

B 10 Maschinelle Übersetzung **Harald H. Zimmermann**

B 10.1 Einführung. Begriffliches

Unter „Maschinellem Übersetzung“ (engl. Machine Translation, im folgenden daher auch kurz: **MT**) versteht man im engeren Sinne die vollautomatische Übersetzung von Texten aus einer natürlichen Sprache in eine andere natürliche Sprache. (Im vorliegenden Beitrag wird die automatische Erkennung und Übersetzung gesprochener Sprache in gesprochene Sprache - Speech Translation - ausgeklammert, da es sich hierbei - v. a. mit Bezug zu fließend gesprochener Sprache - vorwiegend noch um Grundlagenforschung handelt. Allerdings kommt der Umsetzung von gesprochener Sprache in Text wachsende Bedeutung zu.)

Im weiteren Sinne - so auch in diesem Beitrag - werden darunter zusätzlich Systeme verstanden, die den so genannten „Humanübersetzer“ bei seiner entsprechenden Tätigkeit unterstützen (engl. meist Computer Aided Translation, kurz **CAT**; Humanübersetzer ist zugegebenermaßen eine etwas verquere Bezeichnung, die sich aber eingebürgert hat).

Die Qualität und Verfahrensweise am Markt verfügbarer Übersetzungssysteme erlaubt es i.A. nicht, Übersetzungen der Qualität zu erreichen, wie man sie vom menschlichen Übersetzer beziehungsweise der „reinen“ Humanübersetzung (**HT**) erwartet. Die MT hoher Qualität (**HQMT**) kommt ohne spezifische Maßnahmen der Textgestaltung im Vorfeld (Präedition), der Interaktion, d.h. während der Übersetzung, und/oder der Nachbereitung einer Maschinenübersetzung (Post-edition) nicht aus, es sei denn, sie ist hochspezialisiert. (In diesem Beitrag wird nicht auf konkrete Systeme Bezug genommen, um wettbewerbliche Probleme zu vermeiden. Nur soviel sei gesagt: Die „großen“, seit Jahren in Entwicklung stehenden Systeme, vorwiegend im Einsatz bei Behörden bzw. großen oder mittleren Unternehmen, wie LOGOS, METAL oder SYSTRAN (in Europa), haben inzwischen auf dem PC-Sektor „Konkurrenz“ bekommen, v. a. auch mit Bezug zum Massenmarkt (Beispiel Globalink). Dies führte u.a. dazu, dass alle diese Systeme sich heute in einem neuen Outfit präsentieren und alle mindestens einen PC-Zugang haben (Client-Server, wobei für den PC Windows und den Server ggf. ein UNIX-System verwendet wird). Durch eine Portierung der MT Engine auf den PC (Windows '95 bzw. Windows NT) ist weitgehend die gleiche Leistung erreichbar. Mit den neuen Zielgruppen ist auch der Preis „in Bewegung“.

B 10.2 Argumente für den Einsatz von MT

Wie nicht anders zu erwarten, sind es v.a. zwei Ziele, die heute ein Unternehmen dazu bewegen, MT in den Produktionsprozess von Übersetzungen einzubinden:

- Zeitersparnis, d.h. die Erstellung einer Übersetzung in deutlich kürzerer Zeit, als dies bei der HT möglich ist. Als Argument wird v.a. vorgebracht, dass gerade bei der sog. Lokalisierung von kurzlebigen Produkten in über- oder internationalen Märkten die Zeit Geld wert ist („time-to-market“).
- Kostenersparnis, d.h. die Erhöhung der Übersetzungsleistung des einzelnen Übersetzers je Einheit. Sie wird für DIN-A4-Seiten im allgemeinen mit 8-10 Seiten pro Tag bei HT an-

gegeben, mit Hilfe der Textverarbeitung, evtl. auch unter Benutzung von elektronischen Lexika oder Terminologien, v. a. aber auch durch die automatische Übernahme von Textformaten, lässt sich diese Leistung auf etwa 10-12 Seiten pro Tag bei vergleichbarer Qualität steigern, mit der Integration von MT Systemen wird (inkl. Prä- und/oder Postedition) ein Durchsatz in der Größenordnung von etwa 16-20 Seiten pro Tag und Übersetzer erreicht (oder mehr).

