



universität
wien

MASTERARBEIT / MASTER'S THESIS

Titel der Masterarbeit / Title of the Master's Thesis

„Sprachtechnologien und ihre Entwicklung
in der Europäischen Kommission“

verfasst von / submitted by

Borbála Katalin Eke, BA

angestrebter akademischer Grad / in partial fulfilment of the requirements for the degree of
Master of Arts (MA)

Wien, 2016 / Vienna 2016

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

A 060 381 342

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

Masterstudium Übersetzen

Betreut von / Supervisor:

Univ.-Prof. Dr. Gerhard Budin

Inhaltsverzeichnis

1. Einleitung	1
2. Sprachtechnologien	3
2.1. Einführung Sprachtechnologien und Computerlinguistik	3
2.2. Überblick Geschichte der Sprachtechnologien.....	6
2.2.1. Der ALPAC Bericht	7
2.2.2. Aufbau MÜ-Systeme	10
2.2.3. CAT Tools.....	16
3. Europäische Kommission.....	20
3.1. Überblick Geschichte der Europäischen Union.....	20
3.2. Die Kommission	24
3.2.1. Ordentliche Gesetzgebung	27
3.2.2. Besonderheiten der EU-Sprache	29
4. Generaldirektion Übersetzen	32
4.1. Geschichte der DGT	33
4.2. Meilensteine der EU-Sprachtechnologien	37
4.2.1. Systran.....	37
4.2.2. Eurotra.....	38
4.2.3. European Commission Machine Translation	39
4.2.4. Inter-Active Terminology for Europe	39
4.2.5. SDL Translator's Workbench	40
4.3. Qualitätssicherung und Sprachtechnologien.....	41
4.3.1. Ausgangslage	41
4.3.2. Qualitätssicherung in der DGT	43
5. Aktuelle Situation und Projekte der Kommission	45
5.1. Machine Translation at the European Commission (MT@EC)	45
5.1.2. Moses	47

5.1.3. MT@EC außerhalb der DGT	48
5.1.4. MT@EC innerhalb der DGT	56
5.1.5. MT@EC Konferenz	56
5.2. Europäischer Binnenmarkt und Übersetzung	57
5.2.1. Digital Single Market	57
5.2.2. Multilingual Digital Single Market	58
5.3. Connecting Europe Facility	65
5.3.1. CEF Automated Translation	67
5.3.2. Der DSI Bericht	68
5.3.3. European Language Resource Coordination (ELRC)	72
5.4. CAT Tool Projekte	74
5.4.1. Casmacat	74
5.4.2. Matecat	75
6. Interview	76
6.1. Grundlagen	76
6.2. Konklusion und Zusammenfassung	78
7. Ausblick und Schlussfolgerungen	80
7.1. Ausblick ÜbersetzerInnen und Sprachtechnologien	80
7.2. Schlussfolgerungen	89
8. Bibliographie	94
Anhang: Interviewtranskript	101
Curriculum Vitae	108
Abstract	110

1. Einleitung

„Die Sprache Europas ist die Übersetzung.“

Umberto Eco

“Translation is part of the DNA of the European Union.”

Christos Ellinides, Deputy Director-General, DGT

(<https://scic.ec.europa.eu>)

Das Übersetzen und die Verwendung von Computern sind mittlerweile untrennbar miteinander verbunden. Seit der Mitte des 20. Jahrhunderts haben sich die Werkzeuge von ÜbersetzerInnen grundlegend verändert und in mehreren Etappen erstaunliche Innovationen durchlaufen. Die Entwicklungen sind bei Weitem nicht an ihr Ende angekommen – dafür ist gerade die Europäische Union ein gutes Beispiel.

Innerhalb einer Organisation wie der Europäischen Union gelten besondere Bedingungen und Anforderungen an die Tools. Die Auswahl der richtigen Hilfsmittel ist für einen reibungslosen Ablauf entscheidend und muss folglich mit Sorgfalt getroffen werden. Dieser Prozess gestaltet sich nicht immer einfach und so zeigt ein Rückblick auf die Geschichte, dass oftmals mehrere Anläufe notwendig sind und stets Ausschau nach neuer Technologie gehalten werden muss.

Während mit 28 Mitgliedstaaten und 24 Amtssprachen die sprachliche Vielfalt als Kern Europas und der Europäischen Union bezeichnet wird, wird sie im Unionsalltag gleichzeitig auch als eine Hürde gesehen. Um sie zu überwinden, kommen Übersetzungsmethoden jeglicher Art ins Spiel – von Mensch bis hin zur Maschine, samt Hybriden und verschiedenster Arten der Sprachtechnologien.

Die Sprachtechnologie beschäftigt sich „mit der Entwicklung marktreifer Anwendungen der maschinellen Sprachverarbeitung“ (Carstensen 2012: 7). Die Definition ist weit und umfasst über die typischen Übersetzungstools hinausgehend auch traditionelle Schreibmaschinen bis hin zu Programmen zur Spracherkennung. In dieser Arbeit sollen jene wichtigen sprachtechnologischen Entwicklungen dargestellt werden, die die Europäische Kommission und die Generaldirektion Übersetzen (oder eng. *DGT*) am meisten beschäftigt und geprägt haben. Zurzeit umfasst dies die aktuellen Pläne rund um das System MT@EC, welches somit auch einen Schwerpunkt der Masterarbeit bildet. Um die Vorhaben besser nachvollziehen zu können, soll darüber hinaus ein Blick in die Vergangenheit geworfen und Zusammenhänge aufgezeigt werden.

