



universität  
wien

## MASTERARBEIT/ MASTER'S THESIS

Titel der Masterarbeit/ Title of the Master's Thesis

### **„Möglichkeiten und Grenzen Maschiner Übersetzung**

**Die Evaluation der Performanz der MÜ im Vergleich zur Humanübersetzung in Bezug auf die Terminologie, die in den Ausgangs- und Zieltextrn vorkommt“**

verfasst von/ submitted by

Judith Daria Lechthaler BA

angestrebter akademischer Grad/ in partial fulfilment of the requirements for the degree of

Master of Arts (MA)

Wien, 2021 / Vienna 2021

Studienkennzahl lt. Studienblatt/  
degree programme code as it appears on  
the student record sheet:

UA 070 331 351

Studienrichtung lt. Studienblatt/  
degree programme as it appears on  
the student record sheet:

Masterstudium Translation Deutsch  
Spanisch

Betreut von/Supervisor:

Univ.-Prof. Mag. Dr. Gerhard Budin

### **Selbstständigkeitserklärung**

Hiermit bestätige ich, dass ich die vorliegende Arbeit selbstständig verfasst und keine anderen Publikationen, Vorlagen und Hilfsmittel als die angegebenen benutzt habe. Alle Teile meiner Arbeit, die wortwörtlich oder dem Sinn nach anderen Werken entnommen sind, wurden unter Angabe der Quelle kenntlich gemacht. Gleiches gilt für von mir verwendete Internetquellen. Die Arbeit ist weder von mir noch von einem/einer KommilitonIn bereits in einem anderen Seminar vorgelegt worden.

Wien 15.02.2021

Judith Lechner

Ort, Datum, Unterschrift

# I. Inhaltsverzeichnis

II. Abbildungsverzeichnis

III. Abkürzungsverzeichnis

<b>0. Einleitung</b> .....	S.8
<b>1. Forschungsfrage &amp; Hypothesen</b> .....	S.10
1.1 Aufbau & Vorgehensweise.....	S.12
<b>2. Was ist maschinelle Übersetzung?</b> .....	S.15
2.1 Begriffsdefinition.....	S.15
2.2 Klassifizierung nach Anteil des menschlichen Eingreifens in den Übersetzungsprozess.....	S.16
2.2.1 Maschinelle Übersetzungshilfen.....	S.17
2.2.2 Maschinengestützte Übersetzung.....	S.18
2.2.3 Vollautomatische Übersetzung.....	S.19
2.3 Schlussfolgerung und für diese Arbeit relevante Definition.....	S.20
<b>3. Geschichtlicher Hintergrund</b> .....	S.21
3.1 Anfänge der maschinellen Übersetzung.....	S.21
3.2 Die 1950er Jahre und der ALPAC-Bericht.....	S.22
3.3 Weitere Entwicklung maschineller Übersetzung und der gegenwärtige Stand der Forschung.....	S.26
<b>4. Maschinelle Übersetzungssysteme</b> .....	S.31
4.1. Klassifizierung anhand verschiedener Ansätze und Methoden.....	S.31
4.1.1 Regelbasierte Systeme.....	S.31
4.1.2 Statistikbasierte Systeme.....	S.33
4.1.3 Neuronale Maschinelle Übersetzungssysteme.....	S.36
4.2. Aktuelle Herausforderungen neuronaler Maschineller Übersetzung.....	S.39
<b>5. Maschinelle Übersetzung am Beispiel DeepL</b> .....	S.43
5.1 Beschreibung der Software.....	S.43
5.2 Installation und Anwendung.....	S.44
<b>6. Terminologie &amp; Translation</b> .....	S.46
6.1 Abgrenzungen und Überschneidungen .....	S.46
6.2 Die Rolle der Terminologie in der Maschinellen Übersetzung.....	S.47
6.3 Begriffsklärungen: <i>term recognition &amp; termhood</i> .....	S.49

<b>7. Korpusanalyse</b> .....	S.52
7.1 Beschreibung des Korpus.....	S.52
7.2 Korpus.....	S.53
7.2.1 Text 1: Auszug aus <i>Alles könnte anders sein – Eine Gesellschaftsutopie für freie Menschen</i> von Harald Welzer.....	S.53
7.2.2 Text 2: Firmenkatalog Austria- Export, ALSTOM Österreich.....	S.53
7.2.3 Text 3: Auszug aus <i>¿Las ciudades más ecológicas son más desiguales?</i> Von Isabelle Anguelovski.....	S.54
7.2.4. Text 4: Auszug aus <i>Casa tomada</i> von Julio Cortázar.....	S.55
7.3 Methodische Vorgehensweise.....	S.56
7.3.1 SAE Qualitätsmetrik.....	S.57
7.3.2 Fehlerkategorien und Evaluation.....	S.58
7.3.2.1 Fehlerkategorie: Falscher Begriff, WT – Wrong Term.....	S.58
7.3.2.1 Fehlerkategorie: Falscher Begriff, WM – Wrong Meaning.....	S.59
7.3.2.3 Fehlerkategorie Auslassung, OM – Omission.....	S.59
7.3.2.4 Fehlerkategorie: Strukturfehler, SE – Structural Error.....	S.60
7.3.2.5 Fehlerkategorie: Rechtschreibfehler, SP – Misspelling.....	S.60
7.3.2.6 Fehlerkategorie: Interpunktionsfehler, PE – Punctuation Error.....	S.60
7.3.2.7 Fehlerkategorie: Sonstige Fehler, ME – Miscellaneous Error.....	S.61
7.4 Auswertungsbogen.....	S.61
<b>8. Terminologiearbeit</b> .....	S.62
8.1 Struktur der terminologischen Einträge.....	S.62
8.1.1 Benennung.....	S.63
8.1.2 Synonym.....	S.63
8.1.3 Definition.....	S.63
8.1.4 Quelle.....	S.64
8.1.5 Kontext.....	S.64
8.1.6 Fachgebiet.....	S.64
8.2 Mustereintrag.....	S.65
8.3 Präsentation des Glossars.....	S.65
8.3.1. Text 1: Auszug aus <i>Alles könnte anders sein – Eine Gesellschaftsutopie für freie Menschen</i> von Harald Welzer.....	S.65

8.3.2 Text 2: Firmenkatalog Austria- Export, ALSTOM Österreich.....	S.69
8.3.3. Text 3: Auszug aus <i>¿Las ciudades más ecológicas son más desiguales?</i> von Isabelle Anguelovski.....	S.74
8.3.4. Text 4: Auszug aus <i>Casa tomada</i> von Julio Cortázar.....	S.78
<b>9. Auswertung: Gesamte Texte</b> .....	S.82
9.1 Vergleich Human- vs. maschinelle Übersetzung: Text 1 .....	S.82
9.1.1 Übersetzung: Human.....	S.82
9.1.2 Übersetzung: Deep L.....	S.83
9.1.3 Auswertung.....	S.84
9.2 Vergleich Human- vs. maschinelle Übersetzung: Text 2.....	S.85
9.2.1 Übersetzung: Human.....	S.85
9.2.2 Übersetzung: Deep L.....	S.86
9.2.3 Auswertung.....	S.87
9.3 Vergleich Human vs. Maschinelle Übersetzung: Text 3.....	S.88
9.3.1 Übersetzung: Human.....	S.88
9.3.2 Übersetzung: Deep L.....	S.89
9.3.3 Auswertung.....	S.90
9.4 Vergleich Human vs. Maschinelle Übersetzung: Text 4.....	S.92
9.4.1 Übersetzung: Human.....	S.92
9.4.2 Übersetzung: Deep L.....	S.93
9.4.3 Auswertung.....	S.94
<b>10. Auswertung: Gegenüberstellung der ausgewählten Termini</b> .....	S.98
10.1 Fachtext 1.....	S.100
10.2 Fachtext 2.....	S.101
10.3 Fachtext 3.....	S.102
10.4 Fachtext 4.....	S.103
<b>11. Schlussfolgerung</b> .....	S.105
<b>12. Literaturverzeichnis</b> .....	S.109
<b>13. Abstract</b> .....	S.119

## II. Abbildungsverzeichnis

Abbildung 1: <i>Aufwand der drei Phasen der Maschinellen Übersetzung.</i> (Stein 2009:9)...	S.33
Abbildung 2: <i>Berechnung der Übersetzungswahrscheinlichkeiten für jedes Wort im Satz "Das Haus ist klein."</i> (Koehn 2010:84).....	S.34
Abbildung 3: <i>Screenshot DeepL Übersetzer (Spanisch).</i> (DeepL 2020a).....	S.44
Abbildung 4: <i>The interaction between Theory-oriented Terminology, Translation-oriented Terminology and Translation.</i> (Thelen 2015:352).....	S.46
Abbildung 5: Eigene Darstellung. Auswertungsbogen orientiert an SAE J2450 Qualitätsmetrik.....	S.61
Abbildung 6: <i>Eigene Darstellung. Mustereintrag Glossar</i> .....	S.65
Abbildung 7: <i>Eigene Darstellung. Auswertungsbogen Text 1</i> .....	S.85
Abbildung 8: <i>Eigene Darstellung. Auswertungsbogen Text 2</i> .....	S.88
Abbildung 9: <i>Eigene Darstellung. Auswertungsbogen Text 3</i> .....	S.91
Abbildung 10: <i>Eigene Darstellung. Auswertungsbogen Text 4</i> .....	S.96
Abbildung 11: <i>Eigene Darstellung. Gegenüberstellung der Termini</i> .....	S.98
Abbildung 12: <i>Eigene Darstellung. Gegenüberstellung der Evaluation der Fachtexte</i> ...	S.106

## III. Abkürzungsverzeichnis

Abb.	Abbildung
ALPAC	Automatic Language Processing Advisory Committee
CAT	Computer Aided Translation
CEO	Chief Executive Officer
DE	Deutsch
ES	Spanisch
Fach	Fachbereich
FAHQT	Fully Automatic High-Quality Translation
FAMT	Fully Automatic Machine Translation
HAMT	Human-Aided Machine Translation
Ibd	Ibidem
Kap	Kapital
KON	Kontext

MAHT	Machine-Aided Human Translation
ME	Miscellaneous Error
MT	Machine Translation
MÜ	Maschinelle Übersetzung
NMT	Neuronale maschinelle Übersetzung
OM	Omission
PE	Punctuation Error
QUE	Quelle
RBMT	Rule Based Machine Translation
SE	Structural Error
SMT	Statistical Maschine Translation
SMÜ	Statistikbasierte Maschinelle Übersetzung
SP	Misspelling
SYN	Synonym
TM	Translation Memory
WM	Wrong Meaning
WT	Wrong Term

## 0. Einleitung

In Zeiten des stetig wachsenden globalen Austausches sowie der interkulturellen Vernetzung ist das Internet zur Grundlage für die Digitalisierung sozialer Interaktionen geworden. Diese Entwicklung ist dabei aber nicht nur im privaten Bereich, sondern auch im beruflichen Umfeld spürbar.

Der steigende Grad an Technologisierung lässt viele traditionelle Berufe und Tätigkeiten verschwinden. Arbeitskräfte werden schlichtweg durch Maschinen ersetzt, um Zeit und Kosten einzusparen. Diese Veränderungen und Entwicklungen im technischen Bereich wirken sich auch besonders stark auf den Bereich der Translation aus.

Jene Weiterentwicklung bringt dabei Chancen und auch Herausforderungen mit sich. Gerade im Feld der Translation ist der Wert dieses Fortschrittes sehr umstritten. Die Übersetzungen, die ursprünglich von Menschen durchgeführt wurden, sollen nun auch durch Maschinen erfolgen. Dabei haben viele ÜbersetzerInnen Angst um ihre berufliche Zukunft. Gerade heutzutage sollen SprachdienstleisterInnen immer mehr Übersetzungen in den Fachbereichen Technik, Wirtschaft oder Recht anfertigen, die in kürzester Zeit und gleichzeitig in hoher Qualität in eine andere Sprache übertragen werden müssen. Mit der schnellen Entwicklung der Technologien und dem damit verbundenen Entstehen zahlreicher Tools soll wohl die Möglichkeit geschaffen worden sein, die Arbeit der ÜbersetzerInnen zu erleichtern und dadurch auch zu beschleunigen (vgl. Ritschl 2015).

Die Auswahl an Online Wörterbüchern, Übersetzungstools und Sprachassistenten (Spracherkennungstools) ist in den vergangenen Jahren rapide gewachsen. Durch den bequemen und unkomplizierten Zugang zum Internet, dessen kostenlose Verfügbarkeit und dessen einfache Bedienbarkeit wächst die Gruppe an NutzerInnen stetig und umfasst dabei alle Altersgruppen sowie sozialen Schichten. Dadurch, dass diese Tools in solcher Vielzahl benutzt werden, wird fortlaufend an den zu erfüllenden Anforderungen und der Qualität dieser gearbeitet.

Über viele Jahre hinweg entstand der Eindruck, dass die Qualität dieser neuartigen Technologien zu wünschen übrig lassen würde und maximal nur für den privaten Bereich geeignet wäre. Zu Beginn dieser Technologisierung war *Google Translate* wohl das am weitesten verbreitete Übersetzungstool. Mittlerweile gibt es dahingehend zahlreiche Ansätze und Übersetzungsstrategien sowie vielfältige Übersetzungstools. Die Forschung hat sich – vor allem in Bezug auf die neuronale maschinelle Übersetzung – in den letzten Jahren aber so

rasch entwickelt, dass selbst maschinelle Übersetzungstools im beruflichen Umfeld von TranslatorInnen miteingebunden werden. Bereits im Jahre 2017 wurden vom neuronalen Übersetzungssystem *DeepL* Blindtests durchgeführt. Im Rahmen jener Studie wurde die Qualität von *DeepL* durch professionelle ÜbersetzerInnen bewertet und mit anderen bekannten Übersetzungssystemen von *Google* oder *Microsoft* verglichen. Laut den AuftraggeberInnen der Tests wiesen die Ergebnisse auf eine dreifach bessere Leistung von *DeepL* im Vergleich mit anderen MÜ-Tools hin. Daraus lässt sich die Schlussfolgerung ziehen, dass *DeepL* als neuronale maschinelle Übersetzung im Bereich der Fachtexte sehr gute Resultate liefert. Als im Jahr 2020 abermals derartige Tests durchgeführt wurden, konnte *DeepL* seinen Vorsprung sogar noch weiter ausbauen (vgl. DeepL 2020b).

Dass maschinelle Übersetzungstools im privaten Umfeld eine wertvolle Möglichkeit darstellen, sprachliche Schwierigkeiten zu überwinden und interkulturelle Kommunikation zu fördern, ist unumstritten. Selbiges kann jedoch nicht eindeutig für die Nutzung jener Tools im beruflichen Umfeld von TranslatorInnen, die Qualität der Ergebnisse von Übersetzungssystemen in fachspezifischen Bereichen oder die Beeinflussung der Zukunft der übersetzerischen Tätigkeit vorausgesagt werden. Dies soll daher in der vorliegenden Masterarbeit diskutiert werden.

# 1. Forschungsfrage und Hypothesen

Durch die Globalisierung und den dadurch vermehrten interkulturellen Kontakt und Austausch zwischen verschiedenen Menschen und Sprachen werden TranslatorInnen, also DolmetscherInnen und ÜbersetzerInnen, die dabei helfen, diesen Austausch so reibungslos wie möglich zu garantieren, zu einer unabdingbaren Notwendigkeit. Dabei wird in den letzten Jahren nicht nur auf menschliche, sondern auch vermehrt auf maschinelle Dienstleistungen zurückgegriffen. Auch wenn in manchen Fällen das reine Verständnis und eine grobe Übermittlung des Grundgedankens im Mittelpunkt stehen, so ist in den allermeisten Fällen die Qualität dieser Übermittlung keineswegs vernachlässigbar. Dabei soll der Gedanke, die Idee, die Emotion so originalgetreu wie möglich übertragen werden. Die Schwierigkeit bei der Übersetzung beziehungsweise Dolmetschung von beispielsweise Small Talk hält sich in Grenzen. Doch wie sieht es mit der Qualität von modernen maschinellen Übersetzungstools in den verschiedenen Fachbereichen aus? Können Fachtexte in den Bereichen Wirtschaft, Naturwissenschaft, Recht, Geisteswissenschaften und Technik genauso von Maschinen durchgeführt werden und damit menschliche Leistungen und Tätigkeiten ersetzen?

Bis dato existieren bereits mehrere Forschungsarbeiten, die sich mit unterschiedlichen Übersetzungssystemen, deren Qualität und Entwicklungen sowie mit deren Einflussnahme auf das Berufsfeld Translation auseinandersetzen. Um die Fragestellung um eine neuartige Komponente von wissenschaftlichem Interesse zu erweitern, fokussiert die vorliegende Masterarbeit auf den Bereich der Terminologie. In der Regel steht einer Übersetzung in den genannten Fachbereichen eine aufwändige Terminologierecherche bevor. Termini werden dabei genauestens untersucht, um anschließend das richtige zielsprachliche Pendant in der Übersetzung verwenden zu können. Wie also werden Termini bei maschinellen Übersetzungssystemen übersetzt, ohne dabei a priori diese aufwändige Recherche betreiben zu haben? Wie fällt dementsprechend die Qualität dieser Übersetzung in Bezug auf den gesamten Text aber auch in Bezug auf die einzelnen Termini aus? Im Rahmen dieser hier vorliegenden Masterarbeit soll daher die folgende Forschungsfrage beantwortet werden:

*Inwiefern ist die maschinelle Übersetzung ohne Prä- und Postedition im Bereich der Übersetzung von fachlichen Texten in Bezug auf die Terminologie der Ausgangs- und Zieltexte anwendbar?*

Hierbei soll insbesondere auf neuronal-basierte Übersetzungssysteme fokussiert werden und das Sprachenpaar Deutsch-Spanisch behandelt werden.

Es ist anzunehmen, dass zunächst sowohl das maschinelle Übersetzungstool *DeepL* als auch die HumanübersetzerInnen professionell brauchbare Übersetzungen liefern, sich beide Methoden jedoch durch unterschiedliche Stärken und Schwächen in verschiedenen Punkten auszeichnen. Während die Stärken der maschinellen Übersetzung (MÜ) erwartbarerweise in den Bereichen Kosten und Zeitaufwand liegen, so kann die Humanübersetzung in puncto Genauigkeit der übersetzten Fachbenennungen und somit Qualität der Übersetzung punkten. So kann durch die Terminologierecherche vorab genauer auf jeden einzelnen Terminus eingegangen werden. Dadurch, dass neuronale Übersetzungstools im Hintergrund auf verschiedene Korpora zurückgreifen können, die mit unterschiedlichen aber vor allem auch mit wirtschaftlichen Texten gespeist werden, ist naheliegend, dass das Übersetzungstool *DeepL* bei solchen Fachtexten eher mit hoher Qualität punkten kann als bei geisteswissenschaftlichen Texten, bei denen das Einbeziehen des Kontexts und genaueste Terminologierecherche eine große Rolle spielen.

Die zugrundeliegenden Hypothesen lauten daher wie folgt:

*H1: Humanübersetzungen weisen eine höhere Treffsicherheit in Bezug auf die Übersetzung von Fachtermini auf als maschinelle Übersetzungstools.*

*H2: Neuronale maschinelle Übersetzungstools weisen in der Übersetzung von wirtschaftlichen Fachtexten eine höhere Treffsicherheit in Bezug auf die Übersetzung von Fachtermini auf als in der Übersetzung von geisteswissenschaftlichen Fachtexten.*

Ohne konkrete Literatur zu nennen, auf denen meine Überlegungen zugrunde liegen ist die Tatsache, dass bei Geisteswissenschaftlichen Texten der Kontext für die adäquate Übersetzung von Termini besonders wichtig ist, für mich als Übersetzerin mit dem Schwerpunkt Fachübersetzen ganz klar. Auch wenn geisteswissenschaftliche Texte meist auf gemeinsprachlichen Lexemen basieren, so werden ihnen doch auch fachliche Bedeutungen zugeordnet. Beispielsweise bekommen diese durch Metaphern eine ganz neue Bedeutung. Die fachliche Bedeutung der gemeinsprachlichen Lexeme muss also erstmal identifiziert werden. Das kann nur durch das Miteinbeziehen des gesamten Kontextes erfolgen. Daher ist anzunehmen, dass MÜ-Tools bei solchen Texten eine höhere Fehlerquote aufweisen, da sie nur einzelne Satzbauteile aber nicht den ganzen Text bei der Übersetzung miteinbeziehen

können.

Dazu kommt, dass gewisse Fachtermini häufig recherchiert werden müssen, da sie aus anderen sozialen, kulturellen oder fachlichen Bereichen stammen. Auf diese Terminologierecherche können ÜbersetzerInnen kaum verzichten. Daher stelle ich die Hypothese auf, dass die Treffsicherheit neuronaler maschineller Übersetzungstools vor allem in Bezug auf Fachtermini sehr viel geringer sein muss als bei Humanübersetzungen. Neuronale Maschinelle Übersetzungstools, darunter auch DeepL, werden zudem mit großen Textkorpora trainiert. Dabei wird unter anderem auf wirtschaftliche Texte (beispielsweise Dokumente der Europäischen Union) zurückgegriffen. Dadurch kann abgeleitet werden, dass Übersetzungssysteme mit Fachtermini, die dem Trainingskorpus ähneln (beispielsweise aus wirtschaftlichen Texten), weniger Probleme haben als Termini anderer Fachbereiche, die nicht so viel trainiert wurden (beispielsweise aus dem Fachbereich Geisteswissenschaften).

## **1.1 Aufbau & Vorgehensweise**

Im Theorieteil soll zunächst auf die Definition und den geschichtlichen Hintergrund maschineller Übersetzung eingegangen werden, um heutige Entwicklungen hinsichtlich der Forschungsfrage kontextualisieren zu können. Anschließend soll ein allgemeiner Überblick über die Klassifizierung verschiedener Ansätze und Methoden hinsichtlich maschineller Übersetzungssysteme gegeben sowie gängige Übersetzungssysteme und Übersetzungsstrategien und die mit diesen verbundenen Herausforderungen präsentiert werden. Der Fokus dieser Arbeit soll jedoch speziell auf die neuronale Übersetzungsstrategie ausgerichtet sein, weshalb im darauffolgenden Kapitel das neuronale Übersetzungstool *DeepL* näher beschrieben wird. Im Anschluss daran wird auf das Zusammenspiel von Terminologie und Translation in der MÜ eingegangen, bevor der Korpus und das methodologische Vorgehen der Evaluation beschrieben werden.

Diese Masterarbeit ist empirisch aufgebaut. So werden durch die Auswertung und Interpretation von Übersetzungen ausgewählter Fachtexte Erkenntnisse gewonnen, die Aussagen über die Genauigkeit der Übersetzung von Fachbegriffen zulassen. Als Untersuchungskorpus wurden verschiedene Texte aus den Bereichen Geisteswissenschaft, Technik und Naturwissenschaft ausgewählt, um die Übersetzungsleistungen in verschiedenen Fachbereichen miteinander vergleichen zu können. Die Texte werden zum einen von HumanübersetzerInnen (fortgeschrittene Studierende aus dem Master Translation mit dem

Schwerpunkt Fachübersetzen) und zum anderen – ohne Post- und Präedition – von dem neuronal-basierten Übersetzungstool *DeepL* übersetzt. In Kapitel 8 werden die Ergebnisse der Terminologierecherche präsentiert. Die Auswertung der vier Fachtexte in Bezug auf den gesamten Text wird in Form einer tabellarischen Auflistung präsentiert. Zur Evaluierung der gesamten maschinellen Übersetzung wird die Qualitätsmetrik SAE-J2450. Mit der Gründung einer Arbeitsgruppe SAE-J1450 gelang es der Society of Automotive Engineers (SAE) den Status einer SAE Norm zu erlangen, welche zur Errechnung eines gewichteten, numerischen Ergebnisses für die angemessene Darstellung der Translationsqualität verwendet werden kann. (vgl. Dalla-Zuanna 2010:23)

Daraus resultieren auch die folgenden verwendeten Fehlerkategorien: Terminologiefehler, Inhaltlicher Fehler, Auslassung, Strukturfehler, Rechtschreibfehler, Interpunktionsfehler, Sonstige Fehler.

Die Übersetzung der einzelnen Termini und deren Qualität mittels maschineller Übersetzung werden anschließend tabellarisch mit den Ergebnissen der Terminologierecherche und Humanübersetzung verglichen. Abschließend folgt die daraus resultierende Schlussfolgerung mit Zusammenfassung und Beantwortung der Forschungsfrage.

Die beschriebene Vorgehensweise wurde von mir ausgewählt, um die Forschungsfrage bestmöglich beantworten zu können. Durch die Verwendung der genannten Qualitätsmetrik und der anschließenden tabellarischen Auflistung kann auf die verschiedenen Abweichungen einzeln und detailliert eingegangen werden. Dadurch können besonders gelungene Übersetzungen aber auch Fehler hervorgehoben werden. Diese Vorgehensweise entspricht quasi einer menschlichen Postedition und ist dementsprechend sehr gut zur Evaluation der Qualität geeignet.

Die Qualität einer Übersetzung wird in diesem Zusammenhang zunächst als gut empfunden, wenn möglichst wenige Beispiele bei der Evaluation in die unterschiedlichen Fehlerkategorien eingetragen werden. Die Fehler in den verschiedenen Kategorien werden unterschiedlich gewichtet. Nach der Auswertung des gesamten Textes kann ein Fazit gezogen werden. Befindet sich die Übersetzung mit all seinen Fehlerpunkten noch unterhalb der Toleranzgrenze oder überschreitet er diese bereits? Bei der Qualität einer Übersetzung sollte zudem nicht vergessen werden, dass diese natürlich auch immer vom Zweck und der Erwartung der EmpfängerInnen abhängt. Denn nicht jede Abweichung gleicht einem Fehler. Wenn der Anspruch in Bezug auf die Qualität ein sehr hoher ist, da der Text beispielsweise veröffentlicht werden soll, werden bestimmte Äußerungen verbessert oder ausgetauscht

obwohl die Übersetzung an sich eigentlich richtig gewesen wäre. Dadurch verschlechtert sich natürlich der Score der Übersetzung. Weitere Informationen dazu folgen in den entsprechenden Kapiteln.

## 2. Was ist maschinelle Übersetzung?

In diesem Kapitel soll erklärt werden, worum es sich bei der maschinellen Übersetzung handelt. Da seit dem Beginn der Forschung im Bereich maschineller Übersetzung viele verschiedene Definitionen entstanden sind, wird zunächst auf Definitionen der Anfangszeit eingegangen, um abschließend die für diese Arbeit relevante Definition vorzustellen. Im Anschluss daran wird auf drei Unterkategorien der maschinellen Übersetzung näher eingegangen, die zuvor nach Ausmaß des menschlichen Eingreifens eingeteilt wurden.

### 2.1 Begriffsdefinition

Wissenschaftler Jehoschua Bar-Hillel wurde für seine grundlegenden Arbeiten zur maschinellen Übersetzung bekannt und im Jahre 1951 als erster Forscher für seine fachliche Auseinandersetzung ausgezeichnet. Die Idee war anfangs, ein System zur vollautomatischen Anfertigung von Übersetzungen zu entwickeln. Die Qualität einer derartigen Übersetzung sollte einer Übersetzung durch eine/n menschliche/n ÜbersetzerIn entsprechen und weder vor- noch nachbearbeitet werden. Bar-Hillel war dabei zu der Überzeugung gelangt, dass *fully automatic high quality translation* (FAHQT), wie er es nannte, nicht nur in naher Zukunft, sondern auch insgesamt unerreichbar sei (vgl. Bar-Hillel 1960:3). Außerdem müssten menschliche Prä- und Posteditoren zumindest zum Auflösen von Ambiguitäten und zur stilistischen Überarbeitung die Vor- und Nachbearbeitung des Textes übernehmen (vgl. Hutchins 1986:30-33).

Das bekannteste Ereignis in der Geschichte der maschinellen Übersetzung ist zweifelsohne die Veröffentlichung des ALPAC-Berichts. Abermals hoffte man, basierend auf einem maschinenlesbaren Ausgangstext mit Hilfe eines Algorithmus einen brauchbaren Zieltext ohne menschliches Eingreifen erstellen zu können. Die Verfasser des ALPAC-Berichtes richteten jedoch eine klare Botschaft an die Öffentlichkeit und den Rest der wissenschaftlichen Gemeinschaft: Maschinelle Übersetzung sei weiterhin aussichtslos. Im Jahre 1966 wurde die Finanzierung der Forschung im Bereich maschineller Übersetzung in den Vereinigten Staaten vorerst beendet (vgl. ALPAC 1966:19).

Infolgedessen setzte sich ein realistischerer Blick auf das Potential der maschinellen Übersetzung durch und die Akzeptanz eines zusätzlichen menschlichen Eingreifens bei der Vor- und Nachbearbeitung der maschinellen Übersetzung nahm dementsprechend zu. In der

Definition von John Hutchins aus dem Jahre 1995 wird sehr klar geschildert, dass sowohl automatisierte Übersetzungsvorgänge inklusive menschlicher Unterstützung als auch vollkommen automatisierte Computersysteme unter den Begriff *maschinelle Übersetzung* fallen (vgl. Hutchins 1995:431).

“The term ‚machine translation‘ (MT) refers to computerized systems responsible for the production of translations with or without human assistance. It excludes computer-based translation tools which support translators by providing access to on-line dictionaries, remote terminology databanks, transmission and reception of texts, etc.” (Hutchins 1995:431)

An diesem Punkt muss jedoch erwähnt werden, dass sogenannte *CAT Tools* (computer-aided translation) nicht in den Bereich der maschinellen Übersetzung fallen. *CAT-Tools*, wie beispielsweise Translation-Memory-Systeme, Terminologie-Datenbanken, elektronische Wörterbücher, Alignment-Tools oder Textverarbeitungsprogramme, sind Übersetzungshilfen und unterstützen lediglich die ÜbersetzerInnen beim Übersetzungsprozess, während maschinelle Übersetzungsprogramme imstande sind, selbstständig vollständige Übersetzungen anzufertigen. Auch wenn die Qualität der vollständig automatischen maschinellen Übersetzung nicht immer der gewünschten Qualität entspricht, so können sie dennoch ohne menschliches Eingreifen eine Übersetzung erstellen (vgl. Ramlow 2009:118).

## **2.2 Klassifizierung nach Anteil des menschlichen Eingreifens in den Übersetzungsprozess**

Da ForscherInnen vor allem nach der Veröffentlichung des ALPAC Berichts bewusst geworden war, dass eine vollautomatische maschinelle Übersetzung nicht der gewünschten Qualität entsprechen würde und dass menschliches Eingreifen beim Prozess der maschinellen Übersetzung auch weiterhin notwendig sein würde, wurden die verschiedenen Arten der maschinellen Übersetzung nach Grad des menschlichen Eingreifens typisiert. Im Folgenden werden drei Kategorien vorgestellt.

## 2.2.1 Maschinelle Übersetzungshilfen

“We might think of MAHT as a system in which the human translator has control; the machine is simply a tool to be used at the discretion of the translator.” (Lehrberger/Bourbeau 1988:6)

Maschinelle Übersetzungshilfen, auch bekannt als *Machine-Aided Human Translation (MAHT)*, sind im Grunde menschliche Übersetzungen mit nur begrenzter Unterstützung durch die Maschine. Die Maschine besteht aus einem einfachen Textverarbeitungsgerät, das Übersetzungsäquivalente von Wörtern der Ausgangssprache nachschlagen kann. Dies mag zwar schneller vonstattengehen als das Ausschreiben der Übersetzung von Hand (oder das Eintippen mit einer gewöhnlichen Schreibmaschine) und das Durchblättern eines Wörterbuchs nach unbekanntem Begriffen, jedoch wird dem/der ÜbersetzerIn nicht die Last abgenommen, die Übersetzung tatsächlich selbst durchzuführen (vgl. Lehrberger/Bourbeau 1988:6).

Grundsätzlich müssen maschinelle Übersetzungshilfen eine computergestützte linguistische Hilfe beinhalten, wie z.B. ein Programm zur Überprüfung von Rechtschreibung, Grammatik oder Stil der Übersetzung. Die meisten Rechtschreibprüfer basieren auf sehr großen Wörterbüchern, mit denen ein sehr schnelles Nachschlagen von Wörtern ermöglicht wird. Allerdings verfügen derartige Systeme über keine linguistische Intelligenz, um beispielsweise im Englischen *there* von *their* als eine Art Schreibfehler hervorzuheben. Denn Tools zur Grammatik- und Stilüberprüfung sind etwas ausgefeiltere Werkzeuge. Während Grammatikprüfer nach Fehlern, wie der Nichtübereinstimmung von Subjekten und Verben oder Wortwiederholungen, suchen, filtern Stylechecker Merkmale, die als stilistisch unpassend empfunden werden – etwa Klischees oder Sätze, die mit Konjunktionen beginnen oder mit Präpositionen enden – heraus. Online-Nachschlagewerke, wie zweisprachige Wörterbücher und Enzyklopädien, sind Informationsquellen, die ÜbersetzerInnen konsultieren können. Auch die Hinzuziehung interner und externer terminologischer Datenbanken, die Suche nach früheren Übersetzungen zu ähnlichen Themen bzw. derselben KundInnen oder die Möglichkeit, Kontexte anzuzeigen, in denen ein Wort in den zu übersetzenden Texten vorkommt, kann den ÜbersetzerInnen viel Arbeit ersparen: Letztlich verbringen technische ÜbersetzerInnen schätzungsweise bis zu 60 Prozent ihrer Arbeitszeit mit der Konsultation von Wörterbüchern und Nachschlagewerken zur Terminologierecherche (vgl. Hutchins/Somers 1992:149).

## 2.2.2 Maschinengestützte Übersetzung

Bei der maschinengestützten Übersetzung, unter anderem bei Lehrberger/Bourbeau (1998) bekannt als *Human-Aided Machine Translation (HAMT)*, ist der Anteil des menschlichen Eingreifens beim Prozess der Übersetzung im Vergleich zur maschinellen Übersetzungshilfe geringer. Es handelt sich um die Art von Übersetzung, bei der primär das maschinelle Übersetzungssystem die Übersetzung anfertigt (vgl. Ramlow 2009:115).

Bei der HAMT unterstützt der Mensch die Maschine nur im Falle des Auftretens von bestimmten Schwierigkeiten bei der Übersetzungstätigkeit. Dies kann sowohl bei der Prä- und Postedition als auch bei der Interaktion während des Übersetzungsprozesses stattfinden. Bei der Präedition werden beispielsweise vorab Textstellen herausgesucht, die der Übersetzungs-Software Probleme bereiten könnten. Dies betrifft vor allem Eigennamen oder komplizierte syntaktische Strukturen und Mehrdeutigkeiten, die zu den zentralen Problemen der MÜ zählen. Während der Übersetzungstätigkeit kann das System die ÜbersetzerInnen etwa nach einer grammatikalischen Kategorie eines bestimmten Wortes fragen. Nachdem der Ausgangstext hinreichend für die Übersetzungs-Software vorbereitet wurde, wird dieser maschinell übersetzt. Das Ergebnis ist eine Rohübersetzung, welche nun durch HumanübersetzerInnen korrigiert und in eine adäquate zielsprachliche Form gebracht wird, sodass er den Anforderungen des Auftraggebers / der Auftraggeberin gerecht wird (vgl. Schäfer 2002:31 und Ramlow 2009:115).

“The boundaries between Machine-Aided Human Translation (MAHT) and Human-Aided Machine Translation (HAMT) are often uncertain and the term Computer-Aided (or Computer-Assisted) Translation (both CAT) can sometimes cover both.“ (Hutchins/Somers 1992:3)

Die Grenze zwischen HAMT und MAHT ist nur schwer zu ziehen. So könnten die EntwicklerInnen eines interaktiven Systems dieses als *maschinelle Übersetzung* betiteln, während eben jenes System nach wissenschaftlichen Kriterien als *Humanübersetzung mit maschineller Übersetzungshilfe* bezeichnet werden würde, wenn die Maschine ein gewisses Maß an Unterstützung benötigt. In diesem Fall kann das System, ungeachtet der von den DesignerInnen des Systems erhobenen Ansprüche, eher als MAHT denn als HAMT eingestuft werden (vgl. Lehrberger/Bourbeau 1988:7).