Sind die Rahmenbedingungen positiv (stehen v. a. ausreichend große elektronische Lexika für die MT zur Verfügung), so ist die Integration von MT-Verfahren in der Tat von Vorteil. Auch die Rechenleistung ist beachtlich: Übersetzungssysteme auf UNIX-Basis übersetzen heute „spielend“ mindestens 20.000 Wörter in der Stunde Verweilzeit, d.h. 8 DIN-A4-Seiten (Seite zu 2.500 Zeichen), manche sind sogar erheblich schneller.

Als weiteres Argument wird gerne genannt, dass über die MT auch die terminologische Konsistenz eines Textes (v.a. umfangreicher Dokumentationen) gewahrt werden kann. Wenn mehrere Humanübersetzer zeitparallel verschiedene Teile einer Dokumentation bearbeiten, ist die terminologische Kontrolle erschwert; die MT kann hier bis zu einem gewissen Grade unterstützend wirken, da ihr die entsprechenden Fachtermini zur Übersetzung vorgegeben werden können.

Für Außenstehende ist es sehr schwer, die vorwiegend in Firmenreports genannten Zahlen empirisch nachzuvollziehen, neutral erhobene, nachvollziehbare Statistiken liegen m.W. nicht vor. Dennoch kann als Faustregel gelten, dass nach Adaption an das jeweilige Fachgebiet und bei angemessener Integration in den betrieblichen Informationsfluss (s.u.) die Kosten für den „reinen“ Übersetzungsanteil durch die Verwendung am Markt verfügbarer MT Systeme mindestens halbiert werden können und die Aufwandsersparnis mehr als 50% betragen kann.

B 10.3 Zur betrieblichen Integration von MT

Ein wesentlicher Faktor bezüglich der Frage der Nutzung ist die Integration bzw. Integrierbarkeit von MT in den Produktions- und Vertriebsprozess von Produkten. Während man bislang die Übersetzung (mit oder ohne Computer) sozusagen als einen Fremdkörper betrachtete, der im Entwicklungsprozess eher störte, hat die Dokumentation und mit ihr die Übersetzung einen deutlich höheren Stellenwert bekommen (man denke an den Bereich des „Technical Writing“ bzw. das Berufsbild des „Technischen Redakteurs“). In diesem Zusammenhang sucht man heute nach Lösungen, die es erlauben, ein MT System angemessen zu integrieren.

Die Lieferanten und Produzenten von MT Lösungen haben dies inzwischen erkannt und ihr Produkt an die „Produktionsumgebung“ entsprechend angepasst. Dazu gehören:

- Die Anbindung an die verschiedensten Text(system)formate auf den unterschiedlichsten Plattformen (sog. Filter). Tabellen oder Graphiken machen kaum mehr Probleme (es sei denn, die Textteile sind nicht zeichenkodiert, sondern in Pixelform erfasst, auch die bei der Übersetzung erzeugten unterschiedlichen Textlängen können schon einmal ein Kästchen „sprengen“; es kommt auch zu Problemen, wenn der Autor aus technischen Gründen ein Wort „zerhackt“ hat u.a.m.). Filter für gängige Formate wie RTF, FrameMaker, auch SGML und neuerdings HTML (in der Verbindung mit Internet-Applikationen) stehen vielfach bereits zur Verfügung.

