

Maschinelle Übersetzung - Untersuchung moderner Übersetzungssoftware

Bachelorarbeit

**vorgelegt von
Dorothea Brehm**

an der

Universität Konstanz

**Sektion Geisteswissenschaften
Fachbereich Sprachwissenschaft**

- 1. Gutachterin: Prof. Dr. Miriam Butt**
- 2. Gutachterin: Prof. Dr. Nicola Dehé**

Konstanz, Februar 2013

Dorothea Brehm
Tettnanger Str. 208
88214 Ravensburg
Matrikelnr.: 01/746043
dorothea.brehm@uni-konstanz.de

Inhaltsangabe

Die maschinelle oder auch automatische Übersetzung (MÜ) ist aus der Welt der Übersetzungen schon längst nicht mehr wegzudenken und wird in der Praxis erfolgreich eingesetzt. Die zwei attraktivsten Eigenschaften der MÜ sind der Preis und die Geschwindigkeit, beides können menschliche Übersetzer unmöglich liefern. Das größte Problem der MÜ bleibt vorerst die Qualität der Übersetzung, die noch viele Wünsche offen lässt. Aber die Fortschritte der letzten vierzig Jahre waren enorm und MÜ wird sich sicherlich weiter verbessern.

Da die drei untersuchten MÜ-Systeme hauptsächlich mit statistischen Methoden arbeiten, liegt der Fokus dieser Arbeit auf der statistischen maschinellen Übersetzung (SMÜ). Die Funktionsweise wird im Rahmen dieser Arbeit dargestellt. Die Beispiele aus der Untersuchung zeigen die Qualität der Übersetzungen.

Die Untersuchung wird die Frage beantworten, wie gut die Übersetzungen der Systeme inzwischen sind, und welches der drei Systeme das beste ist.

Inhaltsverzeichnis

	Seite
Inhalt	
1. Einführung	1
2. Die maschinelle Übersetzung	2
2.1. Regelbasierte Systeme	2
2.2. Statistische Systeme (SMÜ)	3
2.2.1. Das Übersetzungsmodell	4
2.2.2. Das Sprachenmodell	7
2.2.3. Korpora	9
2.2.4. Modellierung der Wahrscheinlichkeit	11
3. Untersuchung	13
3.1. Methode und Durchführung	13
3.2. Die Systeme	14
3.3. Ergebnis	18
3.3.1. Klassifizierung der Fehler	18
3.3.2. Gesamtüberblick Fehler	21
3.3.3. Fehler nach Themen	23
3.3.4. Verständlichkeit	31
3.4. Interpretation der Ergebnisse	34
Zusammenfassung	35
Bibliographie	36
Anhänge	37

1. Einführung

Wir leben im sogenannten digitalen Zeitalter. Nicht mehr nur die mechanische, physische Arbeit, sondern auch die geistige und intellektuelle Arbeit wird von Maschinen übernommen. Das Denken, die menschliche Intelligenz wird nachgebildet, ausgelagert, und wo möglich durch Computer ersetzt. Waren es zu Beginn die Rechenleistung, die Speicherung und Bereitstellung von Wissen, so sind es mehr und mehr auch komplexe Themen, wie unsere sprachlichen Fähigkeiten. So wie große Teile der mechanischen Arbeit wird auch geistige Arbeit wo möglich von Maschinen erledigt. Wir besitzen alle einen Taschenrechner und wahrscheinlich bald einen kleinen automatischen Übersetzer.

In der Computerlinguistik werden natürliche Sprachen computertechnisch verarbeitet. SMÜ ist ein Ansatz in der MÜ, der auf der Basis von mathematischen Wahrscheinlichkeiten, Statistik und Informatik arbeitet. Dies ist ein Gebiet in der Computerlinguistik, in dem man sogar gänzlich ohne linguistische Kenntnisse auskommen kann, wie weiter unten zu sehen sein wird.

Zuerst folgt eine kurze Einführung in die SMÜ. Die Forschung auf diesem Gebiet ist umfangreich, Koehn behauptet, dass „*about one thousand academic papers have been published on the subject, about half of them in the past three years alone.*“ (Koehn 2011:xi). So werden im Folgenden nur drei Aspekte heraus- gegriffen, das sind vereinfacht dargestellt die Funktionsweise, die Voraussetzungen, unter welchen SMÜ nur funktionieren kann und die grundlegende Mathematik für SMÜ.

Danach werden die Ergebnisse vorgestellt und diskutiert, die drei frei für jeden zugängliche on-line Übersetzungstools geliefert haben. Diese sind in alphabetischer Reihenfolge der Bing Bablefish-Translator, der Google-Translator und der Systran-Translator. Es soll die Frage beantwortet werden, welches der drei SMÜ-Systeme die besten Übersetzungen liefert und wie gut und zuverlässig diese Systeme arbeiten. Mit der Zusammenfassung endet die Arbeit.

2. Die maschinelle Übersetzung

MÜ-Systeme übersetzen Texte nicht in dem Sinn, wie das Wort „übersetzen“ herkömmlich verwendet wird. Wenn ein Herr G. Mounin die Behauptung wagt, dass „*man nicht übersetzen könne, ohne zu verstehen*“ (Georges Mounin 1967:21), so hat er damit eine andere Bedeutung für das Wort „übersetzen“ im Sinn. In diesem Sinne übersetzen MÜ-Systeme nicht, sondern sie suchen, vergleichen und ersetzen. Das Ergebnis ihrer Suche und Ersetzung erscheint dann für die Nutzer als Übersetzung, weswegen das Wort „übersetzen“ in diesem Text weiter verwendet wird.

2.1. Regelbasierte Systeme

Prinzipiell kann man in der MÜ zwei Systeme unterscheiden, die klassischen regelbasierten Systeme und die neueren statistischen Systeme. Linguisten beschäftigen sich bevorzugt mit regelbasierten Systemen und verarbeiten Sprache nach linguistischen Regeln, so selbst zu erfahren an der Universität Konstanz, wo ein regelbasiertes System existiert und kontinuierlich weiter aufgebaut wird. Um die Arbeitsweise kurz zu skizzieren, kann man sagen, dass zuerst die Wörter und die grammatischen Strukturen des Quelltexts auf Satzbasis analysiert werden. Die Syntax wird in einem Baumdiagramm dargestellt, die Wörter werden klassifiziert und in einem zweisprachigen Lexikon abgespeichert. Nach der Analyse werden die Syntax und die Wörter in eine Zielsprachige Struktur überführt. Von dort aus kann der Zielsatz neu generiert werden.

Es ist eine wahre Sisyphosarbeit, die in natürlicher Sprache auf analoge, oft vage, ambige und versteckte Weise verschlüsselten Informationen digital exakt aufzuschlüsseln und in Algorithmen zu beschreiben. Der Programmieraufwand ist enorm, und die ungezählten Regeln der Grammatik mit allen Ausnahmen und Besonderheiten sprengen leicht jeden Zeitrahmen und jedes Budget. Stein schreibt in diesem Zusammenhang, dass in der Praxis die steigende Komplexität der Übersetzungssysteme die Übersetzungsqualität auch nicht mehr weiter verbessert, sondern dass „*interne Konflikte und sich widersprechende Regeln neue Fehler produzieren*“ (Stein 2009:8). Deswegen befinden sich die rein

regelbasierten Systeme eher im wissenschaftlichen Bereich.

2.2. Statistische Systeme (SMÜ)

Informatiker betrachten Sprachen aus einem für Linguisten ungewohnten Blickwinkel, dem rechnerischen. Sie sind der Auffassung, dass es prinzipiell unmöglich ist, alle Regeln für eine natürliche Sprache zu erfassen und zu verarbeiten. Sie haben an Morphologie, Syntax und Semantik schon allein deswegen nicht so viel Interesse, weil sie sich in diesen Gefilden nicht auskennen. Dafür kennen sie die Rechen- und Speicherleistung ihrer Computer, die beständig an Geschwindigkeit und Zuverlässigkeit zunimmt und gleichzeitig immer erschwinglicher wird, und sie wissen, wie sie auszunutzen ist. So gehen sie den Weg der puren Rechengewalt und berechnen Wahrscheinlichkeiten für Übersetzungen sprachenunabhängig, das bedeutet, sie müssen die Fremdsprache, mit der sie arbeiten, erstaunlicherweise nicht zwingend können. „*Vielmehr werden alle benötigten Informationen wie Worthäufigkeit, Grammatikregeln usw. mit Methoden zur Informationsgewinnung aus großen bi- und multilingualen Textkorpora für die beteiligten Sprachen extrahiert*“ (Diehl 2007:16). Informatiker lassen ihre Programme die Grammatikregeln selbst lernen. Die Systeme werden trainiert, das heißt, sie extrahieren aus bilingualen Korpora die Übersetzungsregeln und deren Wahrscheinlichkeiten automatisch. Als Ziel der SMÜ nennt Och die bestmögliche Entscheidung aufgrund von niemals vollständigem Wissen (Och 2004:1), denn theoretisch gesehen werden nie genug Korpora zur Verfügung stehen.

Für Linguisten ist es schwer fassbar, dass „*most systems are largely language-independent, and building a SMT system for a new language pair is mostly a matter of availability of parallel texts.*“ (Koehn 2005:4).

In der Praxis vermischen sich die rein regelbasierten mit den rein statistischen Ansätzen. Die Systeme sind meist hybrid, stochastische Methoden werden im regelbasierten Ansatz eingesetzt und linguistisches Basiswissen in der SMÜ.

Es ist anzunehmen, dass die hier untersuchten SMÜ-Systeme mit einer flachen Regelstruktur für die Syntax ausgestattet sind. Die Untersuchung hat gezeigt, dass die Syntax der Ausgabesätze ziemlich robust ist.

2.2.1. Das Übersetzungsmodell

Entsprechend dem Übersetzungsmodell wird aufgrund von parallel vorliegenden bilingualen Texten die Wahrscheinlichkeit berechnet, mit der Wörter, Phrasen oder Sätze zweier Sprachen Übersetzungen voneinander sind (Somers 2011:435).

