eine Reihe von Problemen überwinden zu helfen verspricht, die in den vorhandenen MT-Systemen noch anstehen - wie auch das Modell der „generativen Semantik“ - meines Wissens noch nicht in der MT-Praxis erprobt. Dies mag daran liegen, dass die derzeit realisierten (modernen) Systeme ihre konzeptionellen Ursprünge bereits etwa Mitte/Ende der 60er Jahre haben, als gerade die generative Transformationsgrammatik voll internalisiert war, Modelle der Kasusgrammatik oder der generativen Semantik aber erst im Entstehen waren.

Daneben entwickelte die MT-Forschung naturgemäß - etwas abseits von der theoretischen Linguistik - ihre eigenen Strategien, die vor allem zur Bewältigung der komplexen Mehrdeutigkeitsprobleme in natürlichen Sprachen dienen. Es entstand dabei u. a. ein langer - und m. E. wenig fruchtbarer - Streit um die Frage der Integration oder Trennung von Grammatik (Regelsystem) und Algorithmus: heute fragt man sich oft, ob es sich bei dieser oder jener „Grammatik“ nicht um ein „Programm“ in einer höheren oder „*highest level*“-Programmiersprache handelt (man betrachte z. B. diesbezüglich das CETA-System); Ähnliches gilt für den inzwischen bei Modellanalysen recht häufig eingesetzten NetworkParser von WOODS, der eine Kombination von Regeltrennung und Regelintegration darstellt.

Allgemein akzeptiert scheint inzwischen - zumindest konzeptionell - die Gliederung der MT-Verfahren in 3 Komponenten zu sein: Analyse - Transfer - Synthese. Die Analyse zerfällt dabei in der Regel wieder in zwei Hauptbereiche: Morphologie (Wörterbuchsuche, Anwendung von Dekompositions- und Derivationsregeln, z. T. Erkennung von idiomatischen Wendungen) und syntakto-semantische Analyse, wobei diese 2. Phase - mehr oder weniger technologisch-analysestrategisch begründet - entweder verzahnt wird (wie bei CETA und den Verfahren in Montreal) oder stufenweise erfolgt (wie bei SUSY oder METALS).

Der Transferteil besteht idealerweise eigentlich nur aus einer lexikalischen Zuordnung der (vereindeutigten) ausgangsprachlichen Begriffe zu (eindeutigen) zielsprachlichen Äquivalenten. Aufgrund mangelnder Analysetiefe (und damit auch unterschiedlicher Basisstrukturen), z. T. auch aus ökonomischen Gründen (Reduktion von Rechenzeiten) wird in der

Transferphase (zielsprachenabhängig) aber noch weitere Analysearbeit nötig. Bei einigen Systemen, z. B. CETA, erfolgt auch ein Transfer der Basis-Strukturbäume der Ausgangssprache in solche der Zielsprache (hier wird auf Erhaltung der Oberflächenstrukturinformationen größerer Wert gelegt).

Die zielsprachliche Synthese - meist wieder zweigeteilt in eine Erzeugung der syntaktischen Oberflächenstruktur und eine nachfolgende morphologische Synthese - erscheint, gemessen an dem Aufwand, der für Analyse und Transfer betrieben werden muss, gering; die eigentlichen Probleme, z. B. zur Generierung eines bestimmten/unbestimmten Artikels oder zur Erzeugung des „richtigen“ Pronomens, sollten dann allerdings spätestens im Transfer gelöst sein.

Wenn man heute aufgrund der Entwicklungen der Linguistik und nicht zuletzt aufgrund der „Erkenntnis“ der Übersetzungsproblematik bei den meisten MT-Systementwicklern auch „gedämpft optimistisch“ in die nähere Zukunft blicken kann - zumindest im Hinblick auf die Entwicklung brauchbar-praktischer computergestützter Übersetzungsverfahren - so scheinen sich andererseits noch keine sprachwissenschaftlich motivierten generellen Lösungen zur Bewältigung des MT-Problems anzubieten. Selbst der Bereich der Semantik bietet bislang keine zuverlässig verwendbaren und damit praktischen Nutzen versprechenden Modelle an.