Die Wahl fiel aus zwei Gründen auf dieses Thema. Zum einen ist es brandaktuell (die Projekte laufen bis 2020) und zum anderen hat es für ÜbersetzerInnen aus mehreren Gründen

Relevanz. Die Europäische Kommission verfügt über eines der größten Übersetzungsdienste weltweit und bietet für ÜbersetzerInnen sowie DolmetscherInnen ein interessantes Tätigkeitsfeld (vgl. <http://ec.europa.eu>^f). Mit ihren Tools vertraut zu sein, ist somit bereits deshalb vorteilhaft, weil die DGT für viele eine potentielle zukünftige Arbeitgeberin sein könnte.

Zudem ist die Union stets darauf bedacht, entsprechend ihrer Sprachenpolitik und ihrem Prinzip der Mehrsprachigkeit neue Projekte im Bereich der Sprachen zu fördern. Während die im Rahmen dessen entwickelten Sprachtechnologien explizit auf die Bedingungen in der Kommission und der Generaldirektion abgestimmt werden und nur einem begrenzten NutzerInnenkreis Zugriff gewährt wird, reicht ihr Einfluss über die Organe der Union hinaus. Die verwendeten Technologien und Standards wirken sich auf die Übersetzung außerhalb der Organisation aus und können ihren Weg auch auf den kommerziellen Markt finden. Sie entfalten somit mittelbar auch auf ÜbersetzerInnen außerhalb des EU-Systems einen Einfluss und bieten Input.

Den geschilderten Aspekten entsprechend bildeten folgende Forschungsfragen den Ausgangspunkt dieser Masterarbeit:

Wie haben sich die Sprachtechnologien in der Europäischen Kommission entwickelt?

Wie ist der aktuelle Stand der neuesten Projekte und welcher Fortschritt konnte erreicht werden? Welchen Einfluss haben diese Entwicklungen auf das Übersetzen und die ÜbersetzerInnen?

An dieser Stelle möchte ich mich zudem für die Unterstützung meiner Familie und Freunde bedanken. Mein Dank gilt auch Herrn Professor Budin, für die Betreuung der Arbeit und seine hilfreichen Tipps.

2. Sprachtechnologien

I rather suspect that, except for extremely urgent or fairly simple texts, people will prefer to continue getting at least the final versions of their work onto paper so that they can carry out a final check, [...] with the good old-fashioned pen, pencil or ball-point, unharassed by modern technology. It is not likely to replace them. After all, translation is in the end a creative activity, not a mechanical chore.

Peter J. Arthern 1979: 97

As we shall argue [...], language technology has been and always will be a way to augment our creative abilities.

Richard Sproat 2009: 240

Das vorliegende Kapitel dient der Darstellung allgemeiner computerlinguistischer Grundlagen, die für das Verständnis der aktuellen Vorgänge im Bereich der Sprachtechnologien in der DGT sowie deren bisherigem Werdegang relevant sind. Im Folgenden sollen weiters jene zwei Sprachtechnologien geschildert werden, in deren Bereich in der Europäischen Kommission Innovationen beobachtbar sind (sprich maschinelle Übersetzung und CAT Tools). Auf andere Technologien, die zu den Arbeitswerkzeugen von ÜbersetzerInnen der Kommission gehören, soll deshalb nur kurz eingegangen werden.

2.1. Einführung Sprachtechnologien und Computerlinguistik

Das Fachgebiet der Computerlinguistik beschäftigt „sich mit der maschinellen Verarbeitung natürlicher Sprache“ und bezieht sich sowohl auf gesprochene als auch auf geschriebene Sprache (Amtrup 2010³: 2).

Anzutreffen ist Computerlinguistik heutzutage bereits in unzähligen Bereichen des Alltages und wird von Menschen regelmäßig (und häufig unbewusst) angewendet. Als Beispiel können etwa jene Funktionen genannt werden, über die jedes Textverarbeitungsprogramm verfügt, seien es etwa die automatische Rechtschreibprüfung, Silbentrennung oder Filter zum gezielten Auffinden bestimmter Informationen. Zum Einsatz kommt die Computerlinguistik ebenso in Form von Textzusammenfassungsprogrammen, die anhand von Schlüsselbegriffen und typischen Phrasen die Kernaussagen eines Textes ermitteln können. Die Sprachsynthese gehört auch zu diesem Gebiet und kommt zum Einsatz, wenn von der Maschine mündliche Kommunikation erkannt werden soll. Dies geschieht etwa bei Navigationssystemen oder als Hilfestellung für blinde und sehbehinderte Menschen (vgl. Menzel 2010³: 24 und Ramlow 2009: 23-26).

