

Übersetzungstheorie und maschinelle Übersetzung

Eva Wikholm

Uppsala University

1. Einleitung

Übersetzen ist eine Tätigkeit, die zum Verständnis zwischen Sprachgebieten beiträgt und jahrtausendlang beigetragen hat. Solange Sprachbarrieren haben überbrückt werden müssen, sind Übersetzungen gemacht worden. Obwohl die Kunst des Übersetzens eine lange und reiche Tradition aufweist, hat die systematisch wissenschaftliche Beschäftigung mit der Theorie erst in den fünfziger Jahren dieses Jahrhunderts begonnen. Unter den Pionieren ist hier E.A. Nida mit seiner Bibelübersetzungsarbeit zu erwähnen.

Aus den ersten theoretischen Ansätzen auf Basis von Erfahrungen praktischer Übersetzungsarbeit ist allmählich ein eigenes Forschungsfeld entwickelt worden, und zwar die *Übersetzungswissenschaft*. In der Natur der Übersetzungswissenschaft liegt es, daß sie Elemente anderer Wissenschaften beinhaltet, z.B. der Sprachwissenschaft, Literaturwissenschaft, Philosophie und Kommunikationsforschung. Dementsprechend weist sie auch ein vielfältiges Spektrum von Teilgebieten auf, unter denen *Übersetzungstheorie* und *linguistisch-sprachenpaargebundene* bzw. *textbezogene Übersetzungswissenschaft* zu finden sind (Koller 1983; Ingo 1991).

Eine der wichtigsten Aufgaben der Übersetzungstheorie ist es, den Übersetzungsprozeß zu beschreiben. Außerdem beschäftigt sie sich damit, Übersetzungsprobleme zu systematisieren und Methoden für deren Lösung zu erarbeiten. Im Rahmen eines linguistisch-sprachenpaargebundenen Ansatzes werden kontrastive Sprachbeschreibungen aufgestellt und die theoretischen Grundlagen dafür gegeben. Die textbezogene Übersetzungswissenschaft betrachtet das Übersetzen aus der Sicht der Textlinguistik und berücksichtigt dabei auch die Textgattung sowie die Stellung des Textes innerhalb einer Texttypologie. Andere Teilgebiete der Übersetzungswissenschaft sind wissenschaftliche Übersetzungskritik, angewandte Übersetzungswissenschaft, theorie-, übersetzungs- und rezeptionsgeschichtliche Übersetzungswissenschaft und Didaktik des Übersetzens (ebd.).

In der folgenden Darstellung werden in erster Linie übersetzungstheoretische und linguistische Aspekte der Übersetzungswissenschaft im Mittelpunkt stehen, indem folgende Fragen erörtert werden: wie sich der Übersetzungsprozeß einteilen läßt, auf welcher sprachlichen Ebene sich die Übersetzung vollzieht und was eine Übersetzungseinheit kennzeichnet. Die Ergebnisse der traditionellen Übersetzungswissenschaft auf diesen Gebieten werden in eine Umgebung maschineller Übersetzung eingegliedert. Dagegen kann hier nicht auf spezifische Methoden für die Auseinandersetzung mit Übersetzungsproblemen infolge von Sprachverschiedenheiten eingegangen werden. Solche Probleme und Methoden werden an anderer Stelle behandelt (Wikholm, in Vorbereitung II).

2. Der Übersetzungsprozeß

Das Wort *Übersetzen* weist viele Verwendungsweisen auf, weswegen es für übersetzungswissenschaftliche Zwecke eine eindeutige Abgrenzung nötig hat. Für diese Zwecke möchten wir weiterhin mit dem Begriff *Übersetzen* den *Vorgang* bezeichnen, bei welchem ein Text der Ausgangssprache (AS) in schriftlicher Form in einen äquivalenten Text der Zielsprache (ZS) von einem Menschen umgesetzt wird (vgl. Koller 1983, S.106). Natürlich bezieht sich der Begriff dabei auch auf sprachliche und textbezogene Aspekte des Übersetzens.

Der Übersetzungsvorgang mit seinen Bestandteilen wird in der übersetzungswissenschaftlichen Literatur beschrieben und modelliert (u.a. Ingo 1991; Koller 1983; Nida 1969; Wilss 1977). Die Beschreibungen unterscheiden sich darin, wie sie den Vorgang gliedern und welche Faktoren sie dabei hervorheben.

