

Hanna Kempgen (Fürth)

**Спирт, конечно, готов, но мясо протухло.
Zu Qualität und Nutzbarkeit maschineller Übersetzung für das
Russische**

Die diesem Beitrag zugrunde liegende Abschlussarbeit entstand im Jahr 2012 an der Justus-Liebig-Universität Gießen. Die auszugsweise Veröffentlichung in der vorliegenden Festschrift geschieht zu Ehren meines Vaters, dessen Einfluss mich auf diesen Weg geführt hat. Gegenstand der Untersuchung sind die Übersetzungsergebnisse verschiedener frei zugänglicher Online-Dienste zur maschinellen Übersetzung russischer Texte ins Deutsche. Betrachtet werden ein linguistischer Fachtext sowie ein publizistischer Alltagstext. Im Rahmen der Evaluation erfolgt die Identifikation von Fehlern und deren Ursachen sowie ihre Analyse und Kategorisierung nach linguistischen Gesichtspunkten. Dabei sind deutliche Qualitätsunterschiede bezüglich der angebotenen Übersetzungsdienste feststellbar.

1. Einführung: Thematik, Ziele und Methodik

Спирт, конечно, готов, но мясо протухло. – Dieser Satz, vor allem in seiner englischen Entsprechung *The vodka was good, but the meat was rotten*, wird seit mehr als 40 Jahren herangezogen, um die schlechte Qualität maschineller Übersetzung (MÜ) zu illustrieren. Er gilt als beispielhaft für deren Unvermögen, die tatsächliche Bedeutung von Sprache zu verstehen und Mehrdeutigkeiten aufzulösen. Die Anekdote, die seit Ende der 1950er Jahre kursiert, besagt, man habe den Bibelspruch *The spirit is willing, but the flesh is weak* maschinell ins Russische und ebenso wieder zurück ins Englische übersetzt und schließlich den genannten Satz als erheiterndes Ergebnis erhalten. Abgesehen davon, dass der Ursprung dieser Geschichte wohl in einem aus dem Kontext entfremdeten Witz eines Journalisten liegt, ist eine solche Begebenheit für die damalige Situation und Entwicklungsstufe der MÜ nicht realistisch (vgl. Hutchins 1995, 17f.). Doch warum wurden diese und ähnliche – darunter auch wahre – Geschichten im Laufe der Jahre nicht vergessen, sondern immer wieder erzählt und ausgeschmückt? Zum einen, weil die Möglichkeiten dieser Technologie damals wie heute begrenzt sind, gerade Mehrdeutigkeiten

oder Idiome bereiten den Systemen Schwierigkeiten. Das macht derartige Fehlübersetzungen glaubhaft und in ihrer Wirkung zum Beispiel für Kritiker unter Journalisten oder Sprachwissenschaftlern selbst attraktiv. Zum anderen hat sich seit der Einführung frei zugänglicher Online-Übersetzungsdienste der Nutzerkreis der MÜ erheblich vergrößert, sodass immer mehr Menschen, zumeist Laien, mittlerweile selbst Erfahrungen mit diesem Phänomen gesammelt haben und amüsante oder bemerkenswerte Erlebnisse gerne weitergeben.

Es ist seit Jahren selbstverständlich, online mit Hilfe einer Suchmaschine zu beinahe jedem Thema die passende Webseite und zu fast jeder Frage die entsprechende Antwort zu finden. Ebenso normal ist es inzwischen für viele, das MÜ-Angebot *Diese Seite übersetzen* in Anspruch zu nehmen, das oft direkt zusammen mit dem Suchergebnis bereitgestellt wird, falls die Informationen nur in einer dem Nutzer unbekannten Sprache verfügbar sind. Ein anderer Weg besteht darin, den gewünschten Abschnitt direkt selbst bei einem MÜ-Dienst einzugeben. Die Technologie ist populär geworden, die Intentionen bei der Anwendung sind vielfältig, ebenso wie Stile und Themenbereiche der Texte. Dass die jeweilige Software die Eingaben nicht immer vollkommen korrekt in eine andere Sprache überführt, ist dabei zu erwarten. Doch lassen sich – trotz gelegentlicher Mängel – auf diese Weise tatsächlich Sprachbarrieren aufheben? Wie brauchbar sind im Allgemeinen die Ergebnisse, wie verlässlich die Übersetzungen? Wo können Online-Übersetzungsdienste sinnvoll zur Anwendung kommen?

Die vorliegende Arbeit will genau diesen Fragen in Bezug auf das Sprachenpaar Russisch-Deutsch nachgehen. Dazu werden ein Presseartikel und ein wissenschaftlicher Text maschinell aus dem Russischen ins Deutsche übersetzt und anschließend auf mögliche Fehler hin analysiert. Weiterhin wird ermittelt, ob sich die verschiedenen Funktionalstile hinsichtlich der Übersetzungsqualität bemerkbar machen und inwiefern sich die Resultate auch je nach Anbieter voneinander unterscheiden. Neben dem russischen Original dient dabei die jeweilige manuelle, menschliche Übersetzung der gewählten Texte als Vergleichsgröße für die Evaluation. Ein Überblick über die Grundlagen, verschiedene Modelle und die historische Entwicklung der MÜ steht der Untersuchung voran, Zusammenfassung und Fazit sollen die Arbeit abschließen.

2. Klassifizierung: Ansätze und Arten der MÜ

„The term Machine Translation (MT) is the now traditional and standard name for computerised systems responsible for the production of translations from one natural language into another, with or without human assistance“ (Hutchins/Somers 1992, 3). Hutchins/Somers liefern hier zwar eine recht allgemein gehaltene Definition von MÜ, tragen damit aber der Tatsache Rechnung, dass dieser Begriff eine Vielzahl von Theorieansätzen und Methoden in sich vereint. Die verschiedenen Herangehensweisen im Einzelnen vorzustellen und ihre Funktionsweise zu erläutern, ist Gegenstand des ersten Teils der Arbeit.

2.1. Sprachliche Ausrichtung

Eine für den Benutzer wesentliche und naheliegende Frage ist, *aus* welchen und *in* welche Fremdsprachen die eigene Muttersprache übertragen werden kann. Bilinguale Systeme sind dabei speziell auf zwei Sprachen ausgelegt und können entweder unidirektional oder bidirektional aufgebaut sein. Eine Maschine, die beispielsweise die Übersetzung Russisch-Englisch beherrscht, ist nicht automatisch auch für die umgekehrte Richtung, Englisch-Russisch, geeignet, da wir es hier mit zwei verschiedenen Sprachenpaaren zu tun haben. Bidirektionalität bedeutet also, dass eben eine solche beidseitige Benutzung möglich ist. Daneben gibt es auch multilinguale Systeme, die ein breiteres Einsatzspektrum bieten. Diese sind dann besonders flexibel, wenn die Analyse- und Generierungsmodule einer jeden Sprache unabhängig von den weiteren beteiligten Sprachen funktionieren und daraus alle Sprachenpaare kombiniert werden können. Ein komplexerer Ansatz ist es, den Analyseprozess der Ausgangssprache (AS) immer direkt auf die Zielsprache (ZS) anzupassen („English-analysis-for-French-as-target, English-analysis-for-German-as-target“, etc.) (Hutchins/Somers 1992, 70) und damit lexikalische und strukturelle Ähnlichkeiten zwischen beiden auszunutzen. Da es sich hier im Endeffekt aber um eine Zusammenstellung vieler unidirektionaler bilingualer Systeme handelt, steigt auch der Arbeitsaufwand. So ist es ungleich mühsamer, neue Entwicklungen und Erkenntnisse in jede einzelne Komponente einzubinden, als es im modulhaften Aufbau „echter“ multilingualer Systeme der Fall wäre (vgl. Hutchins/Somers 1992, 60–71).

2.2. Übersetzungsmethoden

Im Zuge einer Übersetzung soll für jeden Satz des Ausgangstextes (AT) dessen bestmögliche Entsprechung in der ZS gefunden werden. Jedem maschinellen Übersetzungssystem müssen also zu den betreffenden Sprachen ausreichend Informationen zur Verfügung stehen sowie entsprechende Anweisungen, diese Informationen auszuwerten und richtig einzusetzen. Bezogen auf die eigentliche Übertragung eines Satzes oder Textes von der einen in die andere Sprache trennt man traditionelle, regelbasierte Strategien, wie die direkte Übersetzung, die Transferübersetzung und die Interlingua-Methode, von neueren Ansätzen, die korpusbasiert oder wissensbasiert arbeiten. Der Anteil menschlichen Eingreifens reicht dabei von völliger Selbstständigkeit der Maschine bis hin zur computergestützten menschlichen Übersetzung. Das folgende Kapitel liefert einen Überblick über die einzelnen Herangehensweisen, deren Merkmale und die spezifischen Vor- und Nachteile.

Traditionell regelbasiert: Bevor sich in den 1990er Jahren auch andere Methoden durchsetzten, bildeten ausschließlich linguistische Regeln und Gesetzmäßigkeiten die Grundlage für MÜ, weshalb diese Art als *regelbasiert* bezeichnet wird. Dabei erfolgt grundsätzlich eine Analyse der linguistischen Strukturen im AT, jedoch in unterschiedlicher Intensität und Tiefe. Nun folgt anhand festgelegter Regeln die eigentliche Übersetzung, der Sprachtransfer, zum Teil über den Zwischenschritt einer abstrakten Repräsentation der sprachlichen Strukturen und Inhalte. Am Ende steht die Generierung oder Synthese eines entsprechenden Textes in der ZS.¹

¹ Dieses Kapitel wird aus Platzgründen nicht vollständig abgedruckt. Eine ungekürzte Fassung der Arbeit ist unter www.kempgen.ch/mt abrufbar.

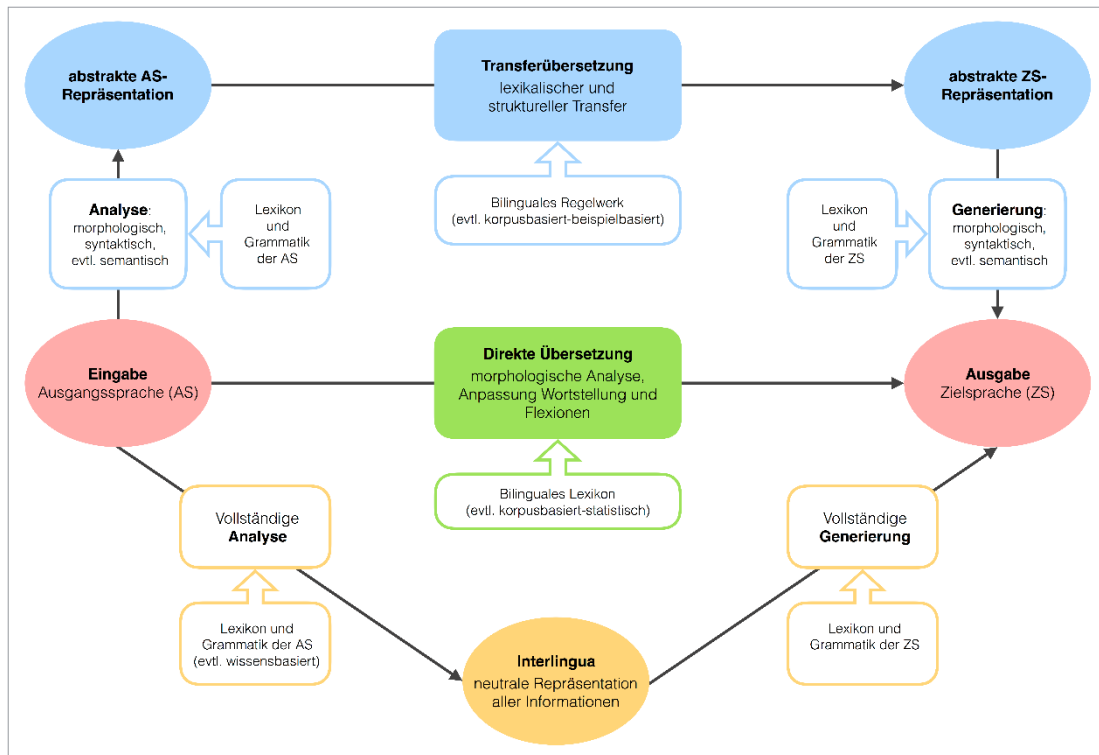


Abbildung 1: Maschinelle Übersetzungsverfahren (weiterentwickelt nach Carstensen 2012, 190)

Neuere Ansätze: Im Modell in Abb. 1 werden zwei Methoden benannt, welche bisher nur in der Einleitung dieses Kapitels erwähnt wurden: der korpusbasierte und der wissensbasierte Ansatz, die – anders als die vorherigen Modelle – nicht auf der Grundlage linguistischer Regeln arbeiten.