Auch auf die Übersetzung übt die Computerlinguistik einen großen Einfluss aus. Als prominentes Beispiel aus dem Alltag können an dieser Stelle etwa Internetsuchmaschinen angeführt werden, die bei ihrer Suche die eingegeben Wörter morphologisch untersuchen und auf diese Art die passendsten Seiten für die gesuchten Begriffe ermitteln. Doch nicht nur hierbei spielt diese Disziplin eine Rolle: Beim Zugriff auf einen durch die Suchmaschine vorgeschlagenen Text besteht meist auch die Option zur Übersetzung, falls die Sprache den NutzerInnen nicht bekannt sein sollte. In diesem Sinne ist es nicht verwunderlich, dass gerade globale Unternehmen wie Google an der Entwicklung neuer Übersetzungstechnologien beteiligt sind (vgl. Amtrup 2010³: 14).

Von ähnlichen Übersetzungssystemen machen auch andere größere internationale Konzerne Gebrauch und fördern die Entwicklung und Forschung von Sprachtechnologie. Im Weiteren können sie solche Systeme für ihre Webseiten, verschiedene Serviceleistungen oder z.B. für Gebrauchsanweisungen heranziehen und das sprachliche Angebot für ihre BenutzerInnen und KundInnen erhöhen (auf die Details und Vorteile eben dieses Vorgehens konzentriert sich auch der Entwurf des mDSM Projektes, das in Kapitel 5.2.2. der vorliegenden Masterarbeit erläutert wird). Es ist nicht zuletzt der Globalisierung zuzuschreiben, dass Firmen ihre Produkte so schnell wie möglich und kosteneffizient mit Informationen in mehreren Sprachen versehen und vertreiben möchten. Auf dem weltweiten Markt sind nicht zuletzt auch Sprachen außerhalb von Europa – zu denken wäre u.a. an Chinesisch – von großer Bedeutung für die Unternehmen (vgl. Amtrup 2010³: 14-15).

Das Fachgebiet ist stark mit Nachbardisziplinen verbunden, aus denen verschiedene Ansätze einfließen. Dazu gehören Methoden und Gedankengänge aus der Linguistik und Informatik, aber teilweise auch aus der Philosophie (z.B. die Suche nach den Zusammenhängen zwischen Handeln, Denken und Sprache), der Kognitionswissenschaften (z.B. die Untersuchung kognitiver Prozesse bei der Sprachbeherrschung des Menschen sowie die Analyse von nicht-sprachlichem Wissen) und der Mathematik (Versuch die natürliche Sprache des Menschen mithilfe von Wahrscheinlichkeiten und Statistiken zu umschreiben) (vgl. Amtrup 2010³: 3-4).

Die zwei wichtigsten Einflüsse stellen die Linguistik und Informatik dar. Mit der Linguistik teilt sich die Computerlinguistik beispielsweise die Struktur und Einteilung der Teilgebiete (z.B. Syntax) sowie einen Großteil der Terminologie. Die linguistische Forschung innerhalb der Computerlinguistik verläuft vor allem auf theoretischer Ebene, d.h. die tatsächliche Umsetzbarkeit der Ideen spielt eine untergeordnete Rolle – dies zeigt sich u.a. bei dem Interlingua Prinzip (vgl. Ramlow 2009: 17).

Eine weitere Verbindung stellt die Anwendung und das Experimentieren mit der praktischen Umsetzung von linguistischen Theorien mit den Mitteln von Computern dar. So untersucht die Computerlinguistik, ob bestimmte Theorien z.B. auch für die maschinelle Übersetzung (MÜ) angewendet werden können. Schließlich beeinflussen sich Linguistik

und Computerlinguistik gegenseitig, denn Hypothesen aus einem der beiden Fachgebiete können zu neuen Erkenntnissen und Forschungen in dem jeweils anderen Gebiet führen. Als Beispiel hierfür kann erwähnt werden, dass für die Ausarbeitung von gesammelten Daten, die beispielsweise für Terminologearbeiten oder Analysen von Textkorpora benötigt werden, bei großen Mengen praktisch nur mehr mit Computerprogrammen möglich ist (vgl. Amtrup 2010³: 3-4).

Die verschiedenen Einflüsse verdeutlichen bereits, dass man sich an das Gebiet der Computerlinguistik sowohl aus praktischem als auch als theoretischem Blickwinkel nähern kann. Zu dem Teilgebiet der praktischen Computerlinguistik gehört die Sprachtechnologie, der nicht nur in dieser Arbeit eine zentrale Bedeutung zukommt. Sie beschäftigt sich mit Sprachsoftwares und legt den Schwerpunkt somit auf die praktische Entwicklung und Anwendung. In der Umsetzung bedeutet dies, dass unter anderem danach geforscht wird, wie Phänomene einer Sprache in Form von Algorithmen in Modelle gefasst werden können, mit denen im Anschluss ein Softwaresystem arbeiten und auf diesem Wege die menschliche Sprache simulieren kann. Dazu gehören z.B. die Erkennung von grammatikalischen Regeln oder einer besondere Lexikographie einzelner Sprachen. Mithilfe der auf diesem Wege entwickelten Maschinen und Computern sollen praktische Probleme der Sprache gelöst werden – wie etwa die Übersetzung. Maschinelle Übersetzung, Translation Memories und eine weitere Reihe von Übersetzungshilfen im Bereich der Terminologie oder bspw. der Dokumentation gehören allesamt in den Bereich der Sprachtechnologie. Aber das Forschungsfeld ist noch weiter: Spracherkennung und Verarbeitung von Informationen (z.B. im Internet) zählen ebenso zu der praktischen Computerlinguistik sowie bestimmte Tests und Methoden, um diese Systeme zu evaluieren (vgl. Amtrup 2010³: 6-7 und Ramlow 2009: 18).