Wie an diesem Beispiel deutlich wird, handelt es sich nur um einen Versuch, die verschiedenen maschinellen Übersetzungstools in Kategorien einzuordnen, deren Grenzen, je nach Text und subjektivem Empfinden des Übersetzers / der Übersetzerin, auch fließend sein können. Der weiter oben erwähnte Begriff der *Computer-Aided / Computer-Assisted Translation (CAT)* kann oftmals beide Kategorien abdecken (vgl. Hutchins/Somers 1992:147).

### 2.2.3 Vollautomatische Übersetzung

Wie weiter oben nachzulesen, war Bar-Hillel schon im Jahre 1951 davon überzeugt, dass eine vollautomatische Übersetzung - er prägte den Begriff *Fully Automatic Machine Translation (FAMT)* - nicht der Qualität einer Humanübersetzung entspreche, da es bei vollautomatischen Übersetzungssystemen unter anderem zu semantischen Problemen kommen könnte (vgl. Hutchins/Somers 1992:148): “We can have either fully automatic translation or high quality (computer-based) translation, but we cannot have both.“ (Hutchins/Somers 1992:149)

Heutzutage versteht man unter der Abkürzung FAMT eine Form der maschinellen Übersetzung, bei welcher der Mensch während des gesamten Übersetzungsprozesses – angefangen von der Texteingabe bis hin zur Rohübersetzung – nicht eingreift. Die Postedition von Rohübersetzungen ist in den meisten Fällen dennoch üblich (vgl. Schäfer 2002:31). Das wirft die Frage auf, wo genau der Unterschied zwischen maschinengestützter und vollautomatischer Übersetzung liegt, da beide Begriffe sehr ähnlich definiert werden. Es ist anzunehmen, dass diese letzte Kategorie der Vollständigkeit halber existiert, um den ForscherInnen rund um maschinelle Übersetzung die Möglichkeit auf ein zukünftig vorhandenes vollautomatisches Übersetzungssystem, welches ohne Prädiction, Postedition und Interaktion während des Übersetzungsprozesses eine qualitativ hochwertige Übersetzung erzeugen kann, offen zu halten.

Das FAMT-System *METEO* aus Kanada, welches vorwiegend Wettervorhersagen übersetzt, ist jedoch ein Beispiel für eine Ausnahme: Hier wird die Rohübersetzung nicht mehr von Menschen nacheditiert. Die Maschine selbst entscheidet, welche der ihr vorgelegten Sätze redigiert werden sollen; alle anderen werden übersetzt und der gebrauchsfertige Text wird als geeignet angesehen (vgl. Lehrberger/Bourbeau 1988:8).

## 2.3 Schlussfolgerung und für diese Arbeit relevante Definition

Die unterschiedlichen Definitionen, die erläutert worden sind, zeigen, dass sich die Auslegung des Begriffs der maschinellen Übersetzung im Laufe der Zeit maßgeblich verändert hat. Auch wenn die drei vorgestellten Kategorien (maschinelle Übersetzungshilfe bzw. Machine-Aided Human Translation / maschinengestützte Übersetzung bzw. Human-Aided Machine Translation / vollautomatische Übersetzung bzw. Fully Automatic Machine Translation) für die Einteilung maschineller Übersetzung vorgeschlagen und auch in der Literatur umgesetzt werden, so kann eine Abweichung von diesen dennoch stattfinden. In der Praxis können mit den momentan existierenden maschinellen Übersetzungssystemen nur die ersten beiden Kategorien erfüllt werden.

Unter maschinellen Übersetzungssystemen werden in dieser Arbeit derartige Systeme verstanden, die in die zweite Kategorie einzuordnen sind (maschinengestützte Übersetzung bzw. Human-Aided Machine Translation). Solche Systeme, die automatisch Rohübersetzungen generieren, die infolge durch HumanübersetzerInnen posteditiert werden müssen, um die gewünschte qualitativ hochwertige Übersetzung zu erhalten, werden demzufolge hier unter die Definition subsumiert. Die Präedition des Textes und Interaktion während des Übersetzungsprozesses ist im Rahmen dieser Definition nicht verpflichtend.

### **3. Geschichtlicher Hintergrund**

Im folgenden Kapitel wird auf den geschichtlichen Hintergrund maschineller Übersetzung eingegangen. Dabei sollen die Anfänge der maschinellen Übersetzung erläutert werden. Unter anderem sollen die Fragen beantwortet werden, wann, wie und warum es das erste Mal überhaupt zur Verwendung von technischen Hilfsmitteln bei der Übersetzung gekommen ist. Ein weiteres Augenmerk liegt auf Rolle des ALPAC-Berichts im Bereich der maschinellen Übersetzung, deren weiterer Entwicklung danach und dem gegenwärtigen Stand der Forschung.

#### **3.1 Anfänge der maschinellen Übersetzung**

Die Verwendung von mechanischen Vorrichtungen, um Sprachbarrieren zu überwinden, wurde erstmals im 17. Jahrhundert vorgeschlagen. Die Idee dahinter war, sowohl die internationale Kommunikation zu verbessern als auch ein logisches und nachvollziehbares Mittel für die Verständigung im wissenschaftlichen Bereich zu erschaffen. Daraufhin kamen Cave Beck, Athanasius Kircher und Johann Joachim Becher auf die Idee, mechanische Wörterbücher mit Hilfe von numerischen Codes zu kreieren. Dabei sollten die lexikalischen Äquivalente aller bekannten Sprachen die gleiche Codenummer erhalten. Es wurde jedoch weder darauf geachtet, die möglichen Äquivalente, die es in anderen Sprachen geben könnte, miteinander zu verknüpfen, noch darauf, Mittel / Systeme zur Bewältigung und Klassifizierung syntaktischer Unterschiede bereitzustellen. Die Vorschläge von Beck, Kircher und Becher gelten mittlerweile als Vorläufer der maschinellen Übersetzung – auch wenn deren Ideen eher als technische Hilfsmittel zu bezeichnen sind, da sie alle erforderten, dass der/die menschliche ÜbersetzerIn die zur Verfügung gestellten Werkzeuge in einer „mechanischen“ Art und Weise benutzte; sprich: dass der Mensch eine Maschine simulierte und der Übersetzungsprozess nicht automatisiert funktionierte (vgl. Hutchins 1986:21f.).

Zu diesem Schritt der Automatisierung kam es erst mit der Erfindung mechanischer Rechenmaschinen im 19. und 20. Jahrhundert. Dadurch wurde ein automatisches Gerät, das einige Übersetzungsvorgänge selbstständig ausführen konnte, erstmals denkbar. So kam es in den 1930er Jahren gleich zu zwei expliziten Vorschlägen, als der Franzose Georges Artsrouni und der Russe Petr Trojanski unabhängig voneinander Patente auf ihre Übersetzungsmaschinen erhielten. Die Erfindung Artsrounis bestand aus einer

motorbetriebenen mechanischen Vorrichtung zum Aufzeichnen und Abrufen von Informationen auf einem breiten Papierstreifen, der hinter einer Tastatur vorbeilief. Das Speichermedium fasste dabei mehrere tausend Zeichen. Jede Zeile sollte das Eintragungswort und zugehörige Äquivalente in mehreren Sprachen enthalten; jedem Eintrag entsprachen dabei kodierte Perforationen auf einem zweiten Band aus Papier oder Metall, das als Selektionsmechanismus fungierte. Der erforderliche Eintrag wurde an der Tastatur gelocht, die Perforationen lokalisiert und der vollständige Eintrag innerhalb von wenigen Sekunden abgerufen. Ein Maschinenprototyp wurde 1937 ausgestellt und vorgeführt. Die französische Eisenbahnverwaltung und der Post- und Telegraphendienst zeigten großes Interesse, der Beginn des Zweiten Weltkriegs verhinderte jedoch die tatsächliche Inbetriebnahme (vgl. Hutchins 2006:1).

Geschichtlich bedeutender war rückblickend das an Trojanski erteilte Patent für eine Maschine zum Auswählen und Drucken von Wörtern beim Übersetzen von einer Sprache in eine oder mehrere andere Sprachen. Dabei waren drei Stufen des Übersetzungsprozesses vorgesehen. Die Maschine war nur in der zweiten Stufe als automatisiertes Wörterbuch entwickelt worden; in der ersten Stufe sollte ein/e menschliche/r RedakteurIn, der/die nur die Ausgangssprache kannte, den eingegebenen Text in eine bestimmte logische Form bringen. Dabei sollten alle flektierten Wörter durch ihre Grundformen ersetzt und ihre syntaktischen Funktionen dem Satz zugeschrieben werden. In der zweiten Stufe wurden sodann die Sequenzen des Ausgangstextes in die Sequenzen der Zielsprache, die ebenfalls in ihrer Grundform angegeben und mit Symbolen versehen wurden, durch die Maschine übersetzt. In der dritten Stufe sollte anschließend ein/e Redakteur/in, der/die nur die Zielsprache kannte, diese Sequenz in die gängigen Ausdrucksformen seiner eigenen Sprache umwandeln. Obwohl die Maschine nur als mechanisches Wörterbuch fungierte, glaubte Trojanski fest daran, dass er eine Maschine entwickeln könnte, die auch den Prozess der logischen Analyse übernehmen könnte (vgl. Hutchins 1986:22f.).

### **3.2 Die 1950er Jahre und der ALPAC-Bericht**

Im Jahre 1942 kam es zur Entwicklung des ersten elektronischen Universalcomputers, der unter anderem gebaut wurde, um Militärkommunikation zu entschlüsseln und ansonsten vorwiegend im Bereich der Mathematik und Physik Verwendung fand. Schon bald wurde das enorme Potential des „elektronischen Gehirns“ erkannt und es kam 1946 zu ersten

Andeutungen von Andrew D. Booth und Warren Weaver über die mögliche Verwendung des elektronischen Computers zur Übersetzung.

Im Weiteren wies Weaver 1947 darauf hin, dass er der Finanzierung eines Computers für die Universität London durch die Rockefeller Stiftung (deren Vizepräsident er war) nur zustimmen würde, wenn zusätzliche Ideen für den Einsatz des Gerätes in einem nicht-numerischen Kontext miteingebracht werden würden, sprich: für die Verwendung des Computers für die Anfertigung von Übersetzungen (vgl. Hutchins 1986:27). Zusammen mit seiner Assistentin Kathleen Britten arbeitete Booth an ersten Ideen, darunter an der Erkennung von Tönen, gedruckten Zeichen und auch an einem geeigneten Code-System. Anschließend entwickelte R. H. Richens, mit dem Booth bereits an Experimenten zu mechanischen Wörterbüchern zusammengearbeitet hatte, ein Übersetzungssystem, welches erstmals eine automatische grammatikalische Analyse von Wortendungen beinhaltete. Er schlug dabei die Segmentierung von Wörtern in ihre Wortstämme und Endungen vor, sowohl um die Größe von Wörterbüchern zu verringern als auch um grammatische Informationen in ein Wörterbuch-Übersetzungssystem einzubringen. Auf diese Weise sollten grammatikalische Anmerkungen die bisher strenge Wort-für-Wort-Wörterbuch-Übersetzung erweitern (vgl. Schwanke 1991:69f.).

Die Validität des Ansatzes wurde von Hand und mit Hilfe von Lochkartenmaschinen an einer Vielzahl von Sprachen getestet und das Ergebnis in einem im Jahre 1949 verfassten Memorandum veröffentlicht. Weavers Memorandum konzentrierte sich dabei mehr auf die allgemeinen Strategien und langfristigen Ziele von maschineller Übersetzung als auf die technischen Probleme, mit denen sich Booth und Richens auseinandergesetzt hatten. Dabei arbeitete er die folgenden vier Hauptprobleme heraus: das Problem der Mehrdeutigkeit, die logische Grundlage der Sprache, die Anwendung der Kommunikationstheorie und kryptographischer Techniken sowie der Einsatz von Sprachuniversalien. Dennoch war sich Weaver sicher, dass die Probleme des Übersetzens mit einem Computer technisch lösbar seien. Aufgrund der Veröffentlichung rückte das Fachgebiet der MÜ plötzlich in das Interesse einer breiten Öffentlichkeit (vgl. Hutchins 1986:28).

Innerhalb weniger Jahre wurden umfassende Forschungsprojekte an einer Reihe von US-Zentren, wie beispielsweise am Massachusetts Institute of Technology (MIT), der University of Washington, der University of California (UCLA), dem National Bureau of Standards (NBS) und der RAND Corporation, losgetreten. Der Sinologe Erwin Reifler erkannte das Problem der Mehrdeutigkeit als Hindernis für eine Wort-für-Wort-Übersetzung, wie sie von Booth und Richens versucht worden war, und führte im Jahre 1950 das Konzept

der Prä- und Postedition ein. Der menschliche „Pre-Editor“ sollte fortan den Text für die Eingabe in den Computer vorbereiten, der „Post-Editor“ Restprobleme lösen sowie den Stil der Übersetzung überarbeiten (ibid:28).

Im Jahre 1951 wurde Yehoshua Bar-Hillel, der am Massachusetts Institute of Technology (MIT) tätig war, zum ersten Vollzeitforscher im Bereich der maschinellen Übersetzung ernannt. Seine Aufgabe bestand darin, die Möglichkeiten von *Machine Translation* (MT) zu untersuchen und die zukünftige Forschung zu planen. Ein Jahr später berief er die erste MT-Konferenz ein, in Zuge derer die Grundzüge der künftigen Forschung diskutiert werden sollten. In Vorbereitung darauf verfasste Bar-Hillel einen Bericht über den Stand der Forschung, der bis heute einen Meilenstein in der Entwicklung der MT darstellt. Obwohl er sich keine große Hoffnung auf hochgenaue und vollautomatische Übersetzung in absehbarer Zeit machte, betonte er, dass der Computer trotzdem als mechanisches Hilfsmittel für die Übersetzung eingesetzt werden könne. Daher prüfte Bar-Hillel verschiedene Optionen für gemischte Übersetzungsprozesse, bei denen menschliche Expertise in Form eines Post-Editors eingreifen sollte, um, unter anderem, semantische Zweideutigkeiten zu beseitigen. Alle TeilnehmerInnen der Konferenz waren sich über die Notwendigkeit einer Nachbearbeitung einig. Viele stimmten mit Reifler sogar darin überein, dass auch eine Art Vorredaktion unerlässlich sei, um Mehrdeutigkeiten und syntaktische Komplexitäten der Ausgangstexte zu minimieren. Als Alternative wurde ein Training der AutorInnen dahingehend, mit MT im Hinterkopf zu schreiben, angedacht (etwa durch Verwendung einer bewusst kontrollierten Sprache) (vgl. Hutchins 1997:23).

Das bedeutendste Ergebnis der MIT-Konferenz im Juni 1952 war zweifellos die positive Entscheidung von Leon Dostert über ein Projekt zur Demonstration der Machbarkeit eines maschinellen Übersetzungssystems in Zusammenarbeit mit IBM. Ein System mit kleinem Umfang zur Übersetzung einiger russischer Sätze ins Englische wurde entwickelt, bei dem ein stark begrenzter Wortschatz von 250 Wörtern und nur sechs Grammatikregeln verwendet wurden. Im Januar 1954 folgte im New Yorker Hauptsitz von IBM die Demonstration. Die breite Öffentlichkeit zeigte sich derartig beeindruckt, dass die MT nun doch als realisierbares Ziel angesehen wurde. Die US-Regierungsbehörden wurden dazu ermutigt, die Forschung im nächsten Jahrzehnt im großen Stil zu unterstützen; in anderen Staaten, insbesondere in der Sowjetunion, wurden zahlreiche MT-Gruppen gegründet (vgl. Ramlow 2009:59). Auch wenn viele Gruppen nun aktiv an verschiedenen Ansätzen forschten und die Zuversicht und Hoffnung auf einen schnellen Erfolg zunächst groß gewesen waren, so traten doch vermehrt technische als auch linguistische Schwierigkeiten auf, die nur schwer

lösbar waren und so den anfänglichen Optimismus bremsen. Die Kritik an der MT wuchs zunehmend und wurde von Jahr zu Jahr lauter (vgl. Bar-Hillel 1960:3).

Im Februar 1959 veröffentlichte Bar-Hillel seinen Bericht über den Stand der maschinellen Übersetzung, der auf Grundlage persönlicher Besuche bei den meisten MÜ-Gruppen in den USA und anhand von Informationen, die er als Antwort auf ein Rundschreiben erhalten hatte, erstellt worden war, in den Vereinigten Staaten und Großbritannien. Er erklärte, dass die MÜ-Forschung mit ihrem Vorhaben, eine vollautomatische, maschinelle und qualitativ hochwertige Übersetzung zu ermöglichen, ein unerreichbares Ziel verfolge, welches weder in naher Zukunft zu einem späteren Zeitpunkt erreichbar sei. Das Hauptargument drehte sich um die mangelhafte Qualität der vollautomatischen Übersetzungen, die der eines/einer professionellen menschlichen ÜbersetzerIn keineswegs entspreche. Stattdessen sollten die Ziele weniger ehrgeizig formuliert werden und auf die Realisierung von Systemen hinarbeiten, die menschliche Zusammenarbeit erfordern. Zwar gebe es viele Situationen, in denen eine eingeschränkte Übersetzungsqualität durchaus zufriedenstellend ist, jedoch stelle sich dabei die Frage, ob sich der vergleichsweise hohe Aufwand lohne, um sich sodann mit einem geringeren Standard als jenem einer durchschnittlichen menschlichen Übersetzung zufrieden zu geben (vgl. Hutchins 1986:28).

Da von nun an viele SponsorInnen nicht mehr bereit waren, die MÜ-Forschung zu unterstützen und Forschungsgelder dementsprechend immer knapper wurden, mussten einige Forschungsgruppen eingestellt werden. Um über die Zukunft der Forschung entscheiden zu können, wurde im Jahr 1964 im Auftrag der amerikanischen Regierung ein Komitee, das *Automatic Language Processing Advisory Committee (ALPAC)*, gegründet, welches den Stand und die allgemeinen Zukunftsperspektiven der maschinellen Übersetzung überprüfen sollte. Der Ausschuss führte unter anderem Studien hinsichtlich Nachfrage, Brauchbarkeit von bereits bestehenden Übersetzungssystemen und Kosten von Übersetzung inkl. Prä- und Postedition durch. Die Ergebnisse waren niederschmetternd. Der im Jahre 1966 veröffentlichte ALPAC-Bericht besagte unter anderem, dass das Angebot an ÜbersetzerInnen die Nachfrage bei weitem übersteige und dementsprechend keine Notwendigkeit für MÜ-Systeme gegeben sei. Auch könne keine Übersetzung von zufriedenstellender Qualität maschinell erstellt werden und auch die Kosten sowie der Posteditationsaufwand waren nachweislich deutlich höher als die der HumanübersetzerInnen. Dabei stütze sich die Studie ausschließlich auf Übersetzungen aus dem Russischen und machte wirtschaftliche Faktoren zum bestimmenden Indikator für das Gutachten (vgl. Ramlow 2009:59f.).

Der ALPAC- Bericht wurde zwar scharf kritisiert und von ForscherInnen als engstirnig, voreingenommen und kurzsichtig verurteilt, hatte aber nichtsdestotrotz schwerwiegende Folgen für die MÜ-Forschung. Die Forschungsgelder seitens der amerikanischen Regierung wurden fast komplett eingestellt, was zur Schließung beinahe sämtlicher US-Forschungseinrichtungen führte. Die MÜ-Forschung verlagerte sich dementsprechend in andere Länder. Abgesehen von den ökonomischen Folgen des Berichts, musste auch das Ansehen des Forschungsbereiches darunter leiden (vgl. Schäfer 2002:23).

### **3.3 Weitere Entwicklung maschineller Übersetzung und gegenwärtiger Forschungsstand**

Infolge des ALPAC-Berichts wurden realitätsnähere Ziele in Bezug auf maschinelle Übersetzungssysteme angestrebt. Hinzu kam die steigende Nachfrage an MÜ-Systemen, ausgelöst vor allem durch die Globalisierung und Europäische Institutionen, weshalb man sich auf neue Forschungsansätze konzentrierte. Dabei setzte man nicht mehr auf hochqualitative Übersetzungen, sondern strebte lediglich eine verständliche Wiedergabe des Inhalts des Ausgangstextes an. In der Anfangsphase dominierten zwar noch direkte Übersetzungssysteme, diese wurden jedoch bald durch indirekte Systeme (interlingua- und transfer-basierte Systeme) ausgetauscht (vgl. Ramlow 2009:63). Zu den direkten Übersetzungssystemen gehört beispielsweise *Logos*, ein für die US-Airforce entwickeltes System, das zur Übersetzung amerikanischer Flugzeughandbücher ins Vietnamesische verwendet wurde. Das bekannteste war jedoch *Systran*, welches zunächst als russisch-englisches System konzipiert wurde. Der interlingua-basierte Ansatz kam etwa bei einem Projekt des CETA (Centre d'études pour la traduction automatique) in Grenoble für Übersetzungen vom Russischen ins Französische und vom Deutschen ins Englische zum Einsatz. Jedoch schaffte es der Ansatz nicht, die gewünschten Erfolge einzubringen.

Daraufhin versuchte man sich mit dem Projekt EURATOM an einem transfer-basierten System. Dies sollte zur Übersetzung vom Russischen ins Englische und von Patenten vom Englischen ins Russische eingesetzt werden. Der Vorteil von transfer-basierten Systemen war, dass sie leicht um weitere Sprachenpaare ergänzt werden konnten (vgl. Hutchins 1986:175f.).

Zudem wurden Systeme entwickelt, die sich auf die Übersetzung von speziellen Fachgebieten und Arbeitsbereichen spezialisierten, sodass die damit verbundene

Terminologie beschränkt war. Als bezeichnendes Beispiel ist das System METEO zu nennen, das 1976 von der kanadischen TAUM-Forschungsgruppe fertiggestellt worden war und von der kanadischen Regierung äußerst erfolgreich zur Übersetzung von Wetterberichten eingesetzt wurde (vgl. Schäfer 2002:23).

Ab Mitte der 1970er Jahre war der Übersetzungsbedarf von Unternehmen und öffentlichen Institutionen durch die zunehmende Globalisierung so groß geworden, dass der Bedarf durch verfügbare ÜbersetzerInnen nicht mehr gedeckt werden konnte. Die Nachfrage nach MÜ-Systemen stieg stetig an (vgl. Krenz/Ramlow 2008:33f.).

Mit der Entwicklung von elektronischen Wörterbüchern und Terminologie-Datenbanken, wurde eine weitere einschneidende Hilfestellung für ÜbersetzerInnen bereitgestellt. Mit dem Einsatz von *Systran* bei der Kommission der Europäischen Union ab Mitte der 70er Jahre, der Anpassung des Systems an den Übersetzungsprozess vom Englischen ins Französische sowie die oben erwähnten neuen Entwicklungen, begann die Forschung wieder an Akzeptanz zu gewinnen (vgl. Ramlow 2009:63).

Das Projekt EUROTRA wurde von der Europäischen Gemeinschaft ins Leben gerufen, mit dem Ziel ein mehrsprachiges transfer-basiertes System zu entwickeln, welches für Übersetzungen aller Sprachen der Europäischen Gemeinschaft einsetzbar sein sollte. Das System war in der Lage, semantische, lexikalische und syntaktische Informationen einzubeziehen, nicht aber außersprachliches Wissen. Auch wenn das Projekt die Forschung in vielen europäischen Ländern ankurbelte, wurde ein Prototyp des Systems nie produziert (vgl. Hutchins 1995:483f.).

Bald schon gab es wieder internationale Konferenzen, Forschungsgelder und Projektgruppen, die über etliche Regionen der Welt verteilt waren. Vor allem kam es auch zu einer umfassenden Kommerzialisierung von Übersetzungssystemen in Japan, wo Forschungsprojekte auch trotz ernüchternder Ergebnisse des ALPAC-Berichts nie unterbrochen worden waren (vgl. Hutchins 1986:327f.).

In den 80er Jahren stütze man sich außerdem auf wissensbasierte Ansätze, um essentielles Weltwissen in den maschinellen Übersetzungsprozess einfließen zu lassen. Jener Ansatz ist stark von der Forschung im Bereich der künstlichen Intelligenz geprägt und unterliegt einer semantischen bzw. pragmatischen Sprachanalysemethode. VertreterInnen dieses Ansatzes sind der Meinung, dass man bei der maschinellen Übersetzung über die linguistische Ebene hinausgehen und sogenannte semantische Netze, also automatische Verstehensprozesse, in die Maschine einbauen solle. Denn erst durch die Fähigkeit, den Inhalt

eines Textes zu erfassen und somit Welt- und Kontextwissen miteinzubeziehen, könne ein Text adäquat übersetzt werden (vgl. Schäfer 2002:24).

Bis zum Ende der 80er Jahre waren vorrangig solche Ansätze verfolgt worden, die auf sprachlichen Regeln für die Syntaxanalyse und dem lexikalischen Transfer basierten (sg. regelbasierte Ansätze). Dieser wurde jedoch durch den statistischen Ansatz abgelöst. Das Forschungsprojekt der Gruppe IBM zeichnete sich dadurch aus, dass sie auf strenge Regeln verzichteten, dafür aber auf statistische Methoden zurückgriffen. So wurde ein Korpus von verschiedenen Berichten verwendet, um anschließend äquivalente Sätze, Wortgruppen oder einzelne Wörter beiden beteiligten Sprachen zuzuordnen. Mittels Wahrscheinlichkeitsrechnung wurde sodann ermittelt, wie hoch die Chance ist, dass ein Wort der Ausgangssprache auch einem oder mehreren Wörtern in der Zielsprache entspricht. Die Ergebnisse des Prototypensystems waren beeindruckend, denn fast die Hälfte aller zu übersetzenden Sätzen stimmte vollständig mit den Segmenten im gespeicherten Korpus überein.

Der beispielbasierte Ansatz ist dabei sehr ähnlich und zählt zusammen mit dem statistischen Ansatz zu dem korpusbasierten Übersetzungsverfahren. Der beispielbasierte Ansatz verzichtet ebenso auf sprachliche Regeln und bezieht sich meist auf bereits angefertigte Übersetzungen, die der aktuellen Übersetzung ähneln (dies geht so weit, dass einzelne Sätze identisch sein können). Dafür müssen vorher Wörter, Sätze oder Wortgruppen als solche klassifiziert werden. Um Übereinstimmungen zu finden, wird sodann auf ein semantisches Netz zurückgegriffen. Durch diesen Ansatz steigt auch die Motivation, Zieltexte in hoher Qualität zu generieren, um sicherzustellen, dass die vorgeschlagenen Phrasen aus dem Korpus ohne große Fehlerbearbeitung übernommen werden können (vgl. Hutchins 1995:440f.).

Rückblickend scheint es, dass die statistischen Ansätze der 80er Jahre ihrer Zeit voraus gewesen waren. So konzentrierte sich ein Großteil der ForscherInnen in den 1990er Jahren weiterhin auf syntaxbasierte Systeme bzw. Interlingua-Systeme, während die Forschung an statistischen Methoden für die maschinelle Übersetzung nur begrenzt weitergeführt wurde. Erst im Jahr 2000 sollte dem Ansatz wieder die volle wissenschaftliche Aufmerksamkeit zukommen (vgl. Koehn 2010:17f.).

Seit den 1990er Jahren kommt es zu einem verstärkten Einsatz von Übersetzungssystemen innerhalb von multinationalen Unternehmen, da die Gesamtheit der Produktinformationen (wie etwa Gebrauchsanweisungen, Produktbeschreibungen oder andere zugehörige Dokumente) schnell und in hoher Qualität sprachlich angepasst werden muss,

wenn das Produkt international vermarktet werden soll. Dazu kommt, dass ÜbersetzerInnen in Unternehmen und Institutionen vermehrt mit elektronischen Glossaren, Wörterbüchern und Terminologie-Datenbanken arbeiten, um, unter anderem, eine möglichst konsistente Terminologie zu verwenden.

Die Verwendung von Übersetzungssystemen ist seit den 90er Jahren jedoch nicht mehr ausschließlich auf professionelle Zwecke beschränkt. Durch die bestehende Nachfrage, auch E-Mails und Texte im privaten Bereich schnell übersetzen zu lassen, bieten seit Mitte der 90er Jahre zahlreiche Hersteller auch Online-Übersetzungsprogramme an. Hierfür wurde zunächst vermehrt auf statistische und beispielbasierte Verfahren gesetzt (vgl. Hutchins 2006:17). Da diese Ansätze nur bedingt zufriedenstellende Ergebnisse lieferten, bildete sich unter anderem die Forschung in Richtung von Hybridverfahren heraus. Die Idee dahinter war es, die besten Eigenschaften aus den regelbasierten, statistischen und beispielbasierten Ansätzen in einem einzigen System miteinander zu kombinieren (vgl. Stein 2009:7).

Während sich die Forschung in den 1980er und 1990er Jahren hauptsächlich um regelbasierte und statistische Ansätze drehte, wurde vereinzelt auch an neuronalen Netzen geforscht:

„Künstliche neuronale Netzwerke sind Rechenmodelle mit der Fähigkeit, sich anzupassen oder zu lernen, zu verallgemeinern und Daten zu gruppieren oder zu organisieren, deren Arbeitsweise auf paralleler Verarbeitung basiert“ (Kröse/van der Smagt 1996:13).

Die damals vorgeschlagenen Modelle ähneln den heutigen stark, jedoch wurde zu dieser Zeit nicht mit Datengrößen gearbeitet, die groß genug waren, um repräsentative Ergebnisse zu generieren. Die erforderliche Rechenkomplexität überstieg bei weitem die Rechenressourcen jener Ära, weshalb die Idee fast zwei Jahrzehnte lang aufgegeben wurde (vgl. Koehn 2010:5).

Die wichtigsten Voraussetzungen für die erfolgreiche Nachahmung der Milliarden an vernetzten Nervenzellen des menschlichen Gehirns (Neuronen) durch einen Computer waren an diesem Punkt erfüllt worden: Eine ausreichend hohe Rechenleistung sowie die Verfügbarkeit großer Datenmengen, mit denen neuronale Netze potentiell trainiert werden konnten, war nun gegeben. Im November 2016 präsentierte Google das erste Übersetzungssystem, welches auf neuronalen Netzen basiert und mit natürlich geformten Sätzen die Lücke zwischen menschlicher und maschineller Übersetzung schließen soll. Google konnte alle befürchteten Schwächen, wie beispielsweise die langsame Trainingsgeschwindigkeit oder die mangelhafte Handhabung von seltenen Wörtern,

erfolgreich ausgleichen. Da die neuronale maschinelle Übersetzung (NM) jedes einzelne Wort in Abhängigkeit zu allen anderen Wörtern im Satz betrachten kann, können auch Satzteile, die grammatikalisch und/oder inhaltlich voneinander abhängig sind, aber nicht nebeneinanderstehen, fehlerfrei übersetzt werden. Allgemein erzeugen NM-Übersetzungssysteme natürlich klingendere Sätze, die eine vergleichsweise kohärentere Satzstruktur, weniger Syntaxfehler und weniger Fehler in der Wortreihenfolge aufweisen (vgl. Krüger 2017:39f).

Im August 2017 trat überraschend das deutsche Unternehmen Linguee GmbH (heute DeepL GmbH) mit der Website *DeepL* in den Markt ein. *DeepL* setzt auf sogenannte konvolutionale Netze, die bisher nur für die Bilderkennung üblich waren. Der Vorteil von diesen Netzwerken ist, dass alle Wörter parallel übersetzt werden können. *DeepL* hatte zuvor bereits die Suchmaschine *Linguee* ins Web gestellt, wodurch vorab viele qualitativ hochwertige Daten gesammelt werden konnten. Auch wenn vereinzelt immer noch auf hybride Verfahren gesetzt wird, wandte sich das gesamte Forschungsgebiet im Zeitraum von ein bis zwei Jahren der neuronalen maschinellen Übersetzung zu (vgl. Wittpahl 2019:198f.).

## 4. Maschinelle Übersetzungssysteme

Im nachstehenden Kapitel werden die unterschiedlichen Arten maschineller Übersetzung, die sich im Laufe der Geschichte entwickelt haben, sowie verschiedene Ansätze und Modelle, auf denen diese basieren, vorgestellt. Dargestellte Überlegungen, wie die theoretische Frage um die Transfermethode, sind als essentiell für die eigentliche Konstruktion von maschinellen Übersetzungssystemen zu sehen und sollen daher im Folgenden veranschaulicht werden (vgl. Hutchins 1992:72).

### 4.1 Klassifizierung anhand verschiedener Ansätze und Methoden

Zu den beiden wichtigsten Ansätzen der Gegenwart zählen zunächst der regelbasierte Ansatz, der stark linguistisch orientiert ist, und der statistikbasierte Ansatz, welcher auf Wahrscheinlichkeiten beruht. Lange wurden die beiden Systeme als konkurrierende Modelle angesehen, doch durch die Weiterentwicklung von Forschung und Technik bewegte sich der Trend immer weiter in Richtung hybrider Modelle. Man ging demnach dazu über, die Vorteile der verschiedenen Ansätze in einem einzigen System miteinander zu kombinieren und kreierte so einen dafür individuellen Ansatz. Im Folgenden werden daher die relevantesten Theorien (regelbasierte, statistikbasierte und neuronale maschinelle Übersetzungssysteme) beschrieben.

Da die neuronale maschinelle Übersetzungsstrategie wesentlicher Gegenstand der Untersuchung der vorliegenden Masterarbeit ist, wird diese im darauffolgenden Kapitel nochmals aufgegriffen und ausführlicher beschrieben.

#### 4.1.1 Regelbasierte Systeme

Die regelbasierte Übersetzungsmethode (*Rule Based Machine Translation RBMT*), galt lange Zeit als die vorherrschende Methode der maschinellen Übersetzung und war in fast allen kommerziellen Systemen wiederzufinden. Die Ergebnisse hatten dabei eine große Bandbreite. Je nach Sprachenpaar, der Einarbeitung von Fachterminologie oder abhängig davon, ob es sich lediglich um ein alltagssprachliches System handelte, reichten die Ergebnisse von merkwürdig bis nützlich. Der Vorgang eines regelbasierten Systems ist dabei in Analyse, Transfer und eine abschließende Synthese aufgebaut (vgl. Stein 2009:8). Grundsätzlich kann

der Bereich der MÜ in zwei regelbasierte Ansätze geteilt werden: den direkten und indirekten Ansatz. Der indirekte Ansatz wiederum wird in die Transfer-Methode und die Interlingua-Methode aufgegliedert (vgl. Schäfer 2002:26). Diese drei Stufen haben jeweils deutliche Auswirkungen auf die Übersetzungsqualität, wobei die direkte Methode den niedrigsten und die Interlingua-Methode den höchsten Grad an Komplexität aufweist (vgl. Stein 2009:8).

Der direkte Ansatz ist historisch gesehen der erste und auch simpelste MÜ-Ansatz. Dabei handelt es sich um die klassische Wort-für-Wort-Übersetzung. Die Wörter der Ausgangssprache werden über eine syntaktische Komponente oberflächlich an die Satzstellung der Zielsprache angepasst. Da es für eine Vielzahl an Wörtern jedoch mehr als nur eine mögliche Übersetzung gibt, sind die Ergebnisse nur bedingt brauchbar. Hinzu kommt, dass viele Mehrwortbenennungen (= Wörter, die durch Leerzeichen voneinander getrennt sind, aber eine syntaktische Einheit bilden), nicht wörtlich übersetzt werden können (ibid:8). Hutchins (1992) vergleicht die Ergebnisse mit dem Vorgang eines/einer Nicht-MuttersprachlerIn, der/die nur sehr grobe Grammatikkenntnisse besitzt und zusätzlich ein eher schlechtes zweisprachiges Wörterbuch verwendet (vgl. Hutchins 1992:72).