Ein *korpusbasiertes* System stützt sich auf große bi- oder multilinguale Korpora menschlicher Übersetzungen, wie sie zum Beispiel in internationalen, mehrsprachigen Organisationen wie der EU entstehen. Für die Verarbeitung ist es wichtig, dass diese Korpora *aligniert* sind, das heißt, dass Textsegmente (Wörter, Phrasen, Sätze) der einen Sprache den entsprechenden Textstücken der anderen Sprache(n) zugeordnet wurden (*Alignments*). Die Übersetzung erfolgt dann entweder *beispielbasiert* oder *statistisch* (vgl. Carstensen 2012, 193).

Im *beispielbasierten* Verfahren wird im Korpus der bereits übersetzten Texte zunächst nach Alignments gesucht, welche den Teilen des zu übersetzenden Textes möglichst ähnlich oder sogar mit ihnen identisch sind. Je größere Teilstücke dabei in der Datenbank gefunden werden, umso besser ist die Übersetzungsqualität. Nun wird – sofern vorhanden – der komplette Satz durch sein Pendant in der ZS ersetzt, oder durch partiellen Abgleich aus kleineren Versatzstücken zusammengefügt. Nach diesem Prinzip funktioniert auch das Hilfsmittel der *Translation Memories*, bei dem allerdings nicht die Maschine, sondern der Übersetzer selbst die jeweils passende Entsprechung wählt, nachdem er das System durch eigene frühere Übersetzungen stetig gefüllt hat (vgl. Carstensen 2001, 519f. und Ramlow 2009, 79f. und 113f.).

Bei der *statistischen* Methode geht es um die Vorhersage der im jeweiligen Kontext wahrscheinlichsten Übersetzungsvariante. Dazu werden die Korpora auf berechenbare Merkmale hin untersucht, etwa das gemeinsame Auftreten von Wörtern, die relative Position der Wörter im Satz oder auch die Länge von Sätzen. Um noch genauere Informationen über den üblichen strukturellen Aufbau einer Sprache zu sammeln, werden zusätzlich einsprachige Korpora herangezogen. Anhand statistischer Berechnung der richtigen Wortentsprechung (*translation model*) und des passenden Wortgefüges (*language model*) sollen wiederkehrende Textmuster für die Übersetzung genutzt werden (vgl. Carstensen 2012, 193 und Ramlow 2009, 80f.).

Doch die reine Korpusbasiertheit birgt einige Nachteile: Zum einen ist das Verfahren stark von bereits vorhandenen, genügend großen und zuverlässigen Korpora abhängig. Diese existieren jedoch nur in begrenzter Zahl und oft nur in Bezug auf bestimmte Domänen. Neue Sammlungen müssten entweder zeit- und kostenintensiv per Hand oder zu Lasten der Qualität automatisch aligniert werden. Zum anderen ist der Verzicht auf linguistische Regeln nicht unbedingt sinnvoll. Ein rein beispielbasiertes System wird nicht in der Lage sein, einen völlig unbekannten Satz (ohne vorhandenes Korpus-Material) zu übersetzen. Daher werden in der Praxis häufig die korpusbasierten mit den regelbasierten Methoden verknüpft, um Lücken in der Alignierung zu überbrücken, es entstehen dann sogenannte Hybrid-Systeme (vgl. Carstensen 2012, 193f., Ramlow 2009, 81f. und Schäfer 2002, 29f.).

Die *wissensbasierte* Herangehensweise setzt bei einem Defizit der Interlingua-Methode an: der Fähigkeit einer Maschine, einen Text tatsächlich zu „verstehen“, das heißt, seine Bedeutung auch in umfassender semantischer und pragmatischer Sicht zu ermitteln. Ein solches Textverstehen gilt im wissensbasierten Ansatz als Voraussetzung für die erfolgreiche MÜ, gleichzeitig wird aber anerkannt, dass allein die syntaktische und teilweise semantische, linguistische Analyse eines Textes dazu nicht ausreicht. Um vor allem das vorherrschende Problem der Mehrdeutigkeiten zu lösen, soll deshalb außersprachliches Wissen, zum Beispiel in Form semantischer Marker und Netzwerke, in den Verarbeitungsprozess und die abstrakte Repräsentation eingebunden werden. So können auch Kontext und pragmatische Faktoren bei der Übersetzung berücksichtigt und die „semantische Barriere“ überwunden werden (vgl. Ramlow 2009, 82 und Schäfer 2002, 28). Durch diese Verknüpfung der MÜ mit Methoden aus dem Bereich der Künstlichen Intelligenz (KI) entstehe im Idealfall ein mit Wissensbasen arbeitendes Interlingua-basiertes MÜ-System, so Ramlow (2008, 43).

Dennoch bleibt die Interlingua-Methode auch mit dieser Erweiterung in der Praxis beschränkt, wie folgendes Beispiel zeigt: *Little Johnny was very upset. He had lost his toy train. Then he found it. It was in his pen.* In diesem Satz „stolpert“ eine Maschine höchstwahrscheinlich über die Mehrdeutigkeit von englisch *pen*. Denn um zu entscheiden, dass hier der Laufstall gemeint ist und nicht ein Stift, müsste sie unter anderem die

Regel beherrschen, dass sich normalerweise nur kleinere Dinge *in* größeren befinden können und zudem Kenntnisse über die relativen Maße der verschiedenen, hier möglichen Objekte (*toy train* gegenüber *train*, Varianten von *pen*) besitzen (vgl. Arnold et al. 1994, 139f.). Dieses Wissen für jeden beliebigen Satz zur Verfügung zu stellen ist nicht machbar. Die in der Forschung bisher entwickelten wissensbasierten Systeme beschränken sich daher auf überschaubare Domänen, für die nur begrenztes Weltwissen relevant ist. Sie beschäftigen sich außerdem hauptsächlich mit der Disambiguierung des ausgangssprachlichen Textes (vgl. Hutchins/Somers 1992, 125).

2.3. Anteil und Art menschlicher Intervention

Bisher haben wir uns damit beschäftigt, auf welche Art eine Übersetzung innerhalb eines MÜ-Systems angefertigt wird, nicht jedoch mit der Frage, ob der Computer dabei völlig selbstständig arbeitet, oder ob ein menschlicher Übersetzer am Prozess beteiligt ist. Anhand von Abb. 2 wollen wir die einzelnen Abstufungen der Mensch-Maschine-Arbeitsteilung betrachten.²

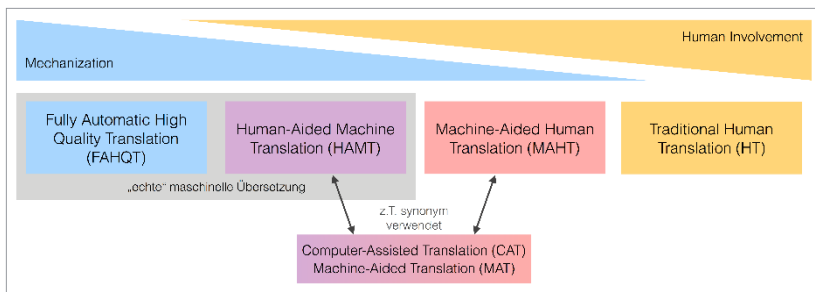


Abbildung 2: Human and Machine Translation (nach Hutchins/Somers 1992, 148)

3. Russisch-Deutsche Übersetzung in Online-MÜ-Diensten

Wir haben uns im ersten Teil der Arbeit mit den Grundlagen der MÜ vertraut gemacht. Daran anschließend widmet sich der zweite Teil nun

² Dieses Kapitel sowie das Kapitel „Von der Idee zur weltweiten Nutzung: Die Geschichte der MÜ“ wird aus Platzgründen nicht vollständig bzw. nicht abgedruckt. Eine ungekürzte Fassung der Arbeit ist unter www.kempgen.ch/mt abrufbar.

der neuesten Erscheinung der MÜ, das heißt den frei verfügbaren Online-Diensten und speziell deren russisch-deutscher Nutzbarkeit. Mit solchen Webseiten wurde während der letzten 10–15 Jahre erstmals einem breiten Publikum die Möglichkeit eröffnet, Texte und Webseiten automatisch übersetzen zu lassen. Als kostenlose und ohne Registrierung oder zusätzliche Installation zugängliche Software wurde die Online-MÜ bald sehr beliebt und ist heute für Internetnutzer selbstverständlich. Allein mit dem *Google Translator* „wird täglich der Textumfang von einer Million Büchern in bis zu 64 Sprachen übersetzt“ (Antosik 2012, o. S.). Dabei beeindrucken vor allem die unmittelbaren, teilweise sogar simultanen Ergebnisse und die Vielzahl an möglichen Sprachenpaaren. Oft findet, nebenbei bemerkt, auch eine zweckentfremdete Nutzung als elektronisches Lexikon statt, ein Verhalten, das eher verwundert, da zum einen die meisten MÜ-Systeme für kontextfreie Wortübersetzungen nicht gemacht sind und zum anderen für diesen speziellen Zweck ja eigenständige, ebenfalls kostenfreie Angebote im WWW existieren und in dieser Hinsicht wesentlich bessere Ergebnisse liefern (vgl. Gaspari/Somers 2007, 15ff.).

So wurden in einer Studie im Jahr 2005 britische Studenten unter anderem gefragt, welche Art von Texten sie online hauptsächlich übersetzen. Es stellte sich heraus, dass „one third of the respondents [...] had used free web-based MT to translate a variety of technical and (semi-)specialised texts, such as business and commercial documents, academic papers and scientific articles [...]. Fairly frequent [...] was the use of free Internet-based MT to translate letters or emails [...], and news reports from online newspapers and magazines [...] were also mentioned“ (Gaspari 2007, 106). Ein Nutzungsverhalten, das sich vermutlich auf Studenten anderer Nationen übertragen lässt, wobei einige der genannten Anwendungsbereiche auf einen Bezug zum Studium hindeuten.

Passend zum Rahmen dieser Arbeit wurden ebenfalls studienrelevante Schriften als Testmaterial für die verschiedenen Online-Dienste gewählt. Sowohl Anbieter als auch Texte werden nun zunächst in ihren Eigenschaften näher beleuchtet, im Anschluss folgen Analyse und Bewertung der Übersetzungsergebnisse.

3.1. Vorstellung der Online-MÜ-Anbieter

Kriterien bei der Auswahl waren die Möglichkeit, Russisch-Deutsch zu übersetzen, die Verfügbarkeit des Angebots auf dem deutschen Markt, d. h. Bedienbarkeit in deutscher Sprache, und schließlich der kostenlose Zugang ohne Registrierung oder Anmeldung. Ursprünglich war es mein Wunsch, auch den Klassiker *Systran* (eines der wenigen Systeme, das die Krise nach dem ALPAC-Report überstanden hat und bis heute genutzt und weiterentwickelt wird) (vgl. Wikipedia 2012d, o. S.) in die Untersuchung mit einzubeziehen, leider beinhaltet das Online-Angebot SYSTRANet (<http://www.systranet.com/de/translate>) jedoch nicht das Sprachenpaar Russisch-Deutsch, lediglich Russisch-Englisch wird angeboten.