Im Mittelpunkt der vorliegenden Arbeit befindet sich die Sprachtechnologie und somit die praktische Computerlinguistik, weshalb die theoretische Computerlinguistik an dieser Stelle nur kurz erwähnt wird. Während sich die Theorie dieses Fachs grundsätzlich mit denselben Fragen wie die Praxis auseinandersetzt, sind den Lösungen keine praktischen Grenzen gesetzt. Es bestehen keine Ansprüche, die Überlegungen unbedingt umsetzen zu müssen, womit die tatsächliche Anwendung der auf Papier entwickelten Systeme eine nebensächliche Rolle einnimmt. Einige Fragen dieses Forschungsgebietes sind: Wie kann ein maschinelles System kreiert werden, das mit einer komplexen Sprache zurechtkommt? Wie muss es gestaltet sein, um in möglichst vielen Sprachen eine adäquate Übersetzung fertigen zu können? Wie schnell kann es neue Eingaben analysieren, berechnen und daraus automatisch für die zukünftige Anwendung lernen? Wie können an sich Fehler vermieden und wenn notwendig, behoben werden? Die hier aufgezählten Fragestellungen zeigen bereits, dass die Lernfähigkeit von Systemen einen zentralen Punkt in der theoretischen Forschung bildet und im Idealfall Input für die Praxis bietet (vgl. Amtrup 2010³: 8-10).

2.2. Überblick Geschichte der Sprachtechnologien

Die ersten Schritte machte die Computerlinguistik bereits Anfang des 20. Jahrhunderts. In den 20er und 30er Jahren lag der Schwerpunkt der Forschung und Anwendung verständlicherweise noch auf anderen Bereichen, so etwa der Anwendung neuer Maschinen „zur Dechiffrierung verschlüsselter Nachrichtentexte“ (Menzel 2010³:18). In den 50er Jahren stellte erstmals Allen Turing Überlegungen zu der Frage nach der maschinellen Verarbeitung von Informationen an und zog dabei auch eine mögliche Sprachverarbeitung in Erwägung.

Im Fokus der Sprachtechnologien stand Anfangs somit die maschinelle Übersetzung. Das ernsthafte Interesse an der MÜ und ihrer umfangreicheren Anwendung etablierte sich nach dem 2. Weltkrieg und wurde (zu Beginn verstärkt in Amerika) durch zwei ausschlaggebende Entwicklungen ausgelöst. Zunächst durch den Kalten Krieg und die damit zusammenhängende Gefahr, die westliche Mächte ausgehend von russischen Technologien, Waffen und wissenschaftlichen Experimenten verspürten. Damit die ForscherInnen, von denen bloß sehr wenige Russisch sprachen, schneller die Informationen über die Vorgänge in Russland verstehen und verarbeiten konnten und um kostspielige Humanübersetzungen zu vermeiden, schien die Entwicklung eines funktionsfähigen maschinellen Übersetzungssystems eine gute Lösung darzustellen. Der zweite Faktor bestand in den soeben genannten Errungenschaften im Bereich des Kodierens bzw. Dekodierens während des Krieges. Nach dem Ende des Krieges begannen WissenschaftlerInnen für ihre neuen Erkenntnisse auch einen Anwendungsbereich außerhalb der militärischen Kreise zu suchen und wandten sich im Zuge dessen vermehrt der möglichen Übersetzung mithilfe von Computern zu (vgl. Sproat 2010: 218).

1954 fand die erste öffentliche Präsentation eines maschinellen Übersetzungssystems statt, das von dem amerikanischen IT-Unternehmen IBM entwickelt wurde. Es handelte sich dabei der Nachfrage entsprechend um ein System mit einer Russisch-Englischen Sprachkombination und war bloß auf eine stark eingeschränkte Anzahl an Wörtern und Grammatikregeln anwendbar – es arbeitete mit insgesamt 250 Wörtern und sechs Grammatikregeln. Während die Präsentation von IBM innerhalb dieses kleinen sprachlichen Anwendungsbereiches ein gelungenes Experiment war und die anwesenden JournalistInnen mit Begeisterung von dem System berichteten, war es bei Weitem nicht auf eine weitreichendere Übersetzung von unterschiedlichen Texten anwendbar. Da die ersten Schritte der MÜ jedoch einen guten ersten Eindruck hinterließen, war im Kreis der EntwicklerInnen im Allgemeinen ein ungebrochenes Selbstbewusstsein vorherrschend: Gemäß der überwiegenden Vorstellung sollte die Technologie in etwa fünf Jahren so weit fortgeschritten sein, dass kommerzielle Programme bis zu zwei Millionen Wörter pro Stunde aufarbeiten können sollten (vgl. Sproat 2010: 221-222).