Wahrscheinlichkeiten können für Wörter, Phrasen oder ganze Sätze berechnet werden.

Am Beispiel des italienischen weiblichen Artikels „la“ soll das Übersetzungsmodell für ein einzelnes Wort veranschaulicht werden. In der Bablefish-Tabelle (siehe Anhang) kommt der italienische Artikel „la“ 26 mal in den italienischen Quellsätzen vor. Für seine Übersetzung ins Deutsche finden sich sechs verschiedene Wörter im Deutschen. Nach dem Übersetzungsmodell kann die Wahrscheinlichkeit für jede Übersetzung leicht selbst berechnet werden, indem man die Zahl der einzelnen Möglichkeiten durch die Gesamtzahl aller Möglichkeiten teilt. Wahrscheinlichkeiten erhalten stets einen Wert zwischen 0 und 1.

(Figur 1)

QS	ZS	Wahrscheinlichkeit
la →	1 x „dem“	(1:26=) 0.04
la →	1 x „sein“	(1:26=) 0.04
la →	5 x „das“	(5:26=) 0.19
la →	14 x „die“	(14:26=) 0.54
la →	3 x „der“	(3:26=) 0.11
la →	2 x NULL	(2:26=) 0.08

In Figur 1 ist leicht abzulesen, dass die Übersetzung „die“ die wahrscheinlichste ist, in über 50% der Fälle wurde „la“ durch „die“ ersetzt. Der Wert 0.54 für die Wahrscheinlichkeit ist ein relativer Wert, der sich nur auf die 100 eingegebenen italienisch-deutschen Beispielsätze bezieht.

Ein Grund für den hohen Wert für „die“ ist möglicherweise, dass deutsche und italienische Nomen haben oft das gleiche Geschlecht haben, was an der nahen Verwandtschaft dieser Sprachen liegt. Nach dieser Wahrscheinlichkeitsberechnung müsste ein System die Übersetzung „die“ auswählen, was wiederum nicht in jedem Fall passt. In der Bablefish-Tabelle findet sich jedoch kein Beispiel für eine falsche „die“- „la“-Übersetzung, alle Übersetzungen sind richtig. Dafür gibt es hauptsächlich zwei Gründe. Der erste ist, dass das System von Bablefish nicht wortbasiert, sondern phrasenbasiert übersetzt. Der zweite Grund ist das Sprachenmodell, das weiter unten in 2.2.2. dargestellt ist.

Die Wahrscheinlichkeit wurde hier auf Wortbasis berechnet. Wortbasierte Modelle haben den großen Nachteil, dass Sprachen sich eben nicht Wort für Wort entsprechen, wie das folgende Beispiel aus der Google-Tabelle (im Anhang) zeigt. (QS) ist in allen Beispielen der Quellsatz, (RS) der Referenzsatz und (ZS) der Zielsatz, der von einem der drei Systeme ausgegeben wurde.

- | | |
|---|-----------------------------|
| (1) a. <u>Il disavanzo del conto corrente</u> si è ridotto. | (QS Nr. 127) |
| b. Das Leistungsbilanzdefizit hat sich verringert. | (RS Nr. 127) |
| c. <u>Das Leistungsbilanzdefizit</u> hat sich verringert. | (ZS Nr. 127 Google-Tabelle) |

Google hat die vier Worte „disavanzo del conto corrente“ (1a) richtig durch ein Wort „Leistungsbilanzdefizit“ (1c) ersetzt, was darauf hindeutet, dass das System die vier Worte des QS nicht einzeln, sondern als Phrase behandelt hat.

Ein weiterer Nachteil wortbasierter Wahrscheinlichkeitsberechnung ist, dass umliegende Kontextinformation für das Wort nicht berücksichtigt werden, wo doch die Übersetzung eines Wortes stark davon abhängen kann (Zens R. 2002:22).

- | | |
|--------------------------------------|--|
| (2) a. Il problema <u>e</u> risolto. | (QS Nr. 162, Wort „e“ falsch geschrieben,
ohne Akzent!) |
| b. Das Problem ist gelöst. | (RS Nr. 162) |
| c. Das Problem <u>wird</u> gelöst. | (ZS Nr. 162 Google-Tabelle) |

In Beispiel (2c) wurde das Wort „e“ richtig übersetzt, obwohl es falsch geschrieben ist und das Wort ohne Akzent „und“ heißt. Es ist hier anzunehmen, dass das System die

umliegende Kontextinformation mit eingerechnet hat, es hätte sonst das Wort „e“ durch „und“ ersetzt.

Mehr Erfolg hat ein System mit der Berechnung der Wahrscheinlichkeit für Phrasen, die aus mehreren zusammengehörenden Wörtern bestehen, sogenannten „n-grams“ (vgl. Koehn 2011:127).

Phrasenbasierte SMÜ-Systeme arbeiten besser. Intuitiv möchte man annehmen, dass die Phrasen in der MÜ den syntaktischen Phrasen entsprechen, was aber nicht immer der Fall ist (Koehn 2011). Möglich sind auch Wortsequenzen, die syntaktisch nicht miteinander verbunden sind. Eine syntaktische Phrase wird in der MÜ als eine Wortsequenz definiert, die in einem Baumdiagramm genau durch einen Ast dargestellt ist (Koehn et al. 2003:50). Das sind gewöhnlicherweise zwei bis drei Wörter, es können auch mehr sein.

Wenn Phrasen eines Satzes getrennt übersetzt werden, dann stehen sie im ZS zunächst in der gleichen Reihenfolge wie im QS. Deswegen ist der nächste Schritt nach der Übersetzung der Phrasen die Neuordnung im ZS nach dem Sprachenmodell, das weiter unten beschrieben wird. Syntaktisch motivierte Phrasen lassen sich einfacher verschieben als nicht syntaktisch motivierte Phrasen.

- | | | | | | | |
|-----|----|----------------------------|---------|------------|--------------------|-----------------------------|
| (3) | a. | Entgegen aller Erwartungen | verlief | die Arbeit | erfolgreich. | (QS Nr. 120) |
| | | 1 | 2 | 3 | 2 | |
| | b. | Contro ogni aspettativa | | il lavoro | ha avuto successo. | (RS Nr. 120) |
| | c. | Contro ogni aspettativa, | | il lavoro | ha avuto successo. | (ZS Nr. 120 Google-Tabelle) |
| | | 1 | | 3 | 2 | |

Der Satz (3a) wurde in die Phrasen 1, 2 und 3 eingeteilt, die Neuordnung der Phrasen in (3c) war erfolgreich.

Wahrscheinlichkeiten für ganze Sätze zu berechnen ist ebenso möglich, aber nicht mehr üblich, denn die Wahrscheinlichkeit, in den Korpora ganze Sätze zu finden, die sich völlig entsprechen, ist sehr gering. In bestimmten Fällen ist es aber sinnvoll, zum Beispiel bei stehenden Redewendungen.

- (4) a. Ich werde ihm das sicherlich mitteilen. (QS Nr. 60)
 b. Gli comunicherò sicuramente questa cosa. (RS Nr. 60)
 c. Non mancherò di dirglielo. (ZS Nr. 60 Google-Tabelle)
 d. Nicht ich werde versäumen es ihm zu sagen. (Wörtliche Übersetzung von (4c))

Die Übersetzung des Satzes (4a) in (4c) ist satzweise geschehen, denn die einzelnen Wörter oder Phrasen in (4c) sind in (4a) nicht zu finden. Satz (4c) ist ein vielbenutzer Standardsatz im Italienischen und wurde vermutlich als solcher abgespeichert.

2.2.2. Das Sprachenmodell

Ein zweiter Baustein in der SMÜ ist das Sprachenmodell, das garantieren soll, dass die übersetzten Phrasen richtig geordnet sind, der Satz flüssig ist, und die Wortstellung in der Zielsprache stimmt. Die durch das Übersetzungsmodell errechnete und gefundene Übersetzung soll in der Zielsprache so gebräuchlich wie möglich sein. Aus diesem Grund erhalten Sätze, die in der Zielsprache oft verwendet werden, einen höheren Wahrscheinlichkeitswert als die weniger üblichen Sätze.

Die Übersetzung passiert also mindestens zwei Filter, das Übersetzungsmodell und das Sprachenmodell. Nach Koehn kann man dies mit der folgenden Formel darstellen (Koehn 2011:181):

$$\text{(Formel 1)} \quad P_{LM}(\text{Der Satz ist sehr einfach.}) > P_{LM}(\text{Die Satz ist sehr einfach.})$$

Das heißt, dass nach dem Sprachenmodell die Wahrscheinlichkeit für den Satz “Der Satz ist sehr einfach” höher ist als die Wahrscheinlichkeit für den Satz “Die Satz ist sehr einfach.”

Wie in 2.2.1 in Figur 1 dargestellt müsste das System entsprechend der errechneten Wahrscheinlichkeiten für das Wort “la” das Wort “die” wählen. Nach dem Sprachenmodell laut Formel 1 würde danach “die” durch “der” ausgetauscht.

Die Wortfolge in einem Satz ist ebenso berechnenbar, ein Wort folgt auf das nächste Wort mit einer bestimmten berechnbaren Wahrscheinlichkeit. Diese hängt mindestens vom

vorhergehenden Wort ab, des öfteren von mehreren vorangestellten Wörtern. Anaphora beispielsweise können durchaus eine lange Geschichte haben und sind deshalb schwierig für die SMÜ. Die Annahme von Markow, ein russischer Mathematiker, ist nun, dass nur eine bestimmte Anzahl von vorangegangenen Wörtern einen Einfluss auf die Wahrscheinlichkeit des nächsten Wortes hat (Koehn 2011:183). Tatsächlich ist es in der Praxis unmöglich, die Geschichte vorangegangener Wörter sehr weit zurückzuverfolgen. Mit der sogenannten Markow-Kette wird die Wahrscheinlichkeit eines Wortes mit dem Auftreten der vorangegangenen Worte in Bezug gesetzt.