W. Wilss (1977) gliedert den Vorgang aus der Sicht des Übersetzers in zwei Hauptphasen, eine *Verstehensphase* und eine *Rekonstruktionsphase*, indem er das Übersetzen als einen Textverarbeitungs- und Textverbalisierungsprozeß auffaßt. In der Verstehensphase analysiert der Übersetzer den AS-Text auf Inhalt und Stil hin. In der Rekonstruktionsphase reproduziert der Übersetzer den AS-Text in der Zielsprache "unter optimaler Berücksichtigung kommunikativer Äquivalenzgesichtspunkte" (ebd. S.72; Koller 1983, S.111). O. Kade (1968) und G. Jäger (1975), die auch den Vorgang in zwei Teile gliedern, versuchen klarzulegen, wie die beiden Phasen miteinander verbunden sind. Die Rekonstruktionsphase kann demgemäß auf zwei verschiedene Weisen vollzogen werden, nämlich durch *Neukodierung* oder *Umkodierung*. Bei Neukodierung sind der gemeinte Sachverhalt und der Bewußtseinsinhalt Ausgangspunkt für die Rekonstruktion. Bei Umkodierung liegt primär eine sprachliche Analyse der Zuordnung von AS- und ZS-Ausdrücken zugrunde (vgl. Koller 1983, S.113). Bei einer zweiphasigen Gliederung umfaßt die Rekonstruktionsphase sowohl die Zuordnung von Äquivalenten als auch die Herstellung und Bearbeitung des ZS-Textes, wobei die beiden Momente so eng miteinander verbunden sind, daß sie nicht getrennt werden können.

Eine andere Betrachtungsweise des Übersetzungsprozesses, die ihren Ursprung in der Tradition der generativen Transformationsgrammatik hat, gliedert dagegen den Prozeß in drei Phasen. In einer solchen dreigliedrigen Beschreibung werden (1) Analyse des AS-Textes, (2) Umsetzung eines AS-Textes in einen äquivalenten ZS-Text und (3) Bearbeitung des ZS-Textes als drei selbständige Schritte betrachtet.

Die drei Phasen haben in verschiedenen Darstellungen unterschiedliche Benennungen. Zu Vertretern eines dreigliedrigen Modells wollen wir hier die Übersetzungswissenschaftler E.A. Nida und R. Ingo machen. Nida (1969) nennt die drei Phasen *analysis*, *transfer* und *restructuring*, wofür Ingo (1991) die Benennungen *analys* (Analyse), *överföring* (Umsetzung) und *bearbetning* (Bearbeitung) verwendet.

Zu dieser Gliederung macht Nida (1969, S.484) folgende Bemerkung: "However, a careful analysis of exactly what goes on in the process of translating, especially in the case of source and receptor languages having quite different grammatical and semantic structures, has shown that, instead of going directly from one set of surface structures to another, the competent translator actually goes through a seemingly roundabout process of analysis, transfer, and restructuring. That is to say, the translator first analyses the message of the source language into its simplest and structurally clearest forms, transfers it at this level, and then restructures it to the level in the receptor language which is most appropriate for the audience which he intends to reach."

Laut Nida (ebd.) und Ingo (1991) ist die *Analyse* als ein komplexes Vorgehen zu verstehen. Dabei müssen verschiedene Aspekte berücksichtigt werden, wie z.B. grammatische, semantische, stilistische und pragmatische (vgl. Ingo 1991, S.39 ff). In der *Analyse* wird der AS-Text auf einfachere

oder generellere Strukturen in der AS hin analysiert, die möglichst wenige einzelsprachspezifische Züge aufweisen und eine Art Tiefenstruktur im generativen Sinne ausmachen.

Auf die Analyse folgt die *Umsetzungsphase*. In dieser Phase wird der analysierte AS-Text zu einer äquivalenten Struktur in der ZS überführt. Die Vorteile der Zuordnung von elementaren Strukturen bestehen darin, daß die Verhältnisse zwischen den Konstituenten auf einer tieferen Ebene einfacher und klarer sind, und daß sich die Sprachen auf dieser Ebene ähnlicher sehen als in der Oberflächenstruktur (vgl. Nida 1969, S.487). Das Ergebnis der Umsetzung soll eine semantisch exakte Wiedergabe der entsprechenden AS-Struktur sein (Ingo 1991, S.165). Die Umsetzung läßt sich allerdings nicht vollständig beschreiben, weil der Umsetzungsprozeß nicht genau beobachtbar ist, da er sich im Gehirn des Übersetzers vollzieht (ebd.). Ingo hebt doch die Unentbehrlichkeit der Umsetzungsphase hervor (freie Übersetzung nach Ingo 1991, S.92, 165): "Die große Bedeutung der Umsetzungsphase bei allen Arten von Übersetzungen spricht jedoch für ihre Selbständigkeit als Phase im Übersetzungsprozeß, ..., obwohl sie teilweise von der Bearbeitungsphase des ZS-Textes schwer zu trennen ist."

Nach der Umsetzung fehlt noch die endgültige Form des ZS-Textes, die ihre Vollendung erst in der *Bearbeitungsphase* erhält. In dieser dritten Phase des Übersetzungsprozesses werden die elementaren ZS-Strukturen stilistisch zu einem ZS-Text bearbeitet, der den Normen der ZS idiomatisch, stilistisch und empfängerbezogen angemessen ist (vgl. Ingo 1991, S.216; Koller 1983, S.119).

