

Aus der Lehre für die Lehre

COMPUTERLINGUISTIK UND MASCHINELLE ÜBERSETZUNG IN EINEM STUDIENGANG FÜR ÜBERSETZER UND DOLMETSCHER*

Johann Haller

1 Einleitung

An den Studiengängen für Übersetzer und Dolmetscher hat die meist im Studienplan vorgeschriebene Linguistik keinen leichten Stand: alles, was über eine grundlegende Einführung hinausgeht, wird von den sehr zielgerichtet studierenden Kommilitonen danach hinterfragt, welchen konkreten Nutzen es für die spätere Berufstätigkeit bringt. Auch die entsprechenden Mitglieder des Lehrkörpers neigen oft zu der Ansicht, daß Ausdauer und Übung wichtiger sind als ein theoretisches Bewußtsein und Reflexion über das eigene Tun. Nicht einmal die ausführlichen Arbeiten kontrastiver Linguisten waren nah genug an der speziellen Vermittlungstätigkeit der Übersetzer, bei deren Aufgabe oft Faktoren wie Auftraggeber, Adressat und Zweck alle theoretischen Ergebnisse einer kontrastiven Analyse überspielen können. Dies war (und ist zum Teil immer noch) jedenfalls die Hauptthese der eigentlichen Übersetzungswissenschaft, die kontrastive Linguistik als eine zwar wichtige, aber eben nur Teile des Problems abdeckende Hilfsdisziplin betrachtete. Erst in den letzten Jahren hat die Übersetzungswissenschaft sich auf empirische Methoden wie die Analyse des lauten Denkens verlegt, die die beim Übersetzen ablaufenden Prozesse einer empirischen Analyse zugänglich

machen soll.

Diese Fakten legen den Schluß nahe, daß die weitaus mehr mathematischen Charakter tragenden Modelle der Computerlinguistik im Kontext der Übersetzerausbildung noch schwieriger zu vermitteln sind als die griffigen Modelle der Zeichentheoretiker, Semantiker und Pragmatiker.

Computer allgemein sind jedoch in der Ausbildung und Berufstätigkeit des Übersetzers ein Thema, das unter vielen verschiedenen Gesichtspunkten schon ausführlich diskutiert wird: in der Form von sogenannten CAT -Systemen (Integration von Terminologieverwaltung und Textverarbeitung), die bereits heute schon mehr oder weniger zum Handwerkszeug gehören, wird ihre Einsatzfähigkeit eigentlich nicht mehr in Frage gestellt. Eine Frage ist hier höchstens der Preis, den ein einzelner Übersetzer für Hard- und Software sowie für Fachwörterbücher zu zahlen hat bzw. in welchen Zeiträumen und bei welchen Umsätzen sich eine solche Investition bezahlt machen kann. Die zunehmende Vernetzung unserer Gesellschaft führt auch dazu, daß man sich in weltweiten Datenbanken über alle Sachgebiete schnell informieren kann, zu denen man eine Übersetzungs- oder Dolmetschaufgabe gestellt bekommt.

Die Sinnfälligkeit von " Translation Memories", die mit statistischen Methoden nach gleichen oder ähnlichen Sätzen suchen, die schon einmal übersetzt wurden, ist vor allem bei technischer Dokumentation sofort

Dieser Beitrag ist eine stark überarbeitete und aktualisierte Fassung eines Vortrags auf dem Symposium "Sprachdatenverarbeitung" an der Universität des Saarlandes 1992 (erschieden in den entsprechenden Proceedings).

