

Kontrastive Analyse von Übersetzungskorpora: ein funktionales Modell

Silvia HANSEN Elke TEICH

1. Einleitung

In den letzten Jahren wird nicht nur in der Linguistik und Computerlinguistik zunehmend mit multilingualen Korpora gearbeitet, sondern auch in der Übersetzungswissenschaft (z.B. BAKER 1995). Wir stellen eine Reihe von Bakers (1996) Hypothesen zu besonderen Eigenschaften von Übersetzungen vor (siehe Abschnitt 2), die sich mit Hilfe von Werkzeugen zur automatischen Textanalyse, wie etwa Alignern, Taggern und Konkordanzen, an Übersetzungskorpora und sogenannten Vergleichskorpora (*comparable corpora*) testen lassen (siehe Abschnitt 3). Unser Hauptanliegen besteht darin, zu zeigen, für welche Fragestellungen im Bereich der Untersuchung der universellen Eigenschaften von Übersetzungen automatische Werkzeuge dienlich sind und für welche Fragestellungen diese Werkzeuge nicht mehr ausreichen, für die dann abstraktere linguistische Kategorien herangezogen werden müssen (siehe Abschnitt 4). Zu deren Interpretation verwenden wir die Systemisch-Funktionale Grammatik (HALLIDAY 1985). Abschnitt 5 enthält eine kurze Zusammenfassung und skizziert unser weiteres Vorgehen.

2. Universelle Eigenschaften von Übersetzungen

In der Geschichte der Korpuslinguistik versuchte man Übersetzungen weitestgehend auszuschließen, da übersetzter Text im Vergleich zur originalen Textproduktion immer als minderwertig und zweitrangig angesehen wurde. SAGER (1984) sah als einer der ersten die Notwendigkeit, Übersetzungen als eine Art der Textproduktion zu analysieren und die besonderen Eigenschaften übersetzter Texte zu identifizieren. Allerdings räumt er ein, daß der Stellenwert einer Übersetzung immer von dem ihres Originals abhängt. Im Gegensatz dazu geht BAKER (1996) einen Schritt weiter, indem sie den Einfluß der Ausgangssprache einer Übersetzung auszuschlie-

ßen versucht, um auf diese Weise typische ausgangssprachenunabhängige Muster in Übersetzungen zu lokalisieren. Zu diesem Zweck erstellte man 1995 an der Universität in Manchester (UMIST) ein monolinguales Korpus, das aus Texten unterschiedlicher Sachgebiete besteht, die aus verschiedenen Sprachen ins Englische übersetzt wurden (TEC: Translational English Corpus). Vergleicht man dieses Korpus mit englischen Originaltexten, ist es möglich, universelle Eigenschaften von Übersetzungen zu analysieren. Diese universellen Eigenschaften von Übersetzungen (BAKER 1996) werden im folgenden kurz erläutert:

Simplification: Von *simplification* spricht man, wenn Übersetzer bewußt oder unbewußt dazu tendieren, Texte zu vereinfachen. Ein Indiz hierfür ist die durchschnittliche Satzlänge. Man kann davon ausgehen, daß Übersetzer lange und komplexe Sätze des ausgangssprachlichen Textes mit mehreren kürzeren Sätzen übersetzen, um die Sachverhalte für den Leser verständlicher zu machen. Das bedeutet aber nicht unbedingt, daß der Text dadurch auch expliziter wird. Generell ist es allerdings möglich, daß die oben erwähnten universellen Eigenschaften von Übersetzungen in kombinierter Form auftreten. Des weiteren kann man *simplification* auch mit Hilfe der lexikalischen Dichte und des Type-Token-Verhältnisses feststellen. Mit der lexikalischen Dichte ist der Prozentsatz gemeint, den man erhält, wenn man die Anzahl der Funktionswörter eines Textes von der Gesamtzahl der Wörter subtrahiert (dann erhält man die Anzahl der Inhaltswörter) und das Ergebnis durch die Gesamtzahl der Wörter dividiert und mit 100 multipliziert. Laut BAKER (1996) ist anzunehmen, daß der Text einfacher zu verstehen ist, wenn mehr Funktionswörter und weniger Inhaltswörter verwendet werden, das heißt, wenn die lexikalische Dichte in den Übersetzungen niedriger ist als in den ausgangssprachlichen oder zielsprachlichen Originaltexten. Mit dem Type-Token-Verhältnis ist der Prozentsatz gemeint, den man erhält, wenn man die Anzahl der Wortstämme (*Types*) eines Textes durch die Gesamtzahl der Wörter (*Tokens*) dividiert und das Ergebnis mit 100 multipliziert. BAKER (1996) nimmt an, daß in englischen Übersetzungen weniger Wortstämme verwendet werden als in englischen Originaltexten. Demzufolge sollte das Type-Token-Ver-