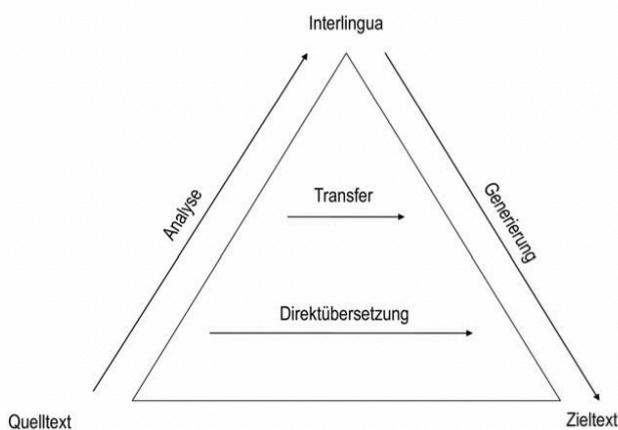
Der indirekte Ansatz hingegen teilt sich in mehreren Phasen auf. Dabei ist als erste die Transferübersetzung zu nennen, die ein wenig komplexer aufgebaut ist als die Wort-für-Wort-Übersetzung, da sie sich nicht nur zusätzlich auf die morphologischen und semantischen Informationen bezieht, sondern auch in Hinblick auf die syntaktische Ebene weiterentwickelt wurde (vgl. Stein 2009:8). Im Übersetzungsprozess wird der Ausgangstext mit Hilfe einer festgelegten Grammatik sowie eines ausgangssprachlichen Wörterbuches analysiert und sodann in eine abstrakte ausgangssprachliche Repräsentation verwandelt. In der zweiten Phase wird das Zwischenergebnis mittels zweisprachiger Transferregeln in eine abstrakte zielsprachliche Repräsentation übertragen. Ausgehend von der dieser Repräsentation wird schlussendlich der Zieltext erstellt, wobei wiederum eine zielsprachliche Grammatik sowie ein zielsprachliches Wörterbuch herangezogen wurden (vgl. Ramlow 2008:40f.).

Die Informationsquellen ließen dabei zunächst tausende neue Regeln und Kombinationen zu. In der Praxis stellte sich jedoch heraus, dass sich die unzähligen Regeln teilweise gegenseitig widersprachen und dadurch viele neue Fehler produzierten. Die Erkenntnis, dass höhere Komplexität nicht mehr automatisch eine bessere Qualität der Übersetzung bedeutete, wurde unausweichlich (vgl. Stein 2009:8).

Den höchsten Grad an Komplexität weist die Interlingua-Übersetzung auf. In zwei Phasen wird der Ausgangstext zunächst analysiert und in eine Zwischensprache (= Interlingua) übertragen sowie ausgehend von dieser zwischensprachlichen Repräsentation der

Zieltext erstellt (vgl. Ramlow 2008:39). Diese Methode beruht auf der Annahme, dass es eine universelle, sprachunabhängige Art gibt, sprachliche Informationen zu kodieren und zu dekodieren. Es wäre dadurch möglich, die Informationen aus einem Text vollständig von der Ausgangssprache zu lösen und in einen gleichwertigen Text in der Zielsprache zu verwandeln. Leider konnte sich diese Methode zu der damaligen Zeit nicht rasch genug weiterentwickeln, so dass man auf andere Methoden umschwenken musste (vgl. Stein 2009:8).

Die folgende Grafik stellt die drei Stufen mit der jeweils verbundenen Komplexität dar und soll abschließend nochmals einen Überblick über die verschiedenen Ansätze geben.



**Abbildung 1:** Stein 2009:9. Aufwand der drei Phasen der Maschinellen Übersetzung

#### 4.1.2 Statistikbasierte Systeme

Im Jahre 1988 stellte der Wissenschaftler Peter Brown seinen Kollegen auf der Second TMI Conference der Carnegie Mellon Universität einen rein statistischen Übersetzungsansatz (SMÜ / SMT) vor (vgl. Brown et al. 1988). Übersetzungsentscheidungen sollten ab sofort von bedingten Wahrscheinlichkeiten abhängen. Dazu waren keine Regelwerke, sondern lediglich große Korpora erforderlich (vgl. Stein 2009:8). Die Trainingskorpora stammen in der Regel aus verschiedenen Quellen und unterscheiden sich nach Thema, Genre, Dialekt, Schreibstil der Autoren o.Ä. (vgl. Chen et al 2016:314). Soll ein Satz von einer Ausgangssprache in eine beliebige Zielsprache übersetzt werden, so wird fortlaufend berechnet, wie hoch die Wahrscheinlichkeit ist, dass bereits übersetzte ausgangssprachliche Begriffe oder Segmente den zielsprachlichen Begriffen entsprechen. Ist dies Probabilität hoch, können jene Begriffe

somit übernommen werden (vgl. Stein 2009:8). Messbare wiederkehrende Charakteristika sind zum Beispiel das simultane Auftreten von mehreren Wörtern in Ausgangs- sowie Zieltext, die Satzlänge oder aber auch die Position der Wörter im Text (vgl. Ramlow 2008:42f.).

Rund um SMÜ-Systeme stellt sich die Frage, welche Art von Statistik konkret nützlich wäre, um eine bestmögliche Übersetzungsentscheidung zu treffen. So bestünde beispielsweise eine Möglichkeit darin, zu zählen, wie oft eine bestimmte Übersetzung in jeder der vorgegebenen Vergleichsübersetzungen vorkommt, wenn es für ein Wort mehrere Übersetzungen gibt. Die zu stellende Frage lautet demnach: Was sind die möglichen Übersetzungen, und wie oft treten diese auf, wenn kein externes Wissen vorhanden ist?

Zur Veranschaulichung wollen wir nun eine lexikalische Übersetzungswahrscheinlichkeitsverteilung aus erfolgten Zählungen abschätzen. Die Funktion soll dabei einen hohen Wert ausgeben, wenn Wort X eine gängige Übersetzung ist und einen niedrigen Wert, wenn Wort X eine seltene Übersetzung ist. Der Wert 0 steht dabei für die Unmöglichkeit der Übersetzung. Um die Wahrscheinlichkeitsverteilung zu ermitteln, wird das Verhältnis der Zählungen herangezogen. Kommt das Wort X beispielsweise 10.000 Mal und das Wort Y 8000 Mal in einer Textsammlung vor und werden beide Zahlen miteinander dividiert, so erhält man ein Verhältnis von 0,8. Die Wahrscheinlichkeit, dass Wort X die bessere Übersetzung wäre, beträgt demnach 0,8 (vgl. Koehn 2010:82f.).

<b>das</b>		<b>Haus</b>		<b>ist</b>		<b>klein</b>	
<i>e</i>	<i>t(e f)</i>	<i>e</i>	<i>t(e f)</i>	<i>e</i>	<i>t(e f)</i>	<i>e</i>	<i>t(e f)</i>
<i>the</i>	0.7	<i>house</i>	0.8	<i>is</i>	0.8	<i>small</i>	0.4
<i>that</i>	0.15	<i>building</i>	0.16	<i>'s</i>	0.16	<i>little</i>	0.4
<i>which</i>	0.075	<i>home</i>	0.02	<i>exists</i>	0.02	<i>short</i>	0.1
<i>who</i>	0.05	<i>household</i>	0.015	<i>has</i>	0.015	<i>minor</i>	0.06
<i>this</i>	0.025	<i>shell</i>	0.005	<i>are</i>	0.005	<i>petty</i>	0.04

**Abbildung 2:** Koehn 2010:84. Berechnung der Übersetzungswahrscheinlichkeiten für jedes Wort im Satz "Das Haus ist klein."

In Abb. 2 wird ersichtlich, wie eine solche Wahrscheinlichkeitsberechnung in Bezug auf einen Satz aussehen könnte. Die Berechnung wird Schritt für Schritt für jedes Wort eines Satzes durchgeführt. Im Anschluss daran wird die Übersetzung des Satzes aus den einzelnen Wörtern, die die höchsten Wahrscheinlichkeiten aufweisen, zusammengesetzt. (ibd:84) Da die Wahrscheinlichkeit jedoch sehr gering ist, dass der komplette Satz in den bereits vorhandenen Korpora wiederzufinden ist, ist es wenig sinnvoll, sich bei der Analyse auf

ganze Sätze zu konzentrieren. Um die Einheiten zu verkleinern und nach allen Ebenen, auf denen die Texte analysiert werden, zu unterscheiden, werden bevorzugt die wortbasierte und die phrasenbasierte Analyse angewandt (vgl. Stein 2009:11).

Bei der wortbasierten SMÜ werden die vorhandenen Daten auf der Wortebene analysiert. Das bedeutet, dass ein Wort der Ausgangssprache einem Wort der Zielsprache entsprechen muss. Es kommt häufig vor, dass sich ein einzelnes Wort der Ausgangssprache nur durch mehrere kombinierte Wörter in der Zielsprache übersetzen lässt, wie beispielsweise *slap* im Englischen und *dar una bofetada* im Spanischen. Dieser Aspekt stellt im Umkehrschluss ein Problem dar, wenn mehrere Wörter in der Ausgangssprache nur einem Wort in der Zielsprache entsprechen; denn jedem Wort in der Ausgangssprache muss mindestens ein Wort in der Zielsprache entsprechen. Ein weiteres Problem ist, dass zusammengehörende Wörter nicht zusammen übersetzt werden können. Dies kommt etwa bei trennbaren Verben vor, da sich die Bedeutung von Verben durch das Weglassen von Präfixen stark ändern kann. Bei Sprachenpaaren, die eine stark voneinander abweichende Syntax aufweisen, tritt dieses Problem ebenso häufig auf (vgl. Stein 2009:12).

Um jene Problematiken zu lösen, entschied man sich dazu, auf der Ebene von Phrasen zu arbeiten; sprich mit Wortgruppen statt ganzen Sätzen. Somit ist es möglich, mehrere zusammengehörende Wörter durch nur eines zu übersetzen und vice versa. Durch die Analyse von Phrasen und dem dadurch entstandenen Kontext ist es außerdem möglich, gewissen sprachlichen Mehrdeutigkeiten vorzubeugen. Für die SMÜ stellt es nun kein Problem mehr dar, zu unterscheiden, welche Übersetzung des Wortes *pretty* in den Fällen *pretty much* und *pretty girl* die richtige ist (vgl. Stein 2009:12).

Die größten Vorteile von statistikbasierten Systemen gegenüber regelbasierten Systemen sind zunächst der Zeit- sowie der Kostenfaktor. Dazu werden lediglich ausreichend alignierte mehrsprachige Korpora benötigt, die für die meisten offiziellen Sprachen der Europäischen Union durch den Paralleltextrkorpus EuroParl, in dem sämtliche Protokolle des Europäischen Parlaments gespeichert sind, zur Verfügung gestellt werden. Dadurch können in kürzester Zeit statistische MÜ-Systeme für Sprachen erstellt werden, bei denen die notwendigen Ressourcen für die Anwendung von regelbasierten Systemen fehlen. Diese können auch qualitativ mit den gängigen regelbasierten maschinellen Übersetzungssystemen mithalten. Auch linguistische Ambiguitäten oder arbiträre Redewendungen können leichter gelöst werden, insofern diese in den Vergleichskorpora ausreichend oft zu finden sind.

Die Nachteile der SMÜ sind unter anderem die möglichen Fehlerquellen im Trainingsmaterial. Daher gilt, dass die Qualität der Übersetzungen stark von der Qualität des

Trainingsmaterials abhängt. Ein weiterer Nachteil ist, dass nicht alle Sprachen über eine solch große Menge an notwendigem Vergleichsmaterial in digitaler Form verfügen (vgl. Stein 2009:12).

### 4.1.3 Neuronale maschinelle Übersetzungssysteme

In den letzten Jahren ist das Interesse an einem neuartigen maschinellen Übersetzungskonzept enorm gestiegen. Die neuronale maschinelle Übersetzung (NMT) ist zum Leitbild geworden und hat begonnen, ihren korpusbasierten Vorgänger, die statistikbasierte maschinelle Übersetzung (SMÜ), zu verdrängen. Damit das Potenzial dieser Technologie im professionellen Bereich voll ausgeschöpft werden kann, war und ist die Einbeziehung von Fachleuten von entscheidender Bedeutung.

Grundsätzlich ist die computergenerierte Übersetzung durch das neuronale MÜ-System flüssiger und bedeutet einen geringeren Aufwand beim Posteditieren im Vergleich zu rein statistischen Verfahren (vgl. Maier 2018: 32). Die neuronale maschinelle Übersetzung ist ein relativ neuer Ansatz im Bereich der Computerlinguistik und der korpusbasierten Übersetzung. Die ersten Untersuchungen dieses maschinellen Übersetzungssystems begannen zwar schon im Jahr 2014; ihr Durchbruch erfolgte jedoch erst zwei Jahre später (vgl. Vashee 2017: 44).

Auch die NMT wird anhand von schier riesigen Korpora aus Übersetzungseinheiten trainiert. Dabei wird jedoch der neuartige rechnerische Ansatz der künstlichen neuronalen Netze angewandt. Diese Netze bestehen aus Tausenden von künstlichen Einheiten, die Neuronen bzw. Nervenzellen insofern ähneln, als dass der Grad, in dem sie aktiviert oder blockiert werden, von den Reizen, die sie von anderen Neuronen erhalten, sowie von der Stärke der Verbindungen, über die diese Reize weitergeleitet werden, abhängt. Die meisten in der NMT verwendeten neuronalen Einheiten arbeiten in zwei Schritten, wenn sie über ihren Zustand oder ihre Aktivierung entscheiden. Dabei wird zunächst die Aktivierung eines Neurons  $X$ , das mit  $N$  Neuronen mit den Nummern  $1, 2, 3, \dots, N$  verbunden ist, wie folgt berechnet: Im ersten Schritt werden die mit Neuron  $X$  verbundenen Neuronen (im Folgenden als  $x_1, x_2, x_3, \dots, x_N$  bezeichnet) addiert. Davor wird jedes mit einem Gewicht (hier:  $w_1, w_2, w_3, \dots, w_N$ ) multipliziert, das die Stärke und Art ihrer Verbindung darstellt. Diese Gewichte können positiv oder negativ sein. Am Ende wird die Komponente  $B$ , die die Tendenz des Neurons, stimuliert zu werden, darstellt üblicherweise zur Gesamtzahl addiert. Wenn der

Impuls also über eine Verbindung mit einem positiven Gewicht empfangen wird, neigt ein stimuliertes Neuron dazu, das Neuron, mit dem es verbunden ist, zu aktivieren; wenn der Impuls über eine Verbindung mit negativem Gewicht empfangen wird, neigt ein stimuliertes Neuron dazu, das Neuron, mit dem es verbunden ist, zu blockieren. Die Berechnung könnte wie folgt aussehen:

$$y = w1 \times x1 + w2 \times x2 + w3 \times x3 + \dots \times \dots + wN \times xN + b$$

Das Ergebnis ist eine Zahl, die jeden möglichen negativen oder positiven Wert annehmen kann, aber noch nicht die Aktivierung  $x$  des Neurons darstellt. Aktivierungen sind normalerweise in irgendeiner Weise gebunden, zum Beispiel zwischen 0 und 1, zwischen -1 und +1 oder immer positiv. Im zweiten Schritt werden die Werte daher von  $y$  durch eine Aktivierungsfunktion auf Werte von  $x$  übertragen. Die logistische Aktivierungsfunktion nimmt Werte zwischen 0 und 1 an. Das heißt, wenn  $y$  immer negativer wird, nimmt die Aktivierung  $x$  einen Wert an, der sich sehr langsam 0 nähert; wenn  $y$  Null ist, nimmt die Aktivierung einen Wert von 0,5 an und wenn  $y$  immer positiver wird, nähert sich  $x$  langsam 1. Die hyperbolische Tangens-Aktivierungsfunktion hat eine sehr ähnliche Form wie die logistische Aktivierungsfunktion, variiert aber stattdessen zwischen -1 und 1. Eine weitere sehr beliebte Aktivierungsfunktion wird als *ReLU* bezeichnet. Wenn  $y$  positiv ist, ist  $x$  gleich  $y$ . Wenn  $y$  negativ ist, ist  $x$  gleich 0. Die Aktivierungen einzelner Neuronen macht erst Sinn, wenn sie mit den Aktivierungen der anderen Neuronen gruppiert werden.

Die Neuronen werden trainiert, die verschiedenen Repräsentationen von Wörtern und ihren Kontexten aufzubauen; sowohl im Kontext des zu verarbeitenden Quellsatzes als auch im Kontext des zu produzierenden Zielsatzes. Eine Repräsentation ist eine Momentaufnahme der Aktivierungszustände jedes Neurons in einer bestimmten Gruppe von Neuronen, die gewöhnlich als *Ebene* bezeichnet wird. Aus den Darstellungen der verschiedenen numerischen Größen / Vektoren wird die eigentliche Übersetzungsausgabe erzeugt. Stellt man sich dies als Diagramm vor, so würden zwei ähnliche Begriffe idealerweise nahe beieinander liegen und daher ähnliche Koordinaten haben; sehr unterschiedliche Begriffe würden weit voneinander entfernt liegen und daher mit unterschiedlichen Koordinaten versehen sein. Für die Kodierungen von Wörtern und Darstellungen von Sätzen werden mehrere Hundert Dimensionen benötigt. Die Darstellungen sind in der Regel tief (Anm.: daher das Schlagwort *deep learning*). Sie werden nicht in einem Zug, sondern in Stufen aus anderen flacheren

Darstellungen oder Schichten aufgebaut. Diese Schichten enthalten in der Regel Hunderte von neuronalen Einheiten (vgl. Forcada 2017:296).

Dank *Deep Learning* kann die Maschine selbstständig und ohne menschliches Eingreifen ihre Qualität verbessern. Dies wird außerdem durch die Extraktion und Klassifizierung der vorhandenen Daten und Informationen ermöglicht. Die gesammelten Trainingsdaten werden in einem breiteren Kontext miteinander verknüpft. Dadurch ist die Maschine fähig, Entscheidungen auf Grundlage dieser Verknüpfungen zu treffen (vgl. Litzel & Luber 2017).

Neuronale maschinelle Übersetzung funktioniert, wie eingangs erwähnt, durch Training mit großen Korpora. Jeder Quellsatz wird dabei gelesen und Repräsentationen werden gebildet, die so nah wie möglich an die entsprechenden Referenzübersetzungen im Trainingssatz (die idealerweise von Übersetzungsfachleuten erstellt wurden) herankommen. Das Gewicht oder die Stärke jeder der Verbindungen zwischen den Neuronen werden ermittelt, um die gewünschten Ergebnisse zu erzielen. Während des Trainings werden die Gewichte so modifiziert, dass der Wert einer spezifischen Fehlerfunktion oder Verlustfunktion, die beschreibt, wie weit die linguistischen Ausgaben der maschinellen Übersetzung von den Referenzübersetzungen entfernt sind, so klein wie möglich ausfällt.

Die meisten NMT-Systeme sind derartig aufgebaut und trainiert, dass sie einem Textvervollständigungsgerät ähneln. Durch die Darstellung des Quellsatzes – oder genauer gesagt durch Darstellungen der einzelnen Wörter des Quellsatzes in ihrem Kontext – wird an jeder Position des zu bildenden Zielsatzes und für jedes mögliche Wort im Zielvokabular die Wahrscheinlichkeit dahingehend, dass das Wort eine Fortsetzung dessen ist, was bereits produziert wurde, angezeigt. Die beste Übersetzung entsteht in der Regel dadurch, dass an jeder Position das wahrscheinlichste Wort ausgewählt wird (vgl. Forcada 2017:297).

Wie bereits erwähnt, wird die neuronale maschinelle Übersetzung seit 2014 aufwendig erforscht. Google präsentierte in einem Bericht aus dem Jahr 2016, dass laut ForscherInnen die neuronale maschinelle Übersetzung eine Schnittstelle zwischen Menschen und Maschine bzw. Computer sei. Aufgrund der herausragenden Ergebnisse von neuronalen maschinellen Übersetzungssystemen ist Microsoft dagegen der Auffassung, dass die neuronale maschinelle Übersetzung mit der Mensch-Maschine-Beziehung bereits auf gleicher Stufe stehe. Diese Erkenntnis hat maßgeblich zur Diskussion über den Einsatz der neuronalen maschinellen Übersetzung und zu einer möglichen Ersatzung menschlicher ÜbersetzerInnen durch diese beigetragen (vgl. Castilho et al. 2017:111f & Hassan et al. 2018 & Moorkens 2018: 5; & Wu et al. 2016).

Neuronale maschinelle Übersetzungssysteme stützen sich auf ein Encoder-Decoder-Modell, um einen Quellsatz zu dekodieren und ihn dann in einen Zielsatz verwandeln zu können. Der Encoder hat zum Ziel, alle sprachlichen Informationen aus einem Ausgangstext zusammen mit ihrem Kontext zu kodieren. Dafür verwendet dieser viele versteckte Einheiten, die eine verborgene Schicht bilden, wodurch wiederum lineare Verbindungen zwischen Ausgangs- und Zielsequenzen verhindert werden sollen. Es werden hoch nicht-lineare Schichten geformt, um dem Decoder genau eine Zielsequenz vorhersagen zu können (vgl. Dziubek 2018). Der Decoder analysiert im Anschluss alle Informationen zur Ausgangssequenz – wie beispielsweise Wörter und ihre Satzstellung, Morphologie sowie Interpunktion – und erzeugt dadurch das beste translatorische Ergebnis dieser Sequenz (vgl. Kostadinov 2019).

#### **4.2. Aktuelle Herausforderungen neuronaler maschineller Übersetzung**

Die neuronale maschinelle Übersetzung hat sich zwar in den letzten Jahren als der vielversprechendste maschinelle Übersetzungsansatz herauskristallisiert, dennoch hat sie signifikante Herausforderungen zu bewältigen, insbesondere wenn es um die Leistung außerhalb der eigenen Domäne geht. Was viele dieser Hürden gemeinsam haben, ist, dass die neuronalen Übersetzungsmodelle kein robustes Verhalten zeigen, wenn sie mit Daten konfrontiert werden, die sich erheblich von den gewohnten Trainingsbedingungen unterscheiden. Dies kann etwa aufgrund einer begrenzten Konfrontation mit Trainingsdaten, ungewöhnlicher Eingaben im Fall von Testsätzen außerhalb der Domäne oder einer unwahrscheinlichen anfänglichen Wortwahl bei einer Textsuche auftreten. Die Ergebnisse neuronaler maschineller Übersetzungssysteme sind zudem wenig interpretierbar. Weshalb und wie die Trainingsdaten dazu führen, dass die Systeme während der Dekodierung bestimmte Wortwahlentscheidungen fällen, liegt in großen Matrizen mit reell nummerierten Werten verborgen (vgl. Köhn 2020:293).

Eine aktuelle Herausforderung ist die multimodale neuronale Maschinenübersetzung. Beim Verständnis der menschlichen Sprache geht es nicht nur um reine Texte, sondern kann das Verständnis der Bedeutung einer Sprache die Hilfe anderer Modalitäten wie Bilder und Videos benötigen. In vielen Szenarien werden ÜbersetzerInnen aufgefordert, eine Rede oder ein Video zu übersetzen. Da beispielsweise die Simultanübersetzung von Reden bei verschiedenen Konferenzen oder internationalen Live-Veranstaltungen immer häufiger

eingefordert wird, ist die Durchführung einer multimodalen Übersetzung unter der Encoder-Decoder-Architektur eine große Herausforderung für die NMT (vgl. Zhang/Zhong:4).

Eine weitere bekannte Herausforderung beim Übersetzen allgemein sowie in Bezug auf neuronale maschinelle Übersetzungssysteme besteht darin, dass Wörter in verschiedenen Bereichen unterschiedlich übersetzt werden und die Bedeutung oftmals in verschiedenen sprachlichen Stilen ausgedrückt wird. Daher stellt die sogenannte Domänenanpassung bei der Entwicklung maschineller Übersetzungssysteme, die auf einen bestimmten Anwendungsfall ausgerichtet sind, einen essentiellen Schritt dar. Denn häufig sind große Mengen an Trainingsdaten nur außerhalb der Domäne verfügbar. Köhn (2020) trainierte beispielsweise verschiedene Systeme unter Verwendung verschiedener Korpora. Während die neuronalen und statistischen maschinellen Übersetzungssysteme innerhalb der Domäne einander qualitativ ähneln (Anm.: neuronale maschinelle Übersetzung ist besser für IT und Untertitel, statistische maschinelle Übersetzung für Recht, Medizin und den Koran), ist die Leistung außerhalb der eigenen Domäne für die NMT-Systeme in fast allen Fällen schlechter, manchmal sogar in dramatischem Ausmaß (vgl. Köhn 2020:295f.).

Eine bekannte Eigenschaft neuronaler maschineller Übersetzungssysteme ist, dass zunehmende Mengen an Trainingsdaten zu besseren Ergebnissen führen. Dies verspricht sowohl eine bessere Verallgemeinerbarkeit als auch verbesserte Kontextualität. Der Kontrast zwischen den Lernkurven der neuronalen und der statistischen maschinellen Übersetzung ist als durchaus auffällig zu beschreiben. Während die NMT in der Lage ist, zunehmende Mengen an Trainingsdaten effektiver auszunutzen, ist es ihr jedoch nicht möglich, mit Trainingskorpora von nur einigen Millionen Wörtern oder weniger zu arbeiten. Die Qualität für die neuronale maschinelle Übersetzung beginnt somit zwar auf niedrigerem Niveau, übertrifft die statistische maschinelle Übersetzung aber mit wachsendem Umfang der Trainingskorpora: Bei etwa 15 Millionen Wörtern schlägt die NMT bereits ein statistisches maschinelles Übersetzungssystem mit großem 2-Milliarden-Wort-in-Domain-Sprachmodell unter ressourcenintensiven Bedingungen (ibid:296).

Obwohl mehrere ForscherInnen berichten, dass NMT-Modelle bei seltenen Wörtern, die zu stark flektierten Kategorien gehören, besonders schlecht abschneiden, stellt Köhn (2020) bei den von ihm durchgeführten Tests das Gegenteil fest (ibid:297f).

Die durchschnittliche Gesamtpräzision ist zwischen dem NMT- und dem SMT-System relativ ähnlich, wobei das SMT-System insgesamt 70,1% und das NMT-System 70,3% erzielt. Dies spiegelt die ähnliche Gesamtqualität der MT-Systeme wider. Eine interessante Beobachtung ist, wie mit unbekanntem Wörtern umgegangen wird, wie zum Beispiel Wörter,

die nie im Trainingskorpus beobachtet wurden. Das SMT-System übersetzt diese in 53,2% der Fälle korrekt, während die Trefferquote beim NMT-System bei 60,1% liegt. Sowohl das SMT- als auch das NMT-System schneidet bei Wörtern, die nur ein einziges Mal im Trainingskorpus beobachtet wurden, am vergleichsweise schlechtesten ab (48,6% bzw. 52,2%).

Eine weitere Herausforderung für die neuronale maschinelle Übersetzung sind störende parallele Daten. ForscherInnen stellten fest, dass bei fast allen Arten von Problemen neuronale maschinelle Übersetzungssysteme stärker betroffen sind als statistische maschinelle Übersetzungssysteme. Kopierte Ausgangssprachsegmente haben dabei besonders katastrophale Auswirkungen auf die Qualität der neuronalen maschinellen Übersetzung: Diese führen dazu, dass das System selbst ein Kopierverhalten erlernt, das es in Folge exzessiv anwendet. Paare deutscher und englischer Sätze, die keine Übersetzungen des jeweils anderen sind, werden als fehlgeleitete Sätze gekennzeichnet. Diese können durch jedes Problem bei Alignierungsprozessen oder durch das Erzwingen der Anpassung von Inhalten, die in realita nicht parallel sind, auftreten. Solche falsch ausgerichteten Sätze sind die größte Fehlerquelle (41%). Ein oder beide Sätze können in einer anderen Sprache als Deutsch und Englisch verfasst sein (3%), beide Sätze können Deutsch (10%) oder beide können Englisch (10%) sein. Auf den ersten Blick scheinen einige Arten von Störungen leichter automatisch identifizierbar zu sein als andere.

Weitere Störfaktoren sind darüber hinaus falsch angeordnete Wörter. Sprache kann in vielerlei Hinsicht missverständlich sein. Dies kann das Ergebnis einer maschinellen Übersetzung, einer schlechten menschlichen Übersetzung oder eines stark spezialisierten Sprachgebrauchs sein, wie beispielsweise Aufzählungszeichen in Produktbeschreibungen. Es kommt dazu, dass ein paralleles Korpus durch Text in einer dritten Sprache verunreinigt sein kann, wenn zum Beispiel Französisch in einem deutsch-englischen Korpus vorkommt. Besonders in parallelen Korpora, die aus dem Web bezogen werden, gibt es oft Sätze, die unübersetzt bleiben. Beispiele dafür sind etwa Navigationselemente oder Copyright-Hinweise in der Fußzeile. Mehrsprachige Websites werden in diesem Fall unter Umständen nur teilweise übersetzt, während ein Teil des Originaltextes kopiert wird. Eine häufige Fehlerquelle in parallelen Korpora ist außerdem die fehlerhafte Ausrichtung von Dokumenten oder Sätzen, indem die Reihenfolge der Sätze auf einer Seite des ursprünglich sauberen parallelen Trainingskorpus zufällig gemischt wird.

Im Rahmen einer weiteren Studie von Köhn (2020) wurden die die Auswirkungen jeder Art von Störung auf die Übersetzungsqualität gemessen, wenn diese zum sauberen

Korpus hinzugefügt wurden. Bei den folgenden Arten von Störungen wird die NMT stärker geschädigt als die SMT: falsch ausgerichtete Sätze, falsch geordnete Wörter oder falsche Sprache. Störungstypen rund um kurze Segmente, unübersetzte Quellsätze und falsche Ausgangssprache haben auf beide Übersetzungssysteme aber nur geringe Auswirkungen (ibd: 301).

Die Aufgabe der Dekodierung besteht darin, die Übersetzung des vollständigen Satzes mit der höchsten Wahrscheinlichkeit zu finden. Ein Merkmal dieser Suchtechnik ist die Verwendung eines so genannten Strahlgrößenparameters, der die Anzahl der Teilübersetzungen pro Eingabewort begrenzt. Die Vergrößerung des Strahls ermöglicht es, eine größere Menge des möglichen Übersetzungsraums zu untersuchen und somit gegebenenfalls Übersetzungen mit besseren Modellbewertungen zu finden. Die Übersetzungsqualität wird jedoch durch eine Erhöhung der Strahlgröße nicht konsequent verbessert. Tatsächlich werden in fast allen Fällen qualitativ schlechtere Translationen jenseits der optimalen Einstellung der Strahlgröße gefunden (vgl. ibd:304).

## 5. Maschinelle Übersetzung am Beispiel DeepL

In den vorherigen Kapiteln wurden bereits einige zentrale Fragestellungen rund um maschinelle Übersetzung, insbesondere bezüglich neuronaler maschineller Übersetzung, behandelt. Bevor genauer auf eine Evaluierung der Übersetzungen anhand von *DeepL* eingegangen werden kann, soll die Technologie zum besseren Verständnis en detail erläutert werden. Der folgende Abschnitt fokussiert daher auf Geschichte, Aufbau und Nutzung der Software *DeepL*.

### 5.1 Beschreibung der Software

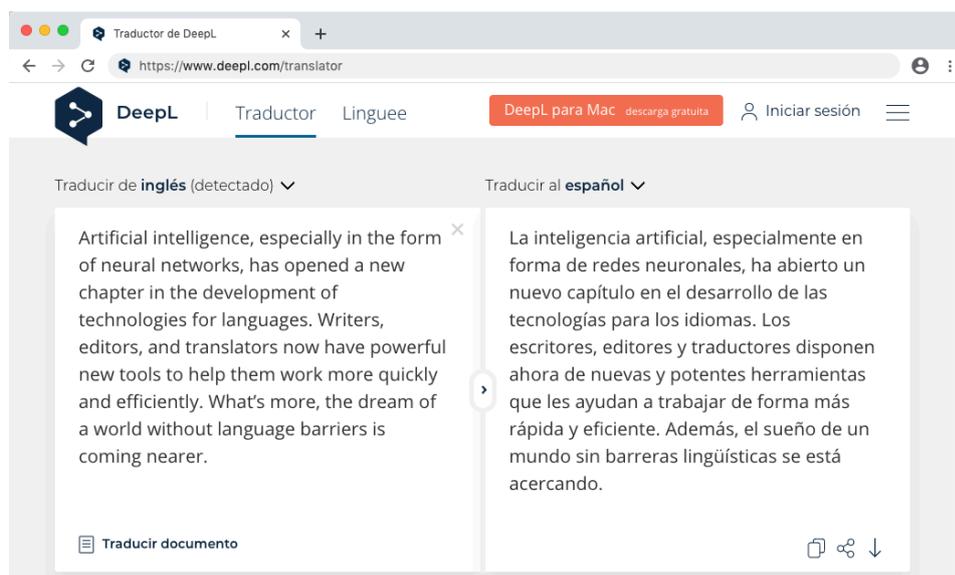
„DeepL's goal is to use neural networks to expand human possibilities, overcome language barriers, and bring cultures closer together.“ (DeepL Firmenprofil)

*DeepL* ist ein deutsches Unternehmen, das sich zum Ziel gesetzt hat, Sprachbarrieren weltweit durch den Einsatz künstlicher Intelligenz zu beseitigen. Seit 2017 bietet es mit dem *DeepL Translator* auf [www.DeepL.com](http://www.DeepL.com) ein maschinelles Übersetzungssystem an, das Blindtests zufolge die weltweit beste Übersetzungsqualität erreichen konnte. Die außergewöhnliche Qualität der maschinellen Übersetzung durch *DeepL* ist das Ergebnis urheberrechtlich geschützter Verbesserungen, die das Team bezüglich der Mathematik und Methodik neuronaler Netze ständig vornimmt.

Das Unternehmen wird von Gründer und CEO Jaroslaw Kutylowski geleitet und von weltbekannten Investoren wie *Benchmark* und *btov* unterstützt. Im Jahre 2016 startete ein Team innerhalb der *Linguee GmbH* die Arbeit an *DeepL Translator*, einem neuartigen Online-Übersetzer, der auf neuronalen Netzen basiert. *Linguee* wiederum ist ein Onlinewörterbuch, das mit einer Suchmaschine verknüpft ist und so den Zugriff auf große Mengen zweisprachiger übersetzter Satzpaare in einer Datenbank ermöglicht. Bereits im August 2017 wurde der *DeepL Translator* lanciert und für die Öffentlichkeit zugänglich gemacht. Geboten werden kostenlose Übersetzungen zwischen den folgenden 7 Sprachen: Englisch, Deutsch, Französisch, Spanisch, Italienisch, Polnisch und Niederländisch. Im März 2018 wurde der bis dato kostenlose Online-Übersetzungsdienst durch eine kostenpflichtige Version mit der Bezeichnung *DeepL Pro* ergänzt. *DeepL Pro* bietet API-Zugang (Anm.: Programmierschnittstelle), die Integration in Übersetzungssoftware und einen optimierten Web-Übersetzer. Im Dezember 2018 wurde *DeepL* um die Sprachen Russisch und

Portugiesisch erweitert. Auch in diesen Sprachen erreichte das System aus dem Stand eine nie dagewesene Übersetzungsqualität. Im September 2019 veröffentlichte *DeepL* eine Anwendung für *Windows* und *MacOS*, die sich in das System integrieren lässt und es den NutzerInnen ermöglicht, Übersetzungen aus jeder Anwendung auf ihrem Computer zu erstellen. Im Februar 2020 erzielte der Online-Übersetzer weitere Forschungsdurchbrüche in der Architektur neuronaler Netze, was einen weiteren Quantensprung in der Übersetzungsqualität auslöste. Die Ergebnisse der Blindversuche zeigten, dass die ÜbersetzerInnen die Übersetzungen von *DeepL* viermal so oft bevorzugten, als von anderen vergleichbaren Systemen. Im März 2020 folgten Chinesisch und Japanisch als neue Sprachen im Programm. Seit Mai 2020 ist *DeepL* der erste kostenlose Übersetzungsdienst, der es den NutzerInnen ermöglicht, genau anzugeben, wie bestimmte Wörter und Ausdrücke übersetzt werden sollen. Die Funktion, die als Glossar fungiert, soll dabei eine innovative Schnittstelle zwischen der künstlichen Intelligenz des Übersetzers *DeepL* und den individuellen Bedürfnissen und Präferenzen der NutzerInnen sein. Das Glossar kann folglich die von den NutzerInnen verwendeten Begriffe sowie deren bevorzugte Übersetzungen speichern. Seit August 2020 bietet *DeepL* zusätzlich eine Lösung für englische Übersetzungen, die die Besonderheiten des amerikanischen und britischen Englisch berücksichtigt (z.B. Unterschiede in der Rechtschreibung: Die amerikanischen *neighbors* werden zu britischen *neighbours*, die sich entweder auf dem *cell phone* oder *mobile phone* anrufen (vgl. DeepL Firmenprofil).