- Dem Nutzer werden in allen Systemen Möglichkeiten geboten, Wörter (z.B. Namen) oder Textteile (z.B. Zitate) durch entsprechende Markierungen vor der Übersetzung zu schützen; dies kann z.T. auch in Regeln und Tabellen festgehalten werden (sog. Pattern Matcher).
- Ein wichtiges Hilfswerkzeug, das v.a. bei der Übersetzung von Wartungs- und Bedienungsanleitungen genutzt wird, bei denen sich bei Produktentwicklungen nicht alle Texte völlig ändern, ist das sog. Translation Memory. Hierbei wird vor der Bereitstellung eines Textes oder Satzes zur maschinellen Übersetzung ein Segment elektronisch mit einem Speicher verglichen, der einmal übersetzte (und geprüfte oder korrigierte) Segmente enthält. Dabei kann unterschieden werden zwischen einer völligen Übereinstimmung und teilweiser Übereinstimmung (sog. Fuzzy Matching). Beim Vergleich werden z.T. mathematische Ähnlichkeitsverfahren benutzt. Es ist hier natürlich festzuhalten, dass Translation-Memory-Verfahren auch bei der CAT verwendet werden, also den Prozess der Humanübersetzung ebenso flankieren können wie den der maschinellen Übersetzung.
- In der Posteditionsphase können zudem elektronische Glossare, die dem Nachredakteur (Übersetzer) zur Verfügung stehen, mit einbezogen werden.
- Schließlich soll nicht unerwähnt bleiben, dass die maschinelle Übersetzung heute noch sehr stark von der Qualität des zu übersetzenden Textes abhängt. Tippfehler und sonstige Oberflächen-Fehler (falsche Endungen usw.) führen bei den meisten Systemen zu Übersetzungsfehlern. Da am Markt inzwischen schon gute Verfahren zur Rechtschreibkontrolle verfügbar sind, ist eine solche Vorüberprüfung dringend zu empfehlen.

Die meisten Übersetzungssysteme sind also nicht besonders „robust“ im Hinblick auf Tolerierung von Fehlern in der Dateneingabe. Die derzeitigen Entwicklungen gehen dementsprechend in diese Richtung, wobei es sich hier allerdings um Komponenten handelt, die zielspracheunabhängig sind und daher auch als „außerhalb“ der MT-Lösung liegend gesehen werden.

Maschinelle Übersetzungssysteme bringen zudem bessere Ergebnisse, wenn die Texte/Sätze nicht zu komplex sind. Unternehmen (bzw. Personen), die - unabhängig von der MT-Anwendung - darauf achten, dass ihre Texte/Dokumentationen lesbar und verständlich sind, profitieren also besonders beim Einsatz von MT-Systemen. Dies gilt natürlich auch für den Einsatz „reduzierter“ Sprache, allerdings ist dieses Konzept in der Regel keine prinzipielle Voraussetzung für die Anwendung. Dem Informations-Management im Betrieb kommt damit die Aufgabe zu, darüber zu entscheiden, wie die Integration der Übersetzung allgemein und die Verwendung von MT-Systemen im Speziellen optimiert werden kann. Dabei stehen die meisten Unternehmen bzw. Organisationen vor einem zusätzlichen Problem: Die am Markt verfügbaren Übersetzungssysteme decken häufig nur einen Teilbereich an Sprachen ab, der in einem Unternehmen gebraucht wird. Die Vorstellung, dass praktisch von einer beliebigen Sprache aus alle (auch „gängigen“) Zielsprachen (etwa des europäischen Marktes) mit MT erreicht werden, ist zumindest vorläufig nicht erfüllt. Dazu sind die erforderlichen Investitionen bei den Entwicklern zu hoch und die Märkte noch zu gering (vgl. Abschnitt B 10.5).

Die breiteste Nutzung ist heute über eine Strategie zu erreichen, die zudem viele Unternehmen sich auch „ohne“ die MT-Problematik zu eigen gemacht haben: die Benutzung von Englisch (amerikanisches Englisch) als Ausgangs- oder zumindest als Zwischensprache. MT-Systeme, die von anderen Sprachen nach Englisch und umgekehrt übersetzen, sind fast zu allen größeren Weltwirtschaftsbereichen verfügbar, etwa Deutsch-Englisch, Französisch-Englisch, Spanisch-

Englisch, Italienisch-englisch, Japanisch-Englisch (wenn auch nicht notwendig bei allen Anbietern). Es bietet sich also für einen deutschen Produzenten, der seinen Markt sowohl in englischsprachigen als auch in französischsprachigen Ländern hat, entweder an, seine Dokumentationen in Englisch als Quellsprache zu erstellen und dann nach Deutsch bzw. Französisch zu übersetzen (anschließend Postedition in beiden Sprachen), oder aber vom Deutschen ins Englische zu übersetzen, eine Postedition vorzunehmen und anschließend das postedierte Dokument ins Französische zu übersetzen und die Übersetzung zu postedieren (bei etwa gleichem Aufwand).