hältnis in Übersetzungen niedriger sein als in ausgangssprachlichen oder zielsprachlichen Originaltexten.

Explicitation: Von *explicitation* spricht man, wenn Übersetzer implizite Sachverhalte des Ausgangstextes in der Übersetzung für den Leser deutlich machen, um sie zu disambiguieren. Dieses Phänomen läßt sich an der Textlänge (Anzahl der Wörter der einzelnen Texte) festmachen. Es ist also anzunehmen, daß Übersetzungen länger sind als ausgangssprachliche oder zielsprachliche Originaltexte und daß darin weniger Ambiguitäten vorhanden sind.

Normalisation: *Normalisation* besagt, daß Übersetzer den typischen sprachlichen Mustern der Zielsprache gerecht werden bzw. daß sie sie sogar übertrieben verwenden. Wenn allerdings der ausgangssprachliche Text oder die Ausgangssprache einen sehr hohen Stellenwert hat verglichen mit der Zielsprache (z.B. wie das Englische gegenüber anderen Sprachen im Softwarebereich), wird die *normalisation* abgeschwächt oder es tritt sogar die gegenläufige Tendenz auf, so daß die sprachlichen Muster der Ausgangssprache in der Übersetzung zu erkennen sind.

Levelling out: In einem Korpus bestehend aus einem Teilkorpus von Übersetzungen, einem Teilkorpus der ausgangssprachlichen Texte und einem Teilkorpus von Texten der Zielsprache, die original in dieser Sprache verfaßt wurden, sind die Texte des Übersetzungskorpus einander ähnlicher als die Texte des ausgangssprachlichen Teilkorpus oder des zielsprachlichen Teilkorpus. Verglichen mit zielsprachlichen oder ausgangssprachlichen Originalen sollten Übersetzungen sich zum Beispiel ähnlicher sein bezüglich ihrer lexikalischen Dichte, ihres Type-Token-Verhältnisses und der durchschnittlichen Satzlänge.

Die hier vorgestellte Liste von universellen Eigenschaften ist keineswegs vollständig, sondern stellt lediglich einen möglichen Ausgangspunkt für die Untersuchung der für Übersetzungen typischen Eigenschaften dar. Außerdem sind lexikalische Dichte, Type-Token-Verhältnis, etc. lediglich Indizien für diese Eigenschaften und lassen sich automatisch mit statistischen Methoden berechnen.

3. Computergestützte Korpusanalyse

Da wir nicht nur die oben genannten universellen Eigenschaften von Übersetzungen testen, sondern auch herausfinden möchten, welche Entscheidungen und Strategien während des Übersetzungsprozesses eine Rolle spielen, haben wir ein multilinguales Korpus, bestehend aus englischen populärwissenschaftlichen Texten und deren deutschen Übersetzungen, gewählt. Das Korpus besteht aus neun englischen und neun deutschen Texten und umfaßt 22.481 englische und 21.204 deutsche, insgesamt also 43.685 Wörter. Die Texte wurden von Monika Doherty, Humboldt-Universität, Berlin, gesammelt und stammen aus *Scientific American* und Spektrum der Wissenschaft. Bis jetzt haben wir nur das Übersetzungskorpus analysiert; unsere weitere Vorgehensweise beinhaltet allerdings den Vergleich unserer Ergebnisse mit in der Zielsprache verfaßten Originaltexten.