## 5.2 Installation und Anwendung



**Abbildung 3:** DeepL 2020a. *Screenshot DeepL Übersetzer (Spanisch).*

Der *DeepL Übersetzer* kann zunächst über die Website *www.DeepL.com* abgerufen werden. Wie in Abb. 3 ersichtlich, ist neben der Übersetzungsfunktion auch das zuvor erläuterte Onlinewörterbuch *Linguee* verknüpft. Bleibt man auf der Seite des Übersetzers, so besteht die Möglichkeit, je eine der angebotenen Ausgangs- und Zielsprachen zu wählen. Sollten bezüglich der Ausgangssprache Unsicherheiten bestehen, kann die Option „Beliebige Sprache (erkennen)“ ausgewählt werden. Der Text kann direkt in das Feld geschrieben oder in dieses kopiert werden. Auch Dokumente in den beiden Dateiformaten Microsoft Word (.docx) und PowerPoint (.pptx) können mit der kostenlosen Version eingefügt und übersetzt werden. Die kostenlose Version ist auf 5000 Zeichen pro Übersetzungsvorgang beschränkt. Sobald die Ausgangssprache, die Zielsprache und der Text ausgewählt / eingegeben wurden, wird das Dokument ohne einen weiteren Mausklick / Tastenbefehl automatisch übersetzt. Abb. 3 zeigt rechts unten drei weitere Funktionen, mit denen beispielsweise die Übersetzung zwischengespeichert oder eine URL erstellt und direkt per Facebook, Twitter oder E-Mail verschickt werden kann. Mit der dritten Funktion können Textdateien erstellt und heruntergeladen werden.

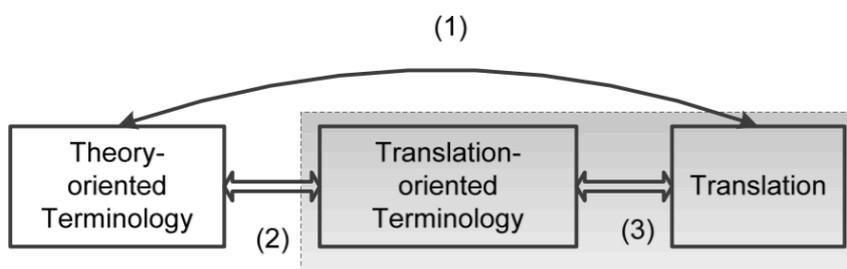
Neben der Webversion besteht auch die Möglichkeit, die Software unter *DeepL.com/app* kostenlos für *Windows* und *MacOS* Betriebssysteme herunterzuladen. Im Prinzip können dadurch alle Funktionen exakt so weitergenutzt werden wie bislang – mit dem Unterschied, dass das Programm in das Betriebssystem integriert wird und so jederzeit und überall auf die Technologie zugegriffen werden kann, ohne die Webseite dafür aufzurufen. Die heruntergeladene Software bleibt dabei so lange im Hintergrund, bis der/die NutzerIn sie mittels einer Tastenkombination aktiviert. Dafür muss die gerade vordergründig benutzte Anwendung, beispielsweise *Word*, gar nicht erst verlassen werden. Der/Die NutzerIn markiert dafür den gewünschten Text und drückt anschließend zweimal hintereinander STRG + C (bzw. Command + C bei *MacOS*). Die *DeepL*-App zeigt nun die Übersetzung an und liefert gleichzeitig ergänzende Alternativvorschläge mit. Durch einen abschließenden Klick auf „Einfügen“, fügt die Software die übersetzten Wörter direkt ins Dokument ein oder kopiert sie in die Zwischenablage.

## 6. Terminologie & Translation

In diesem Kapitel soll gezeigt werden, wie wichtig Terminologie und Translation für die Maschinelle Übersetzung sind und wie diese beiden Komponenten miteinander interagieren. Es folgt zunächst ein Überblick über die drei unterschiedlichen Interaktionsmöglichkeiten zwischen Terminologie und Translation, bevor diskutiert wird, welche bedeutende Rolle Terminologie für die Maschinelle Übersetzung spielt.

### 6.1 Abgrenzungen und Überschneidungen

Ursprünglich beanspruchte es der Fachbereich der Terminologie für sich, eine unabhängige Disziplin zu sein. Tatsächlich ist die Terminologie nach und nach zu einem eher inter- und transdisziplinären Gebiet geworden. Angefangen von Lexikologie, Semantik, Kognitiver Linguistik, über Soziolinguistik, Kommunikation, bis hin zu Philosophie und Sprachplanung steht die Terminologie in enger Verbindung mit einer Reihe von anderen (Teil-)Disziplinen.



**Abbildung 4:** Thelen 2015: 352. *The interaction between Theory-oriented Terminology, Translation-oriented Terminology and Translation.*

Wie in Abb. 4 dargestellt, sind eine Reihe von Interaktionen zwischen Terminologie und Übersetzung möglich. Zum einen (1) zwischen theorieorientierter Terminologie und Übersetzung, (2) zwischen theorieorientierter Terminologie und übersetzungsorientierter Terminologie sowie (3) zwischen übersetzungsorientierter Terminologie und Übersetzung. Die graue Hinterlegung bei (3) soll zeigen, dass in dieser Relation mehr Gemeinsamkeiten bestehen als in Bezug auf die theorieorientierte Terminologie. Der Grund für die Darstellung als zwei getrennte, aber verwandte, Bereiche liegt darin, dass sich zwar beide Disziplinen mit Übersetzung befassen, der Fokus jedoch leicht verschoben ist. Während sich

übersetzungsorientierte Terminologie mit der Übersetzung domänenspezifischer Sprache beschäftigt, fokussiert Übersetzung auf die Übersetzung von domänenspezifischer und allgemeiner Sprache. Das bedeutet, dass professionellen (nicht-literarischen) FachübersetzerInnen weniger Übersetzungsfreiheit zukommt, wenn diese auf einen Fachbegriff stoßen, als wenn sie sich mit einem allgemeinsprachlichen Wort befassen. ÜbersetzerInnen sind sodann mehr oder weniger verpflichtet, in der Übersetzung einen Standard-Äquivalenzbegriff anzugeben. Im Falle eines Wortes mit mehreren synonymen Bedeutungen kann er/sie in der Übersetzung eine Bedeutung wählen, die ihm/ihr gefällt, sofern der Begriff in den Kontext passt.

Andererseits wird die übersetzungsorientierte Terminologie von der theorieorientierten Terminologie getrennt dargestellt, weil erstere nicht dasselbe Maß an Systematik aufweist wie letztere und da ihr Ziel die Übersetzung – und nicht die Terminologie – ist. Außerdem wird in der Literatur zwischen systematischer Terminologie (theorieorientierte Terminologie) und Ad-hoc-Terminologie (übersetzungsorientierte Terminologie) unterschieden. Tatsächlich ist die übersetzungsorientierte Terminologie der Übersetzung viel näher als der theorieorientierten Terminologie. Folglich ist die übersetzungsorientierte Terminologie in Abb. 4 zwischen der theorieorientierten Terminologie und der Übersetzung angesiedelt und kann als „Vermittlerin“ zwischen den beiden betrachtet werden (vgl. Thelen 2015: 353).

## 6.2 Die Rolle der Terminologie in der maschinellen Übersetzung

Durch die rasanten Entwicklungen in sämtlichen linguistischen Bereichen besteht eine fortlaufende Entwicklung an neuen Begriffen. Einem Begriff können je nach Kontext, in dem er auftaucht, mehrere Bedeutungen zugeschrieben sein. Beispielsweise können die Wörter *terminal* (dt. *Busbahnhof* oder *Computerterminal*) und *play* (dt. *Musik spielen* oder *Fußball spielen*) je nach Kontext stark divergierende Bedeutungen haben. Ein polysemischer Begriff wie etwa *terminal* könnte demnach mehrere Übersetzungsäquivalente in einer Zielsprache haben. So beschreibt das englische Wort *charge* bis zu zwanzig Zieläquivalente in diversen Sprachen. Die Zielübersetzung könnte ihre adäquate Bedeutung verlieren, sollte bei der Übersetzung das Domänen- bzw. Kontextwissen nicht berücksichtigt werden.

Dementsprechend ist die Erhaltung des Domänenwissens vom Quell- bis zum Zieltext in jedem Übersetzungsworkflow von zentraler Bedeutung und ist eines der Hauptanliegen der KundInnen in der Übersetzungsbranche, insbesondere in sensiblen Bereichen wie Medizin, Transport, Militär, Recht oder Luft- und Raumfahrt. Auch Übersetzungsdienstleister, die MT

in der Produktion einsetzen, haben entsprechende Erwartungen an die Übersetzungen insofern, dass die Übersetzungen mit dem jeweiligen Kontext und der betreffenden Domäne konsistent sind (vgl. Haque/Hasanuzzaman/Way 2019:1f.).

Eine effektive Übersetzung hängt von der Fähigkeit ab, genaue konzeptuelle Äquivalenzen exakt zu übertragen und nicht nur einzelne Wortassoziationen zusammenzusetzen. Im Hinblick auf die Übersetzung kann die Bedeutung der Terminologie daher nicht hoch genug eingeschätzt werden, insbesondere in Bezug auf die maschinelle Übersetzung. Der Terminologiestatus lässt sich in zwei Hauptkategorien einteilen: a) Genehmigte Begriffe, bei denen ein semantisches Zieläquivalent des Quellterms als gültig anerkannt wird und daher in einer Übersetzung verwendet werden kann und b) Abgelehnte Begriffe, wobei ein genehmigter Zielsatz anstelle eines abgelehnten Begriffs, der entweder in der Quelle oder im Ziel verwendet wird, in der Übersetzung verwendet wird. Nimmt man zum Beispiel an, dass eine Organisation X über eine neu eingeführte Datenbank zum Terminologie-Management (Standardisierung und Verwaltung der Terminologie) verfügt. Diese Terminologiedatenbank enthält die äquivalenten Terminologieeinträge sowohl in der Ausgangssprache als auch in der Zielsprache. Mehrere Begriffe sind gültig, aber nur einer ist für die Verwendung in der Organisation zugelassen. Die Organisation X verfügt auch über ein großes *Translation Memory* (TM), eine Datenbank, in der Segmente gespeichert sind, bei denen es sich in der Regel um Sätze oder satzähnliche Einheiten (Überschriften, Titel oder Elemente in einer Liste) handelt, die zuvor übersetzt worden sind. Jedes Paar von übereinstimmenden Quell- und Zielsegmenten wird hierbei als Übersetzungseinheit bezeichnet. Diese Einheiten entsprechen dem, was für das Training eines MT-Systems benötigt wird. Der größte Teil der derzeit genehmigten Terminologie der Organisation X ist jedoch entweder a) in der Vergangenheit nicht verwendet worden und nicht in Alttexten oder im TM enthalten, die für das maschinelle Übersetzungstraining verwendet werden können, oder b) nur selten verwendet worden und erscheint daher nur in wenigen Texten oder TM mit nur geringer Häufigkeit. Ohne genehmigte Terminologie oder mit nur minimaler Anwendung genehmigter Terminologie als Trainingsdaten ist es dem MT-System deutlich erschwert, qualitativ hochwertige Übersetzungen zu produzieren.

Im Laufe der Zeit werden jedoch neue Übersetzungen, die die genehmigte Terminologie nutzen, dem TM hinzugefügt und werden schließlich Teil der Standard-Trainingsdaten für das MT-System. Dadurch wird sichergestellt, dass die genehmigte Terminologie zu einem integralen Bestandteil künftiger Übersetzungen wird. In der Zwischenzeit möchte die Organisation X jedoch sicherstellen, dass jede maschinelle Übersetzungsausgabe die

genehmigte Terminologie enthält und dass dies unabhängig von den vorhandenen Trainingsdaten immer geschieht. Die Terminologieverwaltung bei der MT weist prinzipiell zwei Haupteinsatzbereiche: die Validierung und die Substitution. Erstere beschreibt die Überprüfung dahingehend, ob die genehmigte Terminologie in jedem geeigneten Fall tatsächlich verwendet wird. Letztere meint das Einfügen oder Ersetzen eines genehmigten Begriffs, wenn dieser nicht in einer Übersetzung auftaucht. Dementsprechend müssen die gewünschten und genehmigten Begriffe in den Übersetzungen manuell ersetzt werden, bis der Übersetzungsspeicher diese Übersetzung auch für zukünftige Translationen übernimmt (vgl. Dougal 2018:6).

### **6.3 Begriffsklärungen: *term recognition & termhood***

Sowohl für TerminologInnen als auch KundInnen, die sich eine qualitativ hochwertige maschinelle Übersetzung wünschen, ist es von großer Bedeutung, dass die Termini einer Übersetzung möglichst repräsentativ für das zugrundeliegende Fachgebiet sind. Begriffe sind sprachliche Repräsentationen von domänenspezifischen Schlüsselkonzepten in einem Teilgebiet, die unser Expertenwissen zu diesem Thema hervorheben können. Die automatische Begriffserkennung (auch: Automatic Term Recognition ATR) spielt eine entscheidende Rolle bei der Verarbeitung und Verwaltung verschiedener Arten von Informationen, Wissen und Dokumenten. (vgl. Chunyu/Xiaoyue 2008:204f.).

Benennungen (= Teile potentieller Termini) können aus struktureller Sicht in zwei Typen kategorisiert werden: Einwortbenennung und Mehrwortbenennung. Entscheidend für die Klassifizierung ist die räumliche Trennung der Wörter, weshalb Komposita und Abkürzungen zu den Einwortbenennungen zählen (vgl. Saß 2004:23).

Die Beschaffenheit der potentiellen Termini wird in Folge hinsichtlich *Termhood* und *Unithood* beurteilt. *Unithood* wird verwendet, um eine Zeichenfolge als lexikalische Einheit zu identifizieren und zu beurteilen, wie wahrscheinlich es ist, dass ein Kandidat eine atomare (nicht unterteilbare) linguistische Einheit darstellt. Komposita und einzelne Wörter weisen beispielsweise eine sehr starke *Unithood* auf, so aber auch Mehrwortbenennungen. *Unithood* geht davon aus, dass nur atomare Einheiten in unserer Sprache als Begriffe qualifiziert sind, wobei sie der Vorstellung folgt, dass Begriffe die Grundeinheiten der mentalen Repräsentation menschlichen Wissens sind. *Termhood* hingegen bezieht sich auf die Relevanz für die Domäne. *Termhood* misst, genauer genommen, den Grad, zu dem eine bestimmte

lexikalische Einheit mit domänenspezifischen Konzepten verbunden ist. Eine niedrige terminologische Relevanz weisen Wörter der Gemeinsprache auf, Fachtermini hingegen verzeichnen eine relativ hohe *Termhood* (vgl. Nakagawa 2001:303f.).

Dennoch ist das Charakteristikum *Einheit* nur die notwendige, aber nicht die ausreichende Bedingung für einen „wahren“ Begriff. So gibt es einerseits etliche linguistische Einheiten – darunter Wörter und eine Vielzahl von Zusammensetzungen und Phrasen – die in keinem Fachgebiet „echte“ Begriffe sind, und andererseits ebenso viele linguistische Einheiten, die nur in einem bestimmten, aber keinem anderen Fachgebiet „echte“ Begriffe sind. Daran wird deutlich, dass ein Einheitsmaß nur als Filter wirkt, um unqualifizierte Begriffskandidaten auszuschließen und einen „wahren“ Begriff zu bestimmen. Nach dem Filtern ist ein Einheitsmaß immer noch erforderlich, um zu bestätigen, ob ein fraglicher Kandidat in einem bestimmten Fachgebiet ein „wahrer“ Begriff ist. In diesem Sinne spielt die *Termhood* als Maß dafür, inwieweit eine linguistische Einheit tatsächlich ein domänenspezifisches Konzept darstellt, bei der Lizenzierung eines Begriffs eine entscheidendere Rolle als die *Unithood* (vgl. Chunyu/Xiaoyue 2008:204f.).

Um die Genauigkeit der Termextraktion zu verbessern, können beispielsweise zusätzliche Domänenwörterbücher verwendet werden. (vgl. Oh et al. 2000:496-503). Die Entwicklung von Werkzeugen für den Aufbau elektronischer lexikalischer Ressourcen bietet eine neue Chance, diese auch im Bereich der Terminologie zu nutzen. Im Rahmen dieser Bemühungen wurde eine Reihe von elektronischen Domänenwörterbüchern entwickelt. Da neue Begriffe in der Regel aus bereits bestehenden Begriffen gebildet werden, können Wörterbücher hierbei als Quelle verwendet werden. Beispielsweise setzt sich die *verteilte Datenbank* aus den Begriffen *verteilt* und *Datenbank* zusammen, die beide Begriffe aus der Informatik darstellen. Ferner werden Konzepte und Begriffe einer Domäne häufig aus verwandten Domänen importiert. Beispielsweise wird der Begriff *Geographisches Informationssystem (GIS)* nicht nur in der Informatik, sondern auch im Bereich der Elektronik verwendet. Um diese Eigenschaften zu nutzen, ist es notwendig, Beziehungen zwischen den Domänen herzustellen. Eine hierarchische Clustering-Methode, die bei der Informationsbeschaffung verwendet wird, stellt ein geeignetes Mittel zum Aufbau einer Wörterbuchhierarchie dar. Eine Hierarchie kann zusätzlich helfen, die Beziehungen zwischen den Domänen abzuschätzen. Darüber hinaus können die geschätzten Beziehungen zwischen Domänen zur Gewichtung von Begriffen in einem gegebenen Korpus verwendet werden. So kann zum Beispiel die Domäne der Elektronik eine tiefe Beziehung zur Informatik haben.

Infolgedessen haben Begriffe in einem Elektronikwörterbuch eine höhere Wahrscheinlichkeit, Informatikbegriffe zu sein, als in anderen Domänenwörterbüchern (vgl. Felber 1984:114f).

## 7. Korpusanalyse

Nach diesem erfolgten ersten Einblick in die verschiedenen theoretischen Bereiche, soll im Anschluss der Korpus vorgestellt werden, anhand dessen die Performanz der MÜ im Vergleich zur Humanübersetzung in Bezug auf die Terminologie, die in den Ausgangs- und Zieltexten vorkommt, evaluiert wird. Im Anschluss wird das methodologische Vorgehen beschrieben und im Zuge dessen die einzelnen Fehlerkategorien sowie der zugrundeliegende Ansatz für die Auswertung dieser Fehler vorgestellt.

### 7.1 Beschreibung des Korpus

Als Untersuchungskorpus wurden Texte aus den vier Fachbereichen Geisteswissenschaft, Technik, Naturwissenschaft sowie Rechts- und Wirtschaftswissenschaften ausgewählt, sodass der gesamte Bereich des Fachübersetzens, aber auch verschiedene Formatvorlagen und Schwierigkeitsgrade umfasst sind. Es werden zwei Fachtexte pro Übersetzungsrichtung untersucht. Das Testmaterial besteht ausschließlich aus Textstichproben, also originalen Texten, die nicht präeditiert wurden.

Für die Übersetzungsrichtung Spanisch - Deutsch fließt zum einen ein Auszug aus dem Artikel *¿Las ciudades más ecológicas son más desiguales?* (2016) von Isabelle Anguelovski in den Korpus ein, der dem Bereich der Naturwissenschaften zugeordnet werden kann und 350 Wörter umfasst. Zum anderen wird ein Auszug aus dem Buch *Casa Tomada* (2016) von Julio Cortázar bearbeitet. Dieser Fachtext fällt dabei in den Bereich der Geisteswissenschaften und umfasst 450 Wörter.

Für die Übersetzungsrichtung Deutsch - Spanisch steht ein Auszug aus dem Artikel *Alles könnte anders sein – Eine Gesellschaftsutopie für freie Menschen* (2019) von Harald Welzer stellvertretend für den Bereich der Wirtschaftswissenschaften und umfasst dabei 217 Wörter. Zum anderen wird aus dem Firmenkatalog *Austria-Export Umwelt* die Firmenbeschreibung von „ALSTOM Austria“ herangezogen, welcher als technischer Text den Korpus mit 250 Wörtern abrundet.

## 7.2. Korpus

### 7.2.1. Text 1: Auszug aus *Alles könnte anders sein – Eine Gesellschaftsutopie für freie Menschen* von Harald Welzer

„Und hier ein Franziskus Zitat: „Eine gerechte Wirtschaft muss die Bedingungen dafür schaffen, dass jeder Mensch eine Kindheit ohne Entbehrungen genießen, während der Jugend seine Talente entfalten, in den Jahren der Aktivität einer rechtlich gesicherten Arbeit nachgehen und im Alter zu einer würdigen Rente gelangen kann. Es ist eine Wirtschaft, in der der Mensch im Einklang mit der Natur das gesamte System von Produktion und Distribution so gestaltet, dass die Fähigkeiten und Bedürfnisse jedes Einzelnen einen angemessenen Rahmen in Gemeinwesen finden“.

Nicht schlecht: Da ist auf den ersten Blick alles drin, worum es geht. Franziskus fängt nicht mit „der Wirtschaft“ an, sondern mit einer „gerechten“ Wirtschaft, setzt sie also in ein Bedingungsverhältnis: Sie ist Mittel, nicht Zweck. Und zwar Mittel zur Ermöglichung einer „Kindheit ohne Entbehrungen“. Man kann unterstellen, dass Franziskus hier nicht das Entbehren eines iPhones oder eines Klassenausflugs nach Mallorca meint, sondern Entbehrungen in jedem substantiellen Sinn, dass Kindheit von Hunger, Arbeit, Krieg und anderen existenziellen Einschränkungen gekennzeichnet ist. Tatsächlich ist das ein Anspruch, der unmittelbar mit der Moderne und den Menschenrechten verknüpft ist: Jedes Kind soll dieselben Möglichkeiten des Aufwachsens und der Entwicklung von Fähigkeiten, Talenten und Vorlieben haben, unabhängig von Geschlecht und Herkunft. Es soll, mit anderen Worten, ein Selbst werden können, das nicht von Mangel und Angst geprägt ist: ein freier Mensch.“ (Welzer 2019)

### 7.2.2. Text 2: Firmenkatalog Austria- Export, ALSTOM Österreich

„ALSTOM Österreich ist Teil des ALSTOM Konzerns, einem weltweiten Marktführer im Energie- und Transportbereich. Mit rund 60 000 Mitarbeitern in über 70 Ländern erzielte das Unternehmen im letzten Geschäftsjahr einen Umsatz von ca. 14 Milliarden Euro.

ALSTOM notiert seit Juni 1998 an der Pariser Börse.

Der Liefer- und Leistungsumfang von ALSTOM Österreich umfasst:

- schlüsselfertige Kraftwerke
- Turbinen, Generatoren, Kraftwerksleittechnik, Umwelttechnologien, Hilfsausrüstungen
- Kundenservice inkl. langfristiger Wartungsverträge
- Projektentwicklung und -finanzierung

Kompetenz und langjährige Erfahrung bilden auch eine solide Basis für die verstärkte Exporttätigkeit des Unternehmens. Der Exportanteil der ALSTOM Österreich beträgt rund 50 % des Gesamtumsatzes. Zu den Hauptmärkten zählen vor allem die südosteuropäischen Staaten, insbesondere Bulgarien, Bosnien-Herzegowina und Türkei, aber auch Bhutan und andere Exportmärkte.

Ein wesentlicher Erfolgsfaktor des Unternehmens ist neben der technischen Kompetenz der Projektleiter auch das umfangreiche Know-how hinsichtlich der Strukturierung von komplexen Finanzierungsmodellen.

Die dem Energiesektor zugeordneten Gesellschaften in Österreich beschäftigen derzeit rund 80 Mitarbeiter und erzielten im vergangenen Geschäftsjahr einen Umsatz von ca. 50 Mio. Euro.” (ALSTOM Österreich- Firmenkatalog Austria- Export)

### **7.2.3. Text 3: Auszug aus *¿Las ciudades más ecológicas son más desiguales?* Von Isabelle Anguelovski**

“¿Las ciudades más ecológicas son más desiguales?

Mejorar la calidad medioambiental en las comunidades urbanas tradicionalmente desfavorecidas.

En la década de 1980, cuando los activistas a favor de la justicia ambiental (EJ, por sus siglas en inglés) de Estados Unidos y otros países se organizaron por primera vez de manera visible para hacer frente al impacto desigual de la contaminación ambiental y los desechos, era evidente quiénes estaban en su punto de mira: las fábricas contaminantes, los infractores de las normativas sobre vertidos o los operadores de las incineradoras de residuos, así como sus cómplices en los Gobiernos, que habían sido incapaces de regular sus actividades o habían hecho la vista gorda. Formaban lo que la población consideraba Usos del Suelo no Deseados por las Comunidades Locales, o LULUs, por sus siglas en inglés.

Desde que se produjeron las primeras protestas de los militantes de la justicia ambiental en Love Canal, en el estado de Nueva York, en 1978, y en el Condado de Warren, en Carolina del Norte, en 1982, pusieron de manifiesto que los residuos tóxicos abundaban sobre todo en las comunidades de bajos ingresos y habitadas por minorías, y auguraron una larga lucha por remediar las desigualdades y el racismo ambientales, una lucha que ha vuelto a estar de actualidad a raíz de la reciente contaminación por plomo y el envenenamiento de más de 6.000 niños, la mayoría negros, ocurridos en Flint, en el estado de Michigan.

Sin embargo, con el tiempo, el activismo a favor de la justicia ambiental en zonas urbanas ha adquirido múltiples vertientes, y engloba a grupos locales que trabajan en proyectos diversos dirigidos a mejorar activamente el medio ambiente, como granjas urbanas, huertos, corredores ecológicos, áreas de juego y parques, que se apoyan mutuamente, atienden a las necesidades en cuanto a salud física y mental, y dan solución a años de abandono de las comunidades y trauma medioambiental. Entre los ejemplos emblemáticos de esta clase de movilización colectiva están los de los barrios de Sant Pere y Santa Caterina en Barcelona, Cayo Hueso en La Habana, o Dudley en Boston.“ (Anguelovski 2016)

#### **7.2.4. Text 4: Auszug aus *Casa tomada* von Julio Cortázar**

“Nos gustaba la casa porque aparte de espaciosa y antigua (hoy que las casas antiguas sucumben a la más ventajosa liquidación de sus materiales) guardaba los secretos de nuestros bisabuelos, el abuelo paterno, nuestros padres y toda la infancia.

Nos habituamos Irene y yo a persistir solos en ella, lo que era una locura pues en esa casa podían vivir ocho personas sin estorbarse. Hacíamos la limpieza por la mañana, levantándonos a las siete, y a eso de las once yo le dejaba a Irene las últimas habitaciones por repasar y me iba a la cocina. Almorzábamos a mediodía, siempre puntuales; ya no quedaba nada por hacer fuera de unos platos sucios. Nos resultaba grato almorzar pensando en la casa profunda y silenciosa y cómo nos bastábamos para mantenerla limpia. A veces llegamos a creer que era ella la que no nos dejó casarnos. Irene rechazó dos pretendientes sin mayor motivo, a mí se me murió María Esther antes que llegáramos a comprometernos. Entramos en los cuarenta años con la inexpresada idea de que el nuestro, simple y silencioso matrimonio de hermanos, era necesaria clausura de la genealogía asentada por los bisabuelos en nuestra casa. Nos moriríamos allí algún día, vagos y esquivos primos se quedarían con la casa y la echarían al suelo para enriquecerse con el terreno y los ladrillos; o mejor, nosotros mismos la voltearíamos justicieramente antes de que fuese demasiado tarde.

Irene era una chica nacida para no molestar a nadie. Aparte de su actividad matinal se pasaba el resto del día tejiendo en el sofá de su dormitorio. No sé porqué tejía tanto, yo creo que las mujeres tejen cuando han encontrado en esa labor el gran pretexto para no hacer nada. Irene no era así, tejía cosas siempre necesarias, tricotas para el invierno, medias para mí, mañanitas y chalecos para ella. A veces tejía un chaleco y después lo destejía en un momento porque algo no le agradaba; era gracioso ver en la canastilla el montón de lana encrespada resistiéndose a perder su forma de algunas horas. Los sábados iba yo al centro a comprarle lana; Irene tenía fe en mi gusto, se complacía con los colores y nunca tuve que devolver madejas. Yo aprovechaba esas salidas para dar una vuelta por las librerías y preguntar vanamente si había novedades en literatura francesa. Desde 1939 no llegaba nada valioso a la Argentina.” (Cortázar 2016:10)

### **7.3 Methodische Vorgehensweise**

Die Bewertung des MÜ-Outputs erfolgt in Form einer linguistischen Performanzanalyse. Zunächst geht dem Übersetzungsprozess eine Terminologiarbeit voran. Im Rahmen der Terminologierecherche werden verschiedene Fachtermini aus den Fachtexten ausgewählt und anhand einer Tabelle inklusive Definitionen, Synonymen, Quellen usw. präsentiert. Eine genaue Ausführung zur Terminologiarbeit und Eintragsstruktur ist in Kapitel 8 zu finden. Die vorab erfolgte Terminologierecherche sowie das daraus entstandene Glossar sollen dabei als Vergleich dienen und untersucht werden, wie genau das maschinelle Übersetzungstool Fachbenennungen übersetzt, wenn vorab keine Terminologierecherche stattfindet. Nach erfolgter Terminologierecherche werden die vier Textproben, welche die Grundlage für die Evaluierung darstellen, von dem Online-Übersetzungsdienst *DeepL* übersetzt und anschließend hinsichtlich etwaiger Fehler auf der lexikalischen, syntaktischen, semantischen und morphologischen Ebene überprüft. Die detektierten Fehler werden sodann in verschiedene linguistische Kategorien eingeteilt. Die verwendeten Textstichproben wurden bereits zuvor von HumanübersetzerInnen ins Deutsche beziehungsweise ins Spanische übersetzt. Diese Übersetzungen werden im Folgenden als Referenz für die Bewertung der maschinellen Übersetzungen verwendet. Die Texte werden weder vor- noch nachbearbeitet. Die Performanz der Übersetzung in Bezug auf die Terminologie ist zwar für das Gesamtergebnis der Evaluierung im Rahmen dieser Masterarbeit am interessantesten, die Evaluierung der anderen Ebenen ist aber dennoch wichtig, um aufzeigen zu können, wo die Stärken und Schwächen des Übersetzungssystems liegen. Zur Evaluierung der gesamten

maschinellen Übersetzung wird die Qualitätsmetrik SAE-J2450 herangezogen. Daraus resultieren die folgenden Fehlerkategorien: Terminologiefehler, Inhaltlicher Fehler, Auslassung, Strukturfehler, Rechtschreibfehler, Interpunktionsfehler, Sonstige Fehler. Eine Auswertung erfolgt anhand einer tabellarischen Auflistung.

### 7.3.1 SAE Qualitätsmetrik

Die SAE J2450 Qualitätsmetrik wurde für die Automobilbranche erstellt, um eine greifbare Methode zur Messung der Qualität von Übersetzungsergebnissen bereitstellen zu können. Das Ziel des vorgeschlagenen Qualitätsmaßstabs ist es dabei, einen einheitlichen Standard zu schaffen, an dem die Qualität der Übersetzungen objektiv gemessen werden kann, unabhängig von Ausgangssprache, Zielsprache oder Art und Weise der Durchführung der Übersetzung (d.h. menschliche Übersetzung oder maschinelle Übersetzung). Auch wenn die Qualitätsmetrik ursprünglich für die Automobilbranche kreiert worden ist, lässt sie sich dennoch auch auf anderwärtige Fachtexte gut anwenden. Die Metrik ermöglicht es einem/einer BegutachterIn, Fehler in einer Übersetzung zu markieren und eine gewichtete, numerische Bewertung zu erstellen, die die Qualität der Übersetzung abbildet.

Die Metrik besteht dabei aus vier Teilen:

1. Sieben Fehlerkategorien
2. Zwei Unterkategorien (schwerwiegende und leichte Fehler)
3. Zwei Meta-Regeln, die bei der Entscheidung über Unklarheiten bei der Zuordnung eines Fehlers zu den Kategorien und Unterkategorien helfen
  - Wenn ein Fehler mehrdeutig ist, soll immer die nächstliegende primäre Kategorie gewählt werden
  - Im Zweifelsfall ist immer *schwerwiegend* vor *leicht* zu wählen
4. Numerische Gewichte

Bei der Anwendung dieser Qualitätsmetrik sollte jeder von der evaluierenden Person gefundene Fehler auf zwei Arten gekennzeichnet werden. Erstens sollte er in eine der sieben unten beschriebenen Kategorien eingestuft werden, z.B. als *falscher Begriff* und sodann eine Angabe darüber erfolgen, ob es sich um einen *schweren* oder *leichten* Fehler handelt. Sowohl die primäre Kategorie als auch die sekundäre Einstufung liegen im Ermessen des/der Beurteilenden. Während im Folgenden relativ vollständige Definitionen vorgegeben werden, um dem/der BegutachterIn dabei zu helfen, die Hauptkategorien mit einer gewissen Konsistenz zu identifizieren, kann es dennoch Raum für Mehrdeutigkeit geben; die endgültige

Auswahl der Kategorie kann fallweise tatsächlich willkürlich erfolgen. Obwohl es praktisch unmöglich ist, den Begriff des *schweren* oder *leichten* Fehlers eindeutig zu definieren, besagt eine allgemeine Richtlinie, dass ein Fehler dann als *schwerwiegend* gekennzeichnet werden soll, wenn dieser gravierende Folgen für den/die FacharbeiterIn (z.B. TechnikerIn) und/oder die Bedeutung der Übersetzung hat. Ein *schwerwiegender* Fehler liegt also dann vor, wenn dieser zu einer unerwünschten Sinnänderung im Zieltext und daher zu einer Irritation des/der AnwenderIn mit der geringen bis hohen Wahrscheinlichkeit einer Fehlhandlung führt (Mögliche Konsequenzen wären etwa ein Sach- oder Personenschaden). Dem gegenüber ist ein *leichter* Fehler ein solcher, der zu keiner bzw. zu einer geringfügigen Irritation des/der AnwenderIn und zu keiner Gefahr auf eine Fehlhandlung führt (dabei entsteht kein Sach- oder Personenschaden).

Es ist eine natürliche und erwartbare Konsequenz, dass es Fälle geben wird, in denen ein/e GutachterIn bei dieser Entscheidung zögert. Wann immer dies eintritt, sollte dieser entgegen der allgemeinen Richtlinie als schwerwiegend gewertet werden. Was zunächst schier willkürlich scheinen mag, ist dennoch die sicherste Alternative in Bezug auf den Diensttext oder die Bedeutung der Übersetzung.

Sobald die Primär- und Sekundärklassifikationen zugewiesen sind, wird ein numerisches Gewicht zwischen 1 und 5 in der Definition nachgeschlagen und dem Fehler zugeordnet. 5 entspricht dem schwersten Fehler, während 1 einen Fehler mit relativ geringen Folgen markiert (vgl. SAE 2001:1f.).

### **7.3.2 Fehlerkategorien und Evaluation**

Im Folgenden wird im Detail auf die Fehlerkategorien eingegangen, um die später folgende Einteilung besser nachvollziehen zu können und um die Qualitätsmetrik transparenter zu gestalten.

#### **7.3.2.1 Fehlerkategorie: Falscher Begriff, WT – Wrong Term**

Diese Kategorie kann erst definiert werden, wenn die Definition eines *Begriffs* feststeht. In dieser Arbeit ist der Begriff wie folgt definiert: Einzelnes Wort, Mehrwortphrase, die als ein einziger, lexikalischer Bestandteil (d.h. als Teil der Rede) verwendet wird, Abkürzung, Akronym, Nummer, Ziffer oder Eigenname, einschließlich Handelsnamen, Markennamen,

eingetragene Warenzeichen, Ortsnamen und Personennamen. Angesichts der Begriffsdefinition wird jeder zielsprachliche Begriff, der gegen das Glossar eines KundInnenbegriffs verstößt, inkonsistent mit anderen Übersetzungen des ausgangssprachlichen Begriffs im gleichen Dokument oder der gleichen Art von Dokument ist (es sei denn, der Kontext für den ausgangssprachlichen Begriff rechtfertigt die Verwendung eines anderen Sprachbegriffs, z.B. aufgrund der Mehrdeutigkeit des ausgangssprachlichen Begriffs), oder der ein Konzept in der Zielsprache bezeichnet, das sich deutlich von dem Konzept des Begriffs der Ausgangssprache unterscheidet, als „falscher Begriff“ definiert. Schwerwiegende Fehler werden demnach mit 5, leichte Fehler mit 2 gewichtet (vgl. SAE 2001:5f.).