Um das obige Beispiel für das Wort “la” und “die” (Tabelle 1) fortzuführen und die Markow-Kette vorzustellen, wird die folgende Rechnung nicht für ein einzelnes Wort, sondern für eine Wortkategorie aufgestellt. Der hier zu Verfügung stehende Korpus ist so klein, dass ein deutsches Folgewort auf den Artikel “die” maximal zweimal vorkommt, die meisten Wörter folgen nur einmal.

Auf das deutsche Wort “die”, das sowohl Artikel als auch Relativpronomen sein kann und in den Sätzen der Bablefish-Tabelle (im Anhang) insgesamt 112 mal vorkommt, folgen 8 verschiedene Wortkategorien mit unterschiedlicher Häufigkeit.

In Tabelle 2 sind die möglichen Folgewörter nach “die” aus der Bablefish-Tabelle (im Anhang) aufgeführt:

(Figur 2)

Def.Artikel od. Rel.Pron.		Folgewort	Anzahl	Wahrscheinlichkeitswert
“die”	+	Nomen fem.Sg.	59	0.53
“die”	+	Nomen fem.Pl., mask. Pl., neutr. Sg., neutr. Pl.,	21	0.19
“die”	+	Poss.Pron. fem. Pl.	1	0.01
“die”	+	Adjektiv fem. Sg., fem. Pl. neutr. Pl.	19	0.17
“die”	+	Verb	2	0.02
“die”	+	Adverb	3	0.03
“die”	+	Präposition	2	0.02
“die”	+	Zahl	1	0.01
“die”	+	Def.Art.	4	0.035

In der Bablefish-Tabelle folgt auf das Wort “die” mit über 50% Wahrscheinlichkeit erwartungsgemäß ein weibliches Nomen im Singular. An der Tabelle ist ausserdem abzulesen, dass nach dem Wort “die” immer ein Wort folgt, das Wort “die” infolgedessen nie am Satzende steht. Die bedingte Wahrscheinlichkeit, dass ein fem. Nomen im Sg. nach dem Wort “die” folgt, lässt sich jetzt mit der Markow-Kette wie folgt berechnen:

$$\begin{aligned} \text{(Formel 2)} \quad P(\text{“die”, “Nomen f.Sg.”}) &= p(\text{“die”}) p(\text{“Nomen f.Sg.”} \mid \text{“die”}) \\ &= 0.54 \times 0.53 = 0.29 \end{aligned}$$

In (Figur 1) wurde der Wert 0.54 für die Wahrscheinlichkeit errechnet, dass “die” eine Übersetzung für “la” ist. In (Figur 2) wurde der Wert 0.53 für die Wahrscheinlichkeit errechnet, dass nach dem Artikel “die” ein weibliches Nomen im Singular folgt. Die Markow-Kette liefert den Wert 0.29 für die bedingte Wahrscheinlichkeit, dass zwei Ereignisse hintereinander eintreten: Erstens wurde “la” durch “die” übersetzt und zweitens folgt im ZS ein weibliches Nomen auf das Wort “die”.

Auf die gleiche Weise können auch Wahrscheinlichkeiten für längere Wortketten berechnet werden.

2.2.3. Korpora

Um funktionieren zu können, brauchen SMÜ-Systeme riesige Mengen an maschinenlesbaren, aufbereiteten und verwertbaren Daten. Was für die regelbasierten Systeme das programmierte Regelwerk ist, das müssen die Korpora für die SMÜ sein.

Basis und wichtigster Datenlieferant der SMÜ ist das Internet mit seinen digital gespeicherten Textdaten. Je umfangreicher diese Basis ist, um so besser wird das Ergebnis. Dieses Phänomen wird in der Mathematik als das Gesetz der großen Zahlen beschrieben.

Die Datenmengen im Netz liegen ungeordnet und nicht immer im passenden Format vor, und sie sind längst nicht alle einfach zugänglich und auswertbar, auch sind die

Übersetzungen nicht alle zuverlässig richtig, es sind einfach nur Rohdaten, die wie Rohstoffe gefunden und bearbeitet werden müssen. Bestimmte Themengebiete, zum Beispiel Informatik und die europäischen Sprachen, allen voran Englisch, sind ausreichend abgedeckt, aber für seltenere Sprachen und besondere Gebiete ist es schwierig bis unmöglich, ausreichend Korpora zu finden.

Für SMÜ-Systeme sind die erste Wahl bilinguale, parallele und alinierte Texte, das heißt Texte, die aus einem Quelltext und seiner Übersetzung bestehen, und parallel geordnet vorliegen. Nach Zinsmeister heißen sie Parallelkorpora (Zinsmeister 2010:6). Ein Beispiel für ein Parallelkorpus ist diese Untersuchung mit den alinierten deutsch-italienischen und italienisch-deutschen Sätzen. Aber auch dieser Korpus ist nicht uneingeschränkt verwertbar, denn darin befinden sich viele vom System falsch übersetzte Sätze. Manuell oder automatisch aufgebaute Übersetzungsspeicher (translation memories) sind ebenfalls begehrte Parallelkorpora.

Alinierung bedeutet, dass Wörter (Wortalinierung), Phrasen (Phrasenalinierung) oder ganze Sätze (Satzalinierung) paralleler bilingualer Korpora direkt aufeinander bezogen sind. Zur Alinierung bilingualer Texte werden Programme geschrieben, die diese Arbeit automatisch und selbständig tun, von Hand dauert es lange. Gale und Church haben zum Beispiel ein Programm zur Alinierung von Sätzen und Phrasen geschrieben, das die einfache Tatsache ausnutzt, dass langen Sätzen in der QS auch lange Sätze in der ZS entsprechen, die Satzlänge wird in Buchstaben gemessen (Gale u. Church 1991:5).

Auch Munteanu und Marcu sprechen von einem „*lack of parallel data*“ (Munteanu und Marcu 2006:1), weshalb sie und viele andere Forscher nach Möglichkeiten suchen, ungeordnete monolinguale Texte automatisch zu finden, zu parallelisieren und zu bi- oder multilingualen Daten zu ordnen.

Die SMÜ-Systeme sollen anhand dieser Korpora trainiert werden. Sie lernen die Übersetzungs- und Sprachregeln automatisch, indem sie aus dem Datenmaterial abgeleitet werden, dann berechnen sie die Wahrscheinlichkeiten aufgrund der extrahierten Regeln.

2.2.4. Modellierung der Wahrscheinlichkeit

Zur Modellierung und Berechnung der Wahrscheinlichkeit für eine Übersetzung finden sich in der Literatur hauptsächlich die folgenden Formeln. Die erste Formel ist das Theorem von Bayes, es wird nach Brown vorgestellt (Brown et al. 1993:2).

$$(Formel 3) \quad \Pr(e|f) = \frac{\Pr(e) \Pr(f|e)}{\Pr(f)}$$

„Pr“ steht für Wahrscheinlichkeit, „e“ für Ausgangssprache und „f“ für Zielsprache. Die Formel ist wie folgt zu interpretieren: Jeder Satz „f“ ist eine mögliche Übersetzung des ausgangssprachlichen Satzes „e“. Jedem Satzpaar (e,f) wird eine Wahrscheinlichkeit $\Pr(e|f)$ zugeordnet, das ist die Wahrscheinlichkeit, mit der ein Übersetzer den Satz „f“ wählen würde, wenn er Satz „e“ vor sich hätte. Jeden Satz „f“ kann man als Übersetzung eines Satzes „e“ ansehen. Das MÜ-System muss somit den Satz „e“ finden, der mit der höchsten Wahrscheinlichkeit die Übersetzung von „f“ ist. Da in der Gleichung der Nenner von „e“ unabhängig ist, kann die Gleichung wie folgt umgeschrieben werden:

$$(Formel 4) \quad \hat{e} = \operatorname{argmax}_e \Pr(e) \Pr(f|e)$$

Dies ist eine fundamentale Gleichung für die SMÜ (Brown 1993:3). „Argmax“ bedeutet, dass unter den gegebenen Umständen, also anhand der vorhandenen Data, die wahrscheinlichste Übersetzung gefunden werden kann, heißt aber nicht, dass es tatsächlich die beste ist, sie wird nur mit einer relativen Wahrscheinlichkeit eine akzeptable Qualität liefern. Es ist nicht realistisch, dass alle Möglichkeiten durchsucht werden, denn niemals werden ausreichend Daten zur Verfügung stehen.

Als weiterer grundsätzlicher Ansatz sei die Formel von Jurafsky (2000:819) zur Modellierung genannt:

(Formel 5) $\text{best-translation } T = \text{argmax}_T \text{fluency}(T) \times \text{faithfulness}(T,S)$

T sei der Zielsatz und S der Quellsatz. Nach Jurafsky ist die beste Übersetzung dann gefunden, wenn sowohl der Textfluss als auch Zuverlässigkeit gegeben sind. Der erste Faktor „ $\text{argmax}_T \text{fluency}(T)$ “ ist nach dem Sprachenmodell berechnet und stellt die Wahrscheinlichkeit dar, mit der die Übersetzung tatsächlich ein wahrscheinlicher Satz in der Zielsprache ist. Der zweite Faktor „ $\text{faithfulness}(T,S)$ “ stellt die Zuverlässigkeit der Übersetzung dar und wird nach dem Übersetzungsmodell berechnet.

Die skizzierten Wahrscheinlichkeitsmodelle werden in der Informatik nicht nur für MÜ, sondern ganz allgemein für Mustererkennung unterschiedlichster Wirklichkeiten eingesetzt, wie zum Beispiel auch für Wirtschaftsprognosen oder die visuelle Mustererkennung. Die Textkorpora werden schlicht als Muster behandelt, dessen zugrundeliegendes Muster erkannt und auf neue, auch unbekannte Muster übertragen werden soll.