#### *4. Perspektiven*

Im Bereich von Hard- und Software wurden in den letzten 15 Jahren ohne Zweifel die größten Fortschritte erzielt. Die Realisierung eines MT-Systems im Heimcomputer ist in den Bereich des ökonomisch und technisch Möglichen gerückt. Damit ist eine Barriere gefallen, die entscheidend zur Haltung der Gutachter des ALPAC-Reports beigetragen hatte. Diese ökonomisch-technischen Entwicklungen beginnen gegenwärtig auch in den „traditionellen“ Übersetzungsbereich einzudringen. Größere Unternehmen, wie z. B. in Deutschland der Übersetzungsdienst bei Siemens und das Bundessprachenamt in Hürth, sind in der Terminologiearbeit - einem überschaubaren Teilbereich des Übersetzungswesens - bereits auf EDV-Unterstützung „umgestiegen“ - und dies mit großem Erfolg. Mit der Einführung von Textsystemen auf Bildschirm-Basis unter Zugrundelegung des gewohnten Zeichenvorrats wird sich die Bereitschaft der Übersetzerdienste und Übersetzer verstärken, Computerhilfe in Anspruch zu nehmen - und sei es zunächst als Ersatz für den Zettelkasten.

Bei den perspektivischen Projekten zur computergestützten Übersetzung wird man gut daran tun, den Übersetzer nicht - wie bisher üblich - zwangsweise vor ein nicht auf seinen Bedarf zugeschnittenes fertiges System zu setzen; vielmehr wird man schrittweise, über die Unterstützung in der Terminologiearbeit, die Hilfe bei der Suche nach Übersetzungsäquivalenten, die Aufbereitung des zu übersetzenden Textmaterials einschließlich der Erfassung des zielsprachlichen Texts und der Revisionshilfen - weitere und interessantere Bereiche der computergestützten Übersetzung angehen - möglichst mittels und unter Mitbestimmung bzw. Mitwirkung der Humanübersetzer.

Es ist möglich, dass sich dabei das Tätigkeitsfeld des Übersetzers (nicht eines jeden übrigens, sondern auf lange Sicht allenfalls in Teilbereichen) verändern wird: Terminologiearbeit (z. B. im Team) wird in den Vordergrund treten, Textuntersuchungen (auf strukturelle Besonderheiten, fachtextbezogene Einschränkungen) treten hinzu, die Interaktion mit einem System während der Texterschließung der Ausgangssprache gewinnt vielleicht gegenüber der zielsprachlichen Kodierung oder Korrektur an Bedeutung.

Beflügelt von der technologischen Entwicklung, wird die computergestützte Übersetzung allmählich in den Alltag hinaustreten, so wie heute der Jumbo-Jet fast Alltag geworden ist. Unter diesem Aspekt sollte auch das gewiss recht ehrgeizige Projekt EUROTRA betrachtet werden: Es gilt, zur Bewältigung der wachsenden Kommunikationsprobleme (sofern sie sprachlicher Natur sind) in einem zusammenwachsenden Europa ein entsprechendes modernes Instrumentarium - neben anderen - zu entwickeln.

Diese technologische Unterstützung sollte dabei nicht zum Selbstzweck werden; gesellschaftliche, soziale und menschliche Probleme sind dabei entscheidend mit zu berücksichtigen.

## Literatur

ALPAC *Language and Machines. Computers in Translation and Linguistics, A Report by the Automatic Language Processing Advisory Committee (ALPAC), National Academy of Sciences - National Council, Washington, D. C. (1966).*

BOITET, C.: *Stand der Arbeiten der G.E.T.A. Anfang 1977*, in: Kommission der Europäischen Gemeinschaften (Hrsg.), *Dritter Europäischer Kongress über Dokumentationssysteme und -netze. Die Überwindung der Sprachbarrieren*, Luxemburg, 3.-6. Mai 1977, Band 2, München, S. 154-184 (1977).

BRUDERER, H. E.: *Handbuch der maschinellen und maschinenunterstützten Sprachübersetzung. Automatische Übersetzung natürlicher Sprachen und mehrsprachige Terminologiedatenbanken*, München, New York (1978).