Der maschinellen Übersetzung wurde somit nach den ersten Versuchen eine hoffnungsvolle Zukunft vorausgesagt. Diese Phase war durch Euphorie und gute Prognosen

gekennzeichnet, aber 1966 brachte der amerikanische ALPAC-Bericht (*Automatic Language Processing Advisory Committee*) einen Einschnitt mit sich und dämpfte die Erwartungen der MÜ-ForscherInnen erheblich (vgl. Sproat 2010: 222).

2.2.1. Der ALPAC Bericht

Der von der *National Academy of Sciences* und der *National Research Council* in Auftrag gegebene ALPAC Bericht (mit dem Untertitel „*Language and machines: computers in translation and linguistics*“), sollte den Stand der maschinellen Übersetzungstechnologie beleuchten sowie feststellen, ob es sich für die amerikanische Regierung weiterhin lohnen würde, neue Forschungen in diesem Bereich finanziell zu unterstützen (Hutchins 1996: 15).

In den Erläuterungen wird MÜ als ein Prozess definiert, bei dem mithilfe von Algorithmen eine Übersetzung angefertigt wird, nämlich „*going by algorithm from machine-readable source text to useful target text, without recourse to human translation or editing*“ (ALPAC 1966: 19). Aus der Definition lässt sich erschließen, dass die GutachterInnen eine notwendige Verbesserung von maschinellen Übersetzungen bereits als große Schwäche der Systeme betrachteten. Ebenso ist es wesentlich zu betonen, dass sich die Analyse auf jene Systeme beschränkte, die innerhalb der Organe der amerikanischen Regierung und ihren militärischen Einrichtungen eingesetzt wurden. Wie bereits erläutert, bedeutete dies, dass die MÜ hauptsächlich zur Übersetzung von russischen Texten von besonderem Interesse war, andere Sprachkombinationen und Einsatzbereiche jedoch nicht analysiert wurden (vgl. Hutchins 1996: 16).

Die ExpertInnen äußerten ihre Zweifel bezüglich der Nützlichkeit von MÜ in mehrerer Hinsicht. Ein wichtiger Aspekt war es, die tatsächliche Übersetzungsnachfrage festzustellen. Dieser stellte sich laut dem Bericht als nicht sonderlich hoch dar, da die russischen Texte primär nur den amerikanischen WissenschaftlerInnen dienten, die bei der US-Regierung angestellt waren und der überwiegende Teil der Originaltexte ohnehin auf Englisch verfasst war. Folglich argumentierte die ALPAC Kommission, dass es sinnvoller wäre, wenn die wenigen WissenschaftlerInnen, die stark auf die russischen Artikel angewiesen waren, einfach Russisch lernen würden, anstatt auf Übersetzungen zu warten. Es werde im Allgemeinen zu viel übersetzt und in Zukunft müsse genauer selektiert werden – der tatsächlich notwendige Arbeitsaufwand könne somit auch mit den in mehr als ausreichender Zahl vorhandenen HumanübersetzerInnen bewältigt werden (vgl. Hutchins 1996: 16-17).

Weitere wesentliche Aspekte, die die Kommission zu einer negativen Entscheidung veranlassten, bestanden in relativ langen Wartezeiten und hohen Kosten der maschinellen Übersetzungen. Die maschinelle Übersetzung eines Journals konnte teilweise bis zu 100 Tage dauern – dies lag unter anderem an den Vor- und Nachbearbeitungsprozessen der von

den MÜ Programmen generierten Übersetzungen. Der ALPAC Bericht bezeichnete die Ergebnisse als ungenau und praktisch unbrauchbar, weshalb die Übersetzungen folglich zu viel Nachbearbeitung und finanzielle Mittel benötigen würden, um effizient und sinnvoll verwendet werden zu können (vgl. Hutchins 1996: 17).

Im Allgemeinen drückte der Bericht die Enttäuschung der ALPAC Kommission über die anfangs vielversprechenden Technologien aus. Ausgehend von den ersten Experimenten mit dem IBM System hatte man sich mehr erwartet, doch die Übersetzungen der wissenschaftlichen Texte erwiesen sich als fehlerhaft und schwer zu lesen. In diesem Sinne deutete die Kommission seinen Standpunkt bereits zu Beginn des Berichtes an:

There is no emergency in the field of translation. The problem is not to meet some nonexistent need through nonexistent machine translation. There are, however, several crucial problems of translation. These are quality, speed, and cost (ALPAC 1966:16).

Im Gegensatz zu diesem negativen Urteil schrieb der Bericht der so genannten *machine-aided translation*, somit den heute gängig als CAT Tools bezeichneten Technologien bzw. ähnlichen Technologien, einen großen Nutzen zu. Im Gegensatz zu der *human-aided machine translation*, wenn also der Mensch als „Stütze“ für die Maschine agiert, sei die Vorgehensweise, bei der die Maschine als Stütze des Menschen dient, der richtige Weg. Unter anderem lobten die ExpertInnen die Terminologiedatenbank der Europäischen Gemeinschaft für Kohle und Stahl, die in den einzelnen Einträgen bereits Kontextsätze enthielt. Hilfswerkzeuge dieser Art seien auf jeden Fall hilfreicher und wertvoller als die von den MÜ-Systemen gebotenen Leistungen, wie der Bericht urteilte (vgl. Hutchins 1996: 18-19).