3. Untersuchung

Das Ziel der Untersuchung ist die Beantwortung der Frage, wie weit die SMÜ inzwischen fortgeschritten ist. Da hier kostenlose und jedem zugängliche Systeme untersucht werden, kann man davon ausgehen, dass sie das Minimum an Leistung anbieten. Sie befinden sich am unteren Ende der Leistungsskala. Alle professionellen Systeme sollten diese Minimalleistung überbieten können.

Der Grund für die Untersuchung ist der Versuch zu zeigen, dass diese minimale Leistung schon sehr beachtlich ist und durchaus bereits genutzt werden kann.

An dieser Stelle sei angemerkt, dass die Übersetzungsleistung der Systeme zwar kostenlos ist, aber möglicherweise nicht ganz uneigennützig. Die Eingabetexte werden urheberrechtlich gesehen zum Eigentum der Anbieter, die die Daten nach ihrem Gutdünken weiter verarbeiten können.

3.1. Methode und Durchführung

Drei bekannte und jedem Internetnutzer einfach zugängliche on-line Übersetzungssysteme, die anschließend kurz vorgestellt werden, übersetzten 200 Sätze, 100 Sätze davon vom Deutschen ins Italienische, die andere Hälfte umgekehrt. Die Sprachen sind beide Mitglieder der indoeuropäischen Sprachfamilie und ihre Sprecher leben im gleichen europäischen Kulturraum, die Sprachen somit liegen nahe beieinander.

Die Anzahl der Sätze wurde durch den Zeitrahmen dieser Arbeit eingeschränkt, mehr Sätze wären hilfreich gewesen. In der Anfangsphase der Untersuchung wurden viele Sätze getestet, die dann nicht in die Bewertung eingingen, weil sie zu schwierig waren, sie brachten insgesamt keine guten Ergebnisse, die Fehleranzahl war zu hoch, um manuell alle Fehler festzustellen und zu klassifizieren. Insofern sind die Sätze nicht beliebig, sondern der Leistungsfähigkeit der Systeme angepasst. Es wurden zunächst einfache und allgemeinsprachliche Sätze ausgewählt, um die Systeme nicht zu überfordern und gute Ergebnisse zu erhalten.

Interrogativsätze und Imperativsätze wurden ausgewählt, um zu sehen, ob die Systeme mit Satz inversionen umgehen können. Fachterminologie aus Wirtschaft und Recht wurde ausgewählt, um zu sehen, wie umfangreich das Vokabular der Systeme ist. Die Sätze mit falsch geschriebenen Wörtern bilden die Realität ab, nämlich dass viele Nutzer die Fremdsprache nicht richtig schreiben können.

Die langen Sätze am Ende der Untersuchung zeigen die Grenzen der Systeme auf.

Es wurde manuell evaluiert, was viel Zeit in Anspruch nimmt, dafür sehr genau ist. Die Qualität der ausgegebenen Übersetzungen wurde anhand von drei Kriterien bewertet. Das erste Kriterium sind die Fehler in den Zielsätzen (ZS), die objektiv gezählt und in 6 Fehlerklassen unterteilt wurden. Das zweite Kriterium sind die Fehler nach Themen. Drittens wurde die Verständlichkeit bewertet, die Bewertung ist subjektiv und beruht auf der menschlichen Einschätzung einer professionellen Übersetzerin. Aus Platzgründen wurden weitere Kriterien hier nicht mehr untersucht, wie zum Beispiel der Textfluss („fluency“), der in der Formel unter Punkt 2.2.4. genannt wird.

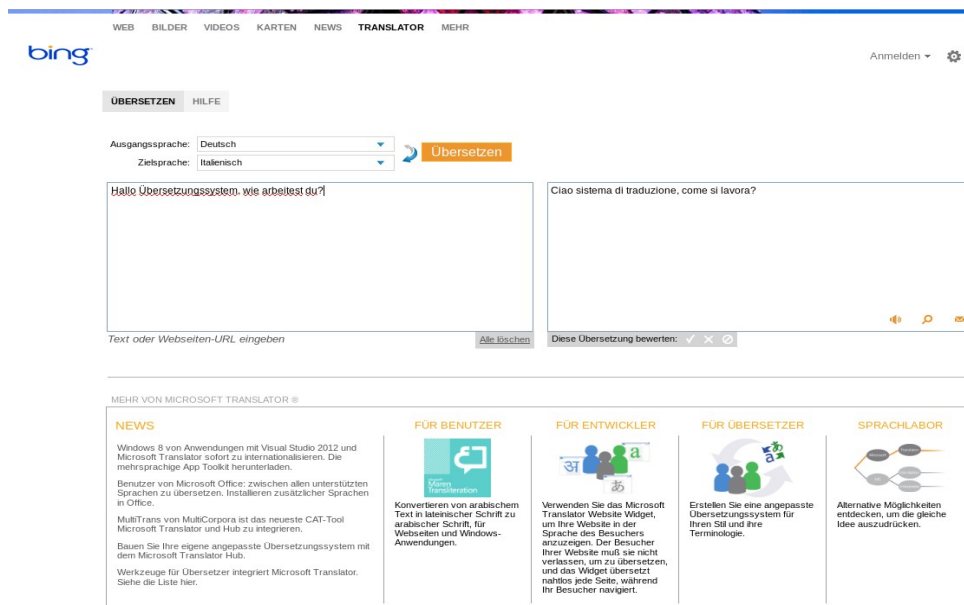
Alle drei Kriterien werden dann zusammen betrachtet, um zu beurteilen, wie gut die Qualität der Übersetzungen insgesamt ist und welches der drei Tools die beste Qualität geliefert hat.

3.2. Die Systeme

Die drei im Folgenden getesteten Systeme sind kostenfrei im Internet verfügbar, es sind der Google-Translator, der Systran-Translator und Bing-Translator. Alle drei sind statistische Systeme (SMÜ), die auf riesige Datenmengen und Übersetzungsspeicher (TMs) im Netz zurückgreifen können. Sie stehen allen Anfragen des breiten Netz-Publikums zur Verfügung und bedienen alle bestmöglichst.

3.2.1. Bing-Translator (<http://www.bing.com/translator>)

Bild Nr. 1



Das System wird von der Microsoft© Corporation, Microsoft Translator bereit gestellt. Es unterstützt derzeit 39 Quellsprachen und 39 Zielsprachen. Mit der Funktion „Auto-Erkennen“ kann der QS in einer beliebigen Sprache eingegeben werden und das System versucht, die Sprache selbst zu erkennen. Die Nutzer können den ausgegebenen ZS anhören und mit „gut“, „falsch“ und „unangemessen“ bewerten.

Eigene Verbesserungsvorschläge können nicht direkt eingegeben werden, man muss dafür auf den Feedback-Knopf klicken, worauf sich ein Fenster für die Mitteilung öffnet.

Wie bei allen drei Systemen ist das Erscheinungsbild insgesamt ansprechend, die Benutzung ist selbsterklärend, Werbung gibt es nicht.

3.2.2. Der Google-Translator <http://translate.google.de/?hl=de#cy/it/>

Bild Nr. 2



Der Google-Translator wird vom 1998 gegründeten Unternehmen Google aus USA zur Verfügung gestellt.

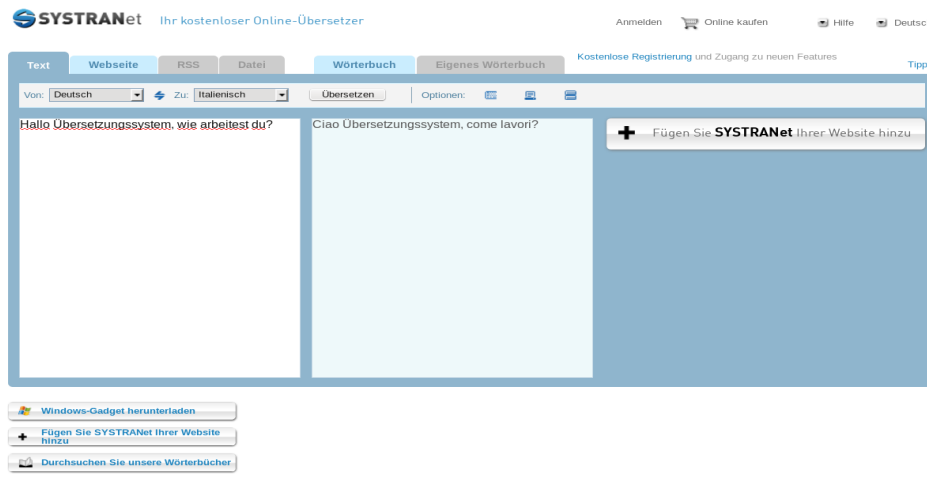
Laut eigenen Angaben unterstützt der Google-Translator derzeit Übersetzungen zwischen 64 Sprachen, zusätzlich gibt es die Funktion „Sprache erkennen“, mit der die Eingabesprache vom System erkannt wird. Die Übersetzung erfolgt während der Eingabe oder sofort danach. Der Google-Translator bietet die „Meinten-Sie“-Funktion, die sich automatisch bei einer Falscheingabe öffnet.

Das ist ein echtes Plus, weil Rechtschreibfehler automatisch verbessert werden. Die Nutzer können sich die Übersetzung vorlesen lassen und bewerten, die Möglichkeiten sind „nützlich“, „nicht nützlich“ und „anstößig“. Klickt man im ZS auf einzelne Worte oder Wortgruppen, werden dafür alternative Vorschläge angezeigt und man kann selbst Alternativen vorschlagen.

Stein ist der Meinung, dass bei Google ein rein statistisches MÜ-System arbeitet (Stein 2009), dasselbe schreibt auch Google auf seiner Webseite.

3.2.3. Systranet (<http://www.systranet.com/de/translate/>)

Bild Nr. 3



Dieses Tool wird vom Unternehmen Systran angeboten, das sich selbst als den „weltweit führenden Hersteller von Übersetzungssoftware“¹ bezeichnet.