B 10.4 Sprachbezogene Fragestellungen

Es sind v.a. Lösungen zu zwei sprachlichen Problemkreisen, die für die Frage der (Nicht-)Anwendung von MT (oder auch CAT) entscheidend sind:

- die Verfügbarkeit fachspezifisch ausgebauter elektronischer Lexika;
- die geeignete Textsorte.

B 10.4.1 Das Lexikonproblem

Die Erfahrung der letzten Jahre, auch entsprechende Experimente, haben gezeigt, dass die Identifikation und Nutzung eines (Fach-)Terminus bzgl. der Akzeptanz (oder Akzeptierbarkeit) von MT-Ergebnissen eine zentrale Rolle spielt.

Dies ist an sich eine triviale Erkenntnis; das Problem liegt jedoch in der Bereitstellung und Verfügbarkeit der Terminologie. Heutige Übersetzungssysteme haben - wenn es hoch kommt - ein (spezifisch erschlossenes, d.h. auch fachspezifisch ausführlich markiertes) Inventar von ein- oder zweihunderttausend Termini, meist jedoch (v.a. bei den Massenprodukten) deutlich weniger. Damit muss sich ein Nutzer - wenn er nicht gerade in einem häufig schon gut oder besser „abgedeckten“ Bereich wie EDV/Software/Telekommunikation oder Maschinenbau arbeitet - sein Inventar selbst erst (z.T. aufwendig) erstellen, ehe er die Vorteile wirklich nutzen kann, also in die eingangs genannten Kosten-Nutzen-Dimensionen gelangt.

Bei allen Anstrengungen, die entwicklerseitig noch unternommen werden müssen, diese Bandbreite der Nutzungen zu erweitern, wird die Terminologie-Erstellung auf Nutzerseite auf längere Sicht ein Problem bleiben.

In diesem Bereich wird der Anwender - sieht man von den rein terminologischen Fragen einmal ab, also der Bereitstellung eines Fachbegriffs und seiner Übersetzung selbst - zunehmend durch systemseitig bereitgestellte Werkzeuge (Software-Tools) unterstützt, die es erlauben, entweder automatisch oder aber über einfache Bedienerführungen die wesentlichen sprachlichen Merkmale mitzugeben, die dem jeweiligen System hinreichen, um die Übersetzung in angemessener Form zu realisieren.

Eine besondere Rolle kommt dabei der Frage zu, bereits beim Anwender vorhandenes lexikalisch-terminologisches Inventar (auf Karteikarten oder - im günstigeren Falle - in maschinenlesbarer Form) rasch und verlässlich und möglichst „automatisch“ in die MT-Anwendung zu überführen. Einige Systeme verfügen inzwischen über derartige „Kodier-Expertensysteme“.

Ein weiteres Beispiel macht besonders deutlich, dass man sich auf Entwicklerseite nicht mehr damit begnügt, ein System als „Black box“ zur Verfügung zu stellen: Ein besonderes Problem stellt die Verarbeitung von Eigennamen (bzw. von Abkürzungen, Produktbezeichnungen ...) dar. Als Hilfslösung kann die Einbeziehung eines der o.a. Pattern-Matcher genutzt werden; es ist hierbei aber immer noch Aufgabe des Nutzers, „seine“ Namen von der Übersetzung auszunehmen, indem er sie entweder lexikalisiert (was für die gängigen, d.h. in vielen Texten auftretenden Wörter in jedem Falle zu empfehlen ist) oder in der aktuellen Anwendung „blockiert“, d.h. im Text entsprechend markiert. Diese Blockierung kann allerdings Nebeneffekte haben, etwa zu einer falschen präpositionalen Übersetzung führen usw.

Die Integration bzw. „Vorschaltung“ eines (monolingualen) Verfahrens, das aus dem Kontext heraus mögliche Eigennamen (usf.) automatisch „erschließt“ und sie für die Übersetzung angemessen „typisiert“, ist daher ein Weg, der den Nutzer von an sich trivialen Fragen entlastet und gleichzeitig die Übersetzungsqualität deutlich steigern kann.