### **7.3.2.2 Fehlerkategorie: Falsche Bedeutung, WM – Wrong Meaning**

Diese Fehlerkategorie ist die treffende, wenn die Bedeutung des Zieltexes eindeutig und in erheblichem Maße von der des Ausgangstextes abweicht. Ein syntaktischer Fehler umfasst dabei auch die folgenden Fälle: a) Einem Quellbegriff wird die falsche Textstelle in seinem zielsprachlichen Gegenstück zugewiesen, b) Der Zieltex enthält eine falsche Phrasenstruktur, z.B. einen Relativsatz, wenn eine Verbalphrase benötigt wird c) die Wörter der Zielsprache sind zwar korrekt, aber in der falschen linearen Reihenfolge gemäß der syntaktischen Regeln der Zielsprache angeordnet. Schwerwiegende Fehler werden mit 4, leichte Fehler mit 2 gewichtet (vgl. ibd:6).

### **7.3.2.3 Fehlerkategorie: Auslassung, OM – Omission**

Diese Fehlerkategorie ist anzuwenden, wenn eine Aussage des Ausgangstextes nicht im Zieltex vorhanden ist, wie etwa das Fehlen eines Wortes oder Satzteils. Dies gilt ebenso für Grafiken, die in der Quellsprache Text enthalten, der im zielsprachlichen Dokument fehlt. Die Auslassung einzelner Wörter im Zieltex ist dabei nicht als Fehler zu bewerten, wenn diese nicht sinnentstellend sind und kein rechtliches Risiko (s.o.) besteht. Dies gilt insbesondere für Füll- und Blähwörter. Bekanntlich lässt sich mit der Weglassprobe herausfinden, welche Wörter entbehrlich sind. Schwerwiegende Fehler werden hier mit 4, leichte Fehler mit 2 gewichtet werden (vgl. ibd:7).

#### **7.3.2.4 Fehlerkategorie: Strukturfehler, SE – Structural Error**

Ein Fehler der falschen Wortstruktur tritt auf, wenn ein ansonsten korrektes Wort (oder Begriff) der Zielsprache in einer falschen morphologischen Form ausgedrückt wird, z.B. Fall, Geschlecht, Zahl, Zeitform, Präfix, Suffix, Infix oder eine andere Beugung. Diese Kategorie deckt folgende Fehler ab: Wortbildungsfehler, Syntaxfehler, Kollokationsfehler sowie Grammatikfehler im Sinne von falschen Konjunktionen, Adverbien und Präpositionen, Modusfehlern, Tempusfehlern, Kongruenzfehlern, Kasusfehlern, falschem Artikelgebrauch, falsche oder unklare Bezüge u.Ä. Styleguides (wenn vorhanden) können Hilfestellung in Zweifelsfällen bieten. Schwerwiegende Fehler werden mit 4, leichte Fehler mit 2 gewichtet (vgl. ibd:7f.).

#### **7.3.2.5 Fehlerkategorie: Rechtschreibfehler, SP – Misspelling**

Ein Rechtschreibfehler tritt auf, wenn ein Begriff der Zielsprache falsch geschrieben wurde, indem er gegen die im Glossar des/der KundIn angegebene Rechtschreibung oder die anerkannten Normen der Rechtschreibung in der Zielsprache verstößt. Ebenso liegt ein solcher Fehler vor, wenn der Begriff in einem für die Zielsprache falschen oder ungeeigneten Schriftsystem geschrieben ist oder wenn die Schreibweise gegen die allgemein geltenden Rechtschreibregeln der jeweiligen Zielsprache verstößt. Alte und neue Rechtschreibung dürfen dabei ebenfalls nicht vermischt werden. Tippfehler (beispielsweisr Buchstabendreher und fehlende Buchstaben) gelten dabei auch als Rechtschreibfehler. Schwerwiegende Fehler werden mit 3, leichte Fehler mit 1 gewichtet werden (vgl. ibd:8).

#### **7.3.2.6 Fehlerkategorie: Interpunktionsfehler, PE – Punctuation Error**

Diese Fehlerkategorie ist zu wählen, wenn die gewählte Zeichensetzung gegen die Interpunktionsregeln verstößt oder die Zeichensetzung nicht einheitlich verwendet wurde. Ein schwerwiegender Interpunktionsfehler könnte auch ein fehlendes Dezimalzeichen oder Komma gemäß der Zielsprache sein. Schwerwiegende Fehler werden mit 2, leichte Fehler mit 1 gewichtet (vgl. ibd:8).

### 7.3.2.7 Fehlerkategorie: Sonstige Fehler, ME – Miscellaneous Error

Jeder sprachliche Fehler in Zusammenhang mit dem zielsprachlichen Text, der sich nicht eindeutig den anderen zuvor aufgeführten Kategorien zuordnen lässt, ist als *sonstiger Fehler* einzustufen. Diese Regel trägt der Tatsache Rechnung, dass nicht alle auftretenden Fehler eindeutig in eine der anderen sechs Fehlerkategorien fallen. Auch kommt es vor, dass einige GutachterInnen eine Übersetzung Korrektur lesen und sie als eindeutig falsch erkennen, jedoch nicht genau sagen können, worin der Fehler liegt. Dazu gehören auch unübersetzte Stellen im Zieltext, sofern die erste Metaregel nicht greift sowie Zusatzinformationen, die im Ausgangstext nicht vorkommen (sofern sie nicht gerechtfertigt sind) oder unbegründete Wort- oder Satzwiederholungen. Schwerwiegende Fehler werden mit 3, leichte Fehler mit 1 gewichtet werden (vgl. ibd: 8f.).

## 7.4 Auswertungsbogen

Im Auswertungsbogen wird die Punktezahl aller Kategorien zusammengezählt. Anschließend wird die Summe aller zusammengezählten Kategorien durch die Gesamtanzahl der Wörter im Text dividiert. Dabei wird jeder Text einzeln gewertet, um eine bessere Übersicht zu erhalten. Der Schwellenwert für die Annehmbarkeit einer Übersetzung steht laut SAE-J2450 dem/der AuftraggeberIn frei festzulegen und wurde beispielsweise bei *VW* und *Audi* bei 0,03 fixiert, was bedeutet, dass maximal drei Fehlerpunkte pro 100 Wörter des Ausgangstextes toleriert werden (vgl. Dalla-Zuanna 2010).

Fehlerkategorie	Anzahl * schwere Fehler	Anzahl * leichte Fehler	GESAMT
WT	... * 5	...*2	
SE	... *4	...*2	
OM	...*4	...*2	
SA	...*4	...*2	
SP	...*3	...*1	
PE	...*2	...*1	
ME	...*3	...*1	

Abbildung 5: Eigene Darstellung. *Auswertungsbogen orientiert an SAE J2450 Qualitätsmetrik.*

## 8. Terminologiearbeit

Im folgenden Kapitel wird die eigens durchgeführte Terminologierecherche anhand eines Glossars präsentiert. Hierfür wurden relevante Begriffe aus den vier bereits vorgestellten Fachtexten ausgewählt. Vorab wurde keine feste Anzahl an Begriffen pro Text festgelegt, da in manchen Texten weniger und in anderen Texten mehr Fachbegriffe zu recherchieren waren. Um die Terminologierecherche so realitätsgetreu wie möglich zu inszenieren, war ein systematisches Durchscannen der Texte nacheinander notwendig, im Zuge dessen jene Termini, die beim Übersetzen potentiell Schwierigkeiten bereiten könnten, ausgemacht wurden. Insgesamt umfasst das erstellte Glossar 23 Fachtermini. Die Termini sind dabei nicht alphabetisch geordnet, sondern wurden der Reihe nach (nach Auftreten) aufgelistet. Zunächst beginnt das Glossar mit fünf Termini aus Text 1; es folgen sieben Termini aus Text 2, sechs Termini aus Text 3 und weitere fünf Termini aus dem letzten Fachtext. Bevor die Ergebnisse der durchgeführten Recherche präsentiert werden, wird kurz auf die Eintragsstruktur eingegangen.

Das Glossar besteht aus terminologischen Einträgen in den beiden Sprachen Spanisch und Deutsch. Bei den ersten beiden deutschsprachigen Ausgangstexten folgt der deutsche Eintrag vor dem spanischen. Bei den letzten beiden spanischsprachigen Ausgangstexten wurden zuerst die Spanischen und danach die deutschen Einträge angeführt. Alle terminologischen Einträge wurden konsistent bezüglich Form und Anordnung aufgebaut, wie im untenstehenden Mustereintrag (Abb. 6) demonstriert wird. Die meisten Datenkategorien mussten zwingend in jedem Eintrag enthalten sein. Die Datenkategorie *Synonym* konnte allerdings nicht bei jedem Eintrag ausgefüllt werden und war demnach frei wählbar. Die folgenden Datenkategorien konnten /mussten ausgefüllt werden: a) Benennung, b) Synonym, c) Definition, d) Quelle der Definition, e) Kontext, f) Quelle des Kontexts, g) Fachgebiet.

### 8.1 Struktur der terminologischen Einträge

Für die Terminologierecherche im Rahmen dieser Arbeit wurde das Eintragsmodell von KÜDES (2002) verwendet. Dabei dienen die KÜDES-Empfehlungen als Orientierung für die Strukturierung der Glossareinträge. Die ausgewählten und für diese Arbeit relevanten Kategorien werden nachfolgend ein eingehend beschrieben.

Einige Kategorien wie beispielsweise die Kurzform oder Anmerkung wurden bewusst ausgespart, da diese im Kontext dieser Terminologierecherche als nicht relevant eingestuft wurden.

### **8.1.1 Benennung**

Die spanischen beziehungsweise deutschen Benennungen wurden nicht immer 1:1 aus den Fachtexten übernommen, sondern teilweise an die von KÜDES vorgeschlagene Grundform angepasst. So wurden beispielsweise Substantive im Nominativ Singular und Verben im Infinitiv eingetragen. Die Forderung nach einer einem eigenen Beitrag für jede einzelne Benennung wurde strikt eingehalten. Es wurde außerdem darauf geachtet, dass bereits festgelegte Benennungen verwendet und keine eigenen Definitionen miteingebracht werden (vgl. Küdes 2002:27f.).

### **8.1.2 Synonym**

Da pro Eintrag nur eine einzelne Benennung angegeben werden darf, wurden die Alternativen als Synonyme angegeben. Dazu zählen sowohl spanische und deutsche alternative Benennungen als auch abweichende Schreibweisen. In diese Datenkategorie wurden nur echte Synonyme aufgenommen. Die Begriffsinhalte sind dabei identisch und können in jedem Kontext ausgetauscht werden (vgl. ibd:35).

### **8.1.3 Definition**

Um eine zuverlässige Terminologearbeit liefern und eine eindeutige Zuordnung von Benennungen zu ihren Begriffen garantieren zu können, wurden die Definitionen sorgfältig gewählt. Dabei sollen diese den Begriff entweder nach Inhalt oder Umfang umschreiben und ihn eindeutig von anderen Begriffen abgrenzen. Im Zuge dieser Masterarbeit wurde großer Wert auf die konsistente Einhaltung der KÜDES-Anforderungen an Definitionen gelegt. Wie bei den Benennungen wurde auch hier darauf geachtet, bereits festgelegte Definitionen zu verwenden. Diese stammen jeweils aus zuverlässigen Quellen (wie zum Beispiel dem Diccionario de la lengua española de la Real Academia und der Tageszeitung El País) und sollten dabei so kurz wie möglich sein, aber dennoch alle wichtigen Informationen beziehungsweise Merkmale beinhalten. Außerdem wurde sichergestellt, dass die Definitionen sich stets auf das jeweilige Fachgebiet beziehen (vgl. ibd:30f.).

#### **8.1.4 Quelle**

Da die Quellen der Definition und des Kontexts einen durchaus wichtigen Aspekt für die Überprüfung der Informationen darstellen und viel über die Zuverlässigkeit, Aktualität und Fachlichkeit dieser aussagen, erfolgte auch in diesem Punkt eine Orientierung an den KÜDES-Empfehlungen. Die Quellenangaben wurden dabei einheitlich angeführt.

Die Quellen wurden allesamt im Internet gefunden und als Volltext angegeben. Durch die ausgewählte Form der Darstellung bot sich genug Platz für die Darstellung des gesamten Links, so dass keine Information gekürzt werden musste. Die Quellen der Definitionen und des Kontexts sind darüber hinaus im Literaturverzeichnis zu finden (vgl. ibd:28f.).

#### **8.1.5 Kontext**

Der Kontext sollte stets die typisch fachsprachliche Verwendung der Benennung zeigen. In gewissen Einträgen kann dieser auch wesentliche Merkmale des Gegenstandes aufzeigen oder Hinweise auf die Bedeutung geben. Besteht der Kontextsatz im Quelldokument aus mehreren semantischen Einheiten, von denen einige für die Verwendung des zu behandelnden Terminus nicht relevant sind, so darf der Kontextsatz auf die wesentliche Aussage verkürzt werden (vgl. ibd:39f.).

#### **8.1.6 Fachgebiet**

Die Datenkategorie *Fachgebiet* soll als Orientierungs- und Verständnishilfe dienen. Die vier Fachtexte stammen aus den Bereichen Wirtschaft, Naturwissenschaft, Technik und Geisteswissenschaft, beinhalten aber auch Termini einiger anderer Fachgebiete. Oft war es auch notwendig, mehrere Sachgebiete für eine Benennung anzugeben, da sich die Begriffe über mehrere unterschiedliche Fachgebieten erstrecken. Durch die gewählte Darstellungsform des Glossars war eine Codierung des Fachgebiets nicht notwendig, da in der Tabelle selbst ausreichend Platz für den Volltext vorhanden ist (vgl. ibd:29f.).

Die Termini des vorliegenden Glossars wurden den nachstehenden Fachgebieten zugeordnet: a) Wirtschaft, b) Umwelt, c) Recht, d) Geisteswissenschaft, e) Technik und f) Industrie.

### **8.2 Mustereintrag**

Der folgende Mustereintrag dient zum einen zur Erklärung der Abkürzungen in der linken Spalte und zum anderen zur Demonstration der Anordnung und der Form des nachstehenden Glossars.

<b>DE</b>	<b>Deutsche Benennung</b>
SYN	Synonym
FACH	Fachbereich
DEF	Definition
QUE DEF	Quelle der Definition
KON	Kontext
QUE KON	Quelle des Kontexts
<b>ES</b>	<b>Spanische Benennung</b>
SYN	Synonym
FACH	Fachbereich
DEF	Definition
QUE DEF	Quelle der Definition
KON	Kontext
QUE KON	Quelle des Kontexts

Abbildung 6: Eigene Darstellung. *Mustereintrag Glossar.*

### 8.3 Präsentation des Glossars

#### 8.3.1. Text 1: Auszug aus *Alles könnte anders sein – Eine Gesellschaftsutopie für freie Menschen* von Harald Welzer

<b>DE</b>	<b>Entbehrung</b>
SYN	Mangel
FACH	Geisteswissenschaft
DEF	„schmerzlich empfundener Mangel, empfindliche Einschränkung“
QUE DEF	<a href="https://www.duden.de/rechtschreibung/Entbehrung">https://www.duden.de/rechtschreibung/Entbehrung</a> <a href="https://www.duden.de/rechtschreibung/Entbehrung">https://www.duden.de/rechtschreibung/Entbehrung</a> (Stand 08.11.2020)
KON	„Sie beinhaltet auch die Entbehrung grundlegender Fähigkeiten und den Mangel an Zugang zu Bildung, Gesundheitsfürsorge, natürlichen Ressourcen, Beschäftigung, Land und Krediten, politischer Beteiligung, Diensten und Infrastrukturen.“

QUE KON	<a href="https://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=COM:2001:0770:FIN:DE:PDF">https://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=COM:2001:0770:FIN:DE:PDF</a> (Stand 08.11.2020)
<b>ES</b>	<b>carencia</b>
SYN	
FACH	Geisteswissenschaft
DEF	„Falta o privación de algo“
QUE DEF	<a href="https://dle.rae.es/carencia">https://dle.rae.es/carencia</a> (Stand 10.11.2020)
KON	“En casi todos los casos de los más ligeros a los más severos que a veces hasta casos de comportamiento suicido, las causas eran muy simplemente una combinación de carencia del amor, carencia de entendimiento y carencia de aceptación del individuo.”
QUE KON	<a href="http://www.isokratia.com/pdf/sp.pdf">http://www.isokratia.com/pdf/sp.pdf</a> (Stand 10.11.2020)

<b>DE</b>	<b>Rente</b>
SYN	Pension (AT)
FACH	Wirtschaft
DEF	„regelmäßiger, monatlich zu zahlender Geldbetrag, der jemandem als Einkommen aufgrund einer [gesetzlichen] Versicherung bei Erreichen einer bestimmten Altersgrenze, bei Erwerbsunfähigkeit o. Ä. zusteht“
QUE DEF	<a href="https://www.duden.de/rechtschreibung/Rente">https://www.duden.de/rechtschreibung/Rente</a> (Stand 10.11.2020)
KON	„Die Rente wird für höchstens zwei Arbeitnehmer je Betrieb bis zum Erreichen des 65. Lebensjahrs gezahlt und beläuft sich auf jährlich 2500 ECU.“
QUE KON	<a href="https://ec.europa.eu/commission/presscorner/detail/de/IP_94_3">https://ec.europa.eu/commission/presscorner/detail/de/IP_94_3</a> (Stand 08.11.2020)
<b>ES</b>	<b>jubilación</b>
SYN	pensión
FACH	Wirtschaft
DEF	“La jubilación es el procedimiento administrativo de abandonar el mercado laboral principalmente por motivos de vejez.”
QUE DEF	<a href="https://economipedia.com/definiciones/jubilacion.html">https://economipedia.com/definiciones/jubilacion.html</a> (Stand 10.11.2020)
KON	„Quizá en algunos países funcione para ciertas personas, si tienen una edad de jubilación relativamente baja.“
QUE KON	<a href="https://www.europarl.europa.eu/sides/getDoc.do?pubRef=-//EP//TEXT+CRE+20010516+ITEMS+DOC+XML+V0//ES&amp;language=ES">https://www.europarl.europa.eu/sides/getDoc.do?pubRef=-//EP//TEXT+CRE+20010516+ITEMS+DOC+XML+V0//ES&amp;language=ES</a> (Stand 10.11.2020)

<b>DE</b>	<b>Gemeinwesen</b>
SYN	Gemeinde, Kommune, Nation
FACH	Geisteswissenschaft
DEF	„Als Gemeinwesen werden in der allgemeinen Begriffsverwendung alle Organisationsformen des menschlichen Zusammenlebens bezeichnet, die über den Familienverband hinausgehen.“
QUE DEF	<a href="https://www.bamf.de/DE/Service/ServiceCenter/Glossar/glossar_nod_e.html">https://www.bamf.de/DE/Service/ServiceCenter/Glossar/glossar_nod_e.html</a> (Stand 08.11.2020)
KON	„Die Krankenpflegeschülerinnen und Krankenpflege- schüler lernen nicht nur, als Mitglieder eines Pflgeteams tätig zu sein, sondern auch ein Pflgeteam zu leiten und die umfassende Krankenpflege einschließlich der Gesundheitserziehung für Einzelpersonen und kleine Gruppen im Rahmen der Gesundheitseinrichtungen oder im Gemeinwesen zu organisieren.“
QUE KON	<a href="https://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:C:2005:058E:0001:0129:DE:PDF">https://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:C:2005:058E:0001:0129:DE:PDF</a> (Stand 08.11.2020)
<b>ES</b>	<b>ser social</b>
SYN	
FACH	Geisteswissenschaften
DEF	“Se refiere a cualquier acción que se relaciona con el aumento del tamaño de una persona o un objeto, pero en relación con lo humano, no solo pasa por el crecimiento físico que un individuo tiene a lo largo de su vida, sino también por lo personal, desarro.”
QUE DEF	<a href="https://sites.google.com/site/358convivencia/comunidades-virtu/la-persona-como-ser-social">https://sites.google.com/site/358convivencia/comunidades-virtu/la-persona-como-ser-social</a> (Stand 10.11.2020)
KON	“Según este filósofo, la familia es la primera comunidad o sociedad formada, que es necesaria para el ser social.”
QUE KON	<a href="https://www.culturagenial.com/es/el-hombre-es-un-ser-social-por-naturaleza/">https://www.culturagenial.com/es/el-hombre-es-un-ser-social-por-naturaleza/</a> (Stand 10.11.2020)

<b>DE</b>	<b>Aufwachsen</b>
SYN	Heranwachsen
FACH	Geisteswissenschaft
DEF	„groß werden, vom Kind zum Erwachsenen heranwachsen“
QUE DEF	<a href="https://www.dwds.de/wb/aufwachsen">https://www.dwds.de/wb/aufwachsen</a> (Stand 08.11.2020)
KON	„Obwohl das Aufwachsen in einem Klima körperlicher und psychischer Gewalt für Kinder gravie- rende Folgen haben kann, stehen Kinder als indirekte Opfer häuslicher Gewalt immer noch am Rande der Wahrnehmung. Auch vor dem Hintergrund der Rechte von Kindern auf ein Leben ohne Gewalt, insbesondere auch auf eine

	gewaltfreie Erziehung, sowie auf Fürsorge und Schutz, darf diese Situation nicht mehr hingenommen werden.“
QUE KON	<a href="https://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:C:2006:325:0060:0064:DE:PDF">https://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:C:2006:325:0060:0064:DE:PDF</a> (Stand 08.11.2020)
<b>ES</b>	<b>crecer</b>
SYN	desarrollarse
FACH	Geisteswissenschaften
DEF	“Se refiere a cualquier acción que se relaciona con el aumento del tamaño de una persona o un objeto, pero en relación con lo humano, no solo pasa por el crecimiento físico que un individuo tiene a lo largo de su vida, sino también por lo personal, desarrollo de conocimiento, el cual puede adquirir a lo largo de los años.”
QUE DEF	<a href="https://conceptodefinicion.de/crecer/">https://conceptodefinicion.de/crecer/</a> (Stand 10.11.2020)
KON	“Desgraciadamente, no todos ellos pueden crecer bajo el cuidado de una familia y tienen que vivir en orfanatos.”
QUE KON	<a href="https://www.europarl.europa.eu/sides/getDoc.do?pubRef=-//EP//TEXT+CRE+20080115+ITEMS+DOC+XML+V0//ES&amp;language=ES">https://www.europarl.europa.eu/sides/getDoc.do?pubRef=-//EP//TEXT+CRE+20080115+ITEMS+DOC+XML+V0//ES&amp;language=ES</a> (Stand 10.11.2020)

<b>DE</b>	<b>Klassenausflug</b>
SYN	Klassenfahrt
FACH	Geisteswissenschaft
DEF	„Ausflug, den eine Klasse gemeinsam unternimmt.“
QUE DEF	<a href="https://www.duden.de/rechtschreibung/Klassenausflug">https://www.duden.de/rechtschreibung/Klassenausflug</a> (Stand 10.11.2020)
KON	„Während eines Klassenausflugs in den nahegelegenen Park hat Frl. Amelia einen schmerzvollen Unfall.“
QUE KON	<a href="http://little-princess-sara.net/episodesD.html">http://little-princess-sara.net/episodesD.html</a> (Stand 08.11.2020)
<b>ES</b>	<b>viaje escolar</b>
SYN	excursión escolar
FACH	Geisteswissenschaft
DEF	“Las excursiones escolares dan oportunidad a los estudiantes de salir del aula y experimentar algo nuevo, como una visita a una ciudad, un museo, lugares con historia o un contacto con la naturaleza. Son visitas a lugares fuera del aula, diseñadas para lograr ciertos objetivos, ya sea de conocimiento o recreación.”
QUE DEF	<a href="https://www.excursionesescolares.com/tipos-excursiones-escolares/">https://www.excursionesescolares.com/tipos-excursiones-escolares/</a> (Stand 10.11.2020)

KON	„Dos niños murieron y decenas resultaron heridos en viaje escolar en Argentina.“
QUE KON	<a href="https://www.elpais.com.uy/mundo/ninos-murieron-decenas-resultaron-heridos-viaje-escolar-argentina.html">https://www.elpais.com.uy/mundo/ninos-murieron-decenas-resultaron-heridos-viaje-escolar-argentina.html</a> (10.11.2020)

### 8.3.2 Text 2: Firmenkatalog Austria- Export, ALSTOM Österreich

<b>DE</b>	<b>schlüsselfertig</b>
SYN	bezugsbereit
FACH	Technik
DEF	„Schlüsselfertig Bauen bedeutet, dass alle Arbeiten vom Beginn bis zur Fertigstellung des Hauses von einem Bauunternehmen durchgeführt werden... Theoretisch sollte die Immobilie dann so fertiggestellt sein, dass Sie direkt nach der Übergabe des Schlüssels in Ihr neues Haus einziehen können.“
QUE DEF	<a href="https://www.drklein.de/schluesselfertig-bauen.html">https://www.drklein.de/schluesselfertig-bauen.html</a> (Stand 09.11.2020)
KON	„So bietet das Unternehmen nicht nur Projektentwicklung und -management für den Bau schlüsselfertiger Anlagen, sondern ist auch know-how-träger und Vertriebspartner für eine breite Palette an Komponenten und Leistungen, die für die Errichtung komplexer Anlagenprojekte benötigt werden.“
QUE KON	<a href="https://www.ferrostaal.com/uploads/tx_mfsmatrix/MF_ECHO_D_D ez09_WEB_2.pdf?PHPSESSID=9178b4853b70516d8acac25c6d424cd8">https://www.ferrostaal.com/uploads/tx_mfsmatrix/MF_ECHO_D_D ez09_WEB_2.pdf?PHPSESSID=9178b4853b70516d8acac25c6d424cd8</a> (Stand 09.11.2020)
<b>ES</b>	<b>de llave en mano</b>
SYN	
FACH	Technik
DEF	“Modalidad de contratación de un bien inmueble en la que el constructor se compromete a entregarla cuando esté totalmente funcional.”
QUE DEF	<a href="https://dpej.rae.es/lema/llave-en-mano">https://dpej.rae.es/lema/llave-en-mano</a> (Stand 10.11.2020)
KON	„Realizamos obras llave en mano en cualquier punto del país.“
QUE KON	<a href="https://www.procons.com.ar/portal/?p=909">https://www.procons.com.ar/portal/?p=909</a> (Stand 10.11.2020)
<b>DE</b>	<b>an der Börse notieren</b>
SYN	
FACH	Wirtschaft
DEF	„zum Handel an der Börse zugelassen und bewertet“
QUE DEF	<a href="https://www.dwds.de/wb/börsennotiert">https://www.dwds.de/wb/börsennotiert</a> (Stand 09.11.2020)

KON	„Alle an der Börse notierten Unternehmen mit mindestens 500 Beschäftigten sind aufgefordert, in ihren Jahresberichten an die Aktionäre eine „dreifache Bilanz“ zu veröffentlichen, in der ihre Leistung anhand von wirtschaftlichen, umwelttechnischen und sozialen Kriterien gemessen wird.“
QUE KON	<a href="https://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=COM:2001:0264:FIN:DE:PDF">https://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=COM:2001:0264:FIN:DE:PDF</a> (Stand 09.11.2020)
<b>ES</b>	<b>cotizar en la bolsa</b>
SYN	
FACH	Wirtschaft
DEF	“Cotizar en bolsa significa que un activo financiero esté admitido a negociación en un mercado bursátil, es decir, que los inversores puedan negociar con él.”
QUE DEF	<a href="https://economipedia.com/definiciones/cotizar-en-bolsa.html">https://economipedia.com/definiciones/cotizar-en-bolsa.html</a> (Stand 10.11.2020)
KON	“Se pretende que Investbx se introduzca en el mercado de las plataformas de negociación para la obtención de capital para las empresas que no pueden cotizar en la bolsa principal del Reino Unido.”
QUE KON	<a href="https://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2008:045:0001:0029:ES:PDF">https://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2008:045:0001:0029:ES:PDF</a> (Stand 10.11.2020)

<b>DE</b>	<b>Leitechnik</b>
SYN	
FACH	Technik
DEF	„Die Leitechnik umfasst sämtliche Methoden, Verfahren und Einrichtungen zur Führung technischer Prozesse. Die an den Prozessen beteiligten Systeme gehen dabei von einem Ausgangszustand in einen Zielzustand über oder bewegen sich zwischen zwei oder mehreren Zielzuständen hin und her.“
QUE DEF	<a href="https://link.springer.com/chapter/10.1007%2F978-3-7091-0562-7_4">https://link.springer.com/chapter/10.1007%2F978-3-7091-0562-7_4</a> (Stand 09.11.2020)
KON	„Durch Erweiterung der Semantik-definition für logische Knoten könnte dieses Modell ebenfalls für Kraftwerksleitechnik, verteilte Energieversorgung, Wasser- und Gassysteme usw. verwendet werden.“
QUE KON	<a href="https://books.google.at/books?id=jfw77xU5yvQC&amp;pg=PA137&amp;lpg=PA137&amp;dq=was+ist+kraftwerksleitechnik+definition&amp;source=bl&amp;ots=mep6EuTjh3&amp;sig=ACfU3U2jLcLrHCISKtsJYFHob_I0QAICDA&amp;hl=de&amp;sa=X&amp;ved=2ahUKEwjP3IvLj_XsAhUIXR0KHRL8AOs4">https://books.google.at/books?id=jfw77xU5yvQC&amp;pg=PA137&amp;lpg=PA137&amp;dq=was+ist+kraftwerksleitechnik+definition&amp;source=bl&amp;ots=mep6EuTjh3&amp;sig=ACfU3U2jLcLrHCISKtsJYFHob_I0QAICDA&amp;hl=de&amp;sa=X&amp;ved=2ahUKEwjP3IvLj_XsAhUIXR0KHRL8AOs4</a>

	ChDoATAAegQICRAC#v=onepage&q=was%20ist%20kraftwerksleitetechnik%20definition&f=false (Stand 09.11.2020)
<b>ES</b>	<b>tecnología de control</b>
SYN	
FACH	Technik
DEF	“La tecnología de control térmico o tecnología de control industrial representa aparatos que controlan, regulan, monitorizan, recopilan datos, comunican y emiten un diagnóstico.”
QUE DEF	<a href="https://www.sew-eurodrive.es/produkte/steuerungstechnik.html">https://www.sew-eurodrive.es/produkte/steuerungstechnik.html</a> (Stand 10.11.2020)
KON	“Nuestras máquinas de aplicación de sleeve están especialmente diseñados para satisfacer las necesidades individuales de los clientes, pero todos se basan en principios bien probados y están equipadas con la última tecnología de control.”
QUE KON	<a href="http://www.sleeve-technology.com/es?lang=es">http://www.sleeve-technology.com/es?lang=es</a> (Stand 10.11.2020)

<b>DE</b>	<b>Finanzierungsmodell</b>
SYN	
FACH	Wirtschaft
DEF	„Modell, nach dem eine Investition finanziert werden kann“
QUE DEF	<a href="https://www.duden.de/rechtschreibung/Finanzierungsmodell">https://www.duden.de/rechtschreibung/Finanzierungsmodell</a> (Stand 09.11.2020)
KON	„Das Unternehmen war zudem maßgeblich an der Entwicklung eines Finanzierungsmodells zur Mobilisierung der Investitionssumme beteiligt.“
QUE KON	<a href="https://www.ferrostaal.com/uploads/tx_mfsmatrix/MF_ECHO_D_D ez09_WEB_2.pdf?PHPSESSID=9178b4853b70516d8acac25c6d424cd8">https://www.ferrostaal.com/uploads/tx_mfsmatrix/MF_ECHO_D_D ez09_WEB_2.pdf?PHPSESSID=9178b4853b70516d8acac25c6d424cd8</a> (Stand 09.11.2020)
<b>ES</b>	<b>modelo de financiación</b>
SYN	modelo financiero
FACH	Wirtschaft
DEF	“Bajo el término modelo financiero se entiende una versión abstracta-matemática que plasma la realidad financiera de una empresa en un set de hojas de cálculo. Un modelo es pues una versión matemática, simplificada e idealizada de la realidad de una empresa.“
QUE DEF	<a href="https://modelandum.com/que-es-un-modelo-financiero/">https://modelandum.com/que-es-un-modelo-financiero/</a> (Stand 10.11.2020)

KON	“Pero llevaba una promesa bajo el brazo: la reforma del siempre complejo modelo de financiación autonómica.”
QUE KON	<a href="https://elpais.com/politica/2019/08/05/actualidad/1565017035_740712.html">https://elpais.com/politica/2019/08/05/actualidad/1565017035_740712.html</a> (Stand 10.11.2020)

<b>DE</b>	<b>Gesellschaft</b>
SYN	Unternehmen
FACH	Wirtschaft
DEF	„Gesellschaft ist die vertragsmäßige Verbindung von zwei oder mehreren Personen zur Erreichung eines gemeinsamen Zweckes mit gemeinsamen Kräften oder Mitteln.“
QUE DEF	<a href="https://www.sign-lang.uni-http://www.rwi.uzh.ch/static/elt/lst-vogt/gesellschaftsrecht/grundlagen/de/html/unit_begriffgesellschaft.php">https://www.sign-lang.uni-http://www.rwi.uzh.ch/static/elt/lst-vogt/gesellschaftsrecht/grundlagen/de/html/unit_begriffgesellschaft.php</a> (Stand 09.11.2020)
KON	„Diese Kapitalzufuhr ist Teil eines Umstrukturierungsplans, mit dem die finanzielle und wirtschaftliche Situation der Gesellschaft bis Ende 1996 wiederhergestellt werden soll.“
QUE KON	<a href="https://ec.europa.eu/commission/presscorner/detail/de/IP_94_426">https://ec.europa.eu/commission/presscorner/detail/de/IP_94_426</a> (Stand 09.11.2020)
<b>ES</b>	<b>empresa</b>
SYN	sociedad
FACH	Wirtschaft
DEF	“Unidad de organización dedicada a actividades industriales, mercantiles o de prestación de servicios con fines lucrativos.”
QUE DEF	<a href="https://dle.rae.es/empresa?m=form">https://dle.rae.es/empresa?m=form</a> (Stand 10.11.2020)
KON	“Dublín permite tributar menos a las grandes empresas de Estados Unidos.”
QUE KON	<a href="https://elpais.com/opinion/2020-08-25/paraiso-irlandes.html">https://elpais.com/opinion/2020-08-25/paraiso-irlandes.html</a> (Stand 10.11.2020)
<b>DE</b>	<b>Leistungsumfang</b>
SYN	Sortiment
FACH	Wirtschaft
DEF	„Leistungsumfang ist die Menge aller Leistungen, die für die Abnahme eines Projekts verpflichtend zu erbringen sind.“
QUE DEF	<a href="https://www.projektmagazin.de/glossarterm/leistungsumfang">https://www.projektmagazin.de/glossarterm/leistungsumfang</a> (Stand 09.11.2020)
KON	„Da sich der Leistungsumfang der gesetzlichen Krankenversicherung vergrößert hat, sind die öffentlichen Gesundheitsausgaben sogar noch rascher angestiegen.“

QUE KON	<a href="https://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=COM:2000:0846:FIN:DE:PDF">https://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=COM:2000:0846:FIN:DE:PDF</a> (Stand 10.11.2020)
<b>ES</b>	<b>gama</b>
SYN	
FACH	Wirtschaft
DEF	“La gama de productos es un conjunto de productos, los cuales abarcan un segmento de mercado determinado.”
QUE DEF	<a href="https://economipedia.com/definiciones/linea-de-productos.html">https://economipedia.com/definiciones/linea-de-productos.html</a> (Stand 10.11.2020)
KON	“La Comunidad también es un mercado tradicional para los exportadores turcos, que poseen unos canales de distribución bien establecidos y una amplia gama de clientes.”
QUE KON	<a href="https://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2008:343:0001:0038:ES:PDF">https://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2008:343:0001:0038:ES:PDF</a> (Stand 10.11.2020)

<b>DE</b>	<b>Hilfsausrüstung</b>
SYN	Sicherheitsausrüstung
FACH	Technik
DEF	„Ausrüstung für Sicherheitszwecke, Brandbekämpfung, Polizei und Verteidigung“
QUE DEF	<a href="https://offenevergaben.at/branchen?node=35000000&amp;anzahl">https://offenevergaben.at/branchen?node=35000000&amp;anzahl</a> (Stand 09.11.2020)
KON	„Außerdem wurde "eine Reinigung der Gasleitung, der Schlammleitung, des Tropfenabscheiders der Gasreinigung und eine Revision der Hilfsausrüstung vorgenommen, die zu dieser dazugehört".“
QUE KON	<a href="https://de.rusbiznews.com/news/n974.html">https://de.rusbiznews.com/news/n974.html</a> (Stand 09.11.2020)
<b>ES</b>	<b>equipo auxiliar</b>
SYN	equipamiento auxiliar
FACH	Technik
DEF	“Los equipos auxiliares son dispositivos que se utilizan para: estabilizar los valores nominales de funcionamiento o ejercer un control sobre la lámpara en el encendido, apagado o regulación de la intensidad.”
QUE DEF	<a href="https://grlum.dpe.upc.edu/manual/sistemasIluminacion-equiposAuxiliares.php">https://grlum.dpe.upc.edu/manual/sistemasIluminacion-equiposAuxiliares.php</a> (Stand 10.11.2020)