Zusammenfassend stellte der Bericht zwar fest, dass maschinellen Übersetzungssysteme nicht gänzlich unnütz seien, aber die Maschinen noch nicht einen Punkt erreicht hätten, an dem es sich lohnen würde, die Projekte finanziell zu unterstützen. Stattdessen empfahlen die ExpertInnen eher in die Erforschung von grundlegenden Fragen der Linguistik und Computerlinguistik zu investieren, um daraus im Späteren möglicherweise nützliche Ansätze für die MÜ zu gewinnen. Als Folge dessen zog die amerikanische Regierung die finanziellen Mittel aus diesem Bereich ab und zu einem Großteil bedeute dies über einen längeren Zeitraum das Ende der großen MÜ-Projekte in den Vereinigten Staaten. In Kanada sowie in Europa blieb das Interesse weiterhin bestehen (vgl. Sproat 2010: 223-224).

Die Veröffentlichung zog eine Reihe von Kritik nach sich: Der ALPAC Kommission wurde vielfach vorgeworfen, dass die Mitglieder bei ihrer Analyse voreingenommen vorgegangen seien und die negativen Aspekte verstärkt hervorgehoben hätten. Aus heutiger Sicht kann vor allem die Meinung vertreten werden, dass der ALPAC Bericht in vielen Bereichen recht hatte, aber letztendlich voreilig und hart geurteilt hatte (vgl. Hutchins 1996:20).

Eine genauere Betrachtung des ALPAC Berichtes erweist sich trotz seiner Fehler aus mehreren Gründen als interessant: einerseits hatte er zweifellos großen Einfluss auf die Entwicklung der maschinellen Übersetzung und förderte technologische Übersetzungshilfsmittel. Andererseits zeigte er bereits damals Aspekte der MÜ auf, die bis heute viel diskutiert werden. Seit der Veröffentlichung des oftmals als berüchtigt bezeichneten Berichtes, sind zwar mehr als vier Jahrzehnte vergangen, die sehr viele Veränderungen mit sich gebracht haben, aber dennoch sind einige grundlegenden Fragen und Beweggründe bis heute gleich geblieben. 1966 stellte die ALPAC Kommission Fragen danach, wie eine hochwertige Übersetzung von teils komplexeren Texten mit angemessenem Zeitaufwand und Kosten erreicht werden kann und welche technologischen Mittel dazu am ehesten beitragen können – Überlegungen, denen auf heutigen Konferenzen weiterhin nachgegangen wird.

Seit den 50er Jahren ist ein großes Problem der MÜ die Diskrepanz zwischen den oftmals sehr hohen Erwartungen und den realistisch erzielbaren Ergebnissen. So wurde zunächst die *FAHQT* Übersetzung als Ziel angesehen: „*fully accurate high quality translation*“ – ein System, das auf Anhieb hochwertige Übersetzungen liefern sollte, mit Texten, die dasselbe Niveau erreichen sollten, wie Humanübersetzungen (Jekat/Volk 2010³: 643). Schon bald zeigte sich deutlich, dass diese hohen Erwartungen nicht erfüllt werden können – und die vollständige Lösung dieser Probleme ist weiterhin ein umfassendes Forschungsgebiet in der Sprachtechnologie. Während neben den bereits existierenden Lösungsansätzen laufend neue Ergebnisse erzielt werden können, bedeutet dies dennoch nicht, dass die Forschung eine endgültige Lösung der Probleme der Sprachtechnologie in Reichweite sieht.

Neben der MÜ stellen jedoch seit mehr als zehn Jahren die so genannten CAT Tools eine wesentliche Hilfe im Übersetzungsprozess dar – genau wie es der ALPAC Bericht beschrieben hatte. *Computer-Aided Translation* (oder eben *machine-aided translation*) wurde speziell für professionelle AnwenderInnen entwickelt und richtet sich dementsprechend stark an deren Bedürfnissen. Zu den oft genannten Vorteilen zählen unter anderem, dass bereits übersetzte Texte leicht wieder angewendet werden können und eine konsistente Übersetzung erleichtert wird. Im Gegensatz zu der MÜ ist die Nützlichkeit und erfolgreiche Anwendung von CAT Tools bereits weitverbreitet und im Bereich der professionellen Übersetzung anerkannt (vgl. Jekat/Volk 2010³: 654, 656 und Hansen-Schirra 2012: 211).

Die Sprachtechnologie liefert somit mehrere praktische Ansätze, die sich in den letzten Jahren durchaus bewährt haben. Einen wesentlichen Fortschritt ermöglichte die Entwicklung von computerlinguistischen Ressourcen. Dazu gehören unter anderem Textkorpora, Baumbanken und lexikalisch-semantische Ressourcen. Dass der Zugriff, die Aufbereitung und die Handhabung einer großen Datenmenge möglich geworden sind, hat für Übersetzungstools einen immensen Vorteil, wie sich in den folgenden Kapiteln zeigen wird (vgl. Carstensen et al. 2010³: 481).