Babel Fish / Bing: In Kooperation mit Systran Software Inc. stellte die Suchmaschine *AltaVista* 1997 als erste Webseite einen kostenlosen Übersetzungsservice (*Babel Fish*) zur Verfügung. Im Zuge einer Komplettübernahme von *AltaVista* durch die Suchmaschine *Yahoo!* wechselte auch *Babel Fish* auf deren Webseite (vgl. Gaspari/Hutchins 2007, o. S. und Wikipedia 2012b, o. S.). Im Mai 2012 ist der Dienst überraschend unter dem Namen *Bing Translator* (<http://www.microsofttranslator.com/>) im Microsoft-Angebot aufgegangen. Seitdem nutzt Microsoft für die 38 angebotenen Sprachen eine „linguistically informed statistical machine translation technology“ (Microsoft Translator Team Blog 2012, o.S.), die sich vom alten Angebot unterscheidet.

Google Translator: Google Translator (<http://translate.google.com/>) arbeitete früher ebenfalls auf Grundlage der regelbasierten Software von Systran. Seit 2007 setzt Google jedoch auf eine eigene, statistische Übersetzungsmaschine, deren zugrundeliegendes Korpus sich unter anderem aus den internationalen Dokumenten der EU und der NATO und den elektronisch erfassten Werken des *Google Books*-Projektes speist (vgl. Adams 2010, 22f.). Die 65 Sprachen, die von Google Translator bedient werden, lassen sich beliebig zu Paaren kombinieren, weshalb theoretisch pro Übersetzungsrichtung je ein statistisches Sprachmodell nötig wäre. Inzwischen verwendet Google allerdings Englisch als zwischengeschaltete Interlingua, sodass jeweils nur *in* und *aus* dieser zentralen Sprache übersetzt werden muss (vgl. Linguatools-Blog 2010a, o. S. und Linguatools-Blog 2010b, o. S.). In die Ergebnisverbesserung und den Lernprozess der Maschine bezieht Google seine Nutzer mit ein (sogenanntes

crowd-sourcing), die bei einer misslungenen Übersetzung die jeweils besser passende Alternative für ein Wort oder Wortgefüge auswählen können, was die Software wiederum speichert (vgl. Adams 2010, 22f. und Helft 2010, o. S.). Umgekehrt macht Google aber auch Verbesserungsvorschläge für den Eingabetext, wenn hier beispielsweise ein Wort falsch geschrieben wurde.

PROMT: Anders als Google und Bing kommt die russische Firma *PROMT* nicht aus dem Suchmaschinen-Bereich, sondern ist ein Hersteller von professioneller Übersetzungssoftware. 1998 startete *PROMT* den ersten russischen Online-Dienst für Text- oder Webseiten-Übersetzungen, später erschien auch eine internationale Version (<http://www.online-translator.com/>) (vgl. Wikipedia 2012c, o. S.). Der Web-Service basiert auf der gleichen Technologie wie die kommerziellen Produkte, weshalb auch online die sieben professionell unterstützten Sprachen sowie verschiedene, wählbare Domänen (Business, Medizin, Autos, u. a.) angeboten werden. Laut eigener Aussage verwendet *PROMT* ein transferbasiertes System mit hierarchischem Lösungsansatz, das heißt mehrere Transferprozesse für jeweils verschieden große Analyseeinheiten werden für den Übersetzungsvorgang hierarchisch miteinander verbunden (vgl. *PROMT* 2005a, o. S., *PROMT* 2005b, o. S. und Epiphantseva 2010, o. S.). Seit 2010 wird im professionellen Bereich auch ein hybrides System vertrieben, das Transfer- und statistischen Ansatz miteinander verbindet (vgl. Svarog Capital 2010, o. S.). Gerade die Herkunft aus Russland macht diesen Anbieter für unsere Fragestellung interessant. Die Texteingabe ist für unregistrierte Nutzer auf 3.000 Zeichen begrenzt.

Reverso: *Reverso* (<http://www.reverso.net>) ist ein Angebot der französischen Firma für Übersetzungssoftware *Softissimo* und wurde zusammen mit *PROMT* entwickelt (vgl. *Softissimo* 2012, o. S. und Wikipedia 2012c, o. S.). In welchem Umfang auch deren Technologie hinter *Reverso* steckt ist unklar, zumindest ist auf der *Softissimo*-Homepage von einer regelbasierten Übersetzung die Rede (vgl. *Softissimo* 2009, o. S.). Der Nutzer kann zwischen 13 Sprachen wählen, das Eingabelimit liegt ohne Anmeldung bei 5.000 Zeichen. *Reverso* ist der einzige der hier vertretenen Anbieter, bei dem keine automatische Spracherkennung des Eingabetextes zur Verfügung steht.

WorldLingo: *WorldLingo* (<http://www.worldlingo.com/>) wurde 1998 gegründet und ist laut Eigenwerbung „the company that pioneered the Internet translation industry“ (Grunwald 2011, o. S.). (Babel Fish war da-gegen das erste *kostenfreie* Angebot.) Zunächst wurde hier ebenfalls Software von Systran verwendet, bis im Jahre 2009 auf das statistische System der Firma *Language Weaver* gewechselt wurde (vgl. Grunwald 2009, o. S.). Ohne Registrierung können bei WorldLingo maximal 500 Wörter übersetzt werden.

3.2. Vorstellung der gewählten Texte

Das Ziel der vorliegenden Arbeit ist ja, die Qualität beziehungsweise den Nutzen der maschinellen deutschen Übersetzung russischer Texte zu testen. Um verschiedene sprachliche und textuelle Phänomene zu berücksichtigen, sollten die Texte jeweils unterschiedliche Funktionalstile repräsentieren. Die Wahl fiel deshalb zum einen auf den wissenschaftlichen, zum anderen auf den publizistischen Stil. Ein weiterer Faktor für die Auswahl der Texte war die Existenz manueller deutscher Übersetzungen, welche für die Beurteilung der Ergebnisse eine wichtige Vergleichsgröße darstellen. Die Texte wurden an manchen Stellen gekürzt, damit sich in beiden Fällen eine vergleichbare Länge von ca. 190 Wörtern ergibt.

Der wissenschaftliche Text stammt von dem russischen Linguisten Jurij Derenikovič Apresjan aus dessen Werk *Идеи и методы современной структурной лингвистики* von 1966. Unter dem Titel *Ideen und Methoden der modernen strukturellen Linguistik* erschien es 1971 in deutscher Sprache. Der Ausschnitt gehört zum Kapitel *Modelle der Redestruktur*, Unterpunkt *Syntaktische Analysemodelle* und wird nachfolgend im Original und als Übersetzung wiedergegeben:

ПРЕДСКАЗУЕМОСТНЫЙ АНАЛИЗ

Чтобы пояснить основную идею предсказуемого автоматического анализа, главные принципы которого были содержательно изложены Ч. Хоккетом несколько лет назад <311>, а затем формализованы И. Родес <114> и А. Эттингером <344>, <114>, проведем следующий эксперимент: попытаемся догадаться, какое слово может заполнить пробел во фразе *Весь декабрь было тепло, и шли дожди; лишь в январе выпал первый* —. Решение, по-видимому, однозначно: на последнем месте в этой фразе, вероятнее всего, будет стоять слово *снег*. Появление другого слова вызвало бы у нас удивление, а появление слова *снег* лишь подтверждает ту гипотезу, которую мы бессознательно сформировали, прочитав уже «поступившие» слова. [...]

Изложенное здесь представление о процессе понимания (анализа) текста формализуется техникой предсказуемого анализа. Средством анализа синтаксической структуры предложения является хранящийся в памяти машины набор синтаксических предсказаний (НСП), содержащий в себе гипотетический перечень возможных в некотором предложении синтаксических структур. [...]

«Процесс анализа заключается в последовательном сравнении информации к словам анализируемого предложения со всеми предсказаниями, перечисленными в НСП. Когда для слова найдено «подходящее» предсказание, слову приписывается пометка соответствующей синтаксической функции, а в НСП вносятся необходимые изменения» <114, 14>. Анализ продолжается до тех пор, пока каждой словоформе предложения не будет приписана определенная синтаксическая функция. (Апресян 1966, ч. IV, гл. 2)

Die Vorhersageanalyse (predictive analysis)

Zur Erklärung der Grundidee der automatischen Vorhersageanalyse, deren Hauptprinzipien von Ch. Hockett vor einigen Jahren inhaltlich erläutert <137> und danach von I. Rhodes <213> und A.G. Oettiger <234; 213> formalisiert worden sind, diene ein kleiner Versuch: Erraten werden soll das letzte Wort im Satz Den ganzen Dezember war es warm und regnete; erst im Januar fiel der erste -. Die Lösung ist, wie es aussieht, eindeutig: Hier wird aller Wahrscheinlichkeit nach das Wort Schnee stehen. Ein anderes Wort würde uns in Verwunderung setzen, während das Wort Schnee nur die Hypothese bestätigt, die wir nach der Lektüre der vorangehenden Wörter bereits unbekannt formuliert haben. [...]

Die hier umrissene Vorstellung über den Prozeß des Verstehens (der Analyse) eines Textes wird durch die Technik der Vorhersageanalyse formalisiert. Die Analyse der syntaktischen Struktur eines Satzes erfolgt mittels einer im Speicher der Maschine bewahrten Menge syntaktischer Vorhersagen (MSV), eines hypothetischen Verzeichnisses der syntaktischen Strukturen, die in einem bestimmten Satz möglich sind.

„[...] Der Analyseprozeß besteht in einem systematischen Vergleich der Informationen über die Wörter des zu analysierenden Satzes mit allen in der MSV gesammelten Vorhersagen. Wenn für ein Wort die ‚passende‘ Vorhersage gefunden ist, wird ihm die entsprechende syntaktische Funktion zugeschrieben, während in der MSV die notwendigen Veränderungen eingetragen werden.“ <114, 14> Die Analyse ist erst dann beendet, wenn jeder Wortform des Satzes eine bestimmte syntaktische Funktion zugeordnet ist. (Апресян 1972, 230f.)

Wie Hoffmann (1999, 1532–1545) beschreibt, weist der schriftliche Wissenschaftsstil einige Besonderheiten auf, die sich in verschiedenen Sprachen durchaus ähnlich sind und sich aus der Bestimmung von Fachtexten (Information, eindeutige Darstellung und Definition, Beweisführung, etc.) ergeben. Zunächst fällt die ungewöhnliche Länge der Sätze auf: So bestehen beispielsweise im Russischen Sätze in geisteswissenschaftlichen Texten aus durchschnittlich 17,57 Wörtern, was deutlich

über dem Wert (4,53) von Dramen, Gedichten und Romanen liegt. Das ergibt sich daraus, dass in wissenschaftlichen Arbeiten häufiger als anderswo komplexere Satzgefüge und damit viele Nebensätze auftreten. So können die geforderte Klarheit und Eindeutigkeit meist besser mithilfe von Relativ- oder Adverbialsätzen erreicht werden als mit einfachen Ausdrücken. Zu dem gleichen Zweck werden auch Wiederholungen in der Lexik oder dem Satzbau in Kauf genommen, während diese in anderen Bereichen wie der künstlerischen Literatur (oder auch der Publizistik) eher gemieden werden. Oft begegnen uns mehrteilige, komplexe Nominalgruppen als wichtigste Träger der Aussage. Typische grammatische Kategorien für (russische) Substantive sind Unbelebtheit, Neutrum oder Singular, bei Verben beispielsweise Passiv, Reflexivität und (im Russischen) imperfektiver Aspekt. Ein unpersönlicher Stil, bei dem die erste Person fast vollkommen zurücktritt, ist vorherrschend. Weitere für uns relevante wissenschaftliche Eigenheiten des vorliegenden Textes sind Zitate und Literaturangaben, Abkürzungen, Parenthesen, die Verwendung von Fachtermini und Fremdwörtern (синтаксическая структура предложения, предсказуемый автоматический анализ, подтверждать гипотезу, словоформа, гипотетический, etc.) sowie das Auftauchen eines typographisch hervorgehobenen Beispielsatzes.