KON	“La oferta de venta también comprende el equipo auxiliar, el personal y los servicios necesarios para el funcionamiento de la red de transporte.”
QUE KON	<a href="https://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:C:2009:133:0010:0011:ES:PDF">https://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:C:2009:133:0010:0011:ES:PDF</a> (Stand 10.11.2020)

### 8.3.3. Text 3: Auszug aus *¿Las ciudades más ecológicas son más desiguales?* von Isabelle Anguelovski

<b>ES</b>	<b>infractor</b>
SYN	infractor de ley
FACH	Recht
DEF	“Persona que quebranta una ley o un precepto.”
QUE DEF	<a href="https://dle.rae.es/infractor?m=form">https://dle.rae.es/infractor?m=form</a> (Stand 10.11.2020)
KON	„Si Argónida tuviese un Estatuto, sus leyes insistirían una y otra vez en el castigo de la gente de orden y en la comprensión moral de los infractores.“
QUE KON	<a href="https://elpais.com/diario/2006/11/11/andalucia/1163200928_850215.html">https://elpais.com/diario/2006/11/11/andalucia/1163200928_850215.html</a> (Stand 10.11.2020)
<b>DE</b>	<b>Rechtsbrecher</b>
SYN	Gesetzesbrecher
FACH	Recht
DEF	„männliche Person, die Rechtsbruch begangen hat; Gesetzesbrecher“
QUE DEF	<a href="https://www.duden.de/rechtschreibung/Rechtsbrecher">https://www.duden.de/rechtschreibung/Rechtsbrecher</a> (Stand 09.11.2020)
KON	„Die Kommission schließt aus diesen Informationen, dass die Sanktionen auf potenzielle Rechtsbrecher nicht wirklich abschreckend wirken. „
QUE KON	<a href="https://ec.europa.eu/commission/presscorner/detail/de/IP_02_1805">https://ec.europa.eu/commission/presscorner/detail/de/IP_02_1805</a> (Stand 09.11.2020)

<b>ES</b>	<b>cómplice</b>
SYN	
FACH	Recht
DEF	“Participante o asociado en crimen o culpa imputable a dos o más personas.”
QUE DEF	<a href="https://dle.rae.es/cómplice?m=form">https://dle.rae.es/cómplice?m=form</a> (Stand 10.11.2020)

KON	“Miembro de las fuerzas de seguridad y cómplice en el diseño o dirección de la política estatal de represión.”
QUE KON	<a href="https://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2009:023:0043:0059:ES:PDF">https://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2009:023:0043:0059:ES:PDF</a> (Stand 10.11.2020)
<b>DE</b>	<b>Handlanger</b>
SYN	Komplizen
FACH	Recht
DEF	„die Absichten oder Aufträge eines Anderen ausführende und erfüllende Person“
QUE DEF	<a href="https://www.verben.de/substantive/Handlanger.htm">https://www.verben.de/substantive/Handlanger.htm</a> (Stand 09.11.2020)
KON	„Wir wissen, wer die anderen Rädelsführer sind, die wir ins Visier nehmen müssen, uns sind die Handlanger und auch die Mitläufer bekannt.“
QUE KON	<a href="https://www.europarl.europa.eu/sides/getDoc.do?pubRef=-//EP//TEXT+CRE+20080709+ITEMS+DOC+XML+V0//DE&amp;language=DE">https://www.europarl.europa.eu/sides/getDoc.do?pubRef=-//EP//TEXT+CRE+20080709+ITEMS+DOC+XML+V0//DE&amp;language=DE</a> (Stand 09.11.2020)

<b>ES</b>	<b>incinerador de residuos</b>
SYN	
FACH	Umwelt
DEF	„En una incineradora se queman los residuos de manera controlada para obtener energía.“
QUE DEF	<a href="https://stopbasura.com/2016/03/30/incineradora/">https://stopbasura.com/2016/03/30/incineradora/</a> (Stand 10.11.2020)
KON	“En el año 2009, se cerró un incinerador de residuos sólidos urbanos que funcionaba en Rotterdam”
QUE KON	<a href="https://www.no-burn.org/wp-content/uploads/Incineración-Enemigo-del-Reciclaje.pdf">https://www.no-burn.org/wp-content/uploads/Incineración-Enemigo-del-Reciclaje.pdf</a> (Stand 10.11.2020)
<b>DE</b>	<b>Müllverbrennungsanlage</b>
SYN	Abfallverbrennungsanlage
FACH	Umwelt
DEF	„Als "Müllverbrennungsanlage" gilt eine stationäre oder mobile technische Einheit und Anlage zur thermischen Müllverarbeitung, mit oder ohne Nutzung der durch die Verbrennung entstandenen Wärme.“
QUE DEF	<a href="https://www.fcc-group.eu/de/osterreich/technologien/behandlung/thermische-">https://www.fcc-group.eu/de/osterreich/technologien/behandlung/thermische-</a>

	abfallbehandlung/definition-der-mullverbrennungsanlage.html (Stand 09.11.2020)
KON	„Dieses Terminal, das vor allem für den Umschlag von Hausmüll aus den Gemeinden bestimmt ist, liegt in unmittelbarer Nähe einer Müllverbrennungsanlage in Alkmaar und ist seit April 2001 in Betrieb.“
QUE KON	<a href="https://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2003:327:0039:0045:DE:PDF">https://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2003:327:0039:0045:DE:PDF</a> (Stand 09.11.2020)

<b>ES</b>	<b>corredor ecológico</b>
SYN	corredor biológico,
FACH	Umwelt
DEF	“una gran región a través de la cual las áreas protegidas existentes (parques nacionales, reservas biológicas), o los remanentes de los ecosistemas originales, mantienen su conectividad mediante actividades productivas en el paisaje intermedio que permiten el flujo de las especies.”
QUE DEF	<a href="https://www.biodiversidad.gob.mx/corredor/corredoresbio.html">https://www.biodiversidad.gob.mx/corredor/corredoresbio.html</a> (Stand 10.11.2020)
KON	“Un corredor ecológico para reciclar el cemento de la costa en Alicante.”
QUE KON	<a href="https://elpais.com/ccaa/2019/02/17/valencia/1550426513_369559.html">https://elpais.com/ccaa/2019/02/17/valencia/1550426513_369559.html</a> (Stand 10.11.2020)
<b>DE</b>	<b>Biotopvernetzung</b>
SYN	Biotopverbindung
FACH	Umwelt
DEF	„räumliche oder funktionale Verbindung oder Anbindung von Biotopen oder Biotopkomplexen, um eine Zerstörung, Zersplitterung oder eine durch zunehmenden Raumwiderstand entstandene Isolation aufzuheben oder ihr entgegenzuwirken.“
QUE DEF	<a href="https://www.spektrum.de/lexikon/biologie/biotopvernetzung/8843">https://www.spektrum.de/lexikon/biologie/biotopvernetzung/8843</a> (Stand 09.11.2020)
KON	„Die Biotopvernetzung setzt mit einem eher kleinräumigen Ansatz im landwirtschaftlichen Umfeld den Biotopverbund auf der lokalen Ebene um und trägt damit zum Gesamtgelingen eines Biotopverbundes bei.“
QUE KON	<a href="https://www.landwirtschaft-bw.info/pb/Len/3650826_3651464_2315361_2316235">https://www.landwirtschaft-bw.info/pb/Len/3650826_3651464_2315361_2316235</a> (Stand 09.11.2020)

<b>ES</b>	<b>residuo tóxico</b>
SYN	desecho tóxico
FACH	Umwelt
DEF	“Los desechos tóxicos son un tipo de residuos peligrosos que pueden causar efectos nocivos sobre la salud de los seres humanos y otras especies de animales.”
QUE DEF	<a href="https://www.hidronor.cl/los-desechos-toxicos/">https://www.hidronor.cl/los-desechos-toxicos/</a> (Stand 10.11.2020)
KON	„El flúor que se añade al agua irlandesa es un residuo tóxico.“
QUE KON	<a href="https://www.europarl.europa.eu/sides/getDoc.do?pubRef=-//EP//TEXT+CRE+20060515+ITEMS+DOC+XML+V0//ES&amp;language=ES">https://www.europarl.europa.eu/sides/getDoc.do?pubRef=-//EP//TEXT+CRE+20060515+ITEMS+DOC+XML+V0//ES&amp;language=ES</a> (Stand 10.11.2020)
<b>DE</b>	<b>Giftmüll</b>
SYN	Sonderabfall
FACH	Umwelt
DEF	„(aus Gewerbe- und Industriebetrieben stammender) giftiger Müll, der die Umwelt schädigt und verseucht“
QUE DEF	<a href="https://www.duden.de/rechtschreibung/Giftmuell">https://www.duden.de/rechtschreibung/Giftmuell</a> (Stand 09.11.2020)
KON	„Wir müssen die Frage stellen, warum das Gelände nicht vom Giftmüll gereinigt wurde, sondern noch immer das Wasser verseucht, von dem die umliegenden Gemeinden abhängen.“
QUE KON	<a href="https://www.europarl.europa.eu/sides/getDoc.do?pubRef=-//EP//TEXT+CRE+20041216+ITEMS+DOC+XML+V0//DE&amp;language=DE">https://www.europarl.europa.eu/sides/getDoc.do?pubRef=-//EP//TEXT+CRE+20041216+ITEMS+DOC+XML+V0//DE&amp;language=DE</a> (Stand 09.11.2020)

<b>ES</b>	<b>contaminación ambiental</b>
SYN	contaminación del medio ambiente
FACH	Umwelt
DEF	“Territorio de extensión y configuración variables que, debido a su disposición y a su estado de conservación, conecta funcionalmente espacios naturales de singular relevancia para la flora o la fauna silvestres,…”
QUE DEF	<a href="https://dpej.rae.es/lema/corredor-ecológico">https://dpej.rae.es/lema/corredor-ecológico</a> (Stand 10.11.2020)
KON	“La contaminación ambiental en Arequipa se triplicó tras el inicio de la fase tres de la reactivación económica en comparación a los meses de marzo y abril informó el jefe de la Dirección Ambiental de la Gerencia Regional de Salud, Zacarías Madariaga.”

QUE KON	<a href="https://rpp.pe/peru/arequipa/arequipa-aumenta-contaminacion-ambiental-tras-la-reactivacion-economica-noticia-1292999">https://rpp.pe/peru/arequipa/arequipa-aumenta-contaminacion-ambiental-tras-la-reactivacion-economica-noticia-1292999</a> (Stand 10.11.2020)
<b>DE</b>	<b>Umweltverschmutzung</b>
SYN	Umweltbelastung
FACH	Umwelt
DEF	„Belastung, Schädigung der natürlichen Umwelt durch Schmutz, schädliche Stoffe o. Ä.“
QUE DEF	<a href="https://www.duden.de/rechtschreibung/Umweltverschmutzung">https://www.duden.de/rechtschreibung/Umweltverschmutzung</a> (Stand 10.11.2020)
KON	„Mit Beginn der Industrialisierung aber erreichte die Umweltverschmutzung eine neue Dimension.“
QUE KON	<a href="https://www.planet-wissen.de/natur/umwelt/umweltverschmutzung/pwieindustriellerevolutionundumweltverschmutzung100.html">https://www.planet-wissen.de/natur/umwelt/umweltverschmutzung/pwieindustriellerevolutionundumweltverschmutzung100.html</a> (Stand 10.11.2020)

#### 8.3.4. Text 4: Auszug aus *Casa tomada* von Julio Cortázar

<b>ES</b>	<b>liquidación</b>
SYN	redención
FACH	Industrie, Wirtschaft, Umwelt
DEF	“Venta al por menor, con gran rebaja de precios, que hace una casa de comercio por cesación, quiebra, reforma o traslado del establecimiento.”
QUE DEF	<a href="https://dle.rae.es/liquidación?m=form">https://dle.rae.es/liquidación?m=form</a> (Stand 10.11.2020)
KON	“Los anteriores costes pueden disminuirse en la cuantía de cualquier ingreso eventual que no se haya incluido entre los ingresos ordinarios del contrato, por ejemplo los ingresos por venta de materiales sobrantes o la liquidación del inmovilizado material, una vez acabado el contrato.”
QUE KON	<a href="https://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2008:320:0001:0481:ES:PDF">https://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2008:320:0001:0481:ES:PDF</a> (Stand 10.11.2020)
<b>DE</b>	<b>Verwertung</b>
SYN	Liquidation
FACH	Industrie, Wirtschaft, Umwelt
DEF	„Verwertung ist die Wertschöpfende Nutzung eines Gegenstands.“

QUE DEF	<a href="http://www.rechtslexikon.net/d/verwertung/verwertung.htm">http://www.rechtslexikon.net/d/verwertung/verwertung.htm</a> (Stand 10.11.2020)
KON	„Verwertung und Beseitigung von Elektro- und Elektronikaltgeräten sollen durch die Hersteller finanziert werden, damit wirtschaftliche Anreize dafür geschaffen werden, dass bei der Konstruktion dieser Geräte den Voraussetzungen für die sichere Entsorgung Rechnung getragen wird.“
QUE KON	<a href="https://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:C:2001:240E:0298:0302:DE:PDF">https://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:C:2001:240E:0298:0302:DE:PDF</a> (Stand 09.11.2020)

<b>ES</b>	<b>matrimonio entre hermanos</b>
SYN	matrimonio de hermanos
FACH	Geisteswissenschaft
DEF	“Matrimonio entre hermanos de ascendencia común.”
QUE DEF	<a href="https://www.ecured.cu/Endogamia">https://www.ecured.cu/Endogamia</a> (Stand 10.11.2010)
KON	„Un matrimonio entre hermanos logró ser legalizado en España“
QUE KON	<a href="https://www.lanacion.com.ar/el-mundo/un-matrimonio-entre-hermanos-logro-ser-legalizado-en-espana-nid66386/">https://www.lanacion.com.ar/el-mundo/un-matrimonio-entre-hermanos-logro-ser-legalizado-en-espana-nid66386/</a> (Stand 10.11.2020)
<b>DE</b>	<b>Geschwisterehe</b>
SYN	Geschwisterheirat
FACH	Geisteswissenschaft
DEF	„Geschwisterehe oder Geschwisterheirat bezeichnet die Verwandtenheirat zwischen leiblichen Geschwistern oder Halbgeschwistern. Sie ist eine seltene Sonderform der Endogamie, der Heirat innerhalb der eigenen sozialen Gruppe.“
QUE DEF	<a href="https://de.wikipedia.org/wiki/Geschwisterehe">https://de.wikipedia.org/wiki/Geschwisterehe</a> (Stand 09.11.2020)
KON	„Ein Beispiel ist die bei den Pharaonen des Alten Ägypten verbreitete Geschwisterehe, die auch nach der griechischen Eroberung unter den hellenistischen Herrschern beibehalten wurde, so war Kleopatra VII. mit ihren Brüdern Ptolemaios XIII. und Ptolemaios XIV. verheiratet.“
QUE KON	<a href="https://www.besthelp.at/lexikon/inzest">https://www.besthelp.at/lexikon/inzest</a> (Stand 09.11.2020)

<b>ES</b>	<b>genealogía</b>
SYN	
FACH	Geisteswissenschaft

DEF	“Serie de progenitores y ascendientes de una persona y disciplina que estudia la genealogía de las personas.”
QUE DEF	<a href="https://dle.rae.es/genealog%C3%ADa?m=form">https://dle.rae.es/genealog%C3%ADa?m=form</a> (Stand 10.11.2020)
KON	“La identificación de Joseph Loveless es por el momento el récord mundial de la genealogía genética.”
QUE KON	<a href="https://elpais.com/elpais/2020/01/08/opinion/1578503270_023319.html">https://elpais.com/elpais/2020/01/08/opinion/1578503270_023319.html</a> (Stand 10.11.2020)
<b>DE</b>	<b>Genealogie</b>
SYN	Ahnenforschung
FACH	Geisteswissenschaft
DEF	„Wissenschaft von den auf Abstammung (Stammbaum) beruhenden Zusammenhängen zwischen Menschen (selten: Organismen allgemein)“
QUE DEF	<a href="https://www.spektrum.de/lexikon/biologie/genealogie/27218">https://www.spektrum.de/lexikon/biologie/genealogie/27218</a> (Stand 10.11.2020)
KON	„Durch die trojanisch-fränkische Genealogie konnte er "alles edle Blut des Kontinents" in seiner Person vereinigen. Mit großem Aufwand wurde die Stammfolge bis Hektor zusammengetragen.“
QUE KON	<a href="http://www.aeiou.at/aeiou.encyclop.g/g235218.htm">http://www.aeiou.at/aeiou.encyclop.g/g235218.htm</a> (Stand 09.11.2020)

<b>ES</b>	<b>mañanita</b>
SYN	
FACH	Geisteswissenschaft
DEF	“Prenda de vestir, de punto o tela, que cubre desde los hombros hasta la cintura y que las mujeres usan principalmente para estar sentadas en la cama.”
QUE DEF	<a href="https://dle.rae.es/mañanita?m=form">https://dle.rae.es/mañanita?m=form</a> (Stand 10.11.2020)
KON	“Camisetas y camisas, combinaciones o forros, faldillas, bragas, camiones, pijamas, "mañanitas", albornoces, batas y artículos análogos, para mujeres o niñas, distintos de los de punto.”
QUE KON	<a href="https://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=COM:2000:0441:FIN:ES:PDF">https://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=COM:2000:0441:FIN:ES:PDF</a> (Stand 10.11.2020)
<b>DE</b>	<b>Bettjacke</b>
SYN	
FACH	Geisteswissenschaft
DEF	„Jacke, die man zum Schlafen über der sonstigen Nachtwäsche trägt.“

QUE DEF	<a href="https://www.wortbedeutung.info/Bettjacke/">https://www.wortbedeutung.info/Bettjacke/</a> (Stand 09.11.2020)
KON	„Eine traditionelle Facette weiblicher Rollenmuster nimmt die Gestalt einer Bettjacke an, deren außen weich aufgeplusterte, Schutz versprechende Wattierung im Inneren mit spitzen Nadeln gespickt den schmerzhaften Panzer in der kleinkindhaften Hülle bloß legt.“
QUE KON	<a href="http://www.evakoestner.de/Text_Maderna_dt.htm">http://www.evakoestner.de/Text_Maderna_dt.htm</a> (Stand 09.11.2020)

<b>ES</b>	<b>pretendiente</b>
SYN	aspirante
FACH	Geisteswissenschaft
DEF	„Un pretendiente es quien pretende alguna cosa. La acción de pretender, por su parte, consiste en desear, anhelar o apetecer algo. El uso más habitual del concepto de pretendiente refiere a la persona que desea convertirse en la pareja sentimental de otro individuo.“
QUE DEF	<a href="https://definicion.de/pretendiente/">https://definicion.de/pretendiente/</a> (Stand 10.11.2020)
KON	“Corte Constitucional reconoce el derecho de las mujeres de rechazar a un pretendiente.”
QUE KON	<a href="https://www.semana.com/nacion/articulo/corte-constitucional-el-derecho-de-las-mujeres-a-rechazar-a-un-hombre/627855/">https://www.semana.com/nacion/articulo/corte-constitucional-el-derecho-de-las-mujeres-a-rechazar-a-un-hombre/627855/</a> (Stand 10.11.2020)
<b>DE</b>	<b>Verehrer</b>
SYN	
FACH	Geisteswissenschaft
DEF	„Mann, der eine Frau verehrt, sich um sie bemüht und von anderen als zu ihr gehörend betrachtet wird.“
QUE DEF	<a href="https://www.duden.de/rechtschreibung/Verehrer">https://www.duden.de/rechtschreibung/Verehrer</a> (Stand 10.11.2020)
KON	„Auch wenn sie jedem von ihrem Freund in Deutschland erzählt, hält das ihre Verehrer nicht ab.“
QUE KON	<a href="https://www.stern.de/familie/beziehung/julia-peirano/meine-verehrer-bedaengen-mich---wie-werde-ich-sie-los--7034992.html">https://www.stern.de/familie/beziehung/julia-peirano/meine-verehrer-bedaengen-mich---wie-werde-ich-sie-los--7034992.html</a> (Stand 10.11.2020)

## 9. Auswertung: Gesamte Texte

In diesem Kapitel werden nun die Ergebnisse der Humanübersetzung mit den Ergebnissen der maschinellen Übersetzung verglichen. Hierfür werden zunächst die gesamten Texte betrachtet und die Fehler beziehungsweise Abweichungen der maschinellen Übersetzung aufgelistet sowie zusätzlich in den oben beschriebenen Bewertungsbogen eingegeben. Anzumerken ist, dass es sich beim Übersetzen stets um eine persönliche Entscheidung bezüglich der Auswahl der Termini handelt, weshalb es schwierig ist, jede einzelne Abweichung zu bewerten. Es wird ein Ergebnis berechnet, mit dessen Hilfe die Übersetzung des Textes in seiner Gesamtheit beurteilt werden kann. Dabei folgt die Präsentation der Humanübersetzung, anschließend jene der zugehörigen maschinellen Übersetzung, welche mit dem Übersetzungstool *DeepL* durchgeführt wurde. An dieser Stelle ist anzumerken, dass bei der maschinellen Übersetzung weder eine Prä- noch eine Postedition stattgefunden haben. Die Humanübersetzungen, die korrigiert wurden, dienen lediglich als Vergleich beziehungsweise als „Lösung“ und werden daher nicht separat bewertet.

Nachdem die Texte als Ganzes beleuchtet worden sind, wird des Weiteren ein genauerer Blick auf die Übersetzungen der recherchierten Terminologie geworfen (siehe Kap. 10). Dabei werden die ausgewählten Termini der Humanübersetzung und der maschinellen Übersetzung mit meinem Glossar verglichen. Auch hier sollen Fehler beziehungsweise Abweichungen aufgezeigt werden.

### 9.1 Vergleich Human- vs. maschinelle Übersetzung: Text 1

#### 9.1.1 Übersetzung: Human

Y aquí una cita del Papa Francisco: “Una economía justa debe crear las condiciones para que cada persona pueda gozar de una infancia sin carencias, desarrollar sus talentos durante la juventud, trabajar con plenos derechos durante los años de actividad y acceder a una digna jubilación en la ancianidad. Es una economía donde el ser humano en armonía con la naturaleza, estructura todo el sistema de producción y distribución para que las capacidades y las necesidades de cada uno encuentren un cauce adecuado en el ser social.”

No está mal: A primera vista, estas palabras recogen todo lo que hay que decir sobre el tema. Francisco no empieza hablando de “la economía”, sino de una economía “justa” y la relaciona

con ciertas condiciones: Es una economía que es el medio, pero no el fin. En concreto, es el medio para posibilitar una "infancia sin carencias". Se puede suponer que Francisco aquí no se refiere a la carencia de un iPhone o de un viaje escolar a Mallorca, sino a carencias en el sentido sustancial, es decir, a una infancia marcada por hambre, trabajo, guerra y otras limitaciones existenciales. De hecho, se trata de una reivindicación conectada con la modernidad y los derechos humanos: Todos los niños y niñas deberían tener las mismas posibilidades respecto al crecer y desarrollar sus capacidades, talentos e intereses, independientemente de su sexo y su origen nacional. Con otras palabras, un niño o una niña debería poder convertirse en un individuo sin carencias ni temor: un ser humano libre.

### **9.1.2 Übersetzung: DeepL**

Y aquí está una cita de Francisco: "Una economía justa debe crear las condiciones para que cada persona disfrute de una infancia sin privaciones, mientras que los jóvenes pueden desarrollar sus talentos, tomar un trabajo con seguridad jurídica durante los años de actividad y recibir una pensión digna en la vejez. Es una economía en la que el hombre, en armonía con la naturaleza, da forma a todo el sistema de producción y distribución de manera que las capacidades y necesidades de cada individuo encuentren un marco apropiado en las comunidades".

No está mal: a primera vista, todo lo que está en juego está ahí. Francisco no empieza con "la economía", sino con una economía "justa", colocándola así en una relación condicional: es un medio, no un fin. Es un medio para hacer posible una "infancia sin privaciones". Se puede suponer que Francisco no significa aquí la privación de un iPhone o un viaje de clase a Mallorca, sino la privación en todo sentido sustancial, que la infancia se caracteriza por el hambre, el trabajo, la guerra y otras restricciones existenciales. De hecho, se trata de una reivindicación que está directamente vinculada a la modernidad y a los derechos humanos: todos los niños deberían tener las mismas oportunidades de crecer y desarrollar sus aptitudes, talentos y preferencias, independientemente de su género y origen. En otras palabras, debería ser capaz de convertirse en un ser que no esté marcado por la falta y el miedo: un ser humano libre.

### 9.1.3 Auswertung

Gleich zu Beginn der Übersetzung finden wir den Satz: *Y aquí está una cita de Francisco*. Bei der Humanübersetzung wurde dieser Teil eigenen Einschätzungen nach sehr viel einfacher, aber auch flüssiger als *Y aquí una cita del Papa Francisco* übersetzt.

Die Passage im deutschen Ausgangstext *Da ist auf den ersten Blick alles drin, worum es geht* wurde mit *a primera vista, todo lo que está en juego está ahí* übersetzt. Vor allem die Übersetzung von *Está en juego* ist nicht passend; auch wenn die Botschaft dadurch übermittelt werden kann, würde ich in Bezug auf mein Sprachgefühl “da ist alles im Spiel” nicht in diesem Kontext verwenden.

Die Passage im Ausgangstext [...], *dass Franziskus hier nicht [...] meint* kann unmöglich mit *Francisco no significa* übersetzt werden, da dies soviel wie *Franziskus bedeutet nicht* heißen würde und die Bedeutung gänzlich verfälscht.

Die Übersetzung des Begriffs *género* für den deutschen Begriff *Geschlecht* ist zwar nicht falsch, aber ich würde in diesem Kontext auf geeignetere Übersetzungen zurückgreifen, da sich der Begriff *género* häufiger auf die deutschen Bezeichnungen *Gattung, Genus, Sorte* bezieht.

Auch wenn der spanische Begriff *privación* soviel wie *Entbehrung, Entzug* etc. bedeuten kann, so sollte im Kontext der Übersetzung eher auf den spanischen Terminus *carencia*, wie er auch in der Humanübersetzung verwendet wurde, zurückgegriffen werden. Dabei handelt es sich um wiederum um keinen Fehler, sondern um ein persönliches Empfinden. Als Übersetzerin würde ich mit meinem Sprachgefühl eher zu der Übersetzung *carencia* tendieren. Diese Abweichung müsste bei einer Korrektur nicht zwingend übernommen werden, anmerken würde ich es aber trotzdem.

Alle oben genannten Beispiele wurden als leichte Fehler in der Kategorie *Wrong Term* gewertet.

Fehlerkategorie	Anzahl * schwere Fehler	Anzahl * leichte Fehler	GESAMT
WT	...*5	5 * 2	10
SE	...*4	...*2	
OM	...*4	...*2	
SA	...*4	...*2	
SP	...*3	...*1	
PE	...*2	...*1	
ME	...*3	...*1	

**Abbildung 7:** Eigene Darstellung. *Auswertungsbogen Text 1.*

Wie bereits weiter oben beschrieben, werden bei VW und Audi maximal drei Fehlerpunkte pro 100 Wörter des Ausgangstextes toleriert. Dies entspricht einem erlaubten Maximalwert von 0,03 (vgl. Dalla-Zuanna 2010:).

Dividiert man nun 160 Fehlerpunkte durch die Wortanzahl des Textes (248 Wörter), erhält man einen Wert von 0,0403. Demnach wäre das Ergebnis dieser maschinellen Übersetzung als unzureichend einzustufen. Dennoch wurde die Botschaft ohne Missverständnisse vermittelt, wobei einige Termini nicht sehr flüssig klingen. Nichtsdestotrotz sind viele andere Termini optimal übersetzt worden; auch in Bezug auf Syntax und Grammatik waren bei diesen keinerlei Fehler zu finden. Ausgelassen wurde ebenfalls nichts. In diesem Fall kann gesagt werden, dass die Verwendbarkeit der Übersetzung ganz klar von ihrem Verwendungszweck abhängt. Sollte der Text veröffentlicht werden, so wäre es ratsam, diesen überarbeiten zu lassen, da die Qualität nicht ausreicht und das Ergebnis der Auswertung für sich spricht. Dient die Übersetzung lediglich zum Verständnis oder für interne Zwecke, so ist die maschinelle Übersetzung durchaus als ausreichend zu werten.

## 9.2 Vergleich Human- vs. maschinelle Übersetzung: Text 2

### 9.2.1 Übersetzung: Human

ALSTOM Austria forma parte del grupo ALSTOM, uno de los líderes mundiales en el sector de la energía y del transporte. Con alrededor de 60.000 empleados en más de 70 países la empresa generó un volumen de ventas de unos 14 millardos de euros el año fiscal pasado.

Desde junio de 1998 ALSTOM cotiza en la Bolsa de París.

La gama de productos y servicios de ALSTOM Austria abarca:

- plantas de energía “llave en mano“
- turbinas, generadores, tecnología de control de centrales eléctricas, tecnología medioambiental, equipo auxiliares
- servicio al cliente incluyendo contratos de mantenimiento a largo plazo
- desarrollo y financiación de proyectos

Su competencia y larga experiencia constituyen también una base sólida para una mayor actividad exportadora de la empresa. La cuota de exportación de ALSTOM Austria representa alrededor de un 50% de la facturación total. La empresa opera principalmente en el mercado de países del sureste de Europa, sobre todo en Bulgaria, Bosnia-Herzegovina y Turquía, pero también en Bután y otros mercados de exportación. El éxito de la empresa se debe, ante todo, a la competencia técnica de los directores de proyecto y su amplio conocimiento y experiencia en la estructuración de modelos complejos de financiación.

Las compañías que trabajan en el sector energético en Austria contratan a casi 80 empleados y generaron un volumen de ventas de cerca de 50 millones de euros en el ejercicio pasado.

### **9.2.2 Übersetzung: DeepL**

ALSTOM Austria forma parte del Grupo ALSTOM, líder mundial en el sector de la energía y el transporte. Con alrededor de 60.000 empleados en más de 70 países, la compañía alcanzó ventas de aproximadamente 14 mil millones de euros en el último año fiscal.

ALSTOM cotiza en la bolsa de París desde junio de 1998.

El alcance de la oferta y los servicios de ALSTOM Austria incluye

- centrales eléctricas llave en mano
- Turbinas, generadores, sistemas de control de centrales eléctricas, tecnologías ambientales, equipos auxiliares
- Servicio de atención al cliente, incluidos los contratos de mantenimiento a largo plazo
- Desarrollo y financiación de proyectos

La competencia y los muchos años de experiencia también constituyen una base sólida para el aumento de las actividades de exportación de la empresa. ALSTOM Las exportaciones de Austria representan alrededor del 50% del volumen de negocios total. Los principales mercados son principalmente los países de Europa sudoriental, especialmente Bulgaria, Bosnia y Herzegovina y Turquía, pero también Bhután y otros mercados de exportación. Además de la competencia técnica de los directores de proyectos, un factor clave del éxito de la empresa son los amplios conocimientos técnicos sobre la estructuración de modelos de financiación complejos.

Las empresas austríacas dedicadas al sector de la energía emplean actualmente a unas 80 personas y generaron ventas de unos 50 millones de euros en el último ejercicio económico.

### 9.2.3 Auswertung

Die Phrase *Der Liefer- und Leistungsumfang* im deutschen Ausgangstext wurde in der maschinellen Übersetzung mit *El alcance de la oferta* übersetzt. Natürlich beschreibt dies den Umfang des Angebots und man kann verstehen, was damit gemeint ist. Nichtsdestotrotz könnte die Idee mit *La gama de productos y servicios* um ein Vielfaches präziser und einfacher beschrieben werden. Da es sich somit um eine Abweichung, man den Sinn aber trotzdem noch versteht, wurde diese als leichter Fehler in der Kategorie *Wrong Term* markiert.

Im dritten Absatz nach der Aufzählung des Liefer- und Leistungsumfangs schien die maschinelle Übersetzung jedoch mit der Firmenbezeichnung *ALSTOM Austria* nicht zurechtzukommen: *La cuota de exportación de ALSTOM Austria representa alrededor de un 50% de la facturación total* wurde mit *ALSTOM Las exportaciones de Austria representan alrededor del 50% del volumen de negocios total* übersetzt. Dabei erkannte *DeepL* nicht, dass die beiden Termini zusammengehören. Dieses Missverständnis wurde als schwerer Fehler gewertet, da dieser sehr auffällig ist und beim Lesen sofort erkannt wird, dass dieser Satz nicht korrekt sein kann. Ich werte diese Abweichung als einen schweren Fehler in der Kategorie *Structural Error*.

Als letzten Punkt wäre bei einem Lektorat die Übersetzung des Begriffs *Gesamtumsatz* anzumerken. Bei der maschinellen Übersetzung wurde *volumen de negocios* verwendet. Auch wenn dieser Begriff in dieser Form in manchen Wörterbüchern zu finden ist, sollte eher zu *facturación total* oder *cifra de ventas* zurückgegriffen werden. Diese Tendenz

basiert einzig und allein auf meinem Sprachgefühl. Da man den Sinn der Botschaft trotz dieser Abweichung versteht, wurde dieser Punkt als leichter Fehler in der Kategorie *Wrong Term* eingetragen.

Fehlerkategorie	Anzahl * schwere Fehler	Anzahl * leichte Fehler	GESAMT
WT	...*5	2 * 2	4
SE	...*4	...*2	
OM	...*4	...*2	
SA	...*4	1 * 2	2
SP	...*3	...*1	
PE	...*2	...*1	
ME	...*3	...*1	

**Abbildung 8:** Eigene Darstellung. *Auswertungsbogen Text 2.*

Dividiert man auch dieses Mal die Anzahl der Fehler (sechs Fehlerpunkte) durch die Anzahl der Wörter dieses Fachtextes (229 Wörter), so erhält man ein Ergebnis von 0,0262 Punkten. Damit liegt der Wert deutlich unter der festgelegten maximalen Toleranzgrenze von 0,3. Dieser Auswertung nach zu urteilen, ist das Ergebnis dieser Übersetzung allgemein betrachtet sehr gut. In diesem Fachtext kamen sehr viele Fachtermini vor, die im Normalfall vorab recherchiert werden sollten. Fast alle dieser Termini wurden dabei von *DeepL* korrekt übersetzt. Es gab weder Auslassungen noch Fehler in der Syntax oder in Bezug auf die Grammatik.

## 9.3 Vergleich Human- vs. maschinelle Übersetzung: Text 3

### 9.3.1. Übersetzung: Human

Sind umweltfreundlichere Städte etwa ungerecht(er)?

Verbesserung der Umweltqualität in den traditionell benachteiligten Stadtteilen

In den 1980er-Jahren, als sich amerikanische und andere AktivistInnen für

Umweltgerechtigkeit (auf Englisch "EJ" für environmental justice) zum ersten Mal öffentlich

sichtbar zusammenschlossen, um die ungleichen Auswirkungen der Umweltverschmutzung der ungleichen Verteilung von Abfallbelastung die Stirn zu bieten, war es offensichtlich, wen sie im Blickfeld hatten: Umweltfeindliche Unternehmen, Rechtsbrecher in der Abwasserentsorgung oder Betreiber von Müllverbrennungsanlagen sowie deren jeweilige Handlanger in der Regierung, die es nicht geschafft hatten, ihre Tätigkeiten ordnungsgemäß zu regeln oder die Probleme nicht als solche erkennen wollten. Es entstanden sogenannte LULUs (englische Abkürzung für Locally Unwanted Land Use), also von der lokalen Bevölkerung unerwünschte Landnutzungen.

Seit den ersten Protesten der UmweltaktivistInnen 1978 in Love Canal, im US-Bundesstaat New York, sowie 1982 in Warren Country, North Carolina, wird darauf hingewiesen, dass Giftmüll vor allem wirtschaftlich schwächere und von Minderheiten bewohnte Communities betreffen. Dies war der Beginn eines langen Kampfes gegen Ungleichheit und Umweltrassismus; ein Kampf der in Anbetracht der neuesten Fälle von Bleiverschmutzung und Vergiftung von mehr als 6000 mehrheitlich farbigen Kindern in Flint, Michigan, aktueller denn je ist.