zu erkennen. Aus diesem Grund werden solche Werkzeuge auch an den meisten Ausbildungsstätten für Übersetzer (und Dolmetscher) während des Studiums präsentiert und diskutiert. Schließlich bieten viele Studiengänge auch eine kurze Einführung in die Grundkonzepte und die Geschichte der eigentlichen maschinellen Übersetzung (fortan MÜ), d.h. in den alten Menschheitstraum von der vollautomatischen Übertragung eines Textes von einer Sprache in die andere; hier sind jedoch die Reaktionen der professionellen und zukünftigen Übersetzer sehr viel zurückhaltender, besonders wenn zuwenig darauf hingewiesen wird, daß es auch bei diesen Systemen heute nur mehr darum geht, eine Rationalisierung des Zeitaufwandes für das Übersetzen in Höhe von einigen Prozentpunkten zu erreichen. Daß dies nur in engbegrenzten Bereichen überhaupt möglich ist, zeigt die aktuelle Orientierung vieler heutiger Forschungsprojekte, die von der EU-Kommission oder US-Unternehmen gefördert werden. "Visionäre" Projekte wie das VERBMOBIL (automatisches Dolmetschen für Terminvereinbarungen) werden von der Gemeinde der Übersetzer mit skeptischen Augen betrachtet; dies liegt zum einen an den stark plakativen Aussagen, die während der Phase der Projekterstellung verwendet und mittlerweile längst durch eine Verlagerung des Akzents auf die erwarteten Spin-Off-Effekte abgelöst wurden. Zum anderen sieht man natürlich auch den hohen finanziellen Aufwand, der für die Erforschung aller nötigen Feinheiten der künstlichen Intelligenz zu treiben ist, wenn ein Dialog in gesprochener Sprache auch nur annähernd verstanden und in einer anderen Sprache akustisch wiedergegeben werden soll.

Hat es denn für den angehenden Übersetzer (und auch für den, der sich neben seiner Arbeit weiterbilden will) überhaupt Sinn, sich mit den computerlinguistischen Grundlagen von Systemen wie SYSTRAN, LOGOS und METAL und von Forschungssystemen wie EUROTRA und CAT2 sowie den Versuchen der Carnegie-Mellon-Universität und der Japaner zu befassen? Welche Erkenntnisse über die Einsetzbar

keit und die Funktionsweise können ihm von Nutzen sein, wenn er an seinem Arbeitsplatz mit einem dieser Systeme konfrontiert wird? Gibt es darüber hinaus andere positive Erkenntnisse für seine Arbeit?

In einem Übersetzerstudiengang wie z.B. an der Universität des Saarlandes geht es zunächst um (kommerzielle und forschungsorientierte) Systeme zur maschinellen Übersetzung, die zwar (möglicherweise) mit Prä- und Postedition, jedoch ohne Intervention des Benutzers während des Übersetzungsvorganges Zielsprachentexte erstellen; diese Systeme werden in einer Einführungsvorlesung vorgestellt und in speziellen Seminaren vor allem als mögliche Arbeitsinstrumente für den Übersetzer behandelt. Beispiele hierfür sind Übungen zur Postedition maschineller Texte, zur Überprüfung der Verständlichkeit von Rohübersetzungen und zu Evaluierungsmethoden allgemein. Auch hier werden wichtige Fragen für die spätere Tätigkeit mit MU-Systemen erörtert: die Eignung bestimmter Textsorten für die maschinelle Verarbeitung, die Integration in Büro-Umgebungen und die Frage der generellen Rentabilität unter bestimmten Bedingungen.

In unserem Beitrag nun soll darüber hinaus die (oft kontrovers diskutierte) Frage untersucht werden, inwiefern der Einblick in die computerlinguistischen Grundlagen, die Funktionsweise und das eigene praktische Arbeiten in Entwicklung und Wartung solcher Systeme einen Nutzen für die universitäre Ausbildung von Übersetzern und anderen mit Sprache befaßten Berufen bringen können.

2 Argumente für computerlinguistische Elemente im Übersetzerstudium

Die Gründe, die für die Beschäftigung mit den computerlinguistischen Grundlagen und der Funktionsweise von MÜ-Systemen während des Übersetzerstudiums sprechen, können unter folgenden Stichworten zusammengefaßt werden:

- > Anforderungen des Arbeitsmarkts
- > Einführung neuer Technologien

- > Ergänzungsfach EDV
- > mathematisch-logische Denkweise
- > Formalisierung von MÜ - Problemen
- > Einfluß auf die Entwicklung neuer MÜ Systeme
- > Verbreiterung der Qualifikation