In der Literatur wird Systran oft genannt, wie zum Beispiel in der Cambridge Enzyklopädie der Sprache, wo Crystal (2010:350) Systran als das bekannteste maschinelle Übersetzungssystem zitiert, möglicherweise auch das älteste.

Der Systran-Translator bietet 15 Quellsprachen und 5 Zielsprachen. Die Funktion „Auto-Erkennen“ hat er nicht. Wie alle drei Systeme hat er einen Knopf, um die Übersetzungsrichtung umzukehren. Die Übersetzung kann nicht bewertet oder verändert werden, dafür hat nur dieses Tool die Funktion „Drucken“ eingebaut.

Systran bietet auf seiner Seite PC-Versionen zum Kauf an, die dann off-line genutzt und den persönlichen Bedürfnissen angepasst werden können.

¹<http://www.systran.de/systran/unternehmensprofil>

3.3. Ergebnis

In Bezug auf Geschwindigkeit und Zuverlässigkeit unterscheiden sich die drei Systeme nicht. Der Ausgabertext kommt immer zuverlässig in Bruchteilen von Sekunden, kein einziges Mal wurde „error“ oder gar kein Text ausgegeben, selbst für die langen, schwierigen Sätze kam eine Übersetzung, wenn auch eine falsche.

3.3.1. Klassifizierung der Fehler

1) Die Wortstellung (WST) im Satz ist falsch, was bedeuten kann, dass die Alinierung im ZS nicht gelang.

- | | | |
|--------|--|-----------------------------|
| (5) a. | Il committente ha diritto di verificare l'opera compiuta. | (QS Nr. 145) |
| b. | Der Kunde hat das Recht, die fertige Arbeit zu überprüfen. | (RS Nr. 145) |
| c. | Der Kunde hat das Recht, <u>zu überprüfen die Arbeit getan</u> . | (ZS Nr. 145 Google-Tabelle) |

In diesem Fall (5c) fand anscheinend keine Alinierung statt, die Wortfolge des QS (5a) ist nicht verändert worden, obwohl sie im ZS nicht mehr passt.

2.1) Ein Wort fehlt im ZS, das im QS vorhanden war.

- | | | |
|--------|-------------------------------|------------------------------|
| (6) a. | La strada <u>non</u> è lunga. | (QS Nr. 2) |
| b. | Der Weg ist nicht lang. | (RS Nr. 2) |
| c. | Der Weg ____ ist lang. | (ZS Nr. 2 Bablefish-Tabelle) |

Im Beispiel (6c) fehlt aus unerklärlichem Grund die Verneinung, was den Sinn des Satzes in sein Gegenteil verkehrt.

Hier sei ein Experiment eingeschoben, um zu zeigen, dass der Grund hierfür an der WST im QS liegen könnte. Gibt man den QS (7a) mit der zielsprachigen WST ein, so erscheint im ZS (7c) erwartungsgemäß das Wort „nicht“:

- | | | |
|--------|--------------------------------|------------------------------|
| (7) a. | La strada è <u>non</u> lunga. | (QS Nr. 2) |
| b. | Der Weg ist nicht lang. | (RS Nr. 2) |
| c. | Der Weg <u>ist nicht</u> lang. | (ZS Nr. 2 Bablefish-Tabelle) |

2.2) Ein Wort fehlt im ZS, das zwar im QS nicht vorhanden war, im ZS aber aus syntaktischen Gründen erforderlich ist.

- (8) a. Ich habe meinem Freund ein Geheimnis anvertraut. (QS Nr. 51)
 b. Ho confidato un segreto al mio amico. (RS Nr. 51)
 c. Ho fidato il mio amico un segreto. (Präp. „a“ fehlt) (ZS Nr. 51 Systran-Tabelle)

Im Fall (8c) entstand der Fehler möglicherweise dadurch, dass die Phrase des indirekten Objekts (8a) der QS keine Präposition hat.

3) Ein Wort aus dem QS ist im ZS vorhanden, aber nicht übersetzt.

- (9) a. Sotto il monte si estendeva il vasto giardino. (QS Nr. 108)
 b. Unter dem Berg erstreckte sich der riesige Garten. (RS Nr. 108)
 c. Sotto il Monte erstreckte sich im riesigen Garten. (ZS Nr. 108 Bablefish-Tabelle)

Im Satz (9c) wurden die drei Wörter einer Präpositionalphrase nicht übersetzt, woraus man schließen könnte, dass die Phrase als ein Ganzes behandelt wurde.

4) Ein Wort aus dem QS ist im ZS vorhanden, aber die morphologische Form des Wortes ist falsch.

- (10)a. Du willst nicht verstehen. (2.Sg.) (QS Nr. 19)
 b. Non vuoi capire. (RS Nr. 19)
 c. Non volete capire. (2.Pl.) (ZS Nr. 19 Bablefish-Tabelle)

5) Ein Wort aus dem QS ist im ZS vorhanden, aber die Bedeutung des Wortes ist falsch.

- (11)a. L'amico conosce tutta la storia. (Def. Art.) (QS Nr.28)
 b. Der Freund kennt die ganze Geschichte. (RS Nr.28)
 c. Mein Freund kennt die ganze Geschichte. (Poss.Pron.) (ZS Nr.28 Bablefish-Tabelle)

Fall 4) und Fall 5) können auch gemeinsam in einem Wort auftreten, wofür dann 2 Fehler gerechnet werden.

6) Im ZS ist ein Wort zu viel vorhanden, das nicht in den Satz gehört.

- (12)a. Robert will die neue Küche nicht. (QS Nr. 34)
 b. Robert non vuole la nuova cucina. (RS Nr. 34)
 c. Robert non vuole che la nuova cucina. (ZS Nr. 34 Bablefish -Tabelle)

Auch für das Beispiel (12) kann experimentell der QS gefunden werden, der den richtigen ZS ausgeben würde.

- (13)a. Robert nicht will neue Küche. (Eingabe)
 b. Robert non vuole la nuova cucina. (Ausgabe)

Welche Mechanismen hier wirken, ist nur zu vermuten. Offensichtlich erscheint das „la“ nicht als Übersetzung von „die“, sondern hängt an der Nominalphrase. Dafür verschwindet das Relativpronomen „che“ bei fehlendem „die“ .

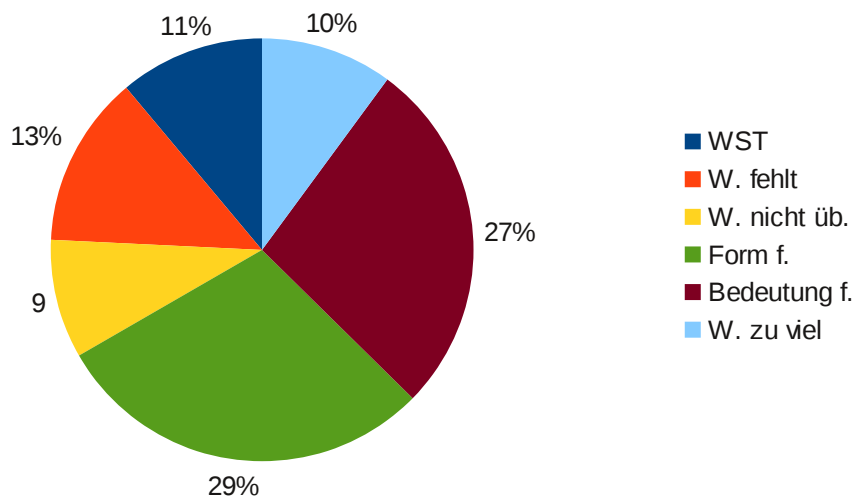
Schreibfehler und Kommafehler wurden nicht mitgezählt.

3.3. Gesamtüberblick Fehler

Tabelle Nr. 1

	BING-Translator	Google-Translator	Systran-Translator	Gesamt
Eingerechnete Sätze insgesamt	180	180	180	= 540
Komplett richtig übersetzte Sätze:	77/180 43%	87/180 48%	55/180 30%	= 219 40%
Rang	2	1	3	
Wörter QS-Sätze insgesamt	1.248	1.248	1.248	= 3.744
Wörter Referenzsätze insgesamt	1.298	1.298	1.298	= 3.894
Wörter ZS-Sätze insgesamt	1.251	1.281	1.334	= 3.866
Fehler insgesamt	215	181	293	= 689
Rang	2	1	3	
davon:				
WST	19	15	45	= 79
1 Wort fehlt ganz	32	22	36	= 90
1 Wort nicht übersetzt	19	13	30	= 62
1 Wort Form falsch	52	64	81	= 197
1 Wort Bedeutung falsch	68	46	71	= 185
1 Wort zu viel	21	17	31	= 69

Diagramm Nr. 1: Darstellung der Fehlertypen insgesamt



180 Sätze sind in die Bewertung eingegangen. Die letzten 20 Sätze sind nicht eingerechnet, weil sie Beispiele für Sätze sind, die nicht mehr übersetzt werden können. Sie dienen nur dem Zweck zu zeigen, wie sich die Systeme an den Grenzen ihrer Funktionsfähigkeit verhalten. Die sehr hohe Fehleranzahl dieser Sätze würde die Fehlerbilanz verfälschen.

Der Google-Translator hat die meisten Sätze fehlerlos übersetzt und insgesamt die wenigsten Fehler gemacht, er liegt somit auf dem ersten Platz.

Auf dem zweiten Platz liegt der Bablefish-Translator mit 43% fehlerlos übersetzten Sätze und 215 Fehlern.

Der Systran-Translator liegt mit 293 Fehlern und 30% vollständig richtig übersetzten Sätzen auf dem dritten Platz. Im Vergleich mit den beiden anderen finden sich bei diesem Translator auch die meisten Fehler in den Wortformen.

Die Anzahl der eingegebenen Wörter und die Anzahl der ausgegebenen Wörter ist in etwa gleich hoch. Systran gab die meisten Wörter aus, sogar mehr als die Referenzsätze Wörter beinhalten, was bei den anderen zwei Systemen nicht der Fall war. Bing gab die wenigsten Wörter aus.