Nutzern, die über (maschinenlesbare) Übersetzungen (in größerer Menge) verfügen, werden darüber hinaus teilweise bereits Werkzeuge angeboten, die eine Adaption an das System erleichtern:

- Mit Hilfe so genannter Alignment-Verfahren werden Texte mit ihren (Human)-Übersetzungen zunächst „parallelisiert“, d.h. in ihre kleinsten Segmente zerlegt. Meist entspricht ein quellsprachiger Satz einem zielsprachigen Satz, es kann aber auch vorkommen, dass ein quellsprachiger Satz in mehrere zielsprachige Sätze aufgelöst wurde und umgekehrt mehrere quellsprachige Sätze zu einem zielsprachigen Satz zusammengefasst wurden. Das Ergebnis kann genutzt werden, um mit den Daten (oder Teilen davon) ein Translation Memory (s.o.) zu „füllen“, das später bei der intellektuellen oder auch der maschinellen Übersetzung mit genutzt werden kann.
Das Verfahren lässt sich daneben unmittelbar nach der Postedition von (neuen) Übersetzungen einbeziehen, also zu dem Zeitpunkt, in dem die Segmente gerade generiert bzw. bearbeitet werden.
- Statistisch-linguistisch basierte Verfahren, ausgehend von einer automatischen Indexierung der quell- und zielsprachigen, parallelisierten Texte, d.h. den Ergebnissen des Alignment-Prozesses, werden genutzt, um aus bestehenden Übersetzungen neue Termini zu extrahieren (Terminologie-Extraktion).
Hierbei handelt es sich in den seltensten Fällen um einfache Wörter (deren Übersetzungen sollten bei leistungsfähigen Systemen schon verfügbar sein); meist sollen komplexe Wörter oder Komposita (im Deutschen) mit ihren passenden (d.h. bisher vom Humanübersetzer verwendeten) Übersetzungsäquivalenten gefunden werden. Auch wenn derartige Verfahren keine optimalen Ergebnisse liefern, d.h. die statistisch-linguistisch ermittelten Terminologie-Vorschläge noch einer intellektuellen Überprüfung und Nachkorrektur bedürfen: Der Weg zur „eigenen“ Terminologiedatenbank und deren Integration in einen Übersetzungsprozess lässt sich dadurch drastisch verkürzen.

B 10.4.2 Das Textsortenproblem

Es ist heute wissenschaftlich grundsätzlich abgeklärt (und auch jedermann „einsichtig“), dass es nicht „die“ natürliche Sprache gibt, sondern dass schriftsprachliche Äußerungen der sprachlich-stilistischen Form nach sog. „Subsprachen“ zugeordnet werden können, die u.a. vorzugsweise in

verschiedensten Textsorten auftreten. Eine E-mail, die an vertraute Personen gerichtet ist oder sich ein bestimmtes Klientel wendet, das eine bestimmte „Etikette“ bevorzugt (nicht umsonst spricht man beispielsweise im Internet heute von einer „Nettikette“), sieht sprachlich deutlich anders aus als etwa ein Vertragstext oder eine Zeitungsmeldung (die sich, wie jeder weiß, auch wieder aufgliedern ließe).

Aus historischen Gründen (manche meinen auch, aus Gründen der generellen Machbarkeit) hat sich die MT schwerpunktmäßig auf die Bereiche der technischen Dokumentation orientiert und setzt auch hier noch - wie oben angesprochen - eine möglichst große Einfachheit im Satzbau voraus. Hierbei handelt es sich nicht um ein lexikalisches Problem, sondern um weitgehend syntaktisch-semantische Eigenschaften.

Ein typisches Beispiel für diese Problematik ist die Art, wie das Problem der semantischen Mehrdeutigkeit (sog. semantische Disambiguierung) in den meisten Systemen derzeit „gelöst“ wird: Die Angabe des Fachgebiets (oder einer Fachgebietsgruppe mit Prioritätskennungen) seitens des Nutzers steuert die Auswahl der Wortbedeutung (oder auch ganz konkret die Wahl der Bezeichnung in der Zielsprache). Dies setzt entsprechende Markierungen (auf Systemseite oder auch seitens des Nutzers) voraus - eine Arbeit, die nicht immer trivial ist. „Nebenbei“ ergibt sich allerdings der eingangs schon angeführte „Vorteil“ der größeren terminologischen Konsistenz (zumindest in der Zielsprache).