Nichtsdestotrotz erlangte der Aktivismus für Umweltgerechtigkeit in urbanen Zonen im Laufe der Zeit mehrere unterschiedliche Ausrichtungen und umfasst etwa auch lokale Gruppen, die sich in verschiedenen Projekten für eine aktive Umweltverbesserung engagieren. Dazu zählen City Farmen, Gärten, Biotopvernetzung, Spielplätze und Parks, die sich gegenseitig unterstützen, körperliche und psychische Gesundheitsbedürfnisse ansprechen und nach Lösungen für jahrelange Vernachlässigung des Gemeinwesens und Umwelttraumata suchen. Zu den Best-Practice-Beispielen dieser Art an kollektiver Mobilisierung zählen jene der Stadtviertel Sant Pere und Santa Caterina in Barcelona, Cayo Huerto in La Habana (Kuba) oder auch Dudley in Boston.

### **9.3.2 Übersetzung: DeepL**

Sind grünere Städte ungleicher?

Verbesserung der Umweltqualität in traditionell benachteiligten städtischen Gemeinden

In den 1980er Jahren, als die Aktivisten der Umweltgerechtigkeit (EJ) in den Vereinigten Staaten und anderen Ländern sich erstmals sichtbar organisierten, um die ungleichen Auswirkungen der Umweltverschmutzung und des Abfalls anzugehen, war klar, wer im Visier war: umweltverschmutzende Fabriken, Deponiebetreiber oder Betreiber von

Müllverbrennungsanlagen sowie ihre Komplizen in den Regierungen, die nicht in der Lage waren, ihre Aktivitäten zu regulieren, oder die ein Auge zugedrückt hatten. Sie bildeten das, was die Bevölkerung als unerwünschte Landnutzung durch örtliche Gemeinschaften (Unwanted Land Use by Local Communities, LULUs) bezeichnete.

Seit den ersten Protesten von Umweltgerechtigkeitsaktivisten 1978 in Love Canal, New York State, und 1982 in Warren County, North Carolina, haben sie offenbart, dass giftige Abfälle besonders in einkommensschwachen Gemeinden und Minderheitengemeinschaften weit verbreitet waren, und sie haben einen langen Kampf zur Wiedergutmachung von Umweltungerechtigkeit und Rassismus eingeleitet, ein Kampf, der nach der jüngsten Bleiverseuchung und Vergiftung von mehr als 6 Personen wieder aktuell geworden ist. Der erste davon war der Fall des Vorfalls in Flint, Michigan, bei dem mehr als 6.000 Kinder, meist Schwarze, vergiftet wurden.

Im Laufe der Zeit ist der Aktivismus für städtische Umweltgerechtigkeit jedoch vielschichtig geworden und umfasst lokale Gruppen, die an verschiedenen Projekten zur aktiven Verbesserung der Umwelt arbeiten, wie z.B. städtische Bauernhöfe, Obstgärten, grüne Korridore, Spielplätze und Parks, die sich gegenseitig unterstützen, körperliche und psychische Gesundheitsbedürfnisse ansprechen und jahrelange Vernachlässigung der Gemeinschaft und Umwelttraumata aufarbeiten. Zu den emblematischen Beispielen für diese Art der kollektiven Mobilisierung gehören die Stadtviertel Sant Pere und Santa Caterina in Barcelona, Key West in Havanna oder Dudley in Boston.

### 9.3.3 Auswertung

Gleich zu Beginn der maschinellen Übersetzung fällt bei einem Lektorat die Überschrift ins Auge. Die Überschrift der Humanübersetzung *Sind umweltfreundlichere Städte etwa ungerecht(er)?* liest sich sehr flüssig, während insbesondere der Begriff *ungleicher* in der maschinellen Übersetzung jedoch sehr holprig klingt. *Sind grünere Städte ungleicher?* wirkt dabei sehr übersetzt und sollte auf jeden Fall überarbeitet werden. Für diese Abweichung wurde ein leichter Fehler in der Kategorie *Wrong Term* eingetragen. Auch wenn die Verwendung dieses Terminus keinerlei schädliche Folgen für das Textverständnis mit sich zieht, so entsteht durch diese Wahl gleich zu Beginn dieser Übersetzung ein eher schlechter Eindruck bezüglich der Qualität der maschinellen Übersetzung.

Im zweiten Absatz gab es bei dem Thema der Bleivergiftung einige Schwierigkeiten. [...]ein Kampf, der nach der jüngsten Bleiverseuchung und Vergiftung von mehr als 6 Personen wieder aktuell geworden ist. Der erste davon war der Fall des Vorfalls in Flint, Michigan, bei dem mehr als 6.000 Kinder, meist Schwarze, vergiftet wurden. Der durchgestrichene Teil des Satzes ergibt keinerlei Sinn. Es ist nicht erklärbar, woher die 6 Personen kommen; der Fall des Vorfalls klingt zusätzlich falsch. Bei der Humanübersetzung wurde diese Passage weitaus verständlicher übersetzt: [...] ein Kampf der in Anbetracht der neuesten Fälle von Bleiverschmutzung und Vergiftung von mehr als 6000 mehrheitlich farbigen Kindern in Flint, Michigan, aktueller denn je ist. Bei der maschinellen Übersetzung müsste dieser Teil dringend überarbeitet werden. Da dieser Fehler deutlich bemerkbar ist, wurde er als leichter Fehler in der Kategorie *Structural Error* gewertet; die Beschaffenheit der fehlerhaften Übersetzung lässt außerdem darauf schließen, dass *DeepL* kein Problem mit der Terminologie, sondern mit der Satzstruktur hatte.

Als letzter Punkt wäre die Übersetzung des spanischen Terminus *corredores ecológicos* anzumerken. *DeepL* hat dabei die eher wortwörtliche Übersetzung *grüne Korridore* gewählt. Auch wenn die Botschaft dennoch verständlich ist, so ist diese Version nicht sonderlich elegant. Stattdessen könnte etwa *Biotopvernetzung* oder *Biotopverbindung* gewählt werden. Für diese eher geringfügige Abweichung wurde ein leichter Fehler in der Kategorie *Wrong Term* eingetragen.

Insgesamt betrachtet bleibt allerdings ein sehr guter Eindruck der maschinellen Übersetzung. Viele Fachtermini wurden korrekt übersetzt und der Text liest sich insgesamt sehr flüssig. Einige schwierige Passagen wurden erstaunlicherweise außerordentlich adäquat übersetzt, vor allem unter Anbetracht dessen, dass diese Übersetzung weder prä- noch posteditiert worden ist.

Fehlerkategorie	Anzahl * schwere Fehler	Anzahl * leichte Fehler	GESAMT
WT	...*5	2 * 2	4
SE	...*4	...*2	
OM	...*4	...*2	
SA	...*4	1 * 2	2
SP	...*3	...*1	
PE	...*2	...*1	
ME	...*3	...*1	

**Abbildung 9:** Eigene Darstellung. *Auswertungsbogen Text 3.*

Dividiert man Umfang des Textes (267 Wörter) durch die Anzahl der Fehler (sechs Fehlerpunkte), so liegt das Ergebnis bei 0,0224, das klar unter dem Maximalwert von 0,3 liegt. Auch wenn in diesem kurzen Textausschnitt drei leichte Fehler beziehungsweise Abweichungen vorkommen, so ist die maschinelle Übersetzung dennoch von guter Qualität. Die Botschaft des Textes wird ohne großartige Änderungen klar vermittelt, was immerhin mit einer der wichtigsten Aspekte einer Übersetzung ist.

## **9.4 Vergleich Human- vs. maschinelle Übersetzung: Text 4**

### **9.4.1 Übersetzung: Human**

Das besetzte Haus

Wir mochten das Haus, weil es nicht nur geräumig und alt war (vor allem da heutzutage die alten Häuser der gewinnbringenderen Verwertung der Materialien erliegen und abgerissen werden), sondern auch, weil es die Geheimnisse unserer Urgroßeltern, unseres Großvaters väterlicherseits, unserer Eltern und unserer gesamten Kindheit beherbergte.

Irene und ich hatten uns schon daran gewöhnt, alleine im Haus verblieben zu sein, was an sich verrückt war, da in diesem Haus acht Personen leben konnten, ohne sich gegenseitig im Weg zu stehen. Am Vormittag putzten wir das Haus. Dafür standen wir um sieben Uhr auf. Um etwa elf Uhr überließ ich Irene die letzten noch zu erledigenden Räume und ging in die Küche. Zur Mittagsstunde aßen wir, immer zur selben Uhrzeit, und abgesehen von etwas schmutzigem Geschirr, das anfiel, gab es nichts mehr zu tun. Wir hatten Gefallen daran gefunden, beim Mittagessen uns Gedanken über das große und ruhige Haus zu machen und wie nur wir beide alleine es schafften, es sauber zu halten. Manchmal dachten wir, es wäre das Haus selbst, welches uns nicht erlauben würde, zu heiraten. Ohne einen wirklichen Grund zu nennen, hatte Irene bereits zwei Verehrern einen Korb verpasst und meine Maria Esther war gestorben, noch bevor wir verlobt waren. Wir waren Anfang vierzig und verfolgten die unausgesprochene Idee, dass unsere schlichte und stille Geschwisterehe das unausweichliche Ende der von unseren Urgroßeltern in diesem Haus etablierten Genealogie war. Eines Tages würden wir hier sterben. Faule, scheue Cousins würden das Haus erben und es niederreißen,

um sich an dem Grundstück und den Gemäuern zu bereichern. Besser wäre, wir würden es selbst im Sinne der Gerechtigkeit zur Gänze auseinandernehmen, noch bevor es zu spät war. Irene war von Natur aus so, dass sie nie jemanden störte. Abgesehen von ihrer morgendlichen Tätigkeit, verbrachte sie den Rest des Tages damit, am Sofa ihres Schlafzimmers zu stricken. Ich weiß nicht, warum sie so viel strickte. Ich glaube ja, dass Frauen dann strickten, wenn sie in dieser Tätigkeit den Vorwand für ihre ansonsten herrschende Untätigkeit gefunden hatten. Irene war anders. Sie strickte immer brauchbare Dinge: Jacken für den Winter, Strümpfe für mich, Bettjacken und Westen für sich selbst. Manchmal strickte sie eine Weste und plötzlich trennte sie wieder alles auf, weil ihr irgendetwas daran nicht gefiel. Es war lustig zu beobachten, wie im Handarbeitskorb die Wollknäuel unruhig herumsprangen und gegen den raschen Formverlust anzukämpfen schienen. Samstags ging ich immer in die Stadt, um neue Wolle zu kaufen. Irene vertraute meinem Geschmack. Ihr gefiel, was ich auswählte und ich musste nie einen Knäuel zurückbringen. Ich nutzte diese kleinen Ausflüge, um die eine oder andere Buchhandlung aufzusuchen und mich vergeblich danach zu erkundigen, ob neue französische Literatur eingetroffen war. Seit 1939 war nichts Lesenswertes mehr nach Argentinien gekommen.

#### **9.4.2 Übersetzung: DeepL**

Das Haus ist besetzt; Julio Cortázar

Das Haus gefiel uns, weil es nicht nur geräumig und alt ist (heute, wo alte Häuser der vorteilhaftesten Liquidierung ihrer Materialien erliegen), sondern auch die Geheimnisse unserer Urgroßeltern, des Großvaters väterlicherseits, unserer Eltern und der ganzen Kindheit bewahrt hat.

Wir haben uns daran gewöhnt, dass Irene und ich allein darin beharrten, was verrückt war, weil acht Leute in diesem Haus leben konnten, ohne im Weg zu stehen. Wir räumten morgens auf, standen um sieben Uhr auf, und um etwa elf Uhr verließ ich die letzten Zimmer, damit Irene hinübergehen und in die Küche gehen konnte. Wir aßen mittags zu Mittag, immer pünktlich; es gab nichts mehr zu tun außer ein paar schmutzigem Geschirr. Es war schön, beim Mittagessen über das tiefe, stille Haus nachzudenken und darüber, dass wir genug waren, um es sauber zu halten. Manchmal kamen wir zu der Überzeugung, dass sie diejenige war, die uns nicht heiraten lassen wollte. Irene lehnte zwei Bewerber ohne jeglichen Grund ab, und ich wurde von Maria Esther vor unserer Verlobung getötet. Wir traten in unsere

vierziger Jahre mit der unausgesprochenen Idee ein, dass unsere einfache und stille Brüderehe den Abschluss, der von unseren Urgroßeltern in unserem Haus eingerichteten Genealogie erforderte. Wir würden dort eines Tages sterben, faule und schwer fassbare Cousins würden das Haus nehmen und es zu Boden werfen, um sich mit dem Land und den Ziegelsteinen zu bereichern; oder besser, wir würden es uns selbst gerecht übergeben, bevor es zu spät ist. Irene war ein Mädchen, das geboren wurde, um niemanden zu stören. Abgesehen von ihrer morgendlichen Aktivität verbrachte sie den Rest des Tages mit Stricken auf dem Sofa in ihrem Schlafzimmer. Ich weiß nicht, warum sie so viel strickte, ich glaube, Frauen stricken, wenn sie in dieser Arbeit den großen Vorwand gefunden haben, nichts zu tun. Irene war nicht so, sie strickte Dinge, die immer notwendig waren, sie strickte für den Winter, Socken für mich, morgens und Westen für sie. Manchmal strickte sie eine Weste und webte sie dann gleich wieder auf, weil ihr etwas nicht gefiel; es war lustig zu sehen, wie sich der Flor krauser Wolle im Korb einige Stunden lang weigerte, seine Form zu verlieren. Samstags ging ich in die Innenstadt, um ihre Wolle zu kaufen; Irene hatte Vertrauen in meinen Geschmack, sie war mit den Farben zufrieden, und ich musste nie Stränge zurückgeben. Ich nutzte diese Ausflüge, um in den Buchläden herumzugehen und vergeblich zu fragen, ob es etwas Neues in der französischen Literatur gäbe. Seit 1939 gelangte nichts Wertvolles nach Argentinien.

### 9.4.3 Auswertung

Gleich zu Beginn dieser Auswertung soll erwähnt werden, dass die maschinelle Übersetzung dieses Fachtextes nicht sonderlich gut gelungen ist und überdurchschnittlich viele Abweichungen und Fehler hervorgehoben werden müssen.

*DeepL* hat den Titel des Buches mit *Das Haus ist besetzt* übersetzt. Da dieses Buch bereits mehrmals ins Deutsche übersetzt wurde, wäre es sinnvoll auch den in der Buchübersetzung benutzten deutschen Titel (*Das besetzte Haus*) zu übernehmen. Dieser Fehler ist als leichter Fehler in der Kategorie *Wrong Term* einzuordnen.

Zu Beginn des zweiten Absatzes kam es zu der Übersetzung: [...] *was verrückt war, weil acht Leute in diesem Haus leben konnten, ohne im Weg zu stehen*. Dabei fehlt ganz klar die Ergänzung: *sich gegenseitig im Weg zu stehen* oder *eineinander im Weg zu stehen*. Dieser Fehler sollte als leichter *Auslassungsfehler* markiert werden.

Einen Satz weiter sollte es heißen [...] *ich überließ Irene die letzten Zimmer* und nicht – wie in der maschinellen Übersetzung – [...] *und um etwa elf Uhr verließ ich die letzten Zimmer*. Diesen Fehler ist ein klassischer leichter Fehler in der Kategorie *Wrong Term*.

Die Übersetzung *Wir aßen mittags zu Mittag* funktioniert ebenso nicht und sollte wie beispielsweise in der Humanübersetzung mit *Zur Mittagsstunde aßen wir, immer zur selben Uhrzeit* gelöst werden. Abermals ist ein leichter Fehler in der Kategorie *Wrong Term* einzutragen.

Im selben Satz gab es einen weiteren Fehler bei der Übersetzung: *[E]s gab nichts mehr zu tun außer ein paar schmutzigem Geschirr* sollte wohl eher in etwa so aussehen: [...] *und abgesehen von etwas schmutzigem Geschirr, das anfiel, gab es nichts mehr zu tun*. Diese Abweichung stellt einen leichten Fehler in der Kategorie *Falsche Bedeutung (Semantical Error/ Wrong Meaning)* dar, da es sich um einen Fehler bezüglich der Phrasenstruktur handelt.

Bei dem Satz *Manchmal kamen wir zu der Überzeugung, dass sie diejenige war, die uns nicht heiraten lassen wollte* ist die Rede vom *Haus*. Demnach ist das Personalpronomen grammatikalisch inkorrekt. Bei der Humanübersetzung wurde diese Schwierigkeit eleganter gelöst: *Manchmal dachten wir, es wäre das Haus selbst, welches uns nicht erlauben würde, zu heiraten*. Dieser leichte Fehler wurde in der Kategorie *Strukturfehler* gewertet.

Der nächste Fehler muss als schwerer Fehler gewichtet werden, da sich die Botschaft komplett verändert. *DeepL* hat die Übersetzung folgendermaßen gelöst: [...] *und ich wurde von Maria Esther vor unserer Verlobung getötet*. Die Botschaft sollte allerdings vermitteln, dass Maria Esther, bevor sie mit dem Sprecher verlobt war, gestorben war: [...] *und meine Maria Esther war gestorben, noch bevor wir verlobt waren*. Dieser schwere Fehler ist in der Kategorie *Falsche Bedeutung* einzutragen.

Zwei Sätze weiter kam es gleich zu drei Fehlern: *Wir würden dort eines Tages sterben, faule und schwer fassbare Cousins würden das Haus nehmen und es zu Boden werfen, um sich mit dem Land und den Ziegelsteinen zu bereichern; oder besser, wir würden es uns selbst gerecht übergeben, bevor es zu spät ist*. Um es kurz zu fassen, kann die Humanübersetzung herangezogen werden, anhand derer elegantere Lösungen zu erkennen sind: *Faule, scheue Cousins würden das Haus erben und es niederreißen, um sich an dem Grundstück und den Gemäuern zu bereichern. Besser wäre, wir würden es selbst im Sinne der Gerechtigkeit zur Gänze auseinandernehmen, noch bevor es zu spät war*. Der erste Fehler findet als leichter Fehler in der Kategorie *Wrong Term* Eingang, die anderen beiden leichten Fehler in der Kategorie *Falsche Bedeutung*.

Im letzten Absatz dieses Auszugs gab es einige Schwierigkeiten bezüglich dessen, welche Dinge die Figur Irene strickt. Die Übersetzung *Irene war nicht so, sie strickte Dinge, die immer notwendig waren, sie strickte für den Winter, Socken für mich, morgens und Westen für sie* sollte so nicht stehen gelassen werden, da *morgens* kein Ding ist, das man stricken kann. Hier hätte man besser auf die deutsche Übersetzung Bettjacke zurückgreifen sollen. Besser wäre etwa: *Sie strickte immer brauchbare Dinge: Jacken für den Winter, Strümpfe für mich, Bettjacken und Westen für sich selbst*. Die Abweichung kann insgesamt als ein leichter Fehler betrachtet und der Kategorie *Falsche Bedeutung* hinzugefügt werden.

*DeepL* hatte auch mit dem folgenden Satz Probleme: *Manchmal strickte sie eine Weste und webte sie dann gleich wieder auf, weil ihr etwas nicht gefiel; es war lustig zu sehen, wie sich der Flor krauser Wolle im Korb einige Stunden lang weigerte, seine Form zu verlieren*. Man *webt* das Gestrickte nicht auf, wenn es einem nicht gefällt, sondern *trennt es auf*. Dieser Fehler wurde leichter Fehler in der Kategorie *Wrong Term* gewertet. Zusätzlich kann der *Flor krauser Wolle* durch den simpleren und treffenderen Terminus *Wollknäuel* ersetzt werden (Wertung ebenfalls als leichter Fehler in der Kategorie *Wrong Term*).

Zusammenfassend lässt sich deutlich erkennen, dass *DeepL* auf große Probleme bezüglich dieses geisteswissenschaftlichen Fachtextes gestoßen ist. Dies liegt vor allem in der Tatsache begründet, dass es nicht ausreichend ist, einzelne Ausdrücke wortwörtlich zu übersetzen, sondern den Gedanken dahinter zu erfassen, insbesondere, weil AutorInnen gerne mit Wörtern spielen und die Bedeutung in der Übersetzung dennoch dieselben bleiben sollten. Dieser Text könnte in dieser Form nicht übernommen werden und müsste dringend prä- und posteditiert werden.

Fehlerkategorie	Anzahl * schwere Fehler	Anzahl * leichte Fehler	GESAMT
WT	... * 5	6 * 2	8
SE	1 * 4	4 * 2	12
OM	...*4	1 * 2	2
SA	...*4	1 * 2	2
SP	...*3	...*1	
PE	...*2	...*1	
ME	...*3	...*1	

Abbildung 10: Eigene Darstellung. Auswertungsbogen Text 4.

Wenn wir nun ein letztes Mal die Fehleranzahl (mindestens 24 Fehler) durch den Textumfang (423 Wörter) dividieren, liegt ein Ergebnis von 0,0567 Punkten vor. Das Ergebnis liegt somit über dem maximalen Toleranzwert. Auch nach persönlicher Einschätzung ist die maschinelle Übersetzung dieses Textes nicht zufriedenstellend, insbesondere in Anbetracht dessen, dass im Rahmen dieser Auswertung viele „unschöne“ Ausdrücke nicht als Fehler gewertet wurden, insodern diese verständlich waren. Dieser Fachtext schneidet somit mit Abstand am schlechtesten im Ranking ab.

## 10. Auswertung: Gegenüberstellung der ausgewählten Termini

Nach erfolgter Auswertung bezüglich der Qualität der Übersetzungen insgesamt wird nun ein Blick auf die Auswahl der Terminologie geworfen. Hierzu werden beispielhaft 23 Termini aus der eigenen Terminologierecherche mit den Übersetzungen der Human- und der maschinellen Übersetzung verglichen. In der untenstehenden Tabelle (Abbildung 11) wird zunächst der Terminus in der Ausgangssprache präsentiert. Die Ausgangssprache der ersten beiden Fachtexte ist Deutsch, jene der letzten beiden Spanisch. Danach folgen die jeweiligen Übersetzungen der Humanübersetzung sowie der maschinellen Übersetzung durch das Übersetzungssystem *DeepL*. Die rechte Spalte der Tabelle (*Terminologierecherche*) beinhaltet jene Begriffe, die mittels eigener Recherche ausfindig gemacht wurden und nach persönlichen Einschätzungen bei der Übersetzung der vorliegenden Fachtexte verwendet werden sollten.

	Original	Human	DeepL	Terminologierecherche
<b>Text 1</b>				
1.	Entbehrung	carencia	privación	carencia
2.	Rente	jubilación	pensión	jubilación, pensión
3.	Gemeinwesen	el ser social	comunidad	el ser social
4.	Aufwachsen	crecer	crecer	crecer, desarrollarse
5.	Klassenausflug	viaje escolar	viaje de clase	viaje escolar, excursión escolar
<b>Text 2</b>				
6.	schlüsselfertig	llave en mano	llave en mano	de llave en mano
7.	an der Börse notieren	cotizar en la bolsa	cotizar en la bolsa	cotizar en la bolsa
8.	Leittechnik	tecnología de control	sistema de control	tecnología de control

9.	Finanzierungsm odell	modelo de financiación	modelo de financiación	modelo financiero, modelo de financiación
10.	Gesellschaft	empresa	empresa	empresa, sociedad
11.	Leistungsumfan g	la gama	alcance de la oferta	la gama
12.	Hilfsausrüstung	equipo auxiliar	equipo auxiliar	equipo auxiliar, equipamiento auxiliar
<b>Text 3</b>				
13.	infractor	Rechtsbrecher	Deponiebetreiber	Rechtsbrecher, Gesetzesbrecher
14.	complice	Handlanger	Komplize	Handlanger, Komplize
15.	incinerador de residuos	Müllverbrennun gsanlage	Müllverbrennungs anlage	Müllverbrennungsanlage, Abfallverbrennungsanlag e
16.	corredor ecológico	Biotopvernetzu ng	grüner Korridor	Biotopvernetzung, Biotopverbindung
17.	residuo tóxico	Giftmüll	giftiger Abfall	Giftmüll, Sonderabfall
18.	contaminación ambiental	Umweltverschm utzung	Umweltverschmut zung	Umweltverschmutzung, Umweltbelastung
<b>Text 4</b>				
19.	liquidación	Verwertung	Liquidierung	Verwertung, Liquidation
20.	matrimonio entre hermanos	Geschwisterehe	Brüderehe	Geschwisterehe, Geschwisterheirat
21.	genealogía	Genealogie	Genealogie	Genealogie, Ahnenforschung
22.	mañana	Bettjackede	morgens	Bettjackede
23.	pretendiente	Verehrer	Bewerber	Verehrer

**Abbildung 11:** Eigene Darstellung. *Gegenüberstellung der Termini.*

## 10.1 Fachtext 1

Der erste Fachtext ist dem Bereich Wirtschaft zuzuordnen. Die Abweichung des ersten Terminus wurde bereits im vorherigen Kapitel näher erläutert. Auch wenn in einigen Wörterbüchern *Entbehrung* als Entsprechung für *privación* angezeigt wird, so sollte in diesem Kontext eher auf *carencia* zurückgegriffen werden. Obwohl die Botschaft der Aussage vermutlich trotzdem richtig übertragen wird, sollte für eine qualitativ hochwertige Übersetzung auf den präziseren Begriff *carencia* zurückgegriffen werden.

Der Terminus *Rente* wurde zum einen mit *jubilación* und zum anderen mit *pensión* übersetzt. Beide Begriffe sind auch Teil der eigenen Terminologierecherche; eine Überarbeitung wäre somit nicht notwendig.

*Gemeinwesen* wurde in der Humanübersetzung mit *el ser social* übersetzt, in der maschinellen Übersetzung allerdings mit *comunidad*. Auch wenn die Begriffe *Gemeinschaft* und *Gemeinwesen* im Deutschen sehr ähnlich sind und ähnliche Konzepte beschreiben, so sind sie dennoch nicht deckungsgleich. Ist eine sehr hohe Qualität der Übersetzung erwünscht, so ist der Begriff *ser social* zu bevorzugen. Die Botschaft würde dennoch auch mit der Benennung *comunidad* verstanden werden.

Der vierte Terminus *Aufwachsen* wurde sowohl in der Human- als auch in der maschinellen Übersetzung mit *crecer* übersetzt und stimmt somit mit der eigenen Terminologierecherche überein.

Der letzte Terminus dieses Fachtextes sollte bei einem Lektorat markiert werden. Die Übersetzung von *Klassenausflug* im Deutschen scheint unpassend mit *viaje de clase*. Auch wenn das Konzept vermutlich verstanden wird, so klingt es dennoch nicht flüssig. Auch wenn diese Abweichung unter Umständen nicht als Fehler markiert wird, sollte trotzdem auf eine mögliche Überarbeitung hingewiesen werden. Dies könnte beispielsweise durch die Benennungen *viaje escolar* oder *excursión escolar* gelöst werden.

Insgesamt ist die Qualität dieses Fachtextes in Bezug auf die Terminologie mit einer guten „Note“ zu versehen. Auch wenn einige Termini noch überarbeitet werden könnten, so stellt dies eine Fehlersuche auf bereits hohem Niveau dar. Die Botschaft des Fachtextes wird sehr gut übertragen und es wurde kein Terminus gewählt, der gänzlich falsch oder unpassend war. Anzumerken ist, dass es sich beim Übersetzen stets um eine persönliche Entscheidung

bezüglich der Auswahl der Termini handelt, weshalb es schwierig ist, jede einzelne Abweichung zu bewerten.

## 10.2 Fachtext 2

Die ersten beiden Termini des Fachtextes (aus dem Bereich Technik) wurden sowohl in der Human- als auch in der maschinellen Übersetzung mit demselben Terminus übersetzt. Der Terminus *schlüssselfertig*“ wurde beide Male mit *llave en mano* und der Terminus *an der Börse notieren* mit *cotizar en la bolsa* übersetzt. Demnach gibt es bezüglich dieser beiden Benennungen nichts hinzuzufügen.

Der Terminus *Leittechnik* wurde allerdings in der Humanübersetzung mit *tecnología de control* und in der maschinellen Übersetzung mit *sistema de control* übersetzt. Da der Begriff die Technik und nicht das System an sich meint, scheint *tecnología de control* aus der Humanübersetzung passender.

Auch die beiden Termini *Finanzierungsmodell* und *Gesellschaft* wurden mit denselben spanischen Benennungen übersetzt: Zum einen mit *modelo de financiación* und zum anderen mit *empresa*. Diese beiden Übersetzungen entsprechen ebenso der eigenen Terminologierecherche im Rahmen dieser Arbeit, weshalb an dieser Stelle nichts auszubessern beziehungsweise zu überarbeiten bleibt.

Bei Terminus 11 war allerdings abermals eine Abweichung zu bemerken. Die Ausgangsbenennung *Leistungsumfang* wurde in der Humanübersetzung mit *la gama* und in der maschinellen Übersetzung mit *alcance de la oferta* übersetzt. Im Zuge der Terminologierecherche fiel die Entscheidung ebenfalls auf die spanische Benennung *la gama*. Auch wenn *alcance de la oferta* den Umfang des Angebots beschreibt, so scheint für eine qualitativ hochwertige Übersetzung die Wahl der Humanübersetzung passender, da mit der spanischen Übersetzung *la gama* die Bedeutung beziehungsweise der Umfang kurz und bündig auf den Punkt gebracht wird.

Auch der letzte angeführte Terminus wurde in beiden Übersetzungen mit der gleichen spanischen Benennung übersetzt. In beiden Versionen steht *Hilfsausrüstung* für *equipo auxiliar*, was gleichermaßen dem gewählten Begriff aus der Terminologierecherche entspricht.

Insgesamt kann diese maschinelle Übersetzung als sehr gelungen beurteilt werden. Fast alle Termini wurden korrekt übersetzt. Die beiden Unregularitäten stellen keine

klassischen Fehler, sondern Abweichungen dar, die nicht unbedingt überarbeitet werden müssen, aber – je nach gewünschter Qualität des Ergebnisses – können.

### 10.3 Fachtext 3

Gleich der erste Terminus des dritten Fachtextes aus dem Fachbereich Naturwissenschaften wurde nicht korrekt übersetzt. Die Benennung *infractor* im Spanischen wurde mit *Rechtsbrecher* ins Deutsche übersetzt (zu diesem Ergebnis kommt auch die eigene Terminologierecherche). *DeepL* wählte dafür jedoch *Deponiebetreiber* aus. Unter Betrachtung des Kontextes, lässt sich darauf schließen, dass der Begriff ausgelassen wurde oder die *Deponiebetreiber*, sprich: die Betreiber der Müllverbrennungsanlagen, als Rechtsbrecher beschrieben werden. Diese Abweichung sollte dringend überarbeitet werden.

Die nächsten beiden Termini *complice* und *incinerador de residuos* wurden in beiden Fällen nicht korrekt übersetzt. Bei dem Terminus *complice* haben zwar die HumanübersetzerInnen und *DeepL* nicht die gleiche Übersetzung (*Handlanger* und *Komplize*) gewählt, dennoch sind diese beiden Termini laut zugrundeliegender Terminologierecherche korrekt. Dahingegen haben beide Versionen den Terminus *Müllverbrennungsanlage* als Übersetzung für den spanischen Begriff *incinerador de residuos* verwendet. Das bedeutet, dass in beiden Fällen keine Überarbeitung notwendig ist.

Das Konzept *corredor ecológico* wurde bei der Humanübersetzung richtigerweise mit *Biotopvernetzung* übersetzt. Bei der maschinellen Übersetzung wurde für den wohl nicht bekannten Begriff *grüner Korridor* verwendet. Neben *Biotopvernetzung* wäre, laut des erarbeiteten Glossars auch *Biotopverbindung* eine Option gewesen. In diesem Fall ist die Abweichung klar zu markieren, so dass die Übersetzung angepasst werden kann.

Die letzten beiden Termini dieses naturwissenschaftlichen Fachtextes wurden wiederum korrekt übersetzt. Der Begriff *residuo tóxico* wurde in der Humanübersetzung mit *Giftmüll* und in der *DeepL*-Übersetzung mit *giftiger Abfall* übersetzt. Auch wenn in der Terminologierecherche *giftiger Abfall* nicht vorkommt, so kann dieser dennoch als korrekt bezeichnet werden. Innerhalb einer Terminologierecherche ist es bei recht umgangssprachlichen Ausdrücken manchmal schwierig alle Synonyme aufzuzählen. Daher kann es vorkommen, dass Termini richtig sind, als Synonyme zählen, aber nicht in der erarbeiteten Terminologiearbeit aufgelistet werden.

Für den spanischen Terminus *contaminación ambiental* wurde in beiden Versionen die deutsche Benennung *Umweltverschmutzung* verwendet, welche auch im angeführten Glossar wiederzufinden ist. Dementsprechend gibt es auch in diesem Fall keine notwendige Überarbeitung.

Insgesamt fällt die maschinelle Übersetzung dieses Fachtextes nach persönlicher Einschätzung und Fehlerquote sehr gut aus. Zwei der fünf ausgewählten Termini sollten gegebenenfalls überarbeitet werden. Dies hängt, wie bei den anderen Texten, von der gewünschten Qualität der Übersetzung ab. Die Übersetzung liest sich sehr flüssig und abgesehen von den zwei terminologischen Abweichungen wird die Botschaft des Textes ohne großartige Änderungen klar übermittelt.

## 10.4 Fachtext

Bei der Übersetzung des vierten Fachtextes (Fachbereich Geisteswissenschaften) wurde der spanische Terminus *liquidación* zum einen bei der Humanübersetzung mit *Verwertung* und bei der maschinellen Übersetzung mit *Liquidierung* übersetzt. Im Glossar stößt man außerdem auf die zwei Benennungen *Verwertung* und *Liquidation*. In persönlicher Einschätzung wäre der Terminus der *DeepL*-Übersetzung nicht zu überarbeiten, da *Liquidierung* und *Liquidation* bedeutungsrechnerisch nicht weit voneinander entfernt liegen und dasselbe Konzept beschreiben.

In der maschinellen Übersetzung wurde *Brüderehe* für den spanischen Ausgangsterminus *matrimonio entre hermanos* verwendet, während die Terminologierecherche *Geschwisterehe* oder *Geschwisterheirat* vorschlägt. Da die *Geschwister* auf Deutsch geschlechtsneutral formuliert sind und somit *Bruder* und *Schwester* miteinschließen, erscheinen *Geschwisterheirat* oder *-ehe* dementsprechend passender. Diese Abweichung sollte in der *DeepL*-Übersetzung überarbeitet werden.

Das Konzept der *genealogía* wurde sowohl in der Human- als auch in der maschinellen Übersetzung mit *Genealogie* übersetzt. Auch in dem der Arbeit zugrundeliegenden Glossar scheinen die Benennungen *Genealogie* und *Ahnenforschung* auf. Daher ist bei diesem Terminus keine Überarbeitung notwendig.

Der Begriff *mañanita* wurde in der Humanübersetzung (vermutlich dank intensiver Terminologierecherche vorab) korrekt mit *Bettjackedecke* übersetzt. Bei der maschinellen Übersetzung wurde daraus stattdessen *morgens*. Diese Abweichung muss definitiv angepasst

werden, ansonsten ergibt der Satz im Deutschen keinen Sinn; eine treffende Übersetzung kann hier nur durch intensive Terminologierecherche erfolgen.

Nicht zuletzt wurde der Terminus *pretendiente* zum einen in der Humanübersetzung mit *Verehrer* und zum anderen in der maschinellen Übersetzung mit *Bewerber* übersetzt. Auch wenn ein *Verehrer* wohl in gewisser Form als *Bewerber* beschrieben werden kann, so ist die Bezeichnung in diesem literarischen Kontext nicht wahrhaft passend. In diesem Umfeld hatte *DeepL* vermutlich Schwierigkeiten, die Bedeutung des Terminus anhand des Kontextes richtig einzuordnen. Eine Überarbeitung dieser Abweichung sollte dementsprechend vorgenommen werden.

Abgesehen von den genannten zwei Abweichungen, die überarbeitet werden sollten, kann die Qualität in Bezug auf die ausgewählten fünf Termini als gut bewertet werden. Bezieht man allerdings nicht nur die ausgewählten Fachtermini mit ein, sondern die übersetzerische Qualität des gesamten Textes, so muss das Ergebnis mit Verweis auf Kapitel 9 relativiert werden.