Die obige Beschreibung beinhaltet für die universellen Eigenschaften *simplification*, *explicitation* und *levelling out* schon konkrete Testmöglichkeiten, wie etwa lexikalische Dichte, Satzlänge, Textlänge, Type-Token-Verhältnis. Die universelle Eigenschaft *normalisation* ist vorerst nur sehr abstrakt und nicht testbar beschrieben und wurde aus diesen Gründen erst einmal außer Acht gelassen. Die Testmöglichkeiten der drei anderen Eigenschaften konnten relativ unproblematisch mit den folgenden Werkzeugen analysiert werden:

- Um die lexikalische Dichte und die durchschnittliche Satzlänge herauszufinden, benutzten wir das Tool "Traduire", entwickelt von Dieter Maas (vgl. IAI-Memos 1999).
- Für das Type-Token-Verhältnis und die Textlänge benutzten wir das Tool "Simple Concordance Program", entwickelt von Alan Reed (vgl. Simple Concordance Program 1999).

Die Analysen wurden für alle deutschen und für alle englischen Texte durchgeführt. Die Gesamtergebnisse für das englische Teilkorpus und für das deutsche Teilkorpus, die im übrigen nicht statistisch ermittelt, sondern nur ausgezählt wurden und für die keine Signifikanztests durchgeführt wurden, werden in der folgenden Tabelle übersichtlich dargestellt:

	Textlänge	Type-Token-V.	Lexikal. Dichte	Satzlänge
gesamt (en)	22.481	17,31	58,21	24,83
gesamt (de)	21.204	23,39	54,35	22,71

Simplification: Betrachtet man die durchschnittliche Satzlänge der Texte der beiden Teilkorpora, wird man feststellen, daß die deutschen Sätze mit 22,71 Wörtern pro Satz im Schnitt kürzer sind als die Englischen mit 24,83 Wörtern pro Satz. Diesen minimalen Unterschied könnte man aber auch darauf zurückführen, daß die deutschen Texte im Schnitt kürzer sind als ihre englischen Originale. Aus diesem Grund ist es notwendig, sich auch die Anzahl der Sätze in den beiden Teilkorpora anzuschauen: Das englische Teilkorpus besteht aus 922 Sätzen, wohingegen das deutsche Teilkorpus 952 Sätze umfaßt. Dies ist in Kombination mit der durchschnittlichen Satzlänge ein deutlicher Indikator dafür, daß während des Übersetzens bewußt oder unbewußt die Entscheidung gefällt wurde, die Sätze zu kürzen und somit für den Leser der Übersetzungen verständlicher zu machen.

Auch die lexikalische Dichte verifiziert unsere oben beschriebene Hypothese der *simplification*, da in den deutschen Übersetzungen mit einer lexikalischen Dichte von 54,35 Prozent mehr Funktionswörter und weniger Inhaltswörter verwendet werden als in den englischen Originalen mit einer lexikalischen Dichte von 58,21 Prozent.

Das Type-Token-Verhältnis bestätigt nicht das erwartete Ergebnis. Laut unserer Hypothese sollten in den Übersetzungen weniger Wortstämme verwendet werden als in den englischen Originalen, das heißt, das Type-Token-Verhältnis müßte in den Übersetzungen niedriger sein. Daß unsere Ergebnisse das Gegenteil besagen, ist darauf zurückzuführen, daß bei der Berechnung des Type-Token-Verhältnisses in dem verwendeten Konkordanzprogramm eine Wortliste als Basis dient, die zwar verschiedene Wörter, aber nicht die Wortstämme voneinander unterscheiden kann. So werden zum Beispiel die Wörter "können" und "kann" als zwei unterschiedliche *Types* gezählt, obwohl sie vom selben Lemma abstammen. Da es im Deutschen erheblich mehr Konjugations- und Deklinationsmöglichkeiten gibt als im Englischen, ist es nicht verwunderlich, daß das Type-

Token-Verhältnis in den deutschen Übersetzungen höher ausfällt als in ihren Ausgangstexten. Diese Analyse müßte erneut mit einem Werkzeug durchgeführt werden, das in der Lage ist, die deutschen Texte korrekt zu lemmatisieren und auf dieser Grundlage das Type-Token-Verhältnis zu berechnen.