Weiterhin untersuchen wir Ausschnitte eines Artikels aus dem Internetauftritt der zweisprachigen Zeitung *Heimat-Родина* aus dem Jahre 2005 mit dem Titel *Поздние переселенцы основывают сеть по оказанию самопомощи в Саксонии*, beziehungsweise *Spätaussiedler gründen Selbsthilfenetzwerk in Sachsen*. Auch hier erfolgt die Wiedergabe des russischen und des deutschen Textes:

Поздние переселенцы основывают сеть по оказанию самопомощи в Саксонии

Интеграция мигрантов проходит лучше благодаря их собственной инициативе

19 и 20 сентября 2005 года по приглашению городской администрации Löbau, Kulturbüro Sachsen e.V., RAA Sachsen e.V. и Общества „Oberlausitz - neue Heimat e. V.“ в Международном центре встреч St. Marienthal встретились более ста человек из 25 обществ и инициатив поздних переселенцев, еврейских иммигрантов и других мигрантов из бывшего Советского Союза для обмена опытом.

В результате открытых и плодотворных обсуждений участники встречи решили создать Netzwerk и поручили инициативной группе координацию первых шагов.

Netzwerk преследует следующие цели:

[...]

Изыскание возможностей вступать в диалог с политиками и администрацией Саксонии по вопросам особых трудностей, которые испытывают поздние переселенцы, еврейские иммигранты и другие мигранты из бывшего Советского Союза, а также поиск предложений по улучшению этой ситуации. Oberbürgermeister города Löbau Dietmar Buchholz приветствовал эту инициативу и обещал активную поддержку Netzwerk'у в период его становления, а также помощь при лоббировании интересов Netzwerk'a в политических кругах Саксонии и страны.

Netzwerk, как гражданская общественная инициатива, приглашает другие инициативы и лица для участия в его работе. Заинтересованные стороны могут зарегистрироваться по адресу: SachsenNetz@t-online.de или позвонить по телефону: 0 35 85/41 53 65 (факс: 0 35 85/41 53 68 или 03 51/495 63 44). (Heimat-Родина 2005b)

Spätaussiedler gründen Selbsthilfenetzwerk in Sachsen

Bessere Integration von Migranten durch Eigeninitiative

Am 19. und 20. September 2005 trafen sich auf Einladung der Stadt Löbau, des Kulturbüro Sachsen e.V., der RAA Sachsen e.V. und des Selbsthilfevereins Oberlausitz - neue Heimat e.V. über 100 Interessierte aus 25 Vereinen und Initiativen von Spätaussiedlern, jüdischen Zuwanderern und anderen Migranten aus der ehemaligen Sowjetunion im Internationalen Begegnungszentrum St. Marienthal zu einem Erfahrungsaustausch über ihre Arbeit.

Im Ergebnis der offenen und fruchtbaren Diskussionen beschlossen die Teilnehmerinnen und Teilnehmer ein gemeinsames Netzwerk zu gründen und beauftragten einen Initiativkreis mit der Koordinierung der ersten Schritte.

Das Netzwerk setzt sich folgende Ziele:

– [...]

– es sucht das Gespräch mit Politik und Verwaltung in Sachsen über die speziellen Problemlagen von Spätaussiedlern, jüdischen Zuwanderern und anderen Migranten aus der ehemaligen Sowjetunion und Vorschlägen (sic) zur Verbesserung dazu.

Der Oberbürgermeister der Stadt Löbau, Dietmar Buchholz, begrüßte die Initiative und versprach, die Gründungsphase sowie die politische Lobbyarbeit in Land und Bund aktiv zu unterstützen.

Das Netzwerk, das sich als eine zivilgesellschaftliche Initiative versteht, lädt weitere Initiativen und Personen ein, sich an seiner Arbeit zu beteiligen. Interessierte können sich melden unter SachsenNetz@t-online.de oder unter Telefon (03585) 41 53 65 bzw. Fax: (03585) 41 53 68 oder (0351) 495 63 44. (Heimat-Родина 2005a)

Zeitungsartikel zeigen sich in Bezug auf ihre sprachlichen Eigenschaften wesentlich heterogener als die wissenschaftliche Literatur. Die prinzipielle Gestaltung hängt hier zunächst stark von der Textsorte ab. So folgt

beispielsweise der Wetterbericht einem Schema aus feststehenden Versatzstücken, Nachrichten sollten möglichst objektiv Antworten auf grundlegende Fragen (Was?, Wer?, Wo?, Wann?) geben, während ein Kommentar bereits größere sprachliche Freiheiten erlaubt (vgl. Lüger 1983, 66–70 und 80–86). Besonders wenn die informative Funktion im Vordergrund steht, können auch in Zeitungstexten fachsprachliche Elemente auftauchen (vgl. Grabowski 1992, 38), deren Verwendung sich allerdings auch an den Vorkenntnissen der Leserschaft orientiert. Besondere Einflussfaktoren sind darüber hinaus zum Beispiel das Verbreitungsgebiet einer Zeitung (regional gefärbte Lexik) oder die Ausrichtung auf eine spezielle Zielgruppe (z. B. leicht verständliche Sprache für Kinder). In unserem Fall kann man von einem allgemeinverständlichen Bericht sprechen (Eroms 2008, 126), der einige längere Nominalphrasen, jedoch wenig komplexe Syntax beinhaltet. Da die Leser sachlich über Ablauf und Ergebnis einer vergangenen Veranstaltung informiert werden sollen, sind die Nennung von Zeit, Ort und beteiligten Personen, die Verwendung des Präteritums und das direkte oder indirekte Zitieren von mündlichen Aussagen charakteristisch. Außerdem begegnen uns im Artikel zahlreiche Eigennamen von Personen, Orten und Institutionen, zum Teil sogar in lateinischer Schreibweise, sowie einige Begriffe aus der Politik (поздние переселенцы, лоббирование интересов, городская администрация, иммигранты, гражданская общественная инициатива, Обер-бургомистр, etc.), die aber keine Fachkenntnisse des Rezipienten erfordern, auch wenn dieser nicht aus dem Kreis der russisch-deutschen Bevölkerung stammt. Den Abschluss bilden Verweise auf weitere Informationsquellen und Kontaktmöglichkeiten.

3.3. Evaluation

Voraussetzungen: Am Anfang der Evaluation steht zwangsläufig die Frage, was eigentlich eine gute oder schlechte Übersetzung ausmacht und auf welche Art man ihre Qualität jeweils ermitteln kann. Hutchins/Somers schlagen dazu folgende Kriterien vor (1992, 163):

„(a) its fidelity or accuracy, the extent to which the translated text contains the ‘same’ information as the original; (b) its intelligibility or clarity, the ease with which a reader can understand the translation; and (c) its style, the extent to which the translation uses the language appropriate to its content and intention. Each factor can be independent.“

Diese Kategorien leuchten ein und finden sich auch bei anderen Autoren, sind aber gleichzeitig problematisch, da ihre Beurteilung stark vom subjektiven Empfinden des Rezipienten abhängt und sie somit schwer messbar sind (vgl. Arnold 1994, 169–173, Carstensen 2012, 194 und Ramlow 2008, 83). Zweifelsohne werden aber Wiedergabetreue, Verständlichkeit und Stil des Ausgabetextes von Häufigkeit und Art der auftretenden Fehlübersetzungen beeinflusst. Solche konkreten sprachlichen Mängel lassen sich objektiver feststellen und ihr Anteil kann als Gradmesser für die Qualität der MÜ dienen, da sie genau den Arbeitsaufwand repräsentieren, der in der Postedition bis zur Druckreife eines Textes anfallen würde. Je nach Fragestellung und Anspruch an das Übersetzungsergebnis ist es bei einer solchen Analyse natürlich möglich, die Fehler unterschiedlich zu gewichten.

Für die Leistungsevaluation der getesteten Angebote können wir als Endnutzer nur von dem ausgehen, was wir tatsächlich sehen. Das heißt, wir haben keine Möglichkeit, die Abläufe im „Inneren“ der Maschine zu beobachten und die einzelnen Übersetzungsschritte nachzuvollziehen, sondern vergleichen nur Ausgangsmaterial und Endergebnis. Bei einer solchen *black-box evaluation* (Hutchins/Somers 1992, 163) können wir über das Zustandekommen eines Fehlers meist nur spekulieren, ob er also beispielsweise durch Schwächen des Analysemoduls oder durch einen lückenhaften Lexikoneintrag verursacht wurde. Zudem ist es oft schwierig, einen einzelnen Fehler von anderen abzugrenzen, da sie sich durch die internen Verarbeitungsprozesse der Software auch gegenseitig bedingen und somit Folgefehler produzieren können (vgl. Flanagan 1994, 65f.).

Mögliche Fehlerquellen: Allein unter dem Gesichtspunkt der Nutzbarkeit eines Dienstes ist die Ursache einer fehlerhaften Ausgabe (da ohnehin unerwünscht) natürlich eher von nachrangiger Bedeutung, aus linguistischer Sicht ist sie aber durchaus interessant, weshalb die Literatur hierzu einige Informationen liefert: Schon mehrfach wurde das Problem der Ambiguität angesprochen. In lexikalischer Hinsicht entsteht diese vor allem durch Homographie und Polysemie von Wörtern, in deren Folge zum Beispiel grammatische Kategorien falsch analysiert werden oder das System Schwierigkeiten hat zu entscheiden, welche Bedeutung des Wortes in der AS gemeint ist und welche der möglicherweise vorhandenen Übersetzungsvarianten dieser Bedeutung in der ZS am nächsten kommt. Durch Mehrdeutigkeit auf struktureller Ebene kann es passieren, dass Satzteile falsch zugeordnet werden (*Der Mann sah den Ballonfahrer mit dem Fernglas.*), Bezüge nicht richtig erkannt werden (*сплочение рабочих бригад вызвало осуждение товарища министра* [Апресян 1966, часть III, глава 2]) oder die Maschine die falsche Analysevariante wählt (vgl. *time flies like an arrow* gegenüber *fruit flies like a banana* [Text Technologies 2006, o. S.]) (vgl. Arnold 1994, 111–115, Austermühl 2001, 170–172 und Hutchins/Somers 1992, 85–94). Während des Transferprozesses ist das Programm oft mit lexikalischen oder strukturellen Nichtentsprechungen (*mismatches*) konfrontiert, die es umgehen muss: Das Konzept von *закусывать* beispielsweise („nach etwas Getrunkenem einen kleinen Happen essen“) lässt sich im Deutschen nur durch eine Paraphrase wiedergeben (auch *lexical hole* oder *lexical gap* genannt). Von *structural mismatches* spricht man, wenn für die Übersetzung Satzglieder getauscht werden müssen (wie in *меня зовут Антон* gegenüber *ich heiße Anton*) oder der Aufbau einer Phrase in der Zielsprache sogar komplett anders ist (*мне 28 лет* gegenüber *ich bin 28 Jahre alt*) (vgl. Arnold 1994, 115–122). Typisch russisch-deutsche Strukturdivergenzen mit Fehlerpotential sind Umformungen wie *обмен опытом* (Nominalgruppe) zu *Erfahrungsaustausch* (Kompositum) oder die verschiedenen Erscheinungsformen der Definitheit: Der Unterschied zwischen *Он послал матери денег* und *Он послал матери деньги* wird im Deutschen durch den unbestimmten (in diesem Fall fehlenden) beziehungsweise bestimmten Artikel angezeigt (vgl. Böttger 2008, 206–209). Auch die Verarbeitung von Idiomen ist nicht zu unterschätzen, da ihre Bedeutung sich nicht aus der

bloßen Summe der Bestandteile ergibt und sie deshalb als Ganzes verstanden und behandelt werden müssen (vgl. Austermühl 2001, 172f. und Arnold 1994, 122–127). Eine wörtliche Übersetzung der Einheit *как снес на голову* würde im Deutschen keinen Sinn ergeben, es müsste eine bedeutungsäquivalente Entsprechung gefunden werden, wie zum Beispiel *völlig unerwartet* oder *aus heiterem Himmel* (vgl. Kammer 1985, 81).