Die oben erörterten Gliederungsvorschläge für einen zweiphasigen bzw. dreiphasigen Übersetzungsvorgang stimmen in ihren Grundzügen mit den Modellen der *maschinellen Übersetzungsansätze* überein. Die ganze prozeßbezogene Betrachtungsweise des Übersetzens ruft eigentlich den Gedanken einer Automatisierung dieses Prozesses hervor. Eine zweiphasige Einteilung kann mit einem *Interlinguamodell* (vgl. z.B. Slocum 1985) verglichen werden, während eine dreiphasige Gliederung grundsätzlich einem *Transfermodell* (ebd.) gleichkommt.

In einem Interlinguamodell sind die zwei Phasen beim Übersetzungsvorgang *Analyse* des AS-Textes und *Synthese* des ZS-Textes. Synthese wird oft bei Interlingua *Generierung* genannt. Der AS-Text wird zuerst zu einer weitmöglichst interlingual konstanten Repräsentation analysiert. Die sprachneutrale Repräsentation soll den Inhalt invariant behalten. Von dieser Repräsentation ausgehend wird dann der ZS-Text generiert. Hier liegt der Vergleich mit der Gliederung von Wilss (1977) nahe, da zwischen *Analyse* und *Verstehensphase* ebenso wie zwischen *Generierung* und *Rekonstruktionsphase* gewisse Ähnlichkeiten bestehen (siehe oben). Bei Kade (1968) und Jäger (1975) kommt der Interlinguaansatz am ehesten dem Gedanken der *Neukodierung* gleich¹.

In einem transferbasierten Übersetzungssystem können dagegen drei Phasen unterschieden werden, und zwar *Analyse*, *Transfer* und *Synthese*. Dieser Ansatz entspricht weitgehend den oben angeführten Einteilungen von Nida (1969) und Ingo (1991). *Analyse* würde *analysis* bei Nida und *analys* (Analyse) bei Ingo entsprechen, wie *Transfer transfer* bzw. *överföring* (Umsetzung) und *Synthese restructuring* oder *bearbetning* (Bearbeitung).

Die Ähnlichkeiten, die zwischen den beiden Forschungsfeldern Übersetzungstheorie und maschinelle Übersetzung vorliegen, deuten darauf hin, daß interessante Beiträge für die Gestaltung interlingua- bzw. transferbasierter maschineller Übersetzungssysteme in der Übersetzungstheorie zu finden sind.

3. Auf welcher Ebene vollzieht sich eigentlich die Übersetzung?

Der Punkt, in welchem im Übersetzungsprozeß der Wechsel von der AS zu der ZS stattfindet und die Art, wie dieser Wechsel vollzogen wird, unterscheiden, wie wir schon gesehen haben, die beiden Ansätze Transfer und Interlingua grundsätzlich voneinander.

Wie sich allerdings Transfer und Interlingua in dieser Hinsicht zueinander verhalten, beschreibt Tucker (1987) wie folgend: "A inherent feature of a transfer-based MT system as opposed to an interlingua system is that the structures to be transferred comprise elements typical of the source language. The deeper the analysis, the fewer the language specific elements, and vice versa, to the point where there are no more language specific features, and the structures representing the meaning of the source language text are interlingual, and, no transfer step (or, as it were, no translation) has to be carried out. What remains is a generation phase."

Die Abgrenzbarkeit der Umsetzungsphase ist also nur bei Transfermodellen vorhanden, da bei einem vollständig durchgeführten Interlinguaansatz keine eigentliche Umsetzung vorkommt, wie auch Tucker betont. Bei transferbasierter Übersetzung stellt sich dann unvermeidlich die Frage, *auf welcher sprachlichen Ebene* sich diese Umsetzung vollzieht. Wir haben schon mit Nida (1969) festgestellt, daß bei einem dreiphasigen Ansatz in der traditionellen Übersetzungstheorie die Zuordnung von äquivalenten AS- und ZS-Ausdrücken nicht direkt in der Oberflächenstruktur der Sprachen stattfindet, sondern auf einer wenig sprachspezifischen Ebene. Die darauf folgenden Fragen sind, wie sprachneutral diese Struktur sein müsste, d.h. wie tief sich diese Ebene befinden würde, und wie eine solche Struktur tatsächlich aussehen könnte.

Ingo (1991, S.165) erkennt hier eine besondere Art von Struktur als Ausgangspunkt für die Umsetzung, die er *zwischenliegende Struktur (mellanstruktur)* nennt. Diese Zwischenstruktur entspricht etwa der Tiefenstruktur in der generativen Grammatik, die von der Basiskomponente einer solchen Grammatik generiert wird. Ingos zwischenliegende Struktur liegt in der Form von solchen primitiven Sätzen, sogenannten *Kernsätzen*, vor. Der Kernsatz generalisiert also die strukturelle Variation und zeigt das vorhandene Satzmuster. Der Sinn des oberflächenstrukturellen Satzes wird aber durch ihn bewahrt, indem er eine Art Paraphrase von dessen Inhalt ausmacht². Die zwischenliegende Struktur entsteht, indem die AS-Sätze auf Kernsätze hin in der AS analysiert (oder rücktransformiert) werden. Die Kernsätze werden definiert, nachdem die Konstituenten des Satzes identifiziert sind, und sie werden dann angemessen gruppiert, damit ihre Verbindungen zueinander explizit werden. Ingo meint, daß die Kernsätze dem Übergang zwischen Sprachen entsprechen, wo es möglichst wenig sprachspezifische Züge gibt, oder, wie er es auf Seite 166 ausdrückt, "wo man die Grenze zwischen den Sprachen an der engsten Stelle überschreitet" (eigene Übersetzung)³.

Nida und Taber (1974) unterscheiden sieben Haupttypen von Kernsätzen im Englischen (Ingo 1991, S.97):

1. John ran quickly.
2. John hit Bill.
3. John gave Bill a ball.
4. John is in the house.
5. John is sick.
6. John is a boy.
7. John is my father.

Die Anzahl der Kernsätze ist von Sprache zu Sprache verschieden, im allgemeinen sind sie aber auf nicht mehr als zehn Typen (ebd.). In der Umsetzungsphase werden die analysierten AS-Strukturen äquivalenten Zwischenstrukturen der ZS auf Kernsatzebene zugeordnet⁴. Aus einem Kernsatz soll es möglich sein, die verschiedenen Satztypen mit Hilfe von Transformationen zu bilden.

Ein Vergleich mit transferbasierten maschinellen Übersetzungssystemen liegt hier nahe, da sich dieser Ansatz auf die Umsetzung von abstrakten Strukturen gründet. Es gibt in den Transfersystemen ein Modul, das die aktuellen Strukturen umsetzt. Dieses Transfermodul macht eine sprachenpaarbezogene Komponente aus. Es gibt aber keine universale Transferstruktur, sondern diese sieht in verschiedenen Systemen unterschiedlich aus. Die Transferstrukturen liegen gewöhnlicherweise in der Form von syntaktisch-semantischen Repräsentationen vor. Die Vermittlung zwischen Transferstrukturen erfolgt mit Hilfe von Transferregeln. Beim Übergang von der einen Sprache zur anderen auf einer generelleren Ebene kann die Anzahl der Transferregeln geringer sein, als wenn der Übergang auf einer sprachspezifischeren Ebene stattfindet.

4. Was für Einheiten werden übersetzt?

Wir haben festgestellt, daß im Übersetzungsprozeß eine Zuordnung von AS- und ZS-Ausdrücken auf eine Art genereller Ebene stattfindet und diese Zuordnung nur in einem transferbasierten Ansatz relevant ist. Aber was für Ausdrücke werden eigentlich übersetzt?

Die historisch gesehen erste Methode für maschinelle Übersetzung, die direkte Methode, gründet sich auf eine Wort-für-Wort-Übersetzung zwischen zwei Sprachen. Unten wird ein Beispiel angeführt, das die Problematik eines solchen Ansatzes beleuchtet:

Schw. Löneförhöjning för nu arbetarna avstå från Dt. Auf Lohnerhöhung müssen nun die Arbeiter verzichten

Nach einer morphologischen Analyse der einzelnen Wörter des schwedischen Satzes sind ohne Berücksichtigung des Kontextes u.a. folgende potentielle Äquivalente vorhanden:

löneförhöjning:	Lohnerhöhung, Gehaltserhöhung
får:	[nn] ⁵ Schaf [vb] dürfen, müssen, mögen, bringen, bekommen, erhalten
nu:	[ab] nun, jetzt; [nn] Augenblick, Nu
arbetarna:	die Arbeiter
avstå:	verzichten
från:	[pp] von, aus, von...her, von...aus, von...an/ab, ab, seit; [ab] aus

Wie aus diesem Beispiel hervorgeht, bringt die Zuordnung auf Wortebene viele Wahlmöglichkeiten mit sich. Die einzelnen Wörter sind erst in ihren Textzusammenhängen, Kontexten, disambiguiert, und eine wortbasierte Übersetzungsmethode stellt sich natürlich als unzulänglich heraus.

Mit Hilfe einer syntaktischen Analyse können die Wörter in gewissem Maße disambiguiert werden. Dies ist der Fall mit *får*, *nu* und *från* in unserem Beispielsatz. Die syntaktische Analyse erkennt *får* als finites Modalverb, das zusammen mit dem Infinitiv *avstå* das Prädikat des Satzes ausmacht. Damit kann die Funktion von *får* als Nomen sowie die nicht modalen Bedeutungen des Verbs, die die deutschen Äquivalente *bringen*, *bekommen* oder *erhalten* tragen, nicht in Frage kommen. *Nu* hat die Funktion als Adverb, und seine unbetonte Stellung deutet darauf hin, daß *nun* lieber als *jetzt* gewählt werden soll. Die Wahl des Äquivalents für *från* unter den Präpositionen erfolgt durch Valenzinformation.

Um die richtige Wahl eines deutschen Äquivalents für *löneförhöjning* und *får*, unter den drei Modalverben *dürfen*, *müssen* oder *mögen*, zu treffen, sind dagegen zusätzliche semantische Informationen notwendig. Das Verb *får* muß im Zusammenhang mit *avstå*, was negativ zu verstehen ist, gesehen werden. Die Wahl des Äquivalents für *lön* in *löneförhöjning* ist davon abhängig, ob hier ein Arbeiter oder ein Beamter gemeint ist. Syntaktische und semantische Analyse des Ausgangssatzes als Ganzes, die ein unumgänglicher Schritt für die Umsetzung ist, ist im Transferansatz zentral.

Wie die lexikalische Äquivalenz in der Transferkomponente beschrieben werden soll, läßt sich nicht generell beantworten. Sowohl Ein-Wort-Lexeme als auch Mehr-Wort-Lexeme (z.B. *avstå från*) müssen ausgedrückt werden können. Die Frage ist, welche sprachliche Größe die angemessene Einheit für die Äquivalenzbeziehungen zwischen lexikalischen Einheiten ausmacht, wenn nun das einzelne Wort sich generell als nicht ausreichend erwiesen hat. Die traditionelle Übersetzungswissenschaft hat sich auf Grund praktischer Übersetzungsarbeit mit der Problematik der Ambiguität und Äquivalenz auseinandergesetzt. Durch solche kontrastiven Studien ist dort eine besondere Art von sprachenpaarbezogener, primär semantischer, Einheit erkannt worden, die grundlegend für die Umsetzung scheint. Diese Einheit wird *Übersetzungseinheit* (ÜE) genannt (z.B. Koller 1983,

S.116). Besonders nützlich hat sie sich bei der Herstellung von Äquivalenzbeziehungen zwischen lexikalischen Einheiten infolge von Sprachverschiedenheiten erwiesen. Laut Koller (ebd.) hat man sich aber "theoretisch wenig mit dem Problem der Übersetzungseinheit beschäftigt, obwohl sie bei sprachenpaarbezogenen Beschreibungen eine wichtige Rolle spielt".

A. Malblanc (1968) und J.-P. Vinay/J. Darbelnet (1971) verwenden für die ÜE den Terminus *unité de pensée*. Damit meinen sie "le plus petit segment de l'énoncé dont la cohésion des signes est telle qu'ils ne doivent pas être traduits séparément" (ebd.). Gemäß dieser Definition behauptet man, daß die ÜE als Sinneinheit in der AS unabhängig von der ZS festgelegt werden kann (Koller 1983, S.116). Es geht jedoch hier um eine Einheit, eine Folge von Zeichen, die dadurch gekennzeichnet ist, daß sie bei Übersetzung *als ein Ganzes* aufgefaßt werden soll, und ist demzufolge nicht sprachenpaarunabhängig. Es scheint unmöglich, diese Einheit festzulegen, ohne eine spezifische oder hypothetische ZS vor Augen zu haben.

O. Kade (1968, S.90) definiert die ÜE gerade unter Berücksichtigung sowohl der AS als auch der ZS: "Die Übersetzungseinheit ist das jeweils kleinste Segment des AS-Textes, für das dank der potentiellen Äquivalenzbeziehungen ein Segment im ZS-Text gesetzt werden kann, das die Bedingungen der Invarianz auf der Inhaltsebene erfüllt."

Da die Einteilung der Sinneinheiten von Sprache zu Sprache unterschiedlich ist (Koller 1983, S.116), müssen wir uns darüber im klaren sein, daß ÜE sprachenpaarbezogen sind, und daß es für jedes Sprachenpaar einen Bestand an ÜE gibt⁶.

Eine ÜE kann aus einem Wort, partiellen oder vollständigen Syntagma, Satz oder Textabschnitt bestehen (vgl. Koller 1983, S.116), wie folgende schwedisch-deutsche Beispiele zeigen:

Wort:

arbetare	Arbeiter
hydraulmatare	hydraulische Vorschubeinrichtung

Partielles Syntagma:

avstå från	verzichten auf
------------	----------------

Syntagma:

slangarnas sträckning	Schlauchstreckung
i enlighet med	laut
be en bön	ein Gebet sprechen
ta under övervägande	in Erwägung ziehen

Satz:

Förbjudet att luta sig ut.	Nicht hinauslehnen.
Nu gäller det!	Es geht um die Wurst.
Tro det eller inte.	Sage und schreibe. / Ob du's glaubst oder nicht.
Allt tar sin tid.	Rom ist auch nicht an einem Tag erbaut worden.

Textabschnitt:

Über allen Gipfeln	Över alla toppar ro.	Alla höjder vilar
Ist Ruh	Knappt ett sus.	i ljus -
In allen Wipfeln	Alla vatten spegla klart	i skogen silar
Spürest du	stillnat ljus.	knappt ett sus
Kaum einen Hauch;		bland träden fram.
Die Vöglein schweigen	Skogens fåglar i sitt bo tystna nu.	Var fågel är tyst när
im Walde.		det svalkas.
Warte nur, balde	Vänta, snart	Vänta - snart nalkas
Ruhest du auch.	vilar också du.	den ro du förnam.
(Goethe: Wanderers	(von Heidenstam)	(Nils Johansson)
Nacht lied)		

Die Notwendigkeit des Begriffes ÜE beschränkt sich nicht auf traditionelle nichtautomatisierte Übersetzung, sondern die ÜE ist eine Erscheinung, die auch bei computergestützter und maschineller Übersetzung von höchster Relevanz zu sein scheint. Eine eindeutige Definition ist gerade für diesen Bereich sogar noch wichtiger, weil die Wahl und Abgrenzung von ÜE Rückwirkungen auf die kontrastive Sprachbeschreibung hat. Die Äquivalenzbeziehungen müssen im Übersetzungssystem so genau beschrieben werden, daß sie mit formalen Transferregeln ausgedrückt werden können. Im folgenden Abschnitt wird ein Beispiel aus aktueller Forschung gegeben, die sich mit solchen Fragestellungen auseinandersetzt.

5. Ein Beispiel

Die Arbeit an dem Projekt *Multilingual Support for Translation and Writing* an der Universität in Uppsala kann als Beispiel dienen, um die Formalisierung der lexikalischen Wahl in einem System für maschinelle Übersetzung zu beleuchten. Dieses Projekt wird von der Professorin in Computerlinguistik Anna Sägval Hein am Institut für Linguistik geleitet. Das Projekt hat zum Ziel, ein Werkzeug für computergestützte Übersetzung zu entwickeln. Es soll die Möglichkeit schaffen, eine automatische Übersetzung der vom menschlichen Übersetzer angegebenen Abschnitte eines Textes zu erhalten. Die Ausgangssprache ist Schwedisch. Als ZS sind in erster Linie Deutsch, Englisch, Französisch und Russisch vorgesehen. Das System, welches das maschinelle Übersetzen besorgt, ist transferbasiert. Die schwedische Transferstruktur liegt als eine Analyse des syntaktischen und morphologischen Parsers, *Uppsala Chart Processor* (Sägval Hein 1987), mit einem Lexikon und einer Grammatik für Schwedisch (Sägval Hein & Sjögreen 1991), vor. Durch Transferregeln wird eine äquivalente Struktur der ZS ermittelt. Transfer erfolgt durch Unifizierung von Transferstrukturen (Beskow, im Druck).

Die Arbeit, ÜE zu identifizieren, ist schon eingeleitet worden. Bisher sind nicht flektierbare funktionale Phrasen (wie in *Übereinstimmung mit*, *in letzter Minute*, *zum Beispiel*) (Sägval Hein, Östling & Wikholm 1990) und Zusammensetzungen (Wikholm, in Vorbereitung I) in kontrastiver Hinsicht studiert worden. Diese Arbeit wird weitergeführt werden. Nachdem die ÜE definiert sind, werden sie und ihre Äquivalenzbeziehungen zu den ZS-Entsprechungen mit Hilfe von Transferregeln ausgedrückt. Der Transferformalismus macht keinen Unterschied zwischen lexikalischen und strukturellen Regeln. Die Beschreibung der ÜE ist demzufolge in diesem Formalismus strukturell, während ihre Definition aber lexikal/semantisch basiert ist. Unten werden einige Beispiele für Regeln gegeben. Die Regeln sind von Anna Sägval Hein definiert worden.

Beispiel 1 bis 3 unten zeigen die Notation von drei lexikalischen Transferregeln. LABEL nennt den Namen der Regel. Im AS-Teil (SOURCE) wird die lexikalische Einheit, das *Lexem*, angegeben, sowie Wortklassenangabe (Sägval Hein 1987). Unter TARGET wird das entsprechende ZS-Lexem genannt, mit notwendigen ZS-bezogenen Auskünften versehen. Für eine ausführlichere Beschreibung des Formats der Transferregeln siehe (Beskow, im Druck).

Die Regeln beschreiben die schwedischen ÜE *hydraulmatare* und *i enlighet med* und die deutschen Äquivalente *hydraulische Vorschubeinrichtungen* bzw. *laut*.

Hydraulmatare ist eine zweiteilige Zusammensetzung und entspricht im Deutschen der nominalen Wortgruppe *hydraulische Vorschubeinrichtungen*, wo dem Nomen ein attributives Adjektiv vorangeht. Die plurale Bedeutung ist hier beabsichtigt. Wir können auch feststellen, daß das Adjektiv starke, oder pronominale (Freund & Sundqvist 1988), Deklination aufweist.

Beispiel 1. Lexikalische Transferregel: Zusammensetzung

LABEL
HYDRAULMATARE
SOURCE
<* LEX> = HYDRAULMATARE.NN.X
<* WORD.CAT > = NOUN
TARGET
<* PHR.CAT> = NG
<* PHRASE > = +
<* ATTR LEX> = HYDRAULISCH.AV.X
<* ATTR WORD.CAT > = ADJ
<* HEAD LEX > = VORSCHUBEINRICHTUNG.NN.X
<* HEAD FEAT GENDER> = FEM
<* HEAD WORD.CAT > = NOUN
TRANSFER

I enlight med ist eine nicht flektierbare funktionale Phrase mit präpositionaler Funktion. Ein potentielles Äquivalent im Deutschen ist die Präposition *laut*, die den Dativ regiert.

Beispiel 2. Lexikalische Transferregel: funktionale Phrase

LABEL
I+ENLIGHET+MED
SOURCE
<* LEX > = I+ENLIGHET+MED.PP.1
TARGET
<* LEX > = LAUT.PP.X
<* WORD.CAT > = PREP
<* CASE > = DAT
TRANSFER

Avstå från, deutsch *verzichten auf*, ist als partielles Syntagma zu betrachten. In dem vollständigen Syntagma sollte auch das Präpositionsobjekt (Freund & Sundqvist 1988) explizit ausgedrückt werden, z.B. *avstå från något*. Da das Objekt austauschbar ist, kann es nicht in eine generelle Transferregel mit einbezogen werden.

Beispiel 3. Lexikalische Transferregel: partielles Syntagma

LABEL
AVSTÅ+FRÅN
SOURCE
<* LEX > = AVSTÅ+FRÅN.VB.1
<* WORD.CAT > = VERB
TARGET
<* LEX > = VERZICHTEN+AUF.VB.X
<* WORD.CAT > = VERB
<* CASE > = ACC
TRANSFER

6. Zusammenfassung der Ergebnisse

Vergleiche von Ergebnissen der beiden Forschungsfelder traditioneller Übersetzungswissenschaft, die von Menschen gemachte Übersetzungen studiert, und maschineller Übersetzung haben

Ähnlichkeiten bei der Beschreibung des Übersetzungsprozesses als ein zweiteiliger bzw. dreiteiliger Vorgang aufgezeigt.

Ausgehend von einem dreigliedrigen Modell hat man sich innerhalb der beiden Felder mit der Frage von Ebene und Form der Transferrepräsentation auseinandergesetzt. Die Notwendigkeit einer genauen Beschreibung einer solchen Repräsentation in maschinellen Übersetzungssystemen spiegelt sich auch in der größeren Anzahl und Ausführlichkeit solcher Vorschläge in dieser Forschungstradition wider.

Als eine besonders große Schwierigkeit bei maschineller Übersetzung hat sich die Formalisierung der lexikalischen Wahl herausgestellt. In der traditionellen Übersetzungswissenschaft ist eine semantisch definierbare Einheit, *Übersetzungseinheit* (ÜE) genannt, identifiziert worden, die grundlegend für die Herstellung von Äquivalenzbeziehungen scheint. Dieser Begriff dürfte auch von Interesse in einer kontrastiven Sprachbeschreibung für maschinelle Übersetzungszwecke sein.

Im Rahmen des Projektes *MULTRA* in Uppsala wird daran gearbeitet, den Formalismus für maschinelle Übersetzung zu gestalten und Transferregeln in diesem Formalismus auszuprobieren anhand konkreter Übersetzungsarbeit. Die Handhabung von ÜE wird in dieser Umgebung studiert. Bisher sind Zusammensetzungen und nicht flektierbare funktionale Phrasen als ÜE untersucht worden, und der Begriff hat sich als nützlich erwiesen. Weitere Arbeit mit solchen Fragestellungen zielt darauf hin, den Begriff genauer zu definieren und, von praktischen Übersetzungsfällen ausgehend, verschiedene Arten von ÜE abzugrenzen und so sorgfältig zu beschreiben, daß es möglich wird, die Äquivalenzbeziehungen mit Transferregeln auszudrücken.

Anmerkungen

- 1) Umkodierung könnte mit der sogenannten direkten Methode für maschinelle Übersetzung verglichen werden
- 2) Ingo (1991) verzichtet darauf, etwas darüber zu sagen, wie z.B. stilistische und thematische Faktoren in der Zwischenstruktur gehandhabt werden.
- 3) Eine tiefgehendere Analyse bis zu einer sogenannten sprachunabhängigen Repräsentation, wie bei Interlingua, scheint aber nicht notwendig. Eine solche Analyse kann stattdessen laut Nida (1969, S.487) zu abweichenden Ergebnissen kommen: "In reality the transfer at the kernel level can generally be made with far less danger of skewing than if one follows the highly involved processes of going to the level of semantic universals and returning again to the kernel level."
- 4) Da die Strukturen der Kemsätze der verschiedenen Sprachen ähnlich sind, sollte demzufolge der Inhalt bei Umsetzung auf Kemsatzebene möglichst invariant bleiben (Nida 1969, S.489). Natürlich sind aber inhaltliche Veränderungen, die mit einem Mangel an semantischer Äquivalenz zusammenhängen, nicht auszuschließen (vgl. ebd.). Hier könnte man auch behaupten, daß lexikalische Ambiguität ein Grund zur inhaltlichen Diskrepanz bei der Umsetzung sein könnte.
- 5) Die Notation der Wortklassenbezeichnungen folgt (Allén et al. 1970): [ab] = Adverb, [nn] = Nomen, [pp] = Präposition, [vb] = Verb.
- 6) Hier kann ein Vergleich mit dem Umgang dieser Problematik innerhalb des Interlinguaansatzes gemacht werden. Dort versucht man, eine Einteilung in sprachunabhängige Sinneinheiten vorzunehmen. Eine universale Einteilung in Sinneinheiten ohne Berücksichtigung existierender Sprachen ist jedoch schwer vorstellbar.

Literaturverzeichnis

- Allén, S. et al. *Nusvensk frekvensordbok 1*. Stockholm. 1970.
- Beskow, B. *Machine Translation in a Unification Based Framework*. Presentation vid de Nordiska Datalogivistikdagarna 1991. Im Druck.
- Catford, J.C. 1965. *A Linguistic Theory of Translation*. An Essay in Applied Linguistics. London (= Language and Language Learning).
- Freund, F. & Sundqvist, B. 1988. *Tysk grammatik*. Natur och Kultur, Arlöv.
- Hutchins, W.J. 1986. *Machine Translation*. Past, Present, Future. Chichester.
- Ingo, R. 1991. *Från källspråk till målspråk*. Introduktion i översättningsvetenskap. Lund.
- Jäger, G. 1975. *Translation und Translationslinguistik*. Halle (Saale) (= Linguistische Studien).
- Kade, O. 1968. *Zufall und Gesetzmäßigkeit in der Übersetzung*. Leipzig (= Beihefte zur Zeitschrift Fremdsprachen, I).
- Koller, W. 1983. *Einführung in die Übersetzungswissenschaft*. Heidelberg.
- Malblanc, A. 1968. *Stylistique comparée du français et de l'allemand*. Essai de représentation linguistique comparée et Etude de traduction. Paris (4. Aufl.) (= Bibliothèque de stylistique comparée, II).
- Nida, E.A. 1969. Science of Translation. In: *Language*. Baltimore. S. 483-498.
- Nida, E.A. & Taber, Ch.R. 1969. *The Theory and Practice of Translation*. Leiden.
- Slocum, J. 1985. A Survey of Machine Translation: Its History, Current Status and Future Prospects. In: *Computational Linguistics*. Volume 11, Number 1, January-March 1985.
- Sågvall Hein, A. 1987. Parsing by means of Uppsala Chart Processor (UCP). In: Bolc, L. (Hrsg). *Natural Language Parsing Systems*. Springer Verlag.
- Sågvall Hein, A., Östling, A. & Wikholm, E. 1990. Phrases in the Core Vocabulary. Center for Computational Linguistics. Uppsala University.
- Sågvall Hein, A. & Sjögreen C. 1991. Ett svenskt stamlexikon för datamaskinell morfologisk analys. En översikt. I: Mats Thelander et al. (utg.) *Förhandlingar vid Artonde sammankomsten för svenskans beskrivning*. Uppsala den 25-26 oktober 1990. Lund. S.348-360.
- Tucker, A. 1987. Current Strategies in Machine Translation Research and Development. In: Nirenburg, S. (ed). *Machine Translation*. Theoretical and Methodological Issues. Cambridge
- Vinay, J.-P./Darbelnet, J. 1971. *Stylistique comparée du français et de l'anglais*. Méthode de traduction (nouvelle édition revue et corrigée, 1. Aufl. 1958). Paris (=Bibliothèque de stylistique comparée, I).
- Wikholm, E. Kontrastivität auf lexikaler Ebene. Die Zusammensetzung als Übersetzungseinheit. Uppsala Universität. Institut für Linguistik. In Vorbereitung I.
- Wikholm, E. Sprachwechsel - Perspektivewechsel? Methoden bei der Auseinandersetzung von Übersetzungsproblemen. Uppsala Universität. Institut für Linguistik. In Vorbereitung II.
- Wilss, W. 1977. *Übersetzungswissenschaft*. Probleme und Methoden. Stuttgart.

Eva Wikholm
Uppsala University
Department of Linguistics
Computational Linguistics
Box 513
S-751 20 Uppsala
E-mail: uduew@seudac21.bitnet