2.1 Anforderungen des Arbeitsmarkts

Seit zwei Jahren führt die Fachrichtung 8.6. unserer Universität eine Statistik der eingegangenen Stellenangebote, die nach verschiedenen Kriterien ausgewertet wird. Die Auswertung der Anforderungen an die Kandidaten zeigt, daß einerseits der Anteil der Stellen, bei denen EDV- und SDV-Kenntnisse explizit gefordert werden, von ca. 30% auf ca. 80% und mehr gestiegen ist, daß aber auch zunehmend explizite Hinweise auf die maschinelle Übersetzung ihren Platz finden. Dies wird zwar meist nur als Wunsch und noch nicht als Voraussetzung genannt, solche Kenntnisse stellen aber in jedem Fall für den Bewerber ein Plus dar.

2.2 Einführung neuer Technologien am Arbeitsplatz

Bei dem Einsatz neuer Softwaretechnologien in der Industrie geht die Initiative meist vom Management aus, das Rationalisierungsmöglichkeiten sieht; beraten werden die Manager dabei vom Rechenzentrum und seinen Mitarbeitern, die schon von ihrer Funktion her an der Anschaffung neuer Soft- (und damit auch Hard-)ware interessiert sind. Gerade bei MÜ - Systemen sind hier einige Beispiele von mißlungenen Einführungen bekannt, die nicht nur zu finanziellen Belastungen für das Unternehmen, sondern auch zur Frustration der Übersetzer beigetragen haben.

Der Übersetzer dagegen, der Kenntnisse über die Funktionsweise solcher Systeme erworben hat, kann die Unternehmensleitung bei Akquisition und organisatorischer Einbettung eines solchen Systems als Fachmann und künftiger Anwender selbst besser

beraten; er kann erkennen, wo geeignete Bedingungen vorliegen oder welche Maßnahmen ergriffen werden müssen, um solche zu schaffen.

Werbeaussagen von MÜ-Herstellern können als solche erkannt und in ihrer Bedeutung relativiert werden; der Dialog mit dem Hersteller kann fachlich besser geführt werden, wenn der Partner auch Kenntnisse über das "Innenleben" des Systems hat.

Der entsprechend ausgebildete Übersetzer kann auch einmal eine fundierte Ablehnung eines solchen Systems formulieren; besitzt er diese Kenntnisse nicht, geschieht die Ablehnung meist nur auf emotioneller Basis.

2.3 Ergänzungsfach EDV

Der zukünftige Übersetzer belegt neben Übungen zu Sprache, Sprachwissenschaft, Landeskunde und Übersetzung ein sogenanntes Ergänzungsfach; er spezialisiert sich (z.B. in Saarbrücken) auf Recht, Wirtschaft oder Technik. Innerhalb des letzteren (meistgewählten) Ergänzungsfaches kann er sich z.B. auf die elektronische Datenverarbeitung spezialisieren.

Bei der Beschäftigung mit MÜ-Systemen erweitert der Übersetzer zugleich seine Kenntnisse in der EDV allgemein; er fügt der DOS-Welt, die er bei der Beschäftigung mit den PC- Werkzeugen ausführlich kennengelernt hat, andere Betriebssysteme hinzu (Großrechner, UNIX); es liegt auf der Hand, daß dies mit Programmen leichter ist, die dem Übersetzer vertraute Aufgaben ausführen als mit abstrakten Tabellen und Zahlenproblemen, wie sie in Einführungen für Informatiker verwendet werden. Diese umfangreichen Kenntnisse können beispielsweise bei Übersetzungsarbeit in der oder für die EDV-Branche nutzbringend angewendet werden. Ein großer Übersetzungsbedarf besteht etwa bei Hard- und Software-Beschreibungen und Manualen. Auch auf anderen Gebieten (z.B. Maschinenbau oder Medizin) dringt die Datenverarbeitung immer weiter vor, und ihr Vokabular muß das entsprechende fachspezifische Wissen ergänzen.