Man würde allerdings den meisten Systemen unrecht tun, wenn man dies so stehen ließe: natürlich wird auch der (Satz-)Kontext mit herangezogen, es werden beispielsweise kontextspezifische Wendungen ermittelt usf.; dennoch bleibt eben ein relativ großer „Rest“ an Ambiguitäten, der über diese Fachgebietskennung „gelöst“ wird.

Für die (nahe) Zukunft bleibt allerdings einiges zu tun, insbesondere wenn man an die Masse an potentiellen Nutzern denkt, die heute bereits international kommunizieren. Die Adaption an weitere Subsprachen und -stile bleibt allerdings heute noch ein Desiderat.

B 10.5 Die Nutzer

Auch wenn man mit Prognosen in einer Zeit, in der sich die Informationstechnik so stark entwickelt, etwas zurückhaltend sein soll, so kann man m.E. doch folgende Eckwerte festhalten:

Es ist eine Unterscheidung möglich zwischen einem Gelegenheitsnutzer und einem häufigen Nutzer. Dies lässt sich dabei ggf. auch je nach Bedarf an einer Institution und/oder Person festmachen:

- Der Gelegenheitsnutzer wird ein System bevorzugen, das er - wie seine Textverarbeitung - preiswert erwirbt (etwa für wenige 100 DM, d.h. das Produkt liegt im Low-cost-Bereich) und für Rohübersetzungen entweder in die Fremdsprache oder aber - was wahrscheinlicher ist - in seine Muttersprache einsetzt. Eine Alternative wird möglicherweise die Nutzung über Telekommunikationsnetze sein, wobei er jedoch je nach Bedarf nicht einen spezifischen Host, sondern den seinem aktuellen Bedürfnis und Preisvorstellungen passenden Service ansprechen wird.

Die derzeitigen Bemühungen (1996) verschiedener Service-Provider, solche Dienste über Internet, CompuServe usf. bereitzustellen, machen dies deutlich. Die Rohübersetzung aus einer (ggf. nicht oder nur rudimentär beherrschten) Fremdsprache in die Muttersprache (oder aber, wenn dies nicht möglich ist, ins Englische) dient hier in erster Linie der Information; daher wird sie gelegentlich als Informations- oder Informativ-Übersetzung bezeichnet. (Sicherlich wird sich der eine oder andere fragen, ob nicht längerfristig - etwa über eine oder zwei Generationen gerechnet - sich manches Problem dadurch „erledigt“, dass praktisch jedermann, der über die Landesgrenzen hinaus sprachlich kommuniziert, die Weltverkehrssprache Englisch so gut beherrscht, dass er keine Übersetzungen in „seiner“ Muttersprache mehr benötigt. Man kann aber auch umgekehrte Überlegungen anstellen: Zwar wird Englisch die Verkehrssprache der Zukunft sein, wenn man aber die Weltmärkte bedienen will, so muss man seine Produkte in jedem Falle lokalisieren. Gäbe es heute schon die hoch qualifizierte automatische Übersetzung für alle gängigen natürlichen Sprachen, so wäre dies außerdem kein Argument mehr.)

- Der häufige Nutzer muss nicht unbedingt ein Übersetzer sein. Es ist eher vorstellbar, dass sich Wissenschaftler oder Ingenieure der MT-Systeme bedienen, um ihre „Produktionen“ sowohl landessprachenspezifisch als auch international (in Englisch) verfügbar zu haben. Selbst wenn man bereits (wie in manchen Wissenschaftsbereichen, etwa der Informatik, der Psychologie, der Chemie und Medizin, um nur einige Beispiele zu nennen) englischsprachige Publikationen als Standard voraussetzt (bei Veröffentlichungen über Telekommunikationsnetze, etwa elektronischen Zeitschriften, wird es in Zukunft nicht anders sein), so bleibt doch noch ein großer Bedarf, vor allem dann, wenn die Zielgruppe nicht der Experte, sondern ein größeres Publikum ist.

Als eine besondere Variante der häufigen Nutzer können die Übersetzungsunternehmen und Übersetzungsabteilungen in Unternehmen und Behörden angesehen werden. Sie stellen bislang die Haupt-Klientel der MT-Lieferanten dar, hier hat man die weitreichendsten Erfahrungen mit diesen Werkzeugen gemacht. Aus Sicht dieser Anwender sind v.a. folgende Fragen von besonderer Bedeutung:

- Entwicklung und Bereitstellung von standardisierten Schnittstellen: Beim Wechsel des Herstellers oder auch bei der Integration von Produkten mehrerer Lieferanten soll der vorhandene Anwendungsrahmen beibehalten werden können.
- Portabilität bzw leichte Portierbarkeit der lexikalischen Daten.