Explicitation: Für die universelle Eigenschaft *explicitation* haben wir die Textlänge der einzelnen Texte betrachtet und festgestellt, daß bis auf eine Ausnahme die deutschen Übersetzungen alle kürzer sind als ihre englischen Originale. Wahrscheinlich ist dieses Ergebnis auf typologische Unterschiede in der Morphologie der beiden Sprachen zurückzuführen. In den deutschen Texten werden zum Beispiel sehr viele Komposita verwendet, die im Englischen als Einzelwörter zählen. Beispielsweise wird der englische Begriff "hydrogen bonding" in den deutschen Texten mit "Wasserstoffbrückenbindung" übersetzt und somit als ein Wort gezählt. Um eine solche Verzerrung der Ergebnisse zu vermeiden, ist es notwendig, vor der Auszählung eine morphologische Analyse auch im Wortbildungsbereich vorzuschalten.

Levelling out: Um diese universelle Eigenschaft zu untersuchen, müssen die Texte im Hinblick auf ihre Extremwerte betrachtet werden. Im folgenden sind die Extremwerte bzw. Spannweiten der lexikalischen Dichte, des Type-Token-Verhältnisses und der durchschnittlichen Satzlänge der verschiedenen Texte aufgelistet:

	englische Texte	deutsche Texte
Type-Token-Verhältnis	28,47 – 35,66	37,13 – 42,33
Lexikalische Dichte	56,44 – 59,83	51,61 – 57,00
Satzlänge	21,09 – 31,22	18,78 – 27,96

Man sieht hier deutlich, daß die Ergebnisse der deutschen Übersetzungen für das Type-Token-Verhältnis und die Satzlänge enger zusammen liegen als die Ergebnisse der englischen Texte. Das bedeutet, daß sich die Übersetzungen in bezug auf diese beiden Kriterien ähnlicher sind als die Originale. In diesen beiden Fällen kann man also von levelling out sprechen. Für die lexikalische Dichte ist das allerdings nicht der Fall. Auch für

diese universelle Eigenschaft wäre es sinnvoll, die Analysen mit lemmatisierten Texten zu wiederholen.

4. Korpusanalyse mit SFG

Die oben beschriebenen Ergebnisse verdeutlichen, daß die Analysen nicht so unproblematisch sind wie anfänglich angenommen. Einige Analysen müßten mit lemmatisierten Texten erneut durchgeführt werden. Allerdings erhalten wir auch bei den Analysen, bei denen die Ergebnisse nicht verzerrt wurden, keine konkrete Beschreibung der linguistischen Phänomene, die diese übersetzungsspezifischen Eigenschaften bewirken. Sie sagen zwar etwas über Quantitäten aus, aber nicht darüber, was sich während des Übersetzungsprozesses auf der sprachlichen Ebene vollzieht. Die oben vorgestellten automatischen Analysen sind nicht aussagekräftig genug, um die universelle Eigenschaft *normalisation* verifizieren oder falsifizieren zu können,.

Um die durch automatische Verfahren bereitgestellten Informationen für den Bereich der Übersetzungswissenschaft nutzen und tiefere Analysen durchführen zu können, bedarf es eines Modells zur Interpretation der durch die Analysen gelieferten Informationen. Zu diesem Zweck schlagen wir vor, ein funktionales Sprachmodell zu verwenden, die Systemisch-Funktionale Grammatik (SFG; HALLIDAY 1985).

In der SFG wird die Sprache in verschiedene Ebenen unterteilt: Man unterscheidet zwischen Semantik, Grammatik, Phonologie und Kontext, wobei Kontext durch Semantik realisiert wird und Semantik durch Grammatik. Des weiteren wird den funktionalen Aspekten von Sprache eine besondere Bedeutung zugemessen. Die Sprache kann drei verschiedene Funktionen erfüllen: die textuelle, interpersonale und ideationale Metafunktion. Die ideationale Metafunktion beschreibt propositionale Inhalte. Die interpersonale Metafunktion hängt mit den Diskursteilnehmern zusammen, und die textuelle Metafunktion bezieht sich auf die sprachlichen Merkmale, die einen Text kohärent und kohäsiv machen. Außerdem unterscheidet man in der SFG zwischen zwei Achsen: Die paradigmatische Achse ist für die Repräsentation der sprachlichen Systeme mit Hilfe von sogenannten *system networks* verantwortlich, wohingegen die syntagmati-

sche Achse für die Realisierung dieser Systeme steht. Die Terme eines grammatischen Systems sind nach der Klassifikationshierarchie geordnet, also vom Allgemeinen zum Spezifischsten. Auf grammatischer und phonologischer Ebene wird die Sprache hierarchisch in Ränge eingeteilt. In der Grammatik unterscheidet man zwischen Sätzen, Phrasen, Wörtern und Morphemen. Die folgende Abbildung skizziert das systemisch-funktionale Modell:

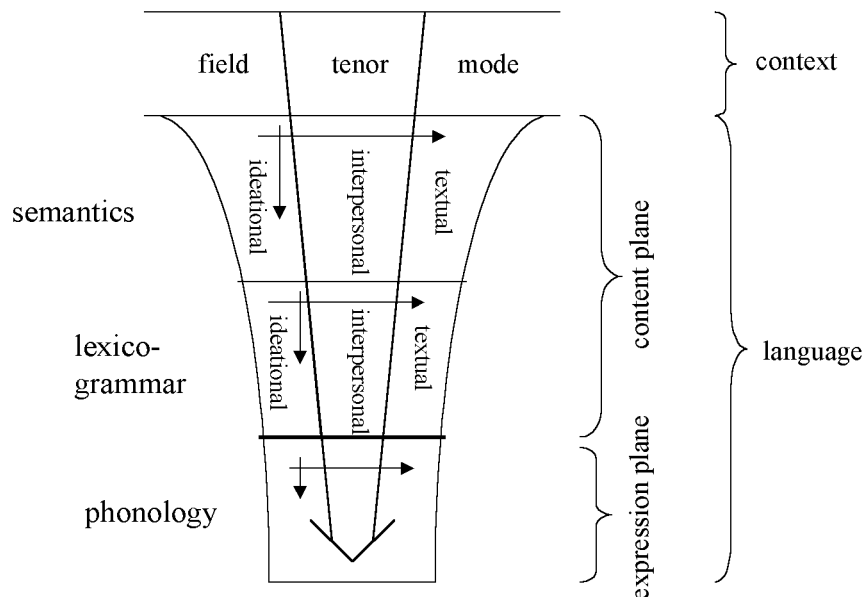


Abbildung 1: Das systemisch-funktionale Modell (vgl. MATTHIESSEN & BATEMAN 1991)

Im folgenden versuchen wir, das systemisch-funktionale Modell für die Präzisierung des Begriffs der *normalisation* anzuwenden. Wie bereits erwähnt, besagt *normalisation*, daß Übersetzer den typischen sprachlichen Mustern der Zielsprache gerecht werden bzw. daß sie sie sogar übertrieben verwenden. Aus diesem Grund betrachteten wir zunächst die typologischen Unterschiede der beiden Sprachen Englisch und Deutsch, da man auf dieser Grundlage feststellen kann, ob die typischen sprachlichen Muster der Zielsprache oder der Ausgangssprache verwendet wurden. Als

besonders frappierende Unterschiede fielen uns dabei in unseren Texten die Subjektrealisierung, die Thema-Rhema-Realisierung, die Wortfolge, die Verwendung agentiver Subjekte und die grammatische Metapher (Funktionsverbgefüge) auf. Deshalb werden im folgenden diese linguistischen Phänomene mit Hilfe der systemisch-funktionalen Parameter zur multilingualen Beschreibung analysiert und erörtert.

Achse: In Bezug auf die Achsen kann man feststellen, daß verschiedene Sprachen zwar ähnliche funktionale Paradigmen haben, aber unterschiedlich in der syntagmatischen Realisierung sein können. In Beispiel (1) wird das englische Subjekt durch die Stellung markiert:

- (1) *Englisch:* "Chemists find hydrogen particularly interesting..."
Deutsch: "Für den Chemiker ist Wasserstoff von besondere Interesse..."