2.4 Mathematisch-logisch Denkweise

Eine schulische und/oder universitäre Ausbildung, die das Schwergewicht nicht auf naturwissenschaftliche Fächer wie Mathematik und Physik legt, führt oft zu einem weniger intensiven Training mathematisch-logischer Denkweisen, wie sie in vielen technischen und naturwissenschaftlichen Berufen zum unbedingt notwendigen Handwerkszeug gehören. Obwohl in manchen Teilen etwa der Übersetzungswissenschaft für eine stark individualisierte und persönlichkeitsbezogene Arbeitsweise plädiert wird (wie sie sicherlich für die Übertragung von poetischen oder auf Werbewirkung ausgerichteten Texten ideal ist), dürfte für einen großen Teil der in der Praxis anfallenden Übersetzungen eine Einsicht auch in mathematisch-naturwissenschaftliche Argumentation sowie formale Logik von Nutzen sein.

Unter anderem kann der Übersetzer sich mit solchen Vorgängen vertraut machen, indem er die Datenverarbeitung intensiver studiert als nur in dem Maße, wie ihm Programme als Werkzeug dienen. Dies geschieht durch die Auseinandersetzung mit Menütechniken, Kommandosprachen, Softwarewerkzeugen beispielsweise zur Stringverarbeitung sowie Einführungen in prozedurale (z.B. BASIC, C) und deklarative (z.B. Prolog) Programmiersprachen. Bei diesen Übungen können einfache Teilprobleme der Textverarbeitung, der Terminologieverwaltung und der maschinellen Übersetzung vom Studenten selbst in algorithmische Form gebracht, programmiert und getestet werden.

2.5 Formalisierung von Übersetzungsproblemen

Die Formulierung von Analyse-, Synthese- und Transferregeln sowie von mono- und bilingualen Lexikoneinträgen in MÜ-Systemen setzt eine starke Formalisierung linguistischer Sachverhalte und die Auseinandersetzung mit linguistischen Theorien voraus, wie sie in den Veranstaltungen zur Sprach- und Übersetzungswissenschaft "auf dem Papier" eingeführt worden sind. .

Dies geschieht in einführenden Seminaren zur Computerlinguistik, die auf die erwähnten theoretischen Vorkenntnisse zugeschnitten sind. Einzelne Themen können durch den Besuch weiter spezialisierter Seminare vertieft werden (Morphologie, Syntax, Semantik, Pragmatik, Dialogsysteme).

Insofern stellt dies eine Ergänzung zur Studienkomponente Sprach- und Übersetzungswissenschaft dar, die den an der Universität ausgebildeten Übersetzer dazu befähigen soll, nicht nur aufgrund langen Trainings und guten Sprachgefühls richtig zu übersetzen, sondern die Schritte seiner Tätigkeit auch mit adäquatem und präzisiertem Vokabular zu begründen und darüber zu reflektieren.

Er kann hierdurch einerseits schneller Rationalisierungsmöglichkeiten in seiner eigenen Tätigkeit erkennen; andererseits kann er auch den gesamten Ablauf des Dokumentationsprozesses im Unternehmen kritisch untersuchen. So kann er z.B. fundierte Kritik an Ausgangstexten üben bzw. Schreibregeln in verständlicher Form für die Autoren entwerfen. In der Regel wird ein (technischer) Text auch für den menschlichen Leser leichter verständlich, wenn gewisse Regeln für Einfachheit, Kürze, Satzbau und logische Abfolge eingehalten werden.

2.6 Mitwirkung des Übersetzers an der Entwicklung von MÜ-Systemen

Seit dem Beginn der Geschichte der MÜ ist festzustellen, daß Systeme in der Hauptsache von Informatikern und später von Linguisten konzipiert und entwickelt wurden; Übersetzer selbst waren nur in sehr geringem Maße beteiligt. Heute scheint sich das Gewicht auf spezialisierte Computerlinguisten, aber auch zu Modellen der Künstlichen Intelligenz zu verschieben.

Bereits während ihres Studiums einschlägig ausgebildete Übersetzer könnten die Lücke füllen, die zur Erstellung von besseren und vor allem auf ihre Arbeitssituation zugeschnittenen Systemen unbedingt geschlossen werden muß.