2.2.2. Aufbau MÜ-Systeme

Bei der Unterscheidung von MÜ-Systemen ist die zugrunde liegende Übersetzungsstrategie ausschlaggebend. Sie hat wesentlichen Einfluss auf die Analyseart, die Funktionsweise und folglich auch auf die gelieferten Ergebnisse der jeweiligen Maschine. Es haben sich zwei komplexe Systeme durchgesetzt, die entweder auf regel-basierten oder daten-basierten Strategien beruhen. Für die Übersetzungsdienste der EU sind beide Ansätze gleichermaßen wichtig: die ersten Gehversuche im Bereich der MÜ machte die DGT mit regel-basierten Systemen, doch in Zukunft werden diese auf daten-basierte Strategien ausgetauscht. In den nächsten Abschnitten sollen beide genauer erläutert werden.

2.2.3.1. Regel-basierte Systeme

Als die einfachste Übersetzungsstrategie der regel-basierten Systeme gilt jene der direkten Übersetzung. Im Zuge dieser wird der zu übersetzende Ausgangstext morphologisch analysiert, jedem Wort seine Wortart zugeordnet und danach mit einem zweisprachigen Wörterbuch in die Zielsprache übertragen. Abschließend werden bei der direkten Übersetzungsmethode noch letzte Anpassungen vorgenommen, damit sichergestellt werden kann, dass der jeweilige Satz über die richtigen Wortendungen und eine korrekte Syntax verfügt. Diese Methode wird jedoch für komplexere Aufgaben nicht verwendet, da sie trotz ihrer Schnelligkeit und niedrigen Kosten keine guten (da oftmals ungrammatische) Ergebnisse liefert. Dafür ist die anfangs durchgeführte Analyse zu oberflächlich und die vom System angewandten grammatikalischen Regeln nicht ausreichend. Besonders bei Sprachen, die sich stark voneinander unterscheiden (starke Abweichung bei z.B. Syntax oder Lexik), ist bloß eine niedrige Qualität zu verzeichnen (vgl. Jekat/Volk 2010³: 644-645 und Ramlow 2009: 73-75).

Gerade diese Schwachstelle macht direkte Übersetzungssysteme für Organisationen wie die Europäischen Union uninteressant: das System wäre bei einer Vielzahl an Sprachkombinationen überfordert und würde bloß einen sehr geringen Nutzen versprechen. Während Übersetzungen aus dem Französischen ins Spanische aufgrund ihrer zahlreichen gemeinsamen Züge in gewisser Hinsicht sicherlich brauchbar wären, könnte dies etwa von äußerst unterschiedlichen Sprachen wie etwa Englisch und Ungarisch wohl nicht mehr behauptet werden.

Verglichen mit anderen regel-basierten MÜ-Ansätzen nimmt die Direktübersetzung aus diesem Grund in der Rangordnung die niedrigste Stufe ein und wird von anderen Herangehensweisen mit elaborierteren Analyseverfahren übertroffen. Diese Hierarchie stellt das Vauquois Dreieck dar.

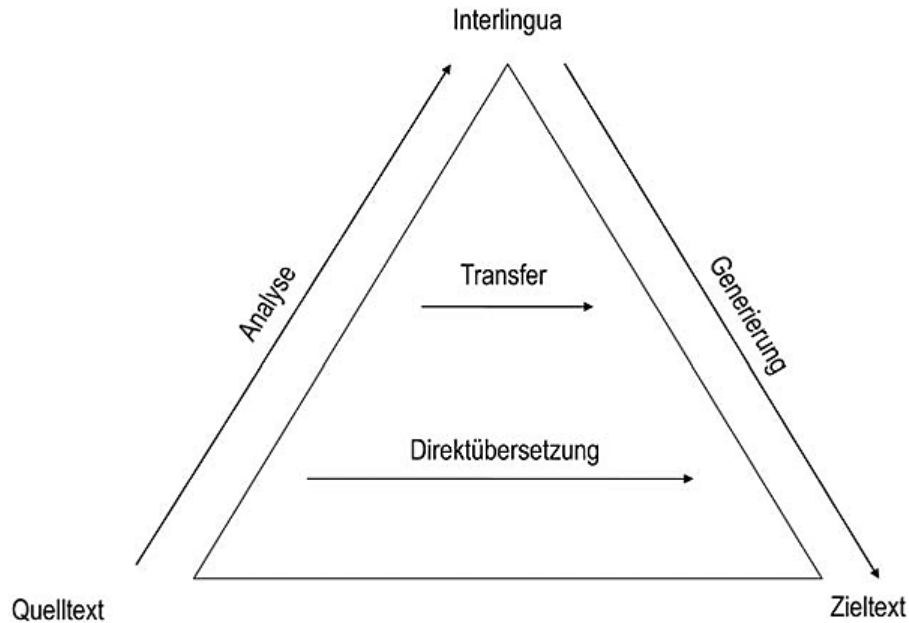


Abbildung 1. Vauquois Dreieck (Seewald-Heeg/Stein 2009: 9)