Vorgehen: Die Vorschläge zu Evaluation und Fehlerkategorisierung von MÜ in der Literatur sind uneinheitlich, auf eine spezielle Fragestellung ausgerichtet oder in Bezug auf kommerzielle Systeme oder bestimmte Forschungsprojekte konzipiert, was ihre Rolle für die geplante Untersuchung einschränkt (vgl. Carstensen 2012, 194, Flanagan 1994, Toral 2011 und Zervaki 2002). In manchen Werken wird versucht, Übersetzungsfehler in ein festgelegtes Raster einzuordnen, sie zu quantifizieren und Ranglisten zu erstellen (vgl. Ramlow 2008 und Wießner 2004). Ein solches Vorgehen erscheint aufgrund der Probleme bei der Fehleridentifizierung und vieler zu erwartender Grenzfälle jedoch nicht optimal.

In der vorliegenden Arbeit sollte die Klassifikation deshalb auf empirischem Wege gefunden werden. Mit groben Zuordnungen wie *Lexik* oder *Syntax* „im Hinterkopf“ erfolgten mehrere Durchsichten der Ergebnisse, um ein Bild der bestehenden Defizite zu gewinnen. Während dieses Prozesses kristallisierten sich häufige Fehlertypen zum Beispiel hinsichtlich der Artikelverwendung oder der Satzstruktur heraus, schrittweise wurden weitere Kategorien erkennbar. So war es nach systematischer Analyse schließlich möglich, die Mängel der Ausgabetexte nach ihrem tatsächlichen Auftreten zu Gruppen zusammenzufassen. Wie das nächste Kapitel zeigt, beziehen sich die Fehlerklassen zum Teil auf unterschiedliche linguistische Dimensionen (vgl. Lexik und Morphologie gegenüber Artikel oder Wortstellung). Diese Mischung mag auf den ersten Blick ungewöhnlich sein, ergibt sich aber aus der oben beschriebenen Vorgehensweise.

3.4. Ergebnisse

Der folgende Abschnitt präsentiert sowohl die maschinell übersetzten Texte (vom 28.7.2012) als auch die Befunde ihrer Fehlerauswertung, im Vergleich zu den russischen Originalen und den manuellen deutschen

Übersetzungen. Es werden jeweils einige geeignete Beispiele angeführt, um einen Eindruck der auftretenden Fehler zu vermitteln.³

3.4.1. Fehler im Bereich Lexik

Einige Wörter verschiedener Kategorien tauchen in den Ausgabetexten nicht mehr auf, sie bleiben also aus Sicht des Nutzers unübersetzt. Auffällig ist, wie viele Ausfälle bei Bing zu verzeichnen sind, hier ist die Verarbeitung offenbar ungenau oder das Lexikon besitzt zahlreiche Lücken. Google und Reverso liegen im Mittelfeld, während WorldLingo nur im sprachwissenschaftlichen Text zwei Leerstellen verzeichnet. Einzig im System von PROMT hat jedes der eingegebenen Wörter den Verarbeitungsprozess bis zum Ende durchlaufen. Dies bedeutet jedoch nicht, dass all die ausgegebenen Wörter auch richtig übersetzt wurden.

Text 1	Bing	Google	Reverso	WL	Text 2	Bing	Google	Reverso
чтобы	x		x		в (центре)	x		
проведем	x				поздних	x		
заполнить	x	x			результате	x		
декабрь	x				участники	x		
шли	x				инициативной	x		
на		x			преследует		x	
стоять		x			изыскание	x		
вызвало	x				испытывают	x		
у	x	x		x	в (период)	x		
нас				x	период	x		
уже	x				позвонить		x	
здесь		x			по (телефону)		x	x
о (процессе)		x			телефону			x
текста	x							
машины		x						
НСП 1			x					
в (предложении)		x						
анализируемого	x							
словам	x							
НСП 2			x					
для	x							
НСП 3			x					
в (НСП 3)	x							
вносятся	x							

Abbildung 3: Fehlende Wörter in den ausgegebenen Texten

In Bezug auf die Lexik ist in den vorliegenden Resultaten das häufigste Phänomen, dass zwar das Wort richtig erkannt wird, als Übersetzung jedoch aus den möglichen deutschen Bedeutungen eine solche gewählt

³ Das folgende Kapitel „Ausgegebene Texte“ wird aus Platzgründen nicht abgedruckt. Eine ungekürzte Fassung der Arbeit ist unter www.kempgen.ch/mt abrufbar.

wird, die im gegebenen Kontext nicht passend ist oder – mal nur sehr wenig, mal etwas mehr – von der ursprünglichen Aussage abweicht. Der Satz bleibt dabei meist verständlich, weshalb hier noch nicht von einer eigentlichen Fehlübersetzung gesprochen werden kann. Es finden sich zahlreiche Beispiele, sie werden jeweils mit dem russischen Ursprungswort auf der einen und den unzutreffenden Varianten der verschiedenen Anbieter auf der anderen Seite präsentiert: *появление* ≈ Entstehung [Bing, Google] / Darstellung [Bing] / Auftauchen [Google] / Aussehen [WorldLingo]; *лишь* ≈ nur [Bing, PROMT] / gerade [Reverso]; *открытый* ≈ geöffnet [PROMT, WorldLingo]; *плодотворный* ≈ fruchtbringend [PROMT]; *испытывать* ≈ erproben [PROMT] / prüfen [Reverso]; *поручить* ≈ anvertrauen [Reverso] / aufladen [WorldLingo]; *возможность* ≈ Gelegenheit [Reverso]; *центр* ≈ Mitte [WorldLingo]; *благодаря* ≈ weil [Google]; *как* ≈ wie [Google, PROMT, Reverso, WorldLingo]; *выпасть* ≈ ausfallen [PROMT] / herausfallen [WorldLingo]; *на* ≈ auf [Bing, PROMT, Reverso]; *место* ≈ Platz [Bing, Google, WorldLingo].

Anders verhält es sich, wenn die Aussage im Ergebnis tatsächlich verfälscht wird oder die gewählte Übersetzung zu weit vom ursprünglichen Sinn entfernt ist. Besonders bei häufigem Auftreten können solche falschen Übersetzungen das Verständnis eines Textes stark beeinflussen. Die Grenzen zwischen „gerade noch verständlich“ und „falsch“ sind natürlich oft fließend und nur im Zusammenhang mit dem jeweiligen Kontext gültig. In einer anderen Situation kann genau dieselbe Übersetzung unter Umständen als passend angesehen werden. An den folgenden Beispielen sind die inhaltlichen Abweichungen zum Originaltext gut erkennbar: *содержательный* ≈ diskret [Bing] / interessanterweise [Reverso] / sinnvoll [Google, WorldLingo]; *выпасть* ≈ aussteigen [Reverso]; *представление* ≈ Vertretung [Reverso]; *средство* ≈ Service [Reverso]; *продолжаться* ≈ Erlös [Reverso]; *обосновывать* ≈ basieren [Bing] / Basis [Google] / niedrig [WorldLingo]; *встретиться* ≈ erfüllen [Bing]; *приглашать* ≈ helfen [Bing] / laden [Google]; *мигрант* ≈ Wanderer [Reverso 3x, WorldLingo 3x]; *гражданский* ≈ höflich [WorldLingo]; *тепло* ≈ Hitze [WorldLingo]; *идти* ≈ schicken [WorldLingo]; *первый* ≈ Erstling [WorldLingo].

Es bietet sich an, solche lexikalischen Mängel, bei denen fachsprachlich relevante Begriffe betroffen sind, gesondert zu betrachten: Im sprachwissenschaftlichen Text wurde für einige Wörter nicht die hier geforderte

Bedeutung aus der Informatik (Maschine, Speicher, speichern) gewählt: машина ≈ Wagen [PROMT], память ≈ Gedächtnis [PROMT, Reverso], храниться ≈ behalten werden [PROMT]. Für den Begriff *предсказуемый* wurde von keinem System der richtige deutsche Ausdruck gefunden. Ein sich wiederholender Fehler ist auch die Wiedergabe von предложение mit *Vorschlag* [Bing 2x, Google, PROMT 4x, Reverso 4x] oder *Antrag* [WorldLingo 4x], statt im linguistischen Kontext mit *Satz*. Was hier falsch ist, wäre allerdings im Zeitungsartikel die richtige Wahl, da es dort tatsächlich um Vorschläge geht. Im zweiten Text liegen die Probleme eher in der treffenden Wiedergabe der politischen Terminologie: лоббирование интересов ≈ Lobbyismus der Interessen [PROMT] / Lobbyismus von Interessen [Reverso] / Beeinflussen der Interessen [WorldLingo]; поздние переселенцы ≈ späte Repatrianten [Bing 2x] / ø Einwanderer [Bing] / neue Einwanderer [Google 2x] / späte Auswanderer [PROMT 3x] / späte Siedler [Reverso 3x] / späte Wanderer [WorldLingo 3x]; Обер-бургомистр ≈ ??? ... Bürgermeister [Reverso] / [Ober]-burgomaster [WorldLingo].

Eine Illustration dafür, wie wichtig es ist, nicht nur richtig zu übersetzen, sondern auch die jeweiligen Kollokationen einer Sprache zu beherrschen sind die Varianten von *вступать в диалог (с политиками)*: den Dialog (mit den Politikern) betreten [PROMT] / im Dialog (mit Politikern) engagieren [Google] / in Dialog (mit Politikern) eingehen [Reverso] / am Dialog (mit den Politikern) teilnehmen [WorldLingo]. Sie sind nicht unbedingt falsch, entsprechen aber auch nicht der üblichen deutschen Ausdrucksweise für „(mit jemandem) in Dialog treten“.

Große Schwierigkeiten bereitet den MÜ-Diensten allgemein die Verarbeitung von Abkürzungen. Im wissenschaftlichen Text betrifft dies das neu eingeführte Akronym *НСП* (набор синтаксических предсказаний) und vor allem die Anfangsbuchstaben der Vornamen von Autoren: Ч. ≈ h. [Bing] / TSCH. [PROMT] / ?. [Reverso] / H. [WorldLingo]; A. ≈ a. [Bing] / ?. [Reverso]; И. ≈ und [Bing] / J. [Google] / ?. [Reverso]; НСП ≈ NSP [Bing, Google 2x] / NRS [Bing 2x, Google]. Besonders „kreativ“ ist WorldLingo bei der Übertragung von *e. V.* zu *E. Das v.* [WorldLingo 3x].

Interessant ist auch, wie Systeme mit unbekannten Wörtern umgehen, sofern sie sie nicht komplett unterschlagen. Manche werden lediglich transkribiert: *предсказуемый* ≈ *predskazuemostnyj* [Bing 3x];

Хоккетом ≈ [Khokketom] [WorldLingo]; Родес ≈ [Rodes] [WorldLingo]; НСП ≈ [NSP] [WorldLingo 3x]; Обер ≈ [Ober] [WorldLingo]. Andere bleiben völlig unangetastet: предска́зуемый [PROMT 3x]; НСП [PROMT 3x]. Dass nicht nur Google Zwischenschritte über das Englische vollzieht, lassen folgende Resultate vermuten: слово ≈ Word [Bing]; словоформа ≈ Wordform [Bing]; предска́зуемый ≈ predictive [Google] / prepredictional [Reverso 3x]; изменения ≈ changess (sic) [Reverso]; самопомощь ≈ Self-help [WorldLingo]; бессознательный ≈ unconsciously [WorldLingo]. Ein solches Vorgehen ist insofern sinnvoll, als der Nutzer womöglich ein englisches Wort verstehen kann, oder aber die Gelegenheit hat, noch manuell in einem Lexikon nachzuschlagen und die Lücken zu füllen. Im Fall von Reverso ist dies nicht möglich, da hier an den betreffenden Stellen Fragezeichen als Platzhalter eingefügt werden und das ursprüngliche Wort nicht mehr erkennbar ist: Ч. ≈ ?.; I. Родес ≈ ?.????; А. ≈ ?.; НСП ≈ ???; Обер ≈ ???.