Hier schließt sich auch der Kreis der Argumentation: wenn er selbst einmal an der Entwicklung neuer Systeme mitwirken

kann, wird der Übersetzer an seinem Arbeitsplatz nicht mehr mit Systemen konfrontiert werden, die seine Belange in der Entwicklung nur am Rande berücksichtigt haben.

2.7 Verbreiterung der Qualifikation

Das früher einheitliche Berufsbild des Übersetzers und Dolmetschers ist heute in der Wandlung begriffen. So werden an mehreren Übersetzerbildungsstätten in Universitäten und Fachhochschulen neue Studiengänge eingerichtet wie Technische Kommunikation, mehrsprachige Wirtschaftsfachleute etc. Diese Auffächerung bzw. Spezialisierung des Berufs zeigt sich auch in der an der in Saarbrücken alle zwei Jahre stattfindenden Veranstaltung "Praxis und Studium"; sie beinhaltet immer einen Arbeitskreis "Fachfremde Tätigkeiten", in denen Absolventen über ihre Tätigkeit berichten, die nicht mehr (oder nicht mehr ausschließlich) die des Übersetzers oder Dolmetschers ist. Terminologie, Technische Normung, Export, Präsentation, Organisation: dies alles sind Tätigkeitsfelder, in denen zu der Orientierung auf sprachliche Kommunikation (die in der heutigen politischen Lage mehrsprachig sein muß) Fähigkeiten wie Flexibilität, technisches Verständnis und das Umgehen mit neuer Technologie allgemein hinzutreten müssen. Diese können u. a. durch die intensive Beschäftigung mit MÜ-Systemen und deren linguistischen und technischen Grundlagen während des Studiums erworben werden.

Zu einer solchen Neuorientierung ist jedoch nicht immer die bürokratisch aufwendige Einrichtung von neuen Studiengängen mit klangvollen Namen, wirtschaftsnah formulierten Ausbildungszielen und schulisch straff organisierten Auslandssemestern mit automatisch anerkannten Scheinen nötig; inhaltlichen Wandel läßt das heutige Universitätsstudium mit Pro- und Hauptseminaren auch jetzt schon zu, Auslandsaufenthalte können durch Kontakte vermittelt und vom Studenten individuell gestaltet werden.

3 Computerlinguistik und Maschinelle Übersetzung in der Übersetzerbildung in Saarbrücken

In diesem letzten Abschnitt soll kurz vorgestellt werden, wie das Gebiet der Maschinellen Übersetzung an unserer Fachrichtung angeboten wird. Die Basiskenntnisse werden in einer Studienkomponente "Sprachdatenverarbeitung" vermittelt, die den Gebrauch von PC- Werkzeugen vermittelt: spezielle Textverarbeitungsprogramme, die den Gebrauch mehrerer Sprachen in einem Dokument erlauben, Terminologiedatenbanken, sogenannte Translation Memories und PC-basierte Übersetzungssysteme.

Wie bereits oben erwähnt, werden in einer Einführungsvorlesung neben einer Übersicht auch kurze Einblicke in die technische Funktionsweise der größeren kommerziellen Systeme gegeben.

Seit mehreren Semestern werden daneben Seminare zu computerlinguistischen Grundlagen, zu forschungsorientierten Systemen wie EUROTRA (mit seiner "sideline" CAT2) und anderen unifiktionsbasierten Systemen angeboten. Hier haben die Studenten eine umfangreiche Grundlektüre zu bewältigen. Nach einer anfangs mühevollen praktischen Ausbildung, bei der auch oft UNIX-Kenntnisse nachgeholt werden müssen, werden jedoch von einzelnen Teilnehmern beachtliche Erfolge erzielt: sie bearbeiten einzelne kontrastive Probleme und erweitern das System selbst um die entsprechenden Regeln und Lexikoneinträge. Systematische Tests und eine ausführliche Dokumentation sind das konkret dokumentierte Ergebnis.