Vergleicht man die Fehleranzahl mit den ausgegebenen Wörtern, so kann man sagen, dass Bing ungefähr jedes sechste Wort falsch ausgegeben hat, Google ungefähr jedes siebte und Systran ungefähr jedes fünfte. Dies sind durchschnittliche Werte, denn es kam auch vor, dass in einem Wort zwei Fehler gezählt wurden.

Von allen drei Systemen wurde nicht einmal die Hälfte aller Sätze vollständig richtig übersetzt, was eine professionelle Nutzung der Systeme ausschließt, denn zu viel Korrektur- und somit Zeitaufwand wäre nötig, um die Sätze weiter zu verwenden.

Klar ersichtlich sind die meisten Fehler aufgrund falscher Wortformen entstanden, dicht gefolgt von den falschen Bedeutungen der Wörter. Die Fehler in allen weiteren Fehlerkategorien sind zusammen so zahlreich wie diese Form- und Bedeutungsfehler.

Fehler nach Themen

1. Kurze Sätze mit einfacher Syntax (ita-de und de-ita)

Eingegeben wurden 20 Aussagesätze, bestehend aus maximal 6 Wörtern, ohne Objekte, ohne Präpositionen, ohne Nebensätze und ohne Fachtermini.

Tabelle Nr. 2

	Bablefish Bing-Translator	Google-Translator	Systran-Translator	Gesamt
Vollständig richtig übersetzte Sätze:	16/20 80%	16/20 80%	17/20 85%	49/60 81%
Wert f. Verständlichkeit:	17,25 86%	17,50 87,5%	20,00 100%	54,75 91,2%
Genau gleich übersetzte Sätze: 25%				

Im Durchschnitt wurden 81% dieser einfachsten Sätze fehlerlos übersetzt. Erstaunlicherweise haben die 3 Systeme nur 25% dieser einfachen Sätze exakt gleich übersetzt, was zeigt, dass die Sprache immer viele Möglichkeiten hat, den gleichen Sachverhalt immer wieder anders auszudrücken.

Der Wert für die Verständlichkeit ist höher als die Anzahl der fehlerlos übersetzten Sätze, was zeigt, dass auch nicht fehlerfreie Sätze völlig verständlich sein können.

Die einfache Syntax wurde nur in 2 von 60 Fällen nicht korrekt in die Zielsprache übertragen.

2. Sätze mit direkten Objekten (ita-de und de-ita)

Direkte Objekte können als Phrasen betrachtet werden und sind von den Systemen anscheinend auch als solche behandelt worden.

Das Interesse liegt jetzt auf der Übersetzung der direkten Objekte. Die direkten Objekte sind in der italienischen Sprache morphologisch nicht gekennzeichnet, dafür ist ihre

Stellung im Satz wichtig, sie stehen hinter dem Prädikat. Im Deutschen sind die direkten Objekte morphologisch durch den Akkusativ markiert und stehen ebenfalls vorwiegend nach dem Prädikat. Die direkten Objekte wurden nur dann als richtig übersetzt gezählt, wenn neben der morphologischen Form auch die Bedeutung des Nomens und das möglicherweise zugehörige Adjektiv richtig waren.

Tabelle Nr. 3

	Bablefish Bing-Translator	Google-Translator	Systran-Translator	Gesamt
Vollständig richtig übersetzte direkte Objekte:	20/20 100%	17/20 85%	17/20 85%	54/60 90%
Wert f. Verständlichkeit:	19/20 95%	17/20 85%	18,75/20 94%	54,75/60 91,2%

In den 60 ausgegebenen ZS stehen alle direkten Objekte regelmäßig nach dem Prädikat, sowohl die deutschen als auch die italienischen. Alle drei Systeme hatten mit den direkten Objekten keine Probleme, alle deutschen Akkusative sind grammatikalisch richtig gebildet. Die in der Tabelle aufgeführten Fehler beziehen sich nicht auf Formfehler.

3. Sätze mit indirekten Objekten (ita-de und de-ita)

Indirekte Objekte sind in der italienischen Sprache mit der Präposition „a“ gekennzeichnet und stehen vorwiegend nach dem Prädikat, im Deutschen stehen sie im Dativ ebenfalls vorwiegend nach dem Prädikat, oder sind zwischen Hilfs- und Vollverb eingeschoben. Die indirekten Objekte wurden nur dann als richtig übersetzt gezählt, wenn sowohl die morphologische Form als auch die Bedeutung des Nomens und das möglicherweise zugehörige Adjektiv richtig waren.

Tabelle Nr. 4

	Bablefish Bing-Translator	Google-Translator	Systran-Translator	Gesamt
Vollständig richtig übersetzte indirekte Objekte:	8/21 38%	6/21 28%	5/21 23%	19/63 30,2%
Wert f. Verständlichkeit für alle 20 Sätze	17/20 85%	12/20 60%	12/20 60%	41/60 68%

- | | | | | | |
|--------|-------------------|------------|-------------------|-------------------------|----------------------------|
| (14)a. | Der Händler | bot | <u>dem Kunden</u> | einen Preisnachlass an. | (QS Nr. 55) |
| b. | Il concessionario | ha offerto | al cliente | uno sconto. | (RS Nr. 55) |
| c. | Il concessionario | ha offerto | <u>ai clienti</u> | uno sconto. | (ZS Nr. 55 Google-Tabelle) |

Beispiel (14a) und (14c) zeigen, dass die Stellung des indirekten Objekts im QS und im ZS gleich ist, es steht nach dem Prädikat. Beide Worte dieser Phrase „ai clienti“ stehen im Plural, statt wie im QS im Singular. Die Phrase ist in sich richtig, die zwei Worte der Phrase sind kongruent. Das lässt darauf schließen, dass sie als Phrase und nicht als Wörter gesucht und gefunden wurden.

Offensichtlich ist der Umgang mit den indirekten Objekten schwieriger als mit den direkten Objekten, denn nicht einmal 1/3 davon wurden richtig übersetzt.

Bablefish übersetzte nur 1 von 10 indirekten Objekten ins Italienische richtig (Satz Nr. 54 Bablefish-Tabelle). Im Satz Nr. 54 steht richtig das italienische Personalpronomen „gli“ („ihm“), das ohne die Präposition „a“ auskommt. In 7 von 11 Fällen hat Bablefish das indirekte Objekt aus dem Italienischen ins Deutsche mit einer Präposition, „an“, „für“ oder „zu“, ergänzt, was mehr den Textfluss als die Verständlichkeit des ZS beeinträchtigt.

Auch Google übersetzte 9 von 11 indirekten Objekte ins Deutsche mit einer Präposition. Ins Italienische hat Google die nötige Präposition „a“ in 2 von 11 Fällen richtig eingefügt. Systran übersetzte 10 der 11 italienischen indirekten Objekte mit der deutschen Übersetzung „zu“, was jedoch nur in 3 Fällen richtig war. Ins Italienische fügte Systran nur in 1 von 9 Fällen die nötige Präposition „a“ ein, allerdings steht dieses „a“ alleine und verloren am Satzende (Satz Nr. 57 Google-Tabelle). Nebenbei angemerkt findet sich ein solches verlorenes „a“ nur bei Systran.

4. Imperativsätze (ita-de und de-ita)

Hier geht es nur um die Verben in der Imperativform, die in beiden Sprachen in der 2.Sg. und der 2.Pl. jeweils in der persönlichen Anredeform oder in der Höflichkeitsform stehen können. Möglich ist in beiden Sprachen auch eine nicht konjugierte Verbform.

Tabelle Nr. 5

	Bablefish Bing-Translator	Google-Translator	Systran-Translator	Gesamt
Vollständig richtig übersetzte Imperativformen:	12/20 60%	13/20 65%	7/20 35%	32/60 53%
Wert f. Verständlichkeit für alle 20 Sätze	16,25/20 81%	18/20 90%	15/20 75%	49,25/60 82%

Im Durchschnitt wurden etwas mehr als die Hälfte aller Imperativformen richtig übersetzt. Die Anredeform (*du/Sie*) ist oft falsch gewählt, was aber die Verständlichkeit nicht stark beeinflusst, der Wert für die Verständlichkeit bleibt hoch.

Bablefish wich in der deutsch-italienischen Übersetzung auf die Infinitivform des Verbs aus und vermied dadurch die Festlegung auf eine persönliche oder höfliche Anredeform und die daraus entstehende Fehlerquelle.

5. Interrogativsätze (ita-de und de-ita)

Der ganze Fragesatz muss richtig übersetzt sein.

Tabelle Nr. 6

	Bablefish Bing-Translator	Google-Translator	Systran-Translator	Gesamt
Vollständig richtig übersetzter Fragesatz:	10/20 50%	11/20 55%	6/20 30%	27/60 45%
Wert für die Verständlichkeit für alle 20 Sätze	13,75/20 69%	14,75/20 74%	16,25/20 81%	44,75 74%

Insgesamt erzielten die Interrogativsätze schlechtere Werte als die Imperativsätze. Die Syntax der Fragesätze wurde mäßig gut aufgelöst. In den 20 Fragesätzen ist Google und BING die Syntax je 6 mal nicht richtig gelungen, Systran 12 mal nicht.

Google übersetzte die meisten Fragesätze vollständig richtig, die Fragesätze von Systran waren trotz mehr Fehler verständlicher.

6. Sätze mit Präpositionalphrase (ita-de und de-ita)

Die 20 Sätze enthalten je eine Präpositionalphrase, die nur dann als richtig zählt, wenn die gesamte Phrase in Form und Bedeutung richtig ist.

Tabelle Nr. 7

	Bablefish Bing-Translator	Google-Translator	Systran-Translator	Gesamt
Komplett richtig übersetzte Präpositionalphrasen:	14/20 70%	18/20 90%	10/20 50%	42/60 70%
Wert f. die Verständlichkeit für alle 20 Sätze	14,75/20 73%	17,5/20 87%	13/20 65%	45,25/60 75%

Präpositionalphrasen machten dem Google-Translator fast gar keine Probleme, dafür den anderen beiden Systemen um so mehr. Am meisten Probleme hatte Systran, das System hat nur 5 dieser 20 Sätze vollständig richtig übersetzt.