Auf Herstellerseite ist dies inzwischen weitgehend erkannt worden. Bis zu einem gewissen Grade wird dieses Problem dadurch gelöst, dass proprietäre Software häufig nur noch für den Kernbereich eines MT Systems verwendet wird, während bei den Systemumgebungen (inkl. der Wörterbücher) Standardsoftware zum Einsatz kommt (etwa relationale Datenbanklösungen). Wie schon angesprochen, sind Filter für alle derzeit gängigen Textverarbeitungssysteme verfügbar, der SGML-Standard wird (inkl. der ISO-Zeichensätze) zunehmend bedient.

Soweit die Systeme nicht „offen“ sind, gibt es doch Portierungsmöglichkeiten, insbesondere für die lexikalischen Daten. Mit den verfügbaren Lösungen auf derzeitigen Plattformen (wie Windows '95 oder Windows NT) bzw Client-Server-Lösungen, die auf der Client-Seite Windows-

orientiert sind und auf der Server-Seite meist eine UNIX-Lösung besitzen, sind die Systeme zumindest äußerlich an die heutigen Betriebsumgebungen angepasst.

Man sollte aber auch darauf hinweisen, dass sich die Systeme durchaus noch in ihren Strategien und Verfahren im Detail unterscheiden können, auch die Frage der herstellerseitigen Betreuung und Erfahrung mag bei Auswahl-Entscheidungen eine Rolle spielen (aber dies gilt natürlich generell).

B 10.6 Stand und Perspektiven

War vor wenigen Jahren die Frage der technischen Anpassung noch ein Problem, so ist dies (auch angesichts der allgemeinen informationstechnischen, aber auch preislichen) Entwicklungen eher sekundär geworden.

Für die professionelle Anwendung spielen in Zukunft weder der Lizenzpreis (der sich für ein Sprachpaar maximal um die 20.000 DM bewegt, aber auch deutlich niedriger sein kann oder sein wird) noch die Hardware die entscheidende Rolle; Prognosen über Preise und Preisentwicklungen sind (und waren immer schon) extrem schwer. Es kann durchaus sein, dass in naher Zukunft PC-Lösungen im Low-Cost-Bereich angeboten werden (CD-ROM vorausgesetzt), die weitgehend „baugleich“ mit professionellen Lösungen sind, sich also nur noch in Bezug auf Betreuung, Hilfssoftware oder den Umfang der mitgelieferten Lexika unterscheiden. Der professionelle Nutzer wird jedoch auf längere Sicht den engen Kontakt mit dem Hersteller brauchen, um von den kontinuierlichen Entwicklungen v. a. im sprachlichen Bereich zu profitieren. Alle Lösungen werden relativ nahe an der Textbe- und -verarbeitung des Anwenders liegen.

Die Adaptionfähigkeit v.a. im lexikalischen Bereich ist deutlich gestiegen, die zusätzlich verfügbaren Hilfslösungen dienen dem gleichen Zweck, wobei ggf. auch Produkte verschiedener Hersteller kombiniert werden können, etwa um Translation Memory, CAT und MT in den Übersetzungsprozess zu integrieren. MT Hersteller und Lieferanten von CAT-Systemen sind zunehmend dabei, Schnittstellen zu schaffen, die es erlauben, proprietäre Lösungen zu kombinieren. Als ein Beispiel kann die Kombination von OPTIMIZER (EuroLang) und LOGOS (MT) genannt werden, andere sind in Entwicklung.

Eine neue Perspektive bietet sich - wie kaum anders zu erwarten - mit der Verfügbarkeit von MT-Lösungen in Telekommunikationsnetzen, seien es nun das so genannte „nicht-kommerzielle“ Internet oder aber kommerzielle Netze wie CompuServe, AOL oder andere. SYSTRAN und LOGOS sind (1996) bereits über eigene URLs im World-Wide-Web (WWW) zu erreichen und bieten derzeit (zumindest noch) kostenlose Übersetzungsproben an.