3.4.2. Fehler im Bereich Morphologie

Im Bereich der Morphologie (Wortart, Kasus, Numerus, etc.) ist es oft schwierig zu entscheiden, wo der eigentliche Fehler liegt, bzw. ob z. B. ein „falscher“ Kasus nur deshalb zustande kommt, weil das System einen abweichenden Satzbau zu verwenden versucht und sich die morphologischen Anpassungen daraus als Folgefehler ergeben. Einige Erscheinungen wurden dennoch als Fehler eingestuft, zum Beispiel bezüglich der grammatischen Kategorie (синтаксических ≈ Syntax [Bing]; основывают ≈ Basis [Google]; помощь ≈ hilft [WorldLingo]), des Kasus (автоматического анализа ≈ automatische Analyse [Bing]; снег ≈ Schnees [Bing]; страны ≈ Land [PROMT]), des Numerus (опытом ≈ Erfahrungen [Bing]; слова ≈ Wort [Google]; поручили ≈ wies [Bing]) oder des Modus (проведем ≈ führen Sie [Google]). Im Zusammenhang mit Genus, Tempus oder Person treten kaum Fehler auf.

Als ein besonderes Problem stellten sich die Instrumental-Formen der Eigennamen *Хоккетом* und *Эттингером* heraus, sie wurden teilweise als Namensbestandteile angesehen und übernommen: Hokketom [Bing, Google]; Ettingerom [Bing]; Chokketom [PROMT]; Etttingeroms [PROMT]. Die Verbindung von Fremdwort und angehängter russischer Endung bei *Netzwerk'a* und *Netzwerk'y* konnte kein System vollständig auflösen: Netzwerk'u [Bing, Google], Netzwerk'a [Bing, Google, PROMT],

Netzwerk'y [PROMT], Netzwerk '? [Reverso 2x], netzwerk'[u] [WorldLingo], netzwerk'[a] [WorldLingo].

Wie bereits erwähnt, ist die Wahl des richtigen Artikels im Deutschen für die Übersetzungssysteme nicht unproblematisch, da diese Kategorie im Russischen fehlt und die Definitheit auf anderem Wege markiert werden muss (vgl. Böttger 2008, 206–209). Es kommt daher relativ häufig dazu, dass entweder der Artikel fehlt (Ø Oberbürgermeister Ø Stadt von Löbau [Bing]), ein eigentlich unnötiger Artikel eingefügt wird (Die späten Auswanderer gründen das Netz nach der Erweisung der Selbsthilfe in Sachsen [PROMT]), fälschlicherweise der definite Artikel verwendet wird (Prozess des Verständnisses ... des Textes [PROMT]), oder umgekehrt eigentlich ein definiter Artikel nötig wäre (wird durch eine Technik der Predskazuemostnogo Analyse formalisiert [Bing]). Abgesehen davon weisen die tatsächlich gesetzten Artikel meist keine morphologischen Fehler auf.

Anders bei den Pronomina, wo die Wahl des Genus (und des richtigen Typs) oft Schwierigkeiten bereitet: (анализ,) которого ≈ von denen [Bing] / die [Google] / der [Reverso], dessen [PROMT]; какое ≈ was [Bing, Google], период его (=Netzwerk) становления ≈ ihrer Bildung [Google].

3.4.1. Fehler im Bereich Syntax

Wie bei vielen anderen bestehen auch innerhalb des Sprachenpaares Russisch-Deutsch einige Strukturdivergenzen, weshalb darauf geachtet werden muss, Aufbau und Konstruktionen der Sätze oder ihrer Bestandteile für die Zielsprache anzupassen. Denn dass eine Eins-zu-eins-Übernahme oft ungrammatisch ist, zeigen die folgenden Beispiele: (1) Появление другого слова вызвало бы у нас удивление ≈ Die Entstehung von einem anderen Wort kann uns Überraschung [Bing] / Die Entstehung eines anderen Wortes verursachen würde uns überraschen [Google] / Das Erscheinen anderen Wortes hätte bei uns das Erstaunen herbeigerufen [PROMT] / Ergebnis anderen Wortes würde an uns verursachen Überraschung [Reverso] / Das Aussehen eines anderen Wortes würde überraschung verursachen [WorldLingo]. (2) зарегистрироваться по адресу ≈ sich anmelden an [Bing] / registriert werden an die Adresse [PROMT] / registriert werden zur Adresse [Reverso] / registriert werden mit der Adresse [WorldLingo]. Obwohl die Aussagen in beiden Fällen einigermaßen verständlich sind, kann keiner der Dienste hier mit seiner Lösung überzeugen.

Ebenso verhält es sich, wenn komplexere Syntax aufgelöst werden soll: Bei dem Versuch, den Abschnitt *вступить в диалог с политиками и администрацией Саксонии по вопросам особых трудностей, которые испытывают поздние переселенцы, еврейские иммигранты и другие мигранты из бывшего Советского Союза* (Zeitungstext) zu übersetzen, geraten in den Ergebnissen Bezüge und Satzstellung stark durcheinander und begonnene Konstruktionen werden meist nicht konsequent zu Ende geführt. Der erste Satz des wissenschaftlichen Textes verbindet sogar besondere Komplexität (finaler Adverbialsatz, von dem ein weiterer Relativsatz abhängt, der wiederum zwei Verbalphrasen beinhaltet) mit Strukturdivergenzen (Konstruktion mit Instrumental), was sich in der fehlenden Kohärenz der ausgegebenen Sätze deutlich niederschlägt.

Im Russischen wird häufig mit Partizipialkonstruktionen gearbeitet, die im Deutschen zum Beispiel als adverbialer Nebensatz aufgelöst werden müssen. Auch dies ist in den genutzten Systemen noch verbesserungswürdig: (... ту гипотезу,) которую мы бессознательно сформировали, прочитав уже «поступившие» слова. ≈ ... die wir unbewusst durch Lesen der Worte „empfangene“ gebildet. [Bing] / ... dass wir unbewusst durch das Lesen der bereits „empfangen“ das Wort gebildet. [Google] / ... die wir bewusstlos gebildet haben, schon „die handelnden“ Wörter gelesen. [PROMT] / ... die wir unbewußt erzeugten, bereits „weitergegangene“ Wörter lesend(anzeigend). [Reverso] / ... die wir unconsciously bildeten, nachdem das Ablesen, das bereits „ist,“ Wörter eintrug. [WorldLingo].

Oft wird die Satzstruktur auch durch Präpositionen, Junktoren und andere Elemente gestört, entweder weil sie fehlen, obwohl die Konstruktion sie verlangt (um ... zu [Bing]; solange bis [PROMT]; Suche nach [Reverso]), oder weil sie unnötigerweise eingefügt wurden (Stadt von Löbau [Bing]; von Sachsen und von Land [WorldLingo]).

Eine weitere Fehlergruppe liegt in der Zuordnung der Satzglieder und dem Herstellen der passenden Bezüge innerhalb der Sätze. Hier werden zum Beispiel Kongruenzen nicht richtig wiedergegeben: в Международном центре встреч St. Marienthal ≈ internationale Tagungen, Zentrum St. Marienthal [Bing]; гипотетический перечень возможных в некотором предложении синтаксических структур ≈ Liste von möglich in einigen

hypothetischen Satz syntaktischen Strukturen [Bing] / Liste der möglichen spekulativen Satz zu einigen syntaktischen Strukturen [Google] / das hypothetische Verzeichnis möglich in einigem Vorschlag der syntaktischen Strukturen [PROMT].

Innerhalb von Nominalphrasen gelingt es in der Übersetzung teilweise nicht, die Bestandteile entsprechend ihrer Funktionen zu beizubehalten, was den Sinn verzerrt: поиск предложений ≈ Suchvorschläge [Google]; хранящийся в памяти машины набор синтаксических предсказаний (НСП) ≈ in der Computer-Speicher Set-Syntax Vorhersagen (NSP) ... gespeichert [Bing]; (заключается в) последовательном сравнении информации ≈ sequenzielle Vergleichsinformationen [Bing]; в Международном центре встреч ≈ internationale Tagungen, Zentrum [Bing] / Treffen im Internationalen Zentrum [Google].

Für das Deutsche ist außerdem die Verwendung von Komposita sehr charakteristisch. Bleiben die Paraphrasen des russischen Originaltextes bestehen, sind diese zwar meist einigermaßen richtig übersetzt, verhindern aber das Erreichen eines adäquaten, natürlichen Sprachstils: (1) для обмена опытом ≈ für den Austausch von Erfahrungen [Bing] / für den Austausch durch Erfahrungen [Reverso] / für den Austausch der Erfahrung [WorldLingo]; (2) сеть по оказанию самопомощи ≈ Basis-Netzwerk, um sich in Sachsen helfen [Google] / das Netz nach der Erweisung der Selbsthilfe [PROMT] / Netzwerk auf Übergabe der Selbsthilfe [Reverso] / Netz ... auf der Übertragung zum Self-help [WorldLingo]; (3) в Международном центре встреч ≈ im Internationalen Zentrum der Treffen [PROMT] / in Internationalem Zentrum der Sitzungen [Reverso] / in der internationalen Mitte von Treffen [WorldLingo]. Manchmal wird allerdings auch erst durch die Bildung eines Kompositums der ursprüngliche Sinn verfälscht: изыскание возможностей ≈ Gelegenheitsvermessen [Reverso]; появление слова снег ≈ das Aussehen eines Wortschnees [WorldLingo] / Wort-Ergebnis-Schnee [Reverso]; основывают сеть ≈ Basis-Netzwerk [Google]; на последнем месте ≈ auf Rücksitz [Reverso].

In einigen Fällen wird durch die falsche Verknüpfung von Satzgliedern sogar die komplette Struktur des Satzes zerstört und das Resultat ist nicht mehr verständlich: Изложенное здесь представление о процессе понимания (анализа) текста формализуется техникой предсказуемост-

ного анализа. ≈ Das Vorstehende ist ein Verständnis des Prozesses (Analyse) des Textes wird formalisiert prädiktive Analyse-Technik. [Google] / Vertretung Festgesetzt hier Prozesses, (Analyse) des Textes zu verstehen, wird durch Technik der prepredicational Analyse formalisiert. [Reverso] / Idee über den Prozeß, der hier von verstehendem dargestellt wird (Analyse) Text ist formalisierte vorbestimmte Analyse Technik. [WorldLingo].

Eng mit der Struktur hängt auch die Wortstellung im Satz zusammen. Es wird ungleich schwerer, einen Text zu verstehen, wenn die Satzteile nicht in der üblichen Reihenfolge auftauchen. Die beiden folgenden Varianten desselben Teilsatzes machen dies deutlich: ... содержащий в себе гипотетический перечень возможных в некотором предложении синтаксических структур ≈ ... in sich selbst hypothetische Liste von syntaktischen in einem Vorschlag (=Satz) möglichen Strukturen enthaltend. [Reverso] / ... mit einer Liste von möglich in einigen hypothetischen Satz syntaktischen Strukturen [Bing]. Es ist jedoch anzumerken, dass es auch bei dieser Fehlergruppe manchmal schwierig ist, sie strikt von anderen (Bezüge, Struktur) zu trennen.

Formales: Die letzte Fehlergruppe bezieht sich nicht mehr auf die sprachlichen Phänomene, sondern auf formale Unregelmäßigkeiten, als da wären fehlende oder überzählige Leerstellen oder Satzzeichen, typographische Fehler bei Bindestrichen, Klammern und Anführungszeichen, die Nichtbeachtung von Absätzen (Reverso, WorldLingo), oder fehlerhafte Orthografie (Folgendes experiment [Bing], initiative [Bing], sachsen [WorldLingo], heimat [WorldLingo], begrüßte [PROMT], unbewußt [Reverso]). Ein grober Fehler unterläuft dem System von Bing, welches aus nicht nachvollziehbaren Gründen die Ziffernfolge innerhalb der Telefonnummern verändert.

3.4.4. Zusammenfassung der Analyse

Wir haben gesehen, dass der Nutzer von frei zugänglichen MÜ-Diensten mit einer Vielzahl von Fehlern konfrontiert wird, möglicherweise mehr, als zu erwarten waren. Nun stellt sich die Frage, ob zwischen den einzelnen Systemen sowie zwischen den beiden Textsorten und Funktionalstilen Qualitätsunterschiede bestehen.