## 11. Schlussfolgerung

Maschinelle Übersetzungssysteme sind zurzeit gefragter denn je und stellen mittlerweile ein sehr bedeutendes Kommunikationsmedium dar. Dies liegt vor allem an der verbesserten Qualität dieser Übersetzungssysteme während der letzten Jahre. Von ÜbersetzerInnen wird dieses Tool unterschiedlich beurteilt und die Meinungen spalten sich bezüglich dieser kontroversen Hilfestellung. Zum einen werden maschinelle Übersetzungssysteme herangezogen, um Inhalte zu verstehen, aber zum anderen auch, um vorab Rohübersetzungen zu generieren, die anschließend lediglich posteditiert werden müssen.

Das Ziel dieser Arbeit war es, die Möglichkeit und Grenzen der maschinellen Übersetzung vor allem im Bereich der neuronalen Übersetzungssysteme anhand der Software *DeepL* aufzuzeigen. Dabei sollte die Qualität der maschinellen Übersetzung vor allem in den Fachbereichen Wirtschaft, Recht, Naturwissenschaften, Geisteswissenschaften und Technik untersucht werden, indem sowohl die Gesamtqualität der jeweiligen Übersetzungen als auch die Terminologie, die in den Ausgangs- und Zieltexten vorkommt, untersucht werden sollte. Dazu wurden vier Fachtexte von HumanübersetzerInnen sowie durch das maschinelle Übersetzungssystem *DeepL* übersetzt und anschließend zur Evaluierung der Qualität einander gegenübergestellt.

Auch wenn zur Evaluierung der Qualität vorab eine Qualitätsmetrik ausgewählt worden ist, ist die Beurteilung der Qualität einer Übersetzung keineswegs ein einfaches Unterfangen. Die Auflistung der Kriterien beziehungsweise Fehlerkategorien kann nicht immer auf die gleiche Art und Weise eingehalten werden und die Entscheidung, in welche Fehlerkategorie die Abweichungen einzuordnen sind, obliegt oftmals der subjektiven Entscheidung des/der WissenschaftlerIn. Anzumerken ist, dass es sich beim Übersetzen stets um eine persönliche Entscheidung bezüglich der Auswahl der Termini handelt, weshalb es schwierig ist, jede einzelne Abweichung zu bewerten. Ebenso verhält es sich mit der Beurteilung dessen, ob es sich im konkreten Fall um leichten oder schwerwiegenden Fehler mit oder ohne Folgen handelt und ob der dieser als stilistische Abweichung oder tatsächlich als voller Fehler einzustufen ist.

Es stellt sich außerdem die Frage, welcher Qualitätsanspruch an die jeweilige Übersetzung gestellt wird. Soll die Übersetzung nur zum allgemeinen Verständnis dienen oder vielleicht sogar veröffentlicht werden? Es lässt sich sagen, dass die Translation in ihrer Funktion als Mittel zur Kommunikation dann als qualitativ hoch angesehen wird, wenn sie

erfolgreich ist, wenn sie funktioniert, sprich: wenn der Inhalt des Ausgangstextes unter Erfüllung seiner bestimmten Funktion oder Intention an das Zielpublikum weitergegeben und von diesem auch verstanden wird.

Die Ergebnisse der in der vorliegenden Arbeit erfolgten Evaluierung sind untenstehend (Abbildung 12) zusammengefasst:

Fachtext	Fachgebiet	Ergebnis Evaluation
Fachtext 1	Wirtschaft	0,0403
Fachtext 2	Technik	0,0262
Fachtext 3	Naturwissenschaft	0,0224
Fachtext 4	Geisteswissenschaften	0,0567

**Abbildung 12:** Eigene Darstellung. *Gegenüberstellung der Evaluation der Fachtexte*

Grundsätzlich kann fällt die Qualität der maschinellen Übersetzungen mit Ausnahme von Text 4 insgesamt durchaus positiv aus. Auch wenn der erste Fachtext aus dem Bereich Wirtschaft mit einem Score von 0,0403 nur wenig über der Toleranzgrenze liegt, so ist die Qualität dennoch akzeptabel. Eine interessante Beobachtung dabei ist, dass die Qualität der Übersetzungen nicht unbedingt mit dem Schwierigkeitsgrad der Texte zusammenhängt. Vor allem die Fachtexte in den Bereichen Wirtschaft, Technik und Naturwissenschaft würden in Bezug auf Grammatik, Terminologie oder Satzstruktur als eher schwierig eingeschätzt werden und gehen weit über den alltäglichen Sprachgebrauch hinaus. So nimmt nicht nur der Schwierigkeitsgrad Schwierigkeitsgrad in Bezug auf Grammatik und Satzstruktur des Ausgangstextes Einfluss auf die Qualität der maschinellen Übersetzung, sondern vor allem auch das jeweilige Fachgebiet und die damit verbunden Terminologie. Demnach haben Übersetzungssysteme wie *DeepL* kleinere oder größere Schwierigkeiten, je nachdem, wie gut oder schlecht sie in Hinblick auf die jeweiligen Fachgebiete trainiert wurden.

Aufgrund der vorliegenden Ergebnisse der Evaluation kann gesagt werden, dass die Qualität der maschinellen Übersetzungen vor allem in den Bereichen Technik und Naturwissenschaften mittlerweile überaus zufriedenstellend sind. Auch in Bezug auf die Terminologie schneiden die Übersetzungen in beiden Fachbereichen gut ab. Wenn dabei zusätzlich die Kosten- und Zeitersparnis in die Überlegungen miteinbezogen werden, wären

Rohübersetzungen mittels maschineller Übersetzung inklusive Prä- und Postedition in der Praxis durchaus eine Überlegung wert.

Bei der Übersetzung im geisteswissenschaftlichen Bereich kann gesagt werden, dass die Qualität der maschinell generierten Übersetzung stark verbesserungswürdig ist. Die Ergebnisse dieser Übersetzung (sowohl in Bezug auf die gesamte Übersetzung als auch auf die vereinzelt Termini) bieten keine Zeitersparnis und wären demnach wahrscheinlich nicht kostengünstiger als eine herkömmliche Humanübersetzung. Gerade im Bereich der geisteswissenschaftlichen Fachtexte fällt auf, dass der Kontext eine sehr große Rolle spielt. Ramlow (2009) erklärt, dass es sich bei einer maschinellen Übersetzung niemals um die Übersetzung eines kohärenten Textes handeln kann, da dieser vom Programm nicht als zusammenhängender Text wahrgenommen wird. Stattdessen wird eher eine Aneinanderreihung von Worten oder im besten Fall eine Aneinanderreihung von Sätzen übersetzt. Die inhaltliche und sprachliche Beziehung, die zwischen den einzelnen Sätzen besteht, und der Kontext, der dadurch entsteht, sind für Übersetzungssysteme bislang nicht ausreichend erkennbar (vgl. Ramlow 2009:110ff.).

Bei der Übersetzung von Fachtext 1 aus dem Bereich Wirtschaft konnten ebengfalls einige Abweichungen gefunden werden, weshalb der Score der Übersetzung nicht herausragend ist. Trotz des „schlechten“ Scores verblieb dennoch ein insgesamt recht positiver subjektiver Eindruck sowohl in Bezug auf den gesamten Text als auch auf die ausgewählten Termini. Geht man davon aus, dass HumanübersetzerInnen viel Zeit in die Terminologierecherche investieren und dennoch oftmals nicht zu 100 Prozent sicher sein können, die richtige Übersetzung des Terminus gewählt zu haben, so spricht es für die neuronale maschinelle Übersetzung, dass diese ganz ohne Recherche ein passables Ergebnis liefert. Dabei sei erwähnt, dass maschinelle Übersetzungen für gewöhnlich meist prä- und posteditiert werden und eine MÜ nie ohne diese übernommen wird (je nach Zweck der Übersetzung und Qualitätsanspruch).

Beide Hypothesen können damit ganz klar bestätigt werden. Humanübersetzungen weisen, den Ergebnissen nach zu urteilen eine höhere Treffsicherheit in Bezug auf die Übersetzung von Fachtermini auf als maschinelle Übersetzungstools. Auch wenn Maschinelle Übersetzungstools in einigen Fachgebieten sehr gut abschneiden, so ist die Humanübersetzung immer noch die sicherste Variante, wenn wir allein den Aspekt der Übersetzung von Fachtermini betrachten. Weiters kann die zweite Hypothese bestätigt werden, dass neuronale maschinelle Übersetzungstools in der Übersetzung von

wirtschaftlichen Fachtexten eine höhere Treffsicherheit in Bezug auf die Übersetzung von Fachtermini aufweisen als in der Übersetzung von geisteswissenschaftlichen Fachtexten.

Auch wenn nicht alle vier Fachtexte eine Zeitersparnis bieten und wahrscheinlich auch nicht kostengünstiger als eine gewöhnliche Humanübersetzung wären, ist doch anzumerken, dass der primäre Sinn von maschineller Übersetzung nicht in Frage gestellt werden sollte; vor allem angesichts dessen, dass menschliche Übersetzung mit nur begrenzter Unterstützung durch die Maschine im Bereich der Machine-Aided Human Translation (MAHT) doch sehr positive Ergebnisse erzielt. Aus der vorliegenden Arbeit geht demnach hervor, dass die Qualität der neuronalen maschinellen Übersetzung zum Teil bereits einen hohen Standard aufweist und diese TranslatorInnen in vielerlei Hinsicht unterstützen kann. Nichtsdestotrotz gibt es weiterhin großen Forschungsbedarf in der Entwicklung von vollautomatischen Übersetzungen, die in keinster Weise der Qualität von Humanübersetzungen nachhinken.

## 12. Literaturverzeichnis

ALPAC (1966). *Language and machines. Computers in translation and linguistics. A Report by the Automatic Language Processing Advisory Committee*. Washington: National Research Council.

Anguelovski, Isabelle (2016). ¿Las ciudades más ecológicas son más desiguales?. *El País*. [https://elpais.com/elpais/2016/09/19/seres\\_urbanos/1474264800\\_147426.html](https://elpais.com/elpais/2016/09/19/seres_urbanos/1474264800_147426.html) (Stand 09.11.2020).

Arnold, Douglas & Balkan, Lorna & Meijer, Siety & Humhreys, R. Lee & Sadler, Louisa (1994). *Machine Translation. An introductory guide*. Manchester: NCC Blackwell.

Bar-Hillel, Yehoshua (1960). The Present Status of Automatic Translation of Languages. In: Alt, Franz L. (Hg.) *Advances in Computers*. Jerusalem: Hebrew University (1), 91-163.

Brown, Peter & Cocke, John & Della Pietra, S. & Della Pietra, Vincent & Jelinek, Frederick & Mercer, Robert & Roossin, P (1988). A statistical approach to french/english translation. In: Vargha, Denes (Hg.) *Proceedings of the 12th conference on Computational linguistics*, 71-76.

Castilho, Sheila & Moorkens, Joss & Gaspari Federico & Calixto, Iacer & Tinsley, John & Way, Andy (2017). Is Neural Machine Translation the New State of the Art? In: *The Prague Bulletin of Mathematical Linguistics*. Berlin: De Gruyter Open (8), 109-120.

Chen, Boxing & Huang, Fei (2016). Semi-Supervised Convolutional Networks for Translation Adaptation with Tiny Amount of In-domain Data. In: Goldberg, Yoav & Riezler, Stefan (Hg.) *Proceedings of the 20th SIGNLL Conference on Computational Natural Language Learning*. Berlin: Association for Computational Linguistics, 314-323.

Chunyu, Kit & Xiaoyue, Liu (2008). Measuring mono-word termhood by rank difference via corpus comparison. In: *Terminology*. Amsterdam: John Benjamins Publishing (14), 204-229.

Cortázar, Julio (2016). Casa tomada. In: *Bestiario*. Barcelona: Debolsillo.

Dalla-Zuanna, Jean-Marc (2010). Direkte Qualitätsmessung. In: *Fachzeitschrift MDÜ*. <https://docplayer.org/1182756-Direkte-qualitaetsmessung.html> (Stand 26.10.2020).

DeepL. Firmenprofil. In: *DeepL*. [https://www.deepl.com/files/press/companyProfile\\_EN.pdf](https://www.deepl.com/files/press/companyProfile_EN.pdf)

(Stand: 18.10.2020).

DeepL (2020 a). Presseinformation– DeepL Übersetzer. In: *DeepL*.  
<https://www.deepl.com/press.html> (Stand: 18.10.2020).

DeepL (2020 b). Pressemitteilung – DeepL Übersetzer. In: *DeepL*.  
<https://www.deepl.com/de/blog/20200206.html> (Stand: 09.12.2020).

Dougal, Duane (2018). *Improving the Quality of Neural Machine Translation Using Terminology Injection*. Brigham Young University: BYU Scholars Archive.

Dziubek, Mateusz (2018). *Let me introduce you to neural networks*. In: *Medium*.  
<https://towardsdatascience.com/let-me-introduce-you-to-neural-networks-fedf4253106a> (Stand: 19.10.2020).

Felber, Helmut (1984). *Terminology Manual*. Vienna: International Information Centre for Terminology Vienna.

Forcada, Mikel (2017). Making sense of neural machine translation. In: *Translation Spaces* (6), 291-303.

Haque, Rejwanul & Hasanuzzaman, Mohammed & Way, Andy (2019). Terminology Translation in Low-Resource Scenarios. In: *Journal of Information* (10), 273.

Hutchins, John & Somers, Harold (Hg.) (1992). *An Introduction to Machine Translation*. London: Academic Press.

Hutchins, John (1986). Machine Translation: past, present, future. In: Horwood, Ellis (Hg.) *Series in Computers and their Application*. Chichester: Halsted Press, 382.

Hutchins, John (1995). Machine Translation: A Brief History. In: Körner, Ernst & Asher, R. (Hg.) *Concise history of the language sciences: from the Sumarians to the cognitivists*. Oxford: Pergamon.

Hutchins, John (1997). First steps in mechanical translation. In: Teller, Virginia & Sundheim, Beth (Hg.) *Proceedings of MT Summit VI: past, present, future*. Washington, D.C.: Association for Machine Translation in the Americas, 14- 23.

Hutchins, John (2006). Machine translation: a concise history. In: Sin Wai, Chad (Hg.) *Journal of Translation Studies*. (13), 29-70.

Koehn, Philipp (2010). *Statistical machine translation*. Cambridge: Cambridge University Press.

- Koehn, Philipp (2020). *Neural machine translation*. Cambridge: Cambridge University Press.
- Kostadinov, Simeon (2019). Understanding Encoder-Decoder Sequence to Sequence Model. In: *Medium*. <https://towardsdatascience.com/understanding-encoder-decoder-sequence-to-sequence-model-679e04af4346> (Stand: 19.10.2020).
- Krenz, Michael & Ramlow, Markus & Seewald-Heeg, Uta (2008) *Maschinelle Übersetzung und XML im Übersetzungsprozess. Prozesse der Translation und Lokalisierung im Wandel. Zwei Beiträge*. Berlin: Frank & Timme.
- Kröse, Ben & Van der Smagt, Patrick (1996). *An Introduction to Neural Networks*. Amsterdam: University of Amsterdam.
- Krüger, Ralph (2017). Von Netzen und Vektoren – Neuronale Maschinelle Übersetzung. In: *MDÜ – Fachzeitschrift für Dolmetscher und Übersetzer*. Berlin: BDÜ (63), 38-44.
- KÜDES (2002). *Empfehlungen für die Terminologearbeit*. Bern: KÜDES. <http://www.bk.admin.ch/dokumentation/sprachen/05078/> (Stand: 30.11.2020).
- Landsbergen, Jan (1992). Können Computer übersetzen? In: *Lebende Sprachen*. Berlin: De Gruyter (37), 7–11.
- Lehrberger, John & Bourbeau, Laurent (1988). *Machine translation. Linguistic characteristics of MT systems and general methodology of evaluation*. Amsterdam: John Benjamins Publishing Company.
- Lennon, Brian (2014). Machine Translation: A Tale of Two Cultures. In: Bermann, Sandra & Porter, Catherine (Hg.) *A Companion to Translation Studies*. Hoboken, New Jersey, Vereinigte Staaten: John Wiley and Sons, 135–46.
- Lizel, Nico & Luber, Stefan (2017). *Was ist Deep Learning?*. In: *Big Data Insider*. <https://www.bigdata-insider.de/was-ist-deep-learning-a-603129/> (Stand: 19.10.2010).
- Maier, Elisabeth (2018). Wohin steuert der Markt? In: *Technische Kommunikation*, (40), 30–34.
- Nakagawa, Hiroshi (2001) Experimental evaluation of ranking and selection methods in term extraction. In: Bourigault, Didier & Jacquemin, Christian & L'Homme, Marie-Claude (Hg.). *Recent Advances in Computational Terminology*. Tokyo: University of Tokyo, 303-325.
- Oh, Jong-Hoon & Lee, KyungSoon & Choi, Key-Sun (2000). Term Recognition Using Technical Dictionary Hierarchy. In: *Proceedings of the 38th Annual Meeting of the*

*Association for Computational Linguistics*. Hong Kong: Association for Computational Linguistics, 496-503.

Reimann, Martina (1991). *Maschinelle Übersetzung: Ein Überblick über Theorie und Praxis*. Berlin: Springer.

Ritschl, Nina (2015). Maschinelle Übersetzung; gute Alternative zur Humanübersetzung? In: *Plattform für Prozess-, Qualitäts-, Risikomanagement und Technische Dokumentation*. <http://www.pqrm.at/2015/07/13/maschinelle-uebersetzung-gute-alternative-zur-humanuebersetzung/> (Stand: 09.12.2020).

Ramlow, Markus (2009). *Die maschinelle Simulierbarkeit des Humanübersetzens. Evaluation von Mensch-Maschine-Interaktion und der Translatqualität der Technik*. Berlin: Frank & Timme.

SAE Translation Quality Metric (2001). In: *SAE International*. <https://www.sae.org/standardsdev/j2450p1.htm> (Stand 26.10.2020).

Saß, Ricarda (2004). *Vergleichende Untersuchung von Terminologie-Extraktions-Tools. Eine computerlinguistische Arbeit mit Englisch und Deutsch*. Saarbrücken: Universität des Saarlandes (21).

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: Peter Lang GmbH.

Schwanke, Martina (1991). *Maschinelle Übersetzung. Ein Überblick über Theorie und Praxis*. Berlin: Springer.

Stein, Daniel (2009). Maschinelle Übersetzung. Ein Überblick. In: *JLCL - Journal for Language Technology and Computational Linguistics* (24), 5-18.

Thelen, Marcel (2015). The Interaction between Terminology and Translation. In: Van Vaerenbergh, Schubert (Hg.) *Trans-kom* (8), 374-381.

Vashee, Kirti (2017). Neural Machine Translation; A Practitioner's Viewpoint. In: Porsiel, Jörg (Hg.) *Maschinelle Übersetzung. Grundlagen für den professionellen Einsatz*. Berlin: BDÜ Fachverlag, 44-59.

Welzer, Harald (2019). *Alles könnte anders sein: Eine Gesellschaftsutopie für freie Menschen*. Berlin: S. Fischer.

Wittpahl, Volker (2019). *Künstliche Intelligenz. Technologien, Anwendung, Gesellschaft*. Berlin: Springer.

Zhang, Jiajun & Zong, Chengqing (2020). Neural Machine Translation: Challenges, Progress and Future. In: *Science China Technological Sciences* (10).

## 12.1 Quellenangaben Glossar:

### **DUDEN:**

[https://www.duden.de/rechtschreibung/Entbehruhttps://www.duden.de/rechtschreibung/Entbe  
hrung](https://www.duden.de/rechtschreibung/Entbehruhttps://www.duden.de/rechtschreibung/Entbe<br/>hrung) (Stand: 08.11.2020).

<https://www.duden.de/rechtschreibung/Rente> (Stand: 10.11.2020).

<https://www.duden.de/rechtschreibung/Klassenausflug> (Stand: 10.11.2020).

<https://www.duden.de/rechtschreibung/Finanzierungsmodell> (Stand: 09.11.2020).

<https://www.duden.de/rechtschreibung/Rechtsbrecher> (Stand: 09.11.2020).

<https://www.duden.de/rechtschreibung/Giftmuell> (Stand: 09.11.2020).

<https://www.duden.de/rechtschreibung/Umweltverschmutzung> (Stand: 10.11.2020).

<https://www.duden.de/rechtschreibung/Verehrer> (Stand: 10.11.2020).

### **EUR-LEX:**

<https://eurlex.europa.eu/LexUriServ/LexUriServ.do?uri=COM:2001:0770:FIN:DE:PDF>  
(Stand: 08.11.2020).

[https://eurlex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:C:2005:058E:0001:0129:DE:PD  
F](https://eurlex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:C:2005:058E:0001:0129:DE:PD<br/>F) (Stand: 08.11.2020).

<https://eurlex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:C:2006:325:0060:0064:DE:PDF>  
(Stand: 08.11.2020).

<https://eurlex.europa.eu/LexUriServ/LexUriServ.do?uri=COM:2001:0264:FIN:DE:PDF>  
(Stand: 09.11.2020).

<https://eurlex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2008:045:0001:0029:ES:PDF>  
(Stand: 10.11.2020).

<https://eurlex.europa.eu/LexUriServ/LexUriServ.do?uri=COM:2000:0846:FIN:DE:PDF>  
(Stand: 10.11.2020).

<https://eurlex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2008:343:0001:0038:ES:PDF>  
(Stand: 10.11.2020).

<https://eurlex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:C:2009:133:0010:0011:ES:PDF>  
(Stand: 10.11.2020).

<https://eurlex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2009:023:0043:0059:ES:PDF>  
(Stand: 10.11.2020).

<https://eurlex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2003:327:0039:0045:DE:PDF>  
(Stand: 09.11.2020).

<https://eurlex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2008:320:0001:0481:ES:PDF>  
(Stand: 10.11.2020).

<https://eurlex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:C:2001:240E:0298:0302:DE:PDF>  
(Stand: 09.11.2020).

<https://eurlex.europa.eu/LexUriServ/LexUriServ.do?uri=COM:2000:0441:FIN:ES:PDF>  
(Stand: 10.11.2020).

### **EUR-PARL:**

<https://www.europarl.europa.eu/sides/getDoc.do?pubRef=-//EP//TEXT+CRE+20010516+ITEMS+DOC+XML+V0//ES&language=ES> (Stand: 10.11.2020).

<https://www.europarl.europa.eu/sides/getDoc.do?pubRef=-//EP//TEXT+CRE+20080115+ITEMS+DOC+XML+V0//ES&language=ES> (Stand: 10.11.2020).

<https://www.europarl.europa.eu/sides/getDoc.do?pubRef=-//EP//TEXT+CRE+20080709+ITEMS+DOC+XML+V0//DE&language=DE> (Stand: 09.11.2020).

<https://www.europarl.europa.eu/sides/getDoc.do?pubRef=-//EP//TEXT+CRE+20060515+ITEMS+DOC+XML+V0//ES&language=ES> (Stand: 10.11.2020).

<https://www.europarl.europa.eu/sides/getDoc.do?pubRef=-//EP//TEXT+CRE+20041216+ITEMS+DOC+XML+V0//DE&language=DE> (Stand: 09.11.2020).

### **DRAE- DICCIONARIO DE LA LENGUA ESPAÑOLA DE LA REAL ACADEMIA ESPAÑOLA:**

<https://dle.rae.es/carencia> (Stand: 10.11.2020).

<https://dpej.rae.es/lema/llave-en-mano> (Stand: 10.11.2020).

<https://dle.rae.es/empresa?m=form> (Stand: 10.11.2020).

<https://dle.rae.es/infractor?m=form> (Stand: 10.11.2020).

<https://dle.rae.es/cómplice?m=form> (Stand: 10.11.2020).

<https://dpej.rae.es/lema/corredor-ecológico> (Stand: 10.11.2020).

<https://dle.rae.es/liquidación?m=form> (Stand: 10.11.2020).

<https://dle.rae.es/genealog%C3%ADa?m=form> (Stand: 10.11.2020).

<https://dle.rae.es/mañanita?m=form> (Stand: 10.11.2020).

### **EUROPÄISCHE KOMMISSION:**

[https://ec.europa.eu/commission/presscorner/detail/de/IP\\_94\\_3](https://ec.europa.eu/commission/presscorner/detail/de/IP_94_3) (Stand: 08.11.2020).

[https://ec.europa.eu/commission/presscorner/detail/de/IP\\_02\\_1805](https://ec.europa.eu/commission/presscorner/detail/de/IP_02_1805) (Stand: 09.11.2020).

[https://ec.europa.eu/commission/presscorner/detail/de/IP\\_94\\_426](https://ec.europa.eu/commission/presscorner/detail/de/IP_94_426) (Stand: 09.11.2020).

### **ECONOMIPEDIA:**

<https://economipedia.com/definiciones/jubilacion.html> (Stand: 10.11.2020).

<https://economipedia.com/definiciones/cotizar-en-bolsa.html> (Stand: 10.11.2020).

<https://economipedia.com/definiciones/linea-de-productos.html> (Stand: 10.11.2020).

### **DWDS: DIGITALES WÖRTERBUCH DER DEUTSCHEN SPRACHE**

<https://www.dwds.de/wb/aufwachsen> (Stand: 08.11.2020).

<https://www.dwds.de/wb/börsennotiert> (Stand: 09.11.2020).

### **FERROSTAAL:**

[https://www.ferrostaal.com/uploads/tx\\_mfsmatrix/MF\\_ECHO\\_D\\_Dez09\\_WEB\\_2.pdf?PHPS ESSID=9178b4853b70516d8acac25c6d424cd8](https://www.ferrostaal.com/uploads/tx_mfsmatrix/MF_ECHO_D_Dez09_WEB_2.pdf?PHPS ESSID=9178b4853b70516d8acac25c6d424cd8) (Stand: 09.11.2020).

[https://www.ferrostaal.com/uploads/tx\\_mfsmatrix/MF\\_ECHO\\_D\\_Dez09\\_WEB\\_2.pdf?PHPS ESSID=9178b4853b70516d8acac25c6d424cd8](https://www.ferrostaal.com/uploads/tx_mfsmatrix/MF_ECHO_D_Dez09_WEB_2.pdf?PHPS ESSID=9178b4853b70516d8acac25c6d424cd8) (Stand: 09.11.2020).

## **EL PAÍS:**

<https://www.elpais.com.uy/mundo/ninos-murieron-decenas-resultaron-heridos-viaje-escolar-argentina.html> (Stand: 10.11.2020).

[https://elpais.com/politica/2019/08/05/actualidad/1565017035\\_740712.html](https://elpais.com/politica/2019/08/05/actualidad/1565017035_740712.html) (Stand: 10.11.2020).

<https://elpais.com/opinion/2020-08-25/paraiso-irlandes.html> (Stand: 10.11.2020).

[https://elpais.com/diario/2006/11/11/andalucia/1163200928\\_850215.html](https://elpais.com/diario/2006/11/11/andalucia/1163200928_850215.html) (Stand: 10.11.2020).

[https://elpais.com/ccaa/2019/02/17/valencia/1550426513\\_369559.html](https://elpais.com/ccaa/2019/02/17/valencia/1550426513_369559.html) (Stand: 10.11.2020).

[https://elpais.com/elpais/2020/01/08/opinion/1578503270\\_023319.html](https://elpais.com/elpais/2020/01/08/opinion/1578503270_023319.html) (Stand: 10.11.2020).

## **SONSTIGE:**

<https://stopbasura.com/2016/03/30/incineradora/> (Stand: 10.11.2020).

<https://definicion.de/pretendiente/> (Stand: 10.11.2020).

<https://www.stern.de/familie/beziehung/julia-peirano/meine-verehrer-bedaengen-mich---wie-werde-ich-sie-los--7034992.html> (Stand: 10.11.2020).

<https://www.semana.com/nacion/articulo/corte-constitucional-el-derecho-de-las-mujeres-a-rechazar-a-un-hombre/627855/> (Stand: 10.11.2020).

<https://www.lanacion.com.ar/el-mundo/un-matrimonio-entre-hermanos-logro-ser-legalizado-en-espana-nid66386/> (Stand: 10.11.2020).

<https://www.ecured.cu/Endogamia> (Stand: 10.11.2010).

<https://de.wikipedia.org/wiki/Geschwisterehe> (Stand: 09.11.2020).

<https://www.besthelp.at/lexikon/inzest> (Stand: 09.11.2020).

<https://www.no-burn.org/wp-content/uploads/Incineración-Enemigo-del-Reciclaje.pdf> (Stand: 10.11.2020).

<http://www.rechtslexikon.net/d/verwertung/verwertung.htm> (Stand: 10.11.2020).

<https://www.fcc-group.eu/de/osterreich/technologien/behandlung/thermische-abfallbehandlung/definition-der-mullverbrennungsanlage.html> (Stand: 09.11.2020).

<https://modelandum.com/que-es-un-modelo-financiero/> (Stand: 10.11.2020).

<https://www.projektmagazin.de/glossarterm/leistungsumfang> (Stand: 09.11.2020)

[https://www.bamf.de/DE/Service/ServiceCenter/Glossar/glossar\\_node.html](https://www.bamf.de/DE/Service/ServiceCenter/Glossar/glossar_node.html) (Stand: 08.11.2020).

<https://www.biodiversidad.gob.mx/corredor/corredoresbio.html> (Stand: 10.11.2020).

<https://sites.google.com/site/358convivencia/comunidades-virtu/la-persona-como-ser-social> (Stand: 10.11.2020).

<https://www.culturagenial.com/es/el-hombre-es-un-ser-social-por-naturaleza/> (Stand: 10.11.2020).

<https://offenevergaben.at/branchen?node=35000000&anzahl> (Stand: 09.11.2020).

<https://de.rusbiznews.com/news/n974.html> (Stand: 09.11.2020).

<https://grlum.dpe.upc.edu/manual/sistemasIluminacion-equiposAuxiliares.php> (Stand: 10.11.2020).

[https://link.springer.com/chapter/10.1007%2F978-3-7091-0562-7\\_4](https://link.springer.com/chapter/10.1007%2F978-3-7091-0562-7_4) (Stand: 09.11.2020).

[https://books.google.at/books?id=jfw77xU5yvQC&pg=PA137&lpg=PA137&dq=was+ist+kr  
aftwerksleittechnik+definition&source=bl&ots=mep6EuTjh3&sig=ACfU3U2jLcLrHCISKtsJ  
YFHob\\_I0QAICDA&hl=de&sa=X&ved=2ahUKewjP3IvLj\\_XsAhUIXRoKHRL8AOs4ChD  
oATAAegQICRAC#v=onepage&q=was%20ist%20kraftwerksleittechnik%20definition&f=fa  
lse](https://books.google.at/books?id=jfw77xU5yvQC&pg=PA137&lpg=PA137&dq=was+ist+kr<br/>aftwerksleittechnik+definition&source=bl&ots=mep6EuTjh3&sig=ACfU3U2jLcLrHCISKtsJ<br/>YFHob_I0QAICDA&hl=de&sa=X&ved=2ahUKewjP3IvLj_XsAhUIXRoKHRL8AOs4ChD<br/>oATAAegQICRAC#v=onepage&q=was%20ist%20kraftwerksleittechnik%20definition&f=fa<br/>lse) (Stand: 09.11.2020).

<https://www.sew-eurodrive.es/produkte/steuerungstechnik.html> (Stand: 10.11.2020).

<http://www.sleeve technology.com/es?lang=es> (Stand: 10.11.2020).

<https://concepto definicion.de/crecer/> (Stand: 10.11.2020).

<http://little-princess-sara.net/episodesD.html> (Stand: 08.11.2020).

<https://www.excursioneescolares.com/tipos-excursioneescolares/> (Stand: 10.11.2020).

<https://www.drklein.de/schluesselfertig-bauen.html> (Stand: 09.11.2020).

[https://www.ferrostaal.com/uploads/tx\\_mfsmatrix/MF\\_ECHO\\_D\\_Dez09\\_WEB\\_2.pdf?PHPS  
ESSID=9178b4853b70516d8acac25c6d424cd8](https://www.ferrostaal.com/uploads/tx_mfsmatrix/MF_ECHO_D_Dez09_WEB_2.pdf?PHPS<br/>ESSID=9178b4853b70516d8acac25c6d424cd8) (Stand: 09.11.2020).

<https://www.procons.com.ar/portal/?p=909> (Stand: 10.11.2020).

[https://www.sign-lang.uni-http://www.rwi.uzh.ch/static/elt/lst-  
vogt/gesellschaftsrecht/grundlagen/de/html/unit\\_begriffgesellschaft.php](https://www.sign-lang.uni-http://www.rwi.uzh.ch/static/elt/lst-<br/>vogt/gesellschaftsrecht/grundlagen/de/html/unit_begriffgesellschaft.php) (Stand: 09.11.2020).

<https://www.verben.de/substantive/Handlanger.htm> (Stand: 09.11.2020).

<https://www.spektrum.de/lexikon/biologie/biotopvernetzung/8843> (Stand: 09.11.2020).

[https://www.landwirtschaft-bw.info/pb/Len/3650826\\_3651464\\_2315361\\_2316235](https://www.landwirtschaft-bw.info/pb/Len/3650826_3651464_2315361_2316235) (Stand: 09.11.2020).

<https://www.hidronor.cl/los-desechos-toxicos/> (Stand: 10.11.2020).

<https://rpp.pe/peru/arequipa/arequipa-aumenta-contaminacion-ambiental-tras-la-reactivacion-economica-noticia-1292999> (Stand: 10.11.2020).

<https://www.planet-wissen.de/natur/umwelt/umweltverschmutzung/pwieindustriellerevolutionundumweltverschmutzung100.html> (Stand: 10.11.2020).

<https://www.spektrum.de/lexikon/biologie/genealogie/27218> (Stand: 10.11.2020).

<http://www.aeiou.at/aeiou.encyclop.g/g235218.htm> (Stand: 09.11.2020).

<https://www.wortbedeutung.info/Bettjacker/> (Stand: 09.11.2020).

[http://www.evakoestner.de/Text\\_Maderna\\_dt.htm](http://www.evakoestner.de/Text_Maderna_dt.htm) (Stand: 09.11.2020).

## 13. Abstract

Die vorliegende Arbeit beschäftigt sich mit der Qualität maschineller Übersetzung ohne Prä- und Postedition. Dazu werden vier Fachtexte im Sprachenpaar Deutsch-Spanisch in Bezug auf die Terminologie, die in den Ausgangs- und Zieldtexten vorkommt, genau analysiert.

Relevante Termini der vier Fachtexte wurden anhand einer Terminologierecherche untersucht. Das daraus entstandene Glossar dient, neben der Humanübersetzung, als Vergleich für die Evaluation der maschinellen Übersetzungen. Hinsichtlich der bestehenden Theorien wurden Hypothesen aufgestellt und Forschungsfragen gestellt.

Es wird zunächst ein Überblick über die maschinelle Übersetzung gegeben.

Begriffsdefinitionen, der geschichtliche Hintergrund und verschiedene maschinelle Übersetzungssysteme werden vorgestellt. Außerdem wird die Rolle der Terminologie in der maschinellen Übersetzung diskutiert. Dabei werden die Begriffe *term recognition* und *termhood* analysiert.

Im empirischen Teil werden die vier Fachtexte und das methodische Vorgehen vorgestellt.

Die Evaluation wird anhand der SAE Qualitätsmetrik durchgeführt. Um die Auswertung nachvollziehen zu können, werden die einzelnen Fehlerkategorien in dem jeweiligen Kapitel einzeln beschrieben. Es folgt die Terminologierecherche und das daraus resultierende Glossar. Für die Auswertung wurden die ausgewählten Termini der maschinell erstellten Übersetzung anhand einer tabellarischen Gegenüberstellung mit der Humanübersetzung verglichen.

Die Ergebnisse zeigen, dass Humanübersetzungen insgesamt eine höhere Treffsicherheit in Bezug auf die Übersetzung von Fachtermini aufweisen als maschinelle Übersetzungstools.

Außerdem hat sich ergeben, dass neuronale maschinelle Übersetzungstools in der Übersetzung von wirtschaftlichen Fachtexten eine höhere Treffsicherheit in Bezug auf die Übersetzung von Fachtermini aufweisen als in der Übersetzung von geisteswissenschaftlichen Fachtexten. Die Qualität der Übersetzungen hängt nicht mit dem Schwierigkeitsgrad der Texte an sich zusammen. Vor allem das Fachgebiet und die damit verbundene Terminologie haben einen großen Einfluss auf die Ergebnisse.