Eines der komplexeren Verfahren ist der transferbasierte Ansatz, bei dem als erster Schritt eine grammatikalische und semantische Analyse des eingegebenen Ausgangstextes erfolgt. Danach kommt es zum Einsatz von Transferregeln, die dazu dienen, den ausgangssprachlichen Text „in eine abstrakte, zielsprachliche Repräsentation“ umzuwandeln (Jekat/Volk 2010³: 646). Transferregeln können beispielsweise Aussagen über die gewöhnliche Stellung von Adjektiven enthalten (z.B. im Englischen vor dem Nomen, im Spanischen hingegen meist dahinter) (vgl. Sproat 2010: 226). So würde für die Umwandlung des englischen Ausdrucks „*the installation configuration*“ ins französische „*la configuration d’installation*“ solch ein System folgende Transferregel enthalten: *N1* (steht für *the installation*), *N2* (für *configuration*) wird zu „*N2 de N1*“ (Jekat/Volk 2010³: 646).

Die dem Programm vorgegebenen Transferregeln dienen dem Zweck, die möglichen Übersetzungen auf jene einzuschränken, die sinnvoll sind. Zum Abschluss fehlt noch eine „natürlichsprachliche Ausgabe“ des Textes, die somit dem fertigen Zieltext entspricht (Jekat/Volk 2010³: 646). Die vom Computer durchgeführte Analyse sowie die vorgegebenen Regeln sind äußerst sprachspezifisch und gelten folglich nur für das gerade verwendete Sprachenpaar. Das ermöglicht zwar ein genaues System, aber es bringt auch einen großen Nachteil mit sich. Sobald die Sprachrichtung geändert wird (z.B. statt ENG-FRA nun FRA-ENG) oder das System weitere Sprachen in sein Repertoire aufnehmen soll, ist die Programmierung mehrerer neuer Komponenten und Transferregeln notwendig. In Anbetracht dessen, dass die neuen Regeln jeweils manuell eingestellt werden müssen, ist es kaum verwunderlich, wie umfangreich und zeitaufwendig sich diese Aufgabe gestaltet (vgl. Jekat/Volk 2010³: 646).

Gerade auch für eine Organisation mit der Größe und Sprachenvielfalt der Union stellt dies keine geringe Herausforderung dar. Für jede mögliche Sprachenkombination sind demnach eigene Transfersysteme mit neuen Regeln und Analysemethoden zu entwickeln. Bereits bei einer kleinen Anzahl an Sprachen bedeutet dies deutlich mehr als zwanzig Systeme – umgerechnet auf die aktuell 24 Amtssprachen der EU steigt diese Zahl um ein Mehrfaches an. Dies ist mit ein Grund, weshalb das regelbasierte System ECMT (*European Commission Machine Translation*) 2010 von der Kommission abgeschafft wurde und mit der Erarbeitung einer praktischeren MÜ-Form begonnen wurde.

Die zweite Form der regel-basierten Systeme ist jene der Interlingua. Ihr liegt die Idee zugrunde, dass der Computer im Zuge einer Analyse des eingegebenen Ausgangstextes zunächst eine von Sprache völlig losgelöste Repräsentation erstellt. Im Unterschied zur direkten Übersetzung und zum Transferansatz spielen somit die Ausgangs- und Zielsprache keine Rolle, da auch keine direkte Verbindung zwischen den beiden hergestellt wird. Wenn der in das System eingegebene Satz „Guten Tag, Herr Professor Klabunde.“ lautet, dann wäre „[Greet]; [Title]; [name: Klabunde]“ die adäquate sprachenunabhängige Repräsentation, die vom Programm generiert wird (Jekat/Volk 2010³: 646). Dies nimmt das Programm als Grundlage für die Erstellung der Übersetzung in der Zielsprache und bezieht sich deshalb nicht unmittelbar auf den Ausgangstext (vgl. Jekat/Volk 2010³: 646-647).

Der soeben beschriebene Ablauf stellt sich jedoch als ein bloß theoretisches Szenario heraus. Bisher konnte keine angemessene Umsetzung eines Interlingua Systems erreicht werden. Welche Bedeutung die einzelnen Ausdrücke verschiedener Sprachen tragen, wie diese auf eine abstrakte Weise korrekt dargestellt und in andere Sprachen übertragen werden können, ist schwer zu erfassen und kaum möglich. Für einen problemlosen Ablauf mit Interlingua müssten sich die Konzepte der jeweiligen Sprachen decken – doch nicht zuletzt jahrzehntelange Forschung und Erfahrung aus der Translationswissenschaft zeigen, dass das keineswegs der Fall ist. Bereits ein „einfaches“ Wort wie das englische „wall“ deckt sich nicht unbedingt mit dem deutschen „Wand“, denn „wall“ kann sich genauso gut auf eine „Mauer“ beziehen – und diesen Unterschied können Maschinen (bislang) nicht erkennen (Jekat/Volk 2010³: 647 und vgl. Sproat 2010: 227).

2.2.3.2. Daten-basierte Systeme

Im Gegensatz zu den soeben beschriebenen Ansätzen, die mit manuell erstellten