7. Sätze mit Fachtermini aus der Wirtschaft (ita-de und de-ita)

Fachtermini sind in ihrer Bedeutung exakt definiert und ihre Übersetzung ist infolgedessen nicht variabel. Hier sind nur die richtigen Fachtermini gezählt. Die 20 Sätze enthalten 38 Fachtermini.

Tabelle Nr. 8

	Bablefish Bing-Translator	Google-Translator	Systran-Translator	Gesamt
Richtig übersetzte Fachbegriffe „Wirtschaft“:	27/38 71%	35/38 92%	19/38 50%	81/114 71%
Wert f. Verständlichkeit der 20 Sätze	11,5/20 57%	17/20 85%	9/20 45%	37,5/60 62%

Fachtermini aus der Wirtschaft werden erwartungsgemäß besser übersetzt als juristische, denn die Welt der Wirtschaft bedient sich des Web viel offener als die juristische Welt, die noch mehr auf dem Papier arbeitet. Infolgedessen sind für den Bereich Wirtschaft viel mehr Korpora im Netz.

8. Fachsprachlicher Text „Recht“ (ita-de und de-ita)

Juristen definieren ihre juristischen Termini exakt und die Übersetzung ist nicht beliebig. Als ein Beispiel für die Genauigkeit juristischer Texte soll die Unterscheidung zwischen dem Begriff „Eigentümer“ und dem Begriff „Besitzer“ gezeigt werden, welche in allgemeinsprachlichen Texten nicht gemacht wird.

(15)a.	Il proprietario di un fondo ...	(QS)
b.	Der Eigentümer eines Grundstücks	(Ref.S.)
c.	Der Besitzer eines Fonds ...	(ZS Nr. 143 Bablefish-Tabelle)
d.	Der Eigentümer eines Grundstücks ...	(ZS Nr. 143 Google-Tabelle)

Zu sehen ist, dass (15c) den Fachterminus zwar allgemeinsprachlich richtig, juristisch und fachsprachlich aber falsch übersetzt hat.

Für rechtsprachliche Übersetzungen sind zuverlässige und von offizieller Seite anerkannte Terminologie-Datenbanken, wie die IATE der Europäischen Union, erforderlich.

Für den Bereich der italienisch-deutschen Rechtsübersetzungen liegen das Italienische Zivilgesetzbuch und das Strafgesetzbuch und weitere in einer offiziell anerkannten zweisprachigen gedruckten Ausgabe vor. Teilweise sind bereits digitale Versionen vorhanden.

Hier interessiert nur die Fachterminologie und deren Übersetzung, andere Fehler sind nicht berücksichtigt. In den 20 Sätzen sind 46 Fachtermini enthalten.

Tabelle Nr. 9

	Bablefish Bing-Translator	Google-Translator	Systran-Translator	Gesamt
Vollständig richtig übersetzte Fachtermini „Recht“:	32/46 69%	31/46 67%	25/46 54%	88/138 64%
Wert f. die Verständlichkeit aller 20 Sätze:	14,25/20 71%	8,5/20 42%	7,25/20 36%	30/60 50%

Der Bablefish-Translator hat die meisten fachsprachlichen Begriffe richtig aufgelöst und mit Abstand die meisten verständlichen Sätze produziert.

9. Sätze mit Schreibfehlern (ita-de und de-ita)

Aus der Praxis ist bekannt, dass besonders in einer Fremdsprache leicht Schreibfehler gemacht werden, weil die Nutzer die Rechtschreibung in der Fremdsprache nicht beherrschen. Die Untersuchung zeigte, dass alle drei Systeme nicht fehlertolerant sind. Vergleicht man den Wert von 42% für die Verständlichkeit für Sätze mit Schreibfehlern in Tabelle 10 mit dem durchschnittlichen Gesamtwert von 69% für alle Sätze (Tabelle 11), so ist zu sehen, dass der Wert stark abgefallen ist.

Tabelle Nr. 10

	Bablefish Bing-Translator	Google-Translator	Systran-Translator	Gesamt
Richtig übersetzte, obwohl falsch geschriebene Wörter:	10/20 50%	7/20 35%	4/20 20%	21/60 35%
Wert f. die Verständlichkeit aller 20 Sätze:	8,25/20 41%	9,75/20 49%	7,25/20 36%	25,25 42%

Das nächste Beispiel (13) zeigt, wie die Systeme mit falsch geschriebenen Wörtern umgehen, sie werden nicht übersetzt, sondern fremdsprachlich übernommen.

- | | | |
|--------|-------------------------------|-----------------------------|
| (16)a. | Il bambino corre. | (QS Nr. 4) |
| b. | Das Kind läuft. | (RS Nr. 4) |
| c. | Das Kind läuft. | (ZS Nr. 4 Google-Tabelle) |
| d. | Il <u>bambbino</u> corre. | (QS Nr. 165) |
| e. | Das Kind läuft. | (RS Nr. 165) |
| f. | Die <u>bambbino</u> verläuft. | (ZS Nr. 165 Google-Tabelle) |

Gleichzeitig hat sich in (16f) auch das Wort vor und das Wort nach dem falsch geschriebenen Wort verändert, wie aus dem Vergleich (16c) und (16f) zu sehen ist. Die Satzstruktur ist jedoch erhalten geblieben, es stimmt auch die Kongruenz zwischen Subjekt und Prädikat, woraus man schließen könnte, dass hier Syntaxregeln unabhängig funktionieren.

Der Google-Translator bietet als einziges Tool die *“Meinten-Sie-Funktion”* an, womit für die fehlerhafte Eingabe automatisch eine berichtigte Version angeboten wird, was wirklich hilfreich ist. Google hat in 18 Sätzen alle falsch geschriebenen Wörter identifiziert und die richtig geschriebene Version angeboten. Zwei Wörter hat Google trotz falscher Eingabe richtig übersetzt.

10. Sätze, die nicht übersetzt werden können (ita-de und de-ita)

Ab einem bestimmten Schwierigkeitsgrad, dessen exakte Definition wohl nicht möglich ist, werden Sätze nicht mehr übersetzt. Ausschlaggebend sind auf jeden Fall die Anzahl der Wörter im Satz, die Wortlänge, die Seltenheit von Wörtern, komplizierte Satzstrukturen, wie eingeschobene Nebensätze, aber auch deutsche Verben, wenn sie erst ganz am Ende eines langen Satzes stehen.

Wie in 2.2.3. dargestellt, bestätigt sich hier die Regel, dass lange Sätze in der Übersetzung wieder lange Sätze produzieren.

Redewendungen und Sprichwörter sind schwierig, weil sie als ganze Sätze auftreten und eine Zerlegung in Phrasen oder Wörter ein falsches Ergebnis bringt. Die Systeme finden einige von ihnen als ganzen Satz in der ZS. Wie das folgende Beispiel (17) zeigt, hat nur Bablefish den Satz zerlegt, und das Nomen „nodi“ wörtlich mit „Knoten“ übersetzt.

Auch Satz Sinn (17c) kann in der Richtung der Bedeutung des QS interpretiert werden, der Sinn liegt aber weiter davon entfernt als Satz (17b).

- | | | |
|--------|--|--------------------------------|
| (17)a. | Tutti i nodi vengono al pettine. | (QS Nr. 183) |
| b. | Alle Fehler kommen ans Licht. | (RS Nr. 183) |
| c. | Wer warten kann, dem kommt alles zur rechten Zeit. | (ZS Nr. 183 Systran-Tabelle) |
| d. | Alle Knoten sind Heimat von Stange. | (ZS Nr. 183 Bablefish-Tabelle) |
| e. | Alle Hühner nach Hause kommen, um schlafen. | (ZS Nr. 183 Google-Tabelle) |

3.3.4. Verständlichkeit

Verständlichkeit ist immer subjektiv, um so mehr, wenn sie nur von einer Person bewertet wird. Viele Personen zu befragen war hier aus Zeitgründen nicht möglich.

In Anlehnung an die Bewertung, die Koehn (2011:219) vorstellt, wurde die Verständlichkeit in fünf Stufen aufgeteilt:

(Figur 3)

0	Satzinhalt unverständlich
0,25	Satzinhalt erahnbar
0,5	Satzinhalt halb verständlich
0,75	Satzinhalt fast verständlich
1	Satzinhalt komplett verständlich

Fehlerfreie Sätze erhalten den Wert 1. Die Verständlichkeit eines Satzes hängt aber nicht linear mit der Fehleranzahl zusammen. Ein Fehler im Satz kann bedeutungslos sein, oder aber den Sinn des Satzes mehr oder weniger bedeutsam verändern.

Wie das nächste Beispiel zeigt, kann ein Satz trotz Fehler völlig verständlich bleiben, weil der Fehler durch unser Sprachverständnis gedanklich repariert wird.

- (18) a. La famiglia vive vicino alla casa della nonna. (QS Nr. 110)
 b. Die Familie lebt neben dem Haus der Großmutter. (RS Nr. 110)
 c. Die Familie wohnt neben ___ Haus der Großmutter. (ZS Nr. 110 aus Bablefish-Tabelle)

Fehlt der Artikel vor dem Nomen (18c), so ist die Verständlichkeit nur minimal beeinträchtigt.

In vielen Fällen ist die Verständlichkeit schwerer entscheidbar, die Bewertung kritisch.

- (19) a. Lo psicologo legge un libro interessante. (QS Nr. 25)
 b. Der Psychologe liest ein interessantes Buch. (RS Nr. 25)
 c. Der Psychologe ein gutes Buch lesen. (ZS Nr. 25 aus Google-Tabelle)

Beispiel (19) zeigt einen Satz (19c) mit falscher WST, die Verständlichkeit ist davon nur wenig betroffen. Die nicht gelungene Neuordnung der Phrasen würde unter dem Kriterium

“Sprachfluss“ sicher negativ bewertet werden.