Zunächst kann man festhalten, dass in allen zehn ausgegebenen Übersetzungen zahlreiche sprachliche und einige formale Mängel gefunden wurden und keiner der Texte ohne Vorbehalte als eine gute deutsche

Wiedergabe der russischen Vorlage zu bezeichnen ist. Bei näherer Betrachtung zeichnen sich jedoch einige Tendenzen ab.

Die Untersuchung wurde anhand zweier in Bezug auf Thematik, Komplexität und Fachsprachlichkeit verschiedenartiger Texte durchgeführt, um sichtbar zu machen, bei welchen sprachlichen Phänomenen die MÜ-Systeme gegebenenfalls bessere oder schlechtere Ergebnisse erzielen. In dieser Hinsicht kommen wir zu einem eher unerwarteten Befund: Das Niveau der Übersetzungen desselben Anbieters ist jeweils für beide Texte ein ähnliches. Eventuell ist ein leicht positiver Ausschlag in Richtung des Zeitungsartikels gegeben, der vermutlich im subjektiven Empfinden noch stärker ausgeprägt ist. Er erscheint möglicherweise zunächst strukturell und inhaltlich besser verständlich. Die Resultate müssen jedoch auch im Verhältnis zum russischen Ausgangsmaterial und den manuellen deutschen Übersetzungen gesehen werden. Vergleicht man diese miteinander, ist auf der einen Seite die lexikalische und syntaktische Gestaltung der wissenschaftlichen Abfassung deutlich anspruchsvoller, und auf der anderen Seite der Pressebericht thematisch und strukturell leichter zugänglich. Unter der Voraussetzung, dass diese Unterschiede unabhängig von der maschinellen Verarbeitung bestehen (bleiben), ergibt sich ein ausgeglicheneres Bild.

Deutliche Unterschiede in der Qualität der Ergebnisse zeigen dagegen einige MÜ-Dienste untereinander. Während Google, Reverso und WorldLingo mit jeweils verschiedenen Stärken und Schwächen im Mittelfeld stehen, fällt Bing im Vergleich deutlich ab. Hier fehlen in den Resultaten mit großem Abstand die meisten Wörter, zudem lassen sich fast alle oben beschriebenen Fehlergruppen durch ein Beispiel aus den Übersetzungen von Bing belegen. Die Änderung der Ziffernfolge der Telefon- und Faxnummern ist als weiterer großer Mangel zu bewerten. Wie zuvor berichtet, benutzt Bing nicht mehr die traditionsreiche und erfolgreiche Maschine von Systran – beziehungsweise Babel Fish – sondern eine eigene Software. Es wäre interessant zu erfahren, inwieweit sich das System dadurch verändert hat. Einige Nutzer, die nach der Umstellung automatisch von Babel Fish zu Bing umgeleitet wurden, üben zumindest Kritik:

- (1) „Babel fish could correctly translate things. Now nothing is 100% accurate.“
- (2) „I entered ‚What is your name‘ in English, and asked for the German translation. It came back as ‚What is your name‘, yes, in English.“
- (3) „I was using yahoo babelfish yesterday on the yahoo page and it was working great. I came

back today to see that it has been changed. So I have done a little testing on the same exact things I was looking at yesterday and today the translation came up incorrectly.“ (Microsoft Translator Team Blog 2012, o. S.)

Die Übersetzungen des Anbieters PROMT erscheinen dagegen sowohl im ersten Eindruck als auch unter Berücksichtigung der Fehlerauswertung vergleichsweise als die besten: Es sind keine fehlenden Wörter zu verzeichnen, fast kein Wort wird mit einer komplett falschen Bedeutung wiedergegeben und Morphologiefehler treten nur selten auf. Ohne Zweifel sind auch diese Texte nicht frei von Mängeln, allerdings werden hier am ehesten grammatikalische Ansprüche erfüllt und Konstruktionen innerhalb eines Satzes konsequent beibehalten. Wir erinnern uns, dass PROMT entgegen des allgemeinen Trends ein regelbasiertes, hierarchisches Transfersystem benutzt. Möglicherweise kann deshalb die grundlegende Satzstruktur öfter aufrecht erhalten werden als bei statistischen oder beispielbasierten Systemen, die einen Satz eher aus einzelnen Bestandteilen zusammensetzen. Beispielsweise hat PROMT als einziges den Grundaufbau des ersten Satzes des wissenschaftlichen Textes richtig wiedergegeben: *Um die Hauptidee [...] zu erklären, [...] werden wir das folgende Experiment durchführen.* Nur bei PROMT finden sich auch einige von Anfang bis Ende sehr gut übersetzte Sätze, in denen zwar vereinzelt nicht die richtige Bedeutungsvariante gewählt wurde, deren Struktur jedoch grammatikalisch korrekt und vollständig ist. Die Markierungen machen deutlich, wie wenig Nachbearbeitung nötig wäre:

(1) „Der Prozess der Analyse besteht im konsequenten Vergleich der Informationen zu den Wörtern des analysierten Vorschlags Satzes mit allen Vorlegenden Vorher-sagen, die in HCH der MSV aufgezählt sind. Wenn für das Wort die „passende“ Voraussage gefunden ist, wird dem Wort der Vermerk der entsprechenden syntaktischen Funktion zugeschrieben, und werden in HCH der MSV die notwendigen Änderungen vorgenommen“ <114, 14> vorgenommen. Die Analyse dauert, bis jeder Wortform des Vorschlags Satzes eine bestimmte syntaktische Funktion zugeschrieben sein wird.

(2) Infolge der geöffneten offenen und fruchtbringenden fruchtbaren Erörterungen haben sich die Teilnehmer des Treffens entschieden, ein Netzwerk zu schaffen und haben der Aktionsgruppe die Koordination der ersten Schritte aufgetragen.

Dass ausgerechnet der russische Anbieter PROMT auch die besten Ergebnisse bei der Übersetzung eines russischen Textes erzielt, weckt naturgemäß die Aufmerksamkeit eines Slavisten. Es wäre interessant, in

weiteren Vergleichen zu überprüfen, ob das von PROMT verwendete System auch bei anderen Sprachenpaaren Vorteile bietet, oder ob es hauptsächlich in Bezug auf das Russische anderen Diensten überlegen ist. In jedem Fall ist das kostenlose Online-Angebot von PROMT zu Unrecht eher unbekannt.

4. Fazit

Aus den bisherigen Betrachtungen gewinnen wir die Erkenntnis, dass die russisch-deutschen Resultate von im Internet frei verfügbaren MÜ-Anbietern erwartungsgemäß keine Konkurrenz zu menschlichen Leistungen, aber auch nicht völlig unbrauchbar sind. Je nach Anspruch an die Übersetzung und in einem bestimmten Rahmen ist die MÜ für das Russische durchaus sinnvoll einsetzbar.

Wer dieser Sprache selbst nicht mächtig ist, erhält durch MÜ relativ einfach die Möglichkeit, sich Informationen zu beschaffen und ungefähr zu verstehen, wovon ein Text handelt und wie die Aussage zu einem Thema ist. Ein Anwendungsbeispiel dazu findet sich im sozialen Netzwerk *Facebook* (<https://www.facebook.com>), das in Zusammenarbeit mit Bing anbietet, direkt innerhalb der Benutzeroberfläche die automatische Übersetzung fremdsprachiger Meldungen zu erhalten. So kann ein Nutzer zum Beispiel verstehen, worüber sich seine russischen Bekannten, mit denen er normalerweise auf Englisch kommuniziert, untereinander unterhalten. Um allerdings tatsächlich auf Russisch eine Konversation zu führen, sind die Ergebnisse zu unzuverlässig und bergen die Gefahr großer Missverständnisse.

Auch im Rahmen des Spracherwerbs ist die Nutzung von Online-MÜ denkbar. Möglicherweise möchte ein Anfänger nach der Lektüre eines Textes aus dem Lehrbuch überprüfen, ob er dessen Aussage verstanden hat, oder sie ist ihm eine willkommene (wenn auch nicht unbedingt geeignete) Hilfe bei Hausaufgaben. Mit fortgeschrittenen Lernern könnte ein Dozent die russischen Resultate einer automatischen Übersetzung analysieren und den Kursteilnehmern die Aufgabe geben, die gefundenen Mängel zu verbessern, sodass sie sich durch die Beschäftigung mit Fehlern gleichzeitig die richtige Sprachverwendung einprägen. Für Lehramtsstudenten wäre dies ebenfalls eine Möglichkeit, Identifizierung und Korrektur von Fehlern zu üben.

Wenn beim Anwender bereits einige Russischkenntnisse vorhanden sind, in unserem Fall etwa ein Student der Slavistik, kann er die MÜ beispielsweise für die Literaturrecherche nutzen, um einen schnellen Überblick über Aufbau und Thematik russischer Abfassungen zu gewinnen. Auf diese Weise kann er entscheiden, welche Passagen für eine genauere manuelle Übersetzung in Frage kommen und welche Texte dagegen für ihn irrelevant sind. Die Ergebnisse können dann zusätzlich als Grundgerüst für die eigene übersetzerische Arbeit dienen. Einem Nutzer mit entsprechenden Sprachkenntnissen ist es außerdem besser möglich, die Defizite der Resultate zu erkennen und sich unverständliche Elemente durch einen Vergleich mit dem Original plausibel zu machen, oder das ein oder andere unbekannte Wort selbst nachzuschlagen.

Sollte es bei manchen Themen auf die inhaltliche Exaktheit einer Information ankommen, zum Beispiel bei den rechtlich relevanten Bestimmungen des russischen Zolls oder der Bewerbung in einem russischen Unternehmen, so sollte der Online-MÜ nicht vertraut werden. In diesem Bewusstsein kann aber die Auffassung vertreten werden, dass in vielen Fällen eine schlechte Übersetzung immer noch besser ist als gar keine Übersetzung.

Maschinelle Übersetzung eröffnet heutzutage neue Wege der Kommunikation und beschleunigt den globalen Informationsfluss. Durch das Massenmedium des Internets wird ihre Verfügbarkeit und Verwendung immer stärker vorangetrieben. Dies ist zum einen eine positive Errungenschaft, zum anderen besteht jedoch das Risiko, dass auch Fehlinformationen und falsche Übersetzungen eine rasche Verbreitung finden. Frei zugängliche MÜ im Internet kann Sprachbarrieren also vermindern, sie aber nicht einreißen. Die Nutzer sollten sich deshalb deren momentan bestehenden Grenzen bewusst sein, ihren Einsatz aber auch nicht scheuen. Bis allerdings die maschinelle Verarbeitung und automatische Übersetzung von geschriebener oder sogar gesprochener Sprache ein Niveau erreicht, das dem des visionären *Universal Translator* aus der Science-Fiction-Serie *Star Trek* nahe kommt, welcher der unmittelbaren Übersetzung jeder beliebigen, auch extraterrestrischen Sprache dient (vgl. Wikipedia 2012a, o. S.), wird durch MÜ-Systeme, wie WorldLingo es ausdrücken würde, vermutlich noch sehr viel „Wortschnee“ produziert.