Das Subjekt "Chemists" ist entsprechend der englischen Satzstruktur Subjekt-Prädikat-Objekt am Anfang des Satzes vor dem Prädikat zu finden. In der deutschen Übersetzung ist der Inhalt des Satzes auf andere Weise realisiert worden: Das Subjekt "Wasserstoff" wird, wie es im Deutschen üblich ist, durch den Kasus Nominativ markiert. Doch obwohl die wörtliche Übersetzung in diesem Fall möglich gewesen wäre, wurde die typisch deutsche Realisierung bevorzugt, indem das englische Objekt zum deutschen Subjekt gemacht wurde.

Klassifikationshierarchie: Je weniger spezifische Typen es bei zwei unterschiedlichen Sprachen in den grammatischen Systemen gibt, desto mehr Ähnlichkeiten kann man bei den beiden Sprachen feststellen. Da die Klassifikationshierarchie im Deutschen ausgeprägter ist als im Englischen, kann man nicht viele Ähnlichkeiten bei der Verwendung der grammatischen Typen feststellen. Im Englischen wird ein unmarkiertes Thema durch das Subjekt dargestellt, wie es Beispiel (2) zeigt:

- (2) *Englisch:* "Text books often write this process as..."
Deutsch: "In Textbüchern wird dieser Prozeß als... dargestellt."

In diesem Fall wird das unmarkierte Thema durch das Subjekt "Text books" dargestellt. Im Deutschen ist die Klassifikationshierarchie, was die

Thema-Rhema-Realisierung angeht, allerdings ausgeprägter, da das unmarkierte Thema die semantisch am engsten gebundene Rolle oder ein Lokaladverbiale sein kann. In Beispiel (2) ist das Lokaladverbiale "In Textbüchern" das unmarkierte Thema. Auch in diesem Fall ist in der Übersetzung die typisch deutsche Struktur zu finden. Eine wörtliche Übersetzung hätte zu einer Personifizierung des Subjekts geführt, die im Deutschen an dieser Stelle nicht angebracht wäre.

Metafunktion: Es ist möglich, daß in verschiedenen Sprachen ein bestimmtes strukturelles Mittel verschiedenen Funktionen dient. Zum Beispiel markiert in den oben genannten Beispielen die Wortfolge im Englischen eine syntaktische Funktion (nämlich das Subjekt), wohingegen sie im Deutschen Thema-Rhema und Informationsverteilung ausdrückt. Im englischen Text ist die Funktion der Wortfolge demnach interpersonal, in der Übersetzung allerdings textuell.

Stratifikation: Verschiedene Sprachen weisen auf einer höheren Ebene des linguistischen Systems mehr Ähnlichkeiten auf als auf einer niedrigeren Ebene. So können identische semantische Inhalte unterschiedliche grammatische Realisierungen in verschiedenen Sprachen haben. Dies veranschaulicht Beispiel (3), in dem das englische Subjekt aktiv und agentiv ist:

- (3) *Englisch:* "...they [the molecules] will readily donate the hydrogen to other molecules."
Deutsch: "...der Wasserstoff sofort an andere Moleküle abgegeben wird."

Die deutsche Übersetzung dieses Satzes hat zwar einen identischen semantischen Inhalt, das englische agentive Subjekt wird allerdings passivisch und nicht agentiv realisiert. Doch obwohl auch in diesem Fall eine wörtliche Übersetzung denkbar wäre, entschied man sich während des Übersetzungsprozesses für *normalisation*.

Rang: Identische semantische Inhalte können in unterschiedlichen Sprachen nicht nur, wie oben gezeigt, unterschiedliche grammatische Realisierungen aufweisen, sondern sie können auch in unterschiedlichen syntaktischen Rängen ausgedrückt werden. In Beispiel (4) wird im Englischen

die Entfernung durch eine Verbalgruppe ausgedrückt, wohingegen sie in der deutschen Übersetzung durch eine Nominalgruppe realisiert wird:

- (4) *Englisch*: "The fluorine atoms are only 226 pm apart."
Deutsch: "Der Bindungsabstand der Fluoratome zueinander beträgt nur 226 pm."

Dies ist wiederum ein gutes Beispiel dafür, daß nicht die wörtliche Übersetzung, sondern eine typisch deutsche Variante bevorzugt wurde.