Auch die nicht deklinierte Verbform wird verstanden. Kritisch bleibt die Bedeutung der Adjektive „interessant“ und „gut“. Wenn sich die Bedeutungsfelder der Adjektive in einem Punkt überschneiden, so ist die Übersetzung richtig, im umgekehrten Fall nicht. Im Duden Synonymwörterbuch überschneiden sich die Bedeutungsfelder nicht, deswegen wurde ein Fehler gezählt. Der Satz (19c) hat also drei Fehler, ist aber dennoch ziemlich verständlich. Deswegen erhielt er den Wert 0,75 für die Verständlichkeit.

Tabelle Nr. 11

	BING-Translator	Google-Translator	Systran-Translator	Gesamt
Verständlichkeit	124,25 /180 69,%	131,25 /180 73%	119,25 /180 66%	=375 /540 69%
0 Satzinhalt unverständlich	37/180 20%	32/180 17%	38/180 21%	107/540 20%
0,25 Satzinhalt erahnbar	11/180 6%	5/180 2%	14/180 8%	30/540 5,5%
0,5 Satzinhalt halb verständlich	11/180 6%	16/180 9%	8/180 4%	25/540 5%
0,75 Satzinhalt fast verständlich	20/180 11%	20/180 11%	28/180 15%	68/540 12%
1 Satzinhalt komplett verständlich	101/180 56%	107/180 60%	90/180 50%	298/540 55%

Fast 70% des Inhalts der Sätze ist verständlich. Den höchsten Wert für die Verständlichkeit erzielte Google, gefolgt vom Bing-Translator.

Etwas mehr als die Hälfte aller Sätze war komplett verständlich, wobei Google den höchsten Wert erzielte und Systran den niedrigsten.

Im Durchschnitt waren 20% aller Sätze unverständlich und somit unbrauchbar. Eine Ausschussquote von 20% ist im Wirtschaftsleben nicht mehr tragbar. Dies ist ein Beweis

für die Unmöglichkeit, diese Systeme kommerziell zu nutzen.

Verbindet man die Bewertung der Fehleranzahl mit der Verständlichkeit, so können folgende Aussagen gemacht werden:

Das System Google, das die wenigsten Fehler gemacht hat, erzielte auch den höchsten Wert in puncto Verständlichkeit.

Das System Systran, das die meisten Fehler gemacht hat, erzielte auch den niedrigsten Wert in puncto Verständlichkeit, es besteht eine nicht lineare Korrelation zwischen Fehleranzahl und Verständlichkeit.

Abschließend soll angemerkt werden, dass die Zahlen für alle drei Systeme nicht sehr stark voneinander abweichen. Vergleicht man die Werte der einzelnen Systeme mit dem Durchschnittswert für alle drei Systeme, so ist ersichtlich, dass sie sich in allen Kategorien um den Durchschnittswert herum bewegen. Diese Tatsache kann als Hinweis gewertet werden, dass die drei Systeme auf ähnliche Weise arbeiten und auch auf ähnliche Quellen im Netz zurückgreifen.

3.4 Interpretation der Ergebnisse

Aus den Ergebnissen folgt eindeutig, dass der Google-Translator am besten abgeschnitten hat. Möglicherweise wird der Google-Translator am häufigsten benutzt.

Für syntaktisch einfache und allgemeinsprachliche Texte sind die Übersetzungen durchaus gut zu gebrauchen. Die Eingabesätze sollten eher kurz gehalten werden, die Rechtschreibung sollte stimmen. Die Ausgabesätze sollten kritisch betrachtet werden. Nutzer sollten sich nicht blind auf die Richtigkeit der Übersetzungen verlassen, sondern selbst nachdenken und versuchen, das Ergebnis richtig zu interpretieren. Deswegen können am ehesten Nutzer profitieren, die die Fremdsprache schon etwas kennen.

Zur Qualität der Übersetzungen lässt sich sagen, dass die Systeme die Übersetzungen schnell, einfach zugänglich und kostenfrei liefern, drei unschlagbare Argumente dafür, sie zu nutzen. Positiv zu bemerken ist, dass die on-line Angebote niederschwellig sind, und eine große Menge Menschen einladen, sich einen ersten Eindruck über die Funktion und den Nutzen der MÜ zu verschaffen. Einem breiten Publikum wird die Welt der Vielsprachigkeit geöffnet, das ist ein positiver und bemerkenswerter Beitrag zur Verständigung der Menschen untereinander.

Zusammenfassung

Die Untersuchung beschrieb die Qualität der Übersetzungen von drei kostenfreien SMÜ-Systemen ausführlich. Aufgrund der Ergebnisse kann die private Nutzung der on-line Translators nicht uneingeschränkt empfohlen werden. Kommerziell kann man die Systeme nur mit größter Vorsicht nutzen, denn die Eingabetexte werden urheberrechtlich zum Eigentum der Anbieter.

Der Google-Translator übersetzte von allen drei Systemen am besten und wird deshalb vor den anderen beiden empfohlen. Dieses System erzielte die besten Ergebnisse. Andererseits weichen alle drei Systeme in ihrer Leistung nicht weit vom Durchschnittswert ab, ihre Leistungen sind folglich nicht sehr unterschiedlich.

Für einen ersten, schnellen Einblick in fremdsprachlichen Text sind die Systeme durchaus schon geeignet. Die Nutzer sollten jedoch vorsichtig sein und sich daran erinnern, dass sie auch ihrem Taschenrechner nicht blind vertrauen, sondern dass eine unerwartete Ergebniszahl sie verwundert. Profitieren können vor allem Nutzer, die sich in der Fremdsprache bereits etwas auskennen.

Die vorgestellten Übersetzungstools werden eingesetzt, um sich den wesentlichen Inhalt eines Textes geben zu lassen, druckreife Übersetzungen liefern sie nicht.

Realität ist, dass der Bedarf und das Interesse an Übersetzungen vor allem durch die Globalisierung der Märkte und durch das Internet enorm angestiegen ist. Viele Unternehmen und internationale Organisationen setzen schon lange mit großem Erfolg professionelle und teure MÜ-Systeme ein, wie als bekanntes Beispiel die Europäischen Union, deren Leistung selbstverständlich mit den on-line Tools nicht vergleichbar ist.

Ohne MÜ-Systeme wäre der Bedarf an Übersetzungen nicht mehr zufriedenzustellen. Der Berufsstand der Übersetzer und Übersetzerinnen wird sich gerne einen Teil der Arbeit abnehmen lassen und selbst MÜ einsetzen.

Bibliographie

- Brown, Peter F., Stephen A. Della Pietra, Vincent J. Della Pietra , Robert L. Mercer. 1993. The Mathematics of Statistical Machine Translation: Parameter Estimation. In *Computational Linguistics Volume 19, Number 2:263-311*. Cambridge USA: MIT Press. S. 2-3
- Crystal, David. 2010. *Die Cambridge Enzyklopädie der Sprache*. Tolkemitt Verlag bei Zweitausendeins. S. 350
- Diehl, Malte. 2007. *Messung maschineller Übersetzbarkeit von Texten*. Diplomarbeit. Oldenburg: Universität Oldenburg. S.16
- Gale William A. und Kenneth W. Church. 1991. A program for aligning sentences in bilingual corpora. In *Proceedings of the 29th annual meeting on Association for Computational Linguistics*. Berkeley, California. S.177-184. S.5
- Jurafsky, Daniel & James H. Martin. (Hg.). 2000. *Speech and Language Processing. An Introduction to Natural Language Processing, Computational Linguistics, and Speech Recognition*. New Jersey: Prentice-Hall, Inc.. Seite 819
- Koehn, Philipp, Franz-Josef Och & Daniel Marcu. 2011. *Statistical Machine Translation*. Cambridge: Cambridge University Press. S. xi, S. 127, S. 181, S.183, S. 219
- Koehn, Philipp. 2005. Europarl: A Parallel Corpus for Statistical Machine Translation. In *Proceedings of Conference, Machine Translation Summit X 2005*. Thailand: Main conference paper. S.4
- Koehn, Philipp, Franz-Josef Och & Daniel Marcu. 2003. Statistical Phrase-Based Translation. In *Proceedings of Human Language Technology (HLT-NAACL) 2003*. Edmonton. Main Papers. S. 50
- Mounin, Goerges. 1967. *Die Übersetzung, Geschichte Theorie Anwendung*. München: Nymphenburger Verlagshandlung. S. 21
- Munteanu Dragos S. und Daniel Marcu. 2006. Improving Machine Translation Performance by Exploiting Non-Parallel Corpora. In *Association for Computational Linguistics*. Stroudsburg USA: Volume 31 (4). S. 1
- Och, Franz J. und Hermann Ney. 2004. The Alignment Template Approach to Statistical Machine Translation. In *Association for Computational Linguistics*. Stroudsburg USA: Volume 30 (4). S.1
- Somers, Harald. 2011. Kirsten Malmkjaer und Kevin Windle (Hg.). *The Oxford Handbook of Translation Studies*. Oxford: University Press. S. 435

Stein, Daniel. 2009. Maschinelle Übersetzung – Ein Überblick. In *Journal for Language Technology and Computational Linguistics (JLCL)*. Vol. 24 (3). S. 8, S. 17

Zens, R., F.J. Och & H. Ney. 2002. Phrase-based statistical machine translation. In *Proceedings German Conference on Artificial Intelligence (KI 2002)*. Aachen. *Springer Verlag, September 2002*. S. 18-32. S.22

Zinsmeister, Heike. 2010. *Korpora*. Universität Konstanz:KOPS. S.6

<http://nbn-resolving.de/urn:nbn:de:bsz:352-opus-108495>

(zuletzt besucht am 31.01.2013)

Anhänge

1. BING Testtabelle
2. GOOGLE Testtabelle
3. SYSTRAN Testtabelle