Literatur

- Adams, Tim (2010): „Finally, you are speaking my language.“ In: *The Observer*, 19.12.2010: *The New Review*. 22–23. Zitiert aus: *Machine Translation Archive*. <http://www.mt-archive.info/Observer-2010-Adams.pdf> (letzter Aufruf 10.8.2012).
- Antosik, Jessica (2012): „Sechs Jahre Google Translate.“ In: *UEPO*. <http://uepo.de/2012/05/04/sechs-jahre-google-translate/> (letzter Aufruf 10.8.2012).
- Apresjan, Ju. D. (²1972): *Ideen und Methoden der modernen strukturellen Linguistik. Kurzer Abriss* (übers. und hg. von Halthof, Brigitte/Mai, Elisabeth). München.
- Апресян, Ю.Д. (1966): „Идеи и методы современной структурной лингвистики. Краткий очерк.“ In: *Classes.ru/Ю.Д. Апресян Идеи и методы современной структурной лингвистики*. http://www.classes.ru/grammar/170.Apresyan/source/worddocuments/__.htm (letzter Aufruf 10.8.2012).
- Arnold, Doug et al. (1994): *Machine translation. An Introductory Guide*. Oxford.
- Austermühl, Frank (2001): *Electronic Tools for Translators*. Manchester.
- Böttger, Katharina (2008): *Die häufigsten Fehler russischer Deutschlerner. Ein Handbuch für Lehrende*. Münster (Mehrsprachigkeit 22).
- Carstensen, Kai-Uwe (2012): „Sprachtechnologie. Ein Überblick.“ In: *Kai-Uwe-Carstensen*. <http://kai-uwe-carstensen.de/Publikationen/Sprachtechnologie.pdf> (letzter Aufruf 10.8.2012).
- Carstensen, Kai-Uwe u. a. (2001): *Computerlinguistik und Sprachtechnologie. Eine Einführung*. Heidelberg.
- Epiphantseva, Julia (2010): „PROMT. Flexible and Efficient Managment of Translation Quality. Präsentation, Berlin.“ In: *Translingual Europe 2010*. <http://www.translingual-europe.eu/slides/JuliaEpiphantseva.pdf> (letzter Aufruf 10.8.2012).
- Eroms, Hans-Werner (2008): *Stil und Stilistik. Eine Einführung*. Berlin (Grundlagen der Germanistik 45).
- Flanagan, Mary A. (1994): „Error Classification for MT Evaluation.“ Original in: *Technology partnerships for crossing the language barrier: Proceedings of the First Conference of the Association for Machine Translation in the Americas, 5–8 October, Columbia, Maryland, USA*. Washington, DC. 65–72. Zitiert aus: *Machine Translation Archive*. <http://www.mt-archive.info/AMTA-1994-Flanagan.pdf> (letzter Aufruf 10.8.2012).
- Gaspari, Federico (2007): *The Role of Online MT in Webpage Translation*. Original: Manchester. Zitiert aus: *Machine Translation Archive*. <http://www.mt-archive.info/Gaspari-2007.pdf> (letzter Aufruf 10.8.2012).
- Gaspari, Federico/Hutchins, John (2007): „Online and Free! Ten Years of Online Machine Translation: Origins, Developments, Current Use and Future Prospects.“ Original in: *MT Summit XI*, Heft 10–14, Kopenhagen. 199–206. Zitiert aus: *Machine Translation Archive*. <http://www.mt-archive.info/MTS-2007-Gaspari-1.pdf> (letzter Aufruf 10.8.2012).
- Gaspari, Federico/Somers, Harold (2007): „Making a sow’s ear out of a silk purse: (mis)using online MT services as bilingual dictionaries.“ Original in: *Translating and the Computer 29. Proceedings of the twenty-ninth international conference on Translating and the*

- Computer, 29–30 November 2007. London. 15ff. Zitiert aus: *Machine Translation Archive*. <http://www.mt-archive.info/Aslib-2007-Gaspari.pdf> (letzter Aufruf 10.8.2012).
- Grabowski, Martina (1992): *Fachsprache und Funktionalstil. Ihr Zusammenwirken, demonstriert am Beispiel russischer Fachtexte des Außenhandels*. Frankfurt am Main (Beiträge zur Slavistik 19).
- Grunwald, Dave (2009): „Language Weaver Boosts WorldLingo Website Traffic.“ In: *GTS Translation Services*. <http://blog.gts-translation.com/2009/10/26/language-weaver-boosts-worldlingo-website-traffic/> (letzter Aufruf 10.8.2012).
- Grunwald, Dave (2011): „Can Transperfect save Worldlingo?“ In: *GTS Translation Services*. <http://blog.gts-translation.com/2011/11/15/can-transperfect-save-worldlingo/> (letzter Aufruf 10.8.2012).
- Hausser, Roland (2000): *Grundlagen der Computerlinguistik. Mensch-Maschine-Kommunikation in natürlicher Sprache*. Berlin.
- Heimat-Родина (2005a): „Spätaussiedler gründen Selbsthilfenetzwerk in Sachsen.“ In: *Heimat-Родина*. <http://www.zeitung-heimat.de/integration/in97.html> (letzter Aufruf 10.8.2012).
- Heimat-Родина (2005b): „Поздние переселенцы основывают сеть по оказанию самопомощи в Саксонии.“ In: *Heimat-Родина*. <http://www.zeitung-heimat.de/integration/in96.html> (letzter Aufruf 10.8.2012).
- Helft, Miguel (2010): „Google’s Computing Power Refines Translation Tool.“ In: *New York Times*, 8.3.2010. Zitiert aus: *Machine Translation Archive*. <http://www.mt-archive.info/NYT-2010-Helft.pdf> (letzter Aufruf 10.8.2012).
- Hoffmann, Lothar (1999): „Die russischen Fachsprachen im 20. Jahrhundert und ihre Erforschung: eine Übersicht.“ In: Hoffmann, Lothar et al. (Hrsg.), *Fachsprachen. Ein internationales Handbuch zur Fachsprachenforschung und Terminologiewissenschaft*. Berlin. 1532–1545 (Handbücher zur Sprach- und Kommunikationswissenschaft 14.2.).
- Hutchins, John (1995): „«The whisky was invisible», or Persistent myths of MT.“ Original in: *MT News International* 11. Zitiert aus: *John Hutchins personal website*. <http://www.hutchinsweb.me.uk/MTNI-11-1995.pdf> (letzter Aufruf 10.8.2012).
- Hutchins, John/Somers, Harold (1992): *An introduction to machine translation*. Original: London. Zitiert aus: *John Hutchins personal website*. <http://www.hutchinsweb.me.uk/IntroMT-TOC.htm> (letzter Aufruf 10.8.2012).
- Kammer, Gerlinde (1985): *Probleme bei der Übersetzung Phraseologischer Einheiten aus dem Russischen ins Deutsche*. München (Slavistische Beiträge 183).
- Linguatools-Blog (2010a): „Google-Übersetzer arbeitet mit Interlingua.“ In: *Linguatools-Blog*. <http://www.linguatools.de/blog/?p=102> (letzter Aufruf 10.8.2012).
- Linguatools-Blog (2010b): „Interview mit Franz Josef Och, Teamleiter von Google Translate.“ In: *Linguatools-Blog*. <http://www.linguatools.de/blog/?p=285> (letzter Aufruf 10.8.2012).
- Lüger, Heinz-Helmut (1983): *Pressesprache*. Tübingen.

- Marchuk, Yu. M. (1989): „Machine-Aided Translation: A Survey of Current Systems.“ In: Batori, István S. u. a. (Hrsg.), *Computerlinguistik. Ein internationales Handbuch zur computergestützten Sprachforschung und ihrer Anwendungen*. Berlin. 682–688 (Handbücher zur Sprach- und Kommunikationswissenschaft 4.)
- Melby, Alan (1989): „Machine Translation: General Development.“ In: Batori, István S. et al. (Hrsg.), *Computerlinguistik. Ein internationales Handbuch zur computergestützten Sprachforschung und ihrer Anwendungen*. Berlin. 622–629 (Handbücher zur Sprach- und Kommunikationswissenschaft 4.).
- Microsoft Translator Team Blog (2012): „Welcoming Yahoo! Babel Fish users!“ In: *Microsoft Translator Team Blog*. <http://blogs.msdn.com/b/translation/archive/2012/05/30/welcoming-yahoo-babel-fish-users.aspx> (letzter Aufruf 10.8.2012).
- PROMT (2005a): „Kurze Einführung in die PROMT Machine Translation Technologie.“ In: PROMT. <http://www.promt.de/media/pdf/promt-machine-translation-technologie.pdf> (letzter Aufruf 10.8.2012).
- PROMT (2005b): „Wie der Computer übersetzt.“ In: PROMT. <http://www.promt.de/media/pdf/wie-der-computer-ubersetzt.pdf> (letzter Aufruf 10.8.2012).
- Ramlow, Markus (2008): „Maschinelle Übersetzungssysteme im Vergleich.“ In: Seewald-Heeg, Uta (Hrsg.), *Maschinelle Übersetzung und XML im Übersetzungsprozess. Prozesse der Translation und Lokalisierung im Wandel*. Berlin. 15–149.
- Ramlow, Markus (2009): *Die maschinelle Simulierbarkeit des Humanübersetzens. Evaluation von Mensch-Maschine-Interaktionen und der Translatqualität der Technik*. Leipzig.
- Ruge, Nina (1990): *Russisch-französische maschinelle Übersetzung. Darstellung der Problematik anhand des in Grenoble (GETA) entwickelten Prototypen*. Hagen (Bochumer Slavistische Beiträge 17).
- Рябцева, Н. К. (1986): *Информационные процессы и машинный перевод. Лингвистический аспект*. Moskau.
- Schäfer, Falko (2002): *Die maschinelle Übersetzung von Wirtschaftstexten. Eine Evaluierung anhand des MÜ-Systems der EU-Kommission, SYSTRAN, im Sprachenpaar Französisch-Deutsch*. Frankfurt am Main (Europäische Hochschulschriften, Reihe XXI Linguistik 251).
- Schwanke, Martina (1991): *Maschinelle Übersetzung. Ein Überblick über Theorie und Praxis*. Berlin.
- Softissimo (2009): „White Paper: Translation Software.“ In: *Reverso*. <http://reverso.softissimo.com/en/white-paper> (Menü: *What is a translation software?*) (letzter Aufruf 10.8.2012).
- Softissimo (2012): „Einige Ratschläge, um die von Reverso Online gelieferten Übersetzungen zu verbessern.“ In: *Reverso*. <http://www.reverso.net/help.aspx?lang=DE-section6> (letzter Aufruf 10.8.2012).
- Svarog Capital (2010): „PROMT announces groundbreaking PROMT DeepHybrid MT Solution.“ In: *Svarog Capital News* <http://www.svarogcapital.com/news/?id=43> (letzter Aufruf 10.8.2012).

- Text Technologies (2006): „That great linguist, Groucho Marx, and other stories.“ In: *Text Technologies*. <http://www.texttechnologies.com/2006/06/09/that-great-linguist-groucho-marx-and-other-stories/> (letzter Aufruf 10.8.2012).
- Toral, Antonio et al. (2011): „Comparative Evaluation of Research vs. Online MT Systems.“ Original in: Forcada, Mikel L. (Hrsg.), *Proceedings of the 15th conference of the European Association for Machine Translation, 30–31 May 2011*. Leuven. 13–20. Zitiert aus: *Machine Translation Archive*. <http://www.mt-archive.info/EAMT-2011-Toral-1.pdf> (letzter Aufruf 10.8.2012).
- Wießner, Ellen (2004): „Automatic WWW-Translations: an Empirical Comparison“. In: *Technische Universität Chemnitz*. <http://www.tu-chemnitz.de/phil/english/chairs/linguist/independent/kursmaterialien/translation/Automatic%20www-translations> (letzter Aufruf 10.8.2012).
- Wikipedia (2012a): „Universal Translator: Star Trek.“ In: *Wikipedia*. http://en.wikipedia.org/wiki/Universal_translator#Star_Trek (letzter Aufruf 10.8.2012).
- Wikipedia (2012b): „Babel Fish.“ In: *Wikipedia*. http://de.wikipedia.org/wiki/Babel_Fish (letzter Aufruf 10.8.2012).
- Wikipedia (2012c): „PROMT.“ In: *Wikipedia*. <http://en.wikipedia.org/wiki/PROMT> (letzter Aufruf 10.8.2012).
- Wikipedia (2012d): „SYSTRAN.“ In: *Wikipedia*. <http://en.wikipedia.org/wiki/SYSTRAN> (letzter Aufruf 10.8.2012).
- Zervaki, Thei (2002): „Online free translation services.“ Original in: *Translating and the Computer 24: proceedings from the Aslib conference held on 21-22 November 2002*. London. Zitiert aus: *Machine Translation Archive*. <http://www.mt-archive.info/Aslib-2002-Zervaki.pdf> (letzter Aufruf 10.8.2012).
- Zimmer, Dieter E. (1997): *Die Elektrifizierung der Sprache. Über Sprechen, Schreiben, Computer, Gehirne und Geist*. München.