Da EUROTRA und seine Folgeprojekte im Förderprogramm LE (Linguistic Engineering) der Kommission der Europäischen Union mit allen offiziellen Sprachen der Gemeinschaft arbeiten und auch einige Experimentierversionen mit Sprachen wie Russisch, Koreanisch und Japanisch existieren, können Studenten aller Sprachen, die in Saarbrücken angeboten werden, hier Erfahrung erwerben. Für besonders Interessierte stehen Praktikumsplätze sowohl in

Saarbrücken (am Institut für Angewandte Informationsforschung, IAI) als auch an den Partneruniversitäten in den anderen westeuropäischen Ländern, aber auch z.B. in Moskau und Mexico, zur Verfügung. Die bisher durchgeführten Praktika zeigten eine große Akzeptanz sowohl bei den Studenten als auch bei den aufnehmenden Institutionen. Ins besondere ist das Ziel einer Integration der Forschungsarbeiten am IAI mit der Ausbildung an der Universität erreicht.

Kleine Gruppen nehmen an Spezialseminaren zu einzelnen Aspekten wie syntaktische Theorien (GB, HPSG u.a.), Semantik, Wissensrepräsentation und Formalismen teil; Ergänzungsmöglichkeiten bestehen hier auch in Zusammenarbeit mit der Fachrichtung 8.7. Computerlinguistik. Die Studenten dieser Fachrichtung können ihrerseits Maschinelle Übersetzung als Ergänzungsfach wählen, was bereits häufig wahrgenommen wird und woraus sich interessante Zusammensetzungen von Seminaren ergeben.

Durch Lehraufträge an Vertreter von kommerziellen MÜ-Systemen wird periodisch die Möglichkeit gegeben, solche Systeme näher kennenzulernen; auch hieraus entwickeln sich oft Praktikumsmöglichkeiten.

Schließlich steht seit SS 92 das System Logos in einer UNIX-Version zur Verfügung. In einer regelmäßig stattfindenden Übung werden in Gruppen Testsatzbatterien in verschiedenen Sprachen entwickelt und in die Hauptfremdsprache der Teilnehmer übersetzt. Durch den Studiengang mit Französisch als Muttersprache, durch die jährlich ca. 60 ERASMUS-Studenten aus allen europäischen Ländern sowie Austauschstudenten aus der ganzen Welt können meist alle in den Zielsprachen erzeugten Texte auch durch Muttersprachler überprüft werden. Nach einer ersten Übersetzung entsteht eine linguistische Fehlerbeschreibung; das Kodieren unbekannter oder falsch analysierter Wörter wird geübt sowie die Regelmäßigkeit auftretender Fehler durch Variieren der Sätze überprüft. Hierzu muß die Handhabung des Gesamtsystems sowie der Kodiersysteme ALEX (für einfache

Wörter und morphosyntaktische Informationen) und SEMANTHA (für semantische Informationen und Wortverbindungen) erlernt werden. Schlußtests dokumentieren die Entwicklung des Übersetzungsergebnisses sowie die hierzu notwendigen Schritte.

Obwohl durch die politisch bedingten Sparmaßnahmen an der Universität des Saarlandes schon seit dem Wintersemester 1994 weniger Studienanfänger im Übersetzen und Dolmetschen aufgenommen werden, wächst die kleine Gruppe der an dieser neuen Technologie interessierten Studenten kontinuierlich weiter an; auch wenn sich mittelfristig nur ein kleiner Teil der Übersetzer mit dem "Innenleben" solcher Systeme beschäftigen wird, kann gesagt werden, daß sich das 1990 eingerichtete Fach Maschinelle Übersetzung an der Universität des Saarlandes erfolgreich etabliert hat. Neben die Übersetzern des Diplom-Studienganges treten derzeit mehrere Magisterkandidaten für dieses Fach, das als Haupt- und Nebenfach studiert werden kann. Inzwischen liegen auch schon einige Diplomarbeiten sowie zwei Dissertationen im Fach Maschinelle Übersetzung vor, die sich mit den angegebenen Themen ausführlich auseinandersetzen; weitere stehen kurz vor der Fertigstellung.