CHAUCHÉ, J.: *Présentation du système C.E.T.A., Grenoble (1976), Transducteurs et arborescences - Etude et réalisation des systèmes appliqués au Treatment automatique de langues naturelles*, Thèse d'Etat, Grenoble (1976).

CHEVALIER, M., DANSEREAU, J., POULIN, G.: *TAUM-METEO. Description du système*, Montréal.

CHOMSKY, N.: *Aspects of the theory of syntax*. Cambridge, Mass., M.I.T. Pr. (1965).

DIETRICH, R., KLEIN, W.: *Computerlinguistik. Eine Einführung*, Stuttgart (1974).

FREIGANG, K. H., THOME, G., WILSS W.: *Der Stand der Forschung auf dem Gebiet der maschinellen Übersetzung*, in: Linguistische Arbeiten Neue Folge, Heft 2, Universität d. Saarlandes, Sonderforschungsbereich 100, Elektronische Sprachforschung, Saarbrücken (1979).

GARVIN, P. L.: *On machine translation: selected papers*. The Hague, Mouton (1972).

HUTCHINS, W. J.: *Machine translation and machine-aided translation*, in: Journal of Documentation 34, 1978, 119-159 (Progress in Documentation).

ISABELLE, P. et al.: *TAUM-AVIATION: un système pur la traduction automatique de manuels techniques*. Vortrag COLING '78, Bergen, 14.-18. August (1978).

KLEIN, W.: *Organisation des Wissens durch Sprache. Konsequenzen für die maschinelle Sprachanalyse*, in: IBM Nachrichten 27, H. 234, S. 11-17 (1977).

LEHMANN, W. P., STACHOWITZ, R. A.: *Development of German-English Machine Translation System*. Final Technical Report 6 July 1973 - 29 June 1974, Report No RADC-TR-75-2, Linguistics Research Center, The University of Texas at Austin (1975).

LENDERS, W.: *Semantische und argumentative Textdeskription*, Hamburg (1975).

LOGOS *Corporate Profile*, New Hampton, N. Y. (1977).

MAAS, H. D.: *Zum Stand der automatischen Übersetzung im Sonderforschungsbereich (SFB) 100 - Das Übersetzungssystem SUSY*, in: D. KRALLMANN (Hrsg.), *Kolloquium zur Lage der Linguistischen Datenverarbeitung*, 22.-24. Februar 1978, Essen, LVD-Fittings e. V., Essen, S. 82-88 (1978).

MEL'CHUK, I. A. and ZHOLKOVSKII, A. K.: *Towards a functioning 'meaning-text' model of language*, *Linguistics* 57, 1970, 10-47 (1970).

MELBY, A. K.: *The BYU computer-assisted translation System*, Vortrag COLING '78, Bergen 14.-18. August 1978.

METALS *German-English MT*. University of Texas at Austin. *Characteristics of "METAL" System*, Austin (o. J.).

SCHANK, R. C.: (Hrsg.): *Conceptual Information Processing*, Amsterdam.

STAMMERJOHANN, H. (ed): *Handbuch der Linguistik*, München (1975).

TOMA P. P.: *SYSTRAN - Ein System für die mehrsprachige automatische Übersetzung*, in: Kommission der Europäischen Gemeinschaften (Hrsg.), *Dritter Europäischer Kongress über Dokumentationssysteme und -netze. Die Überwindung der Sprachbarrieren*, Luxemburg, 3.-6. Mai 1977, Band 1, München, S. 513-524 (1977).

VAUQUOIS, B.: *La traduction automatique à Grenoble*, Grenoble (1975).

WILKS, Y.: *Grammar, Meaning, and the Machine Analysis of Language*, London (1972).

WILKS, Y.: *An artificial intelligence approach to machine translation*, in: SCHANK, R. C. & COLBY, K. M. (eds.) *Computer models of thought and language* (San Francisco, Freeman, 1973), 114-151.

WINOGRAD, T.: *Understanding Natural Language*, in: *Cognitive Psychology* 3, S. 1-191 (1972).

WOODS, W. A.: *Transition network grammars for natural language analysis*, in: *Communications of the ACM* 13, 1970, 591-606.