Diese Beispiele zeigen, daß *normalisation* an relativ abstrakten linguistischen Kategorien (wie etwa Subjekt, Thema-Rhema etc.) festgemacht wird. Die gängigen Korpuswerkzeuge, wie z.B. Tagger oder Konkordanztools, können hier nicht zu Hilfe genommen werden. Wir annotieren deshalb die Texte in Bezug auf *normalisation* per Hand mit dem TATOE-System (entwickelt von Melina ALEXA und Lothar ROSTEK, vgl. TATOE 1999), das erlaubt, eigene Annotierungsschemata zu definieren.

5. Zusammenfassung und Ausblick

Das Hauptanliegen des vorliegenden Artikels bestand darin, zu zeigen, wie Hypothesen zu den universellen Eigenschaften von Übersetzungen überprüft werden können. Die von BAKER (1996) beobachteten Eigenschaften von Übersetzungen können größtenteils mit gängigen Korpusanalysewerkzeugen untersucht werden. Eine Ausnahme stellt die Eigenschaft *normalisation* dar (vgl. Abschnitt 4). *Normalisation* muß an abstrakteren linguistischen Kategorien festgemacht werden und kann nicht von oberflächlichen syntaktischen Merkmalen oder gar Wortanzahl pro Satz oder Text abgeleitet werden. Wir haben zur Interpretation der sprachlichen Reflexe von *normalisation* vorgeschlagen, ein funktionales Sprachmodell einzusetzen, die Systemisch-Funktionale Grammatik. Für die weitere Ausarbeitung der universellen Eigenschaften von Übersetzungen ist es erforderlich, nicht nur quantitative Untersuchungen durchzuführen (wie wir es für *simplification*, *explicitation* und *levelling out* gezeigt haben), sondern die Resultate der quantitativen Analysen auch linguistisch zu interpretieren, das heißt, qualitativ zu analysieren. Außerdem ist es notwendig, die Übersetzungen mit in der Zielsprache verfaßten Original-

texten zu vergleichen. Fragen, denen wir in der zukünftigen Arbeit nachgehen werden, schließen die folgenden ein:

- Welche sprachlichen Mittel bzw. übersetzerischen Entscheidungen lassen einen zielsprachlichen Text länger werden (*explicitation*)?
- Welche sprachlichen Mittel bzw. übersetzerischen Entscheidungen verursachen Verschiebungen in der Satzlänge, der lexikalischen Dichte oder im Type-Token-Verhältnis (*simplification*)?
- Welche sprachlichen Mittel lassen eine Menge von Übersetzungen ähnlicher erscheinen als eine Menge von Vergleichstexten, die original in einer Sprache verfaßt worden sind (*levelling out*)?

Um diese Fragen zu bearbeiten, müssen sowohl im Übersetzungskorpus als auch in einem zielsprachlichen Vergleichskorpus abstraktere linguistische Kategorien betrachtet werden.

Literatur

- BAKER, M. (1995): Corpora in translation studies: An overview and some suggestions for future research. In: *Target* 7(2), 223-243.
- BAKER, M. (1996): Corpus-based translation studies: The challenges that lie ahead. In: H. SOMERS (ed.), *Terminology, LSP and Translation: Studies in Language Engineering in Honour of Juan C. Sager*, 175-186. Amsterdam: John Benjamins.
- HALLIDAY, M.A.K. (1985): *An introduction to Functional Grammar*. London: Edward Arnold.
- MATTHIESSEN, C. and BATEMAN, J. (1991): *Text generation and systemic-functional linguistics: Experiences from English and Japanese*. London: Pinter Publishers.
- SAGER, J.C. (1984): Reflections on the didactic implications of an extended theory of translation. In: W. WILSS und G. THOME (eds.), *Die Theorie des Übersetzens und ihr Aufschlußwert für die Übersetzungs- und Dolmetschdidaktik / Translation theory and its implementation in the teaching of translating and interpreting*, 333-343. Tübingen: Gunter Narr.

Web-Seiten

- IAI-Memos (1999). <http://www.iai.uni-sb.de/memos.html> [Stand: 10.9.1999].
- Simple Concordance Program (1999). <http://www.textworld.freemove.co.uk/scp> [Stand: 10.9.1999].
- TATOE (1999): Text Analysis Tool with Object Encoding. <http://www.darmstadt.gmd.de/~rostek/tatoe.htm> [Stand: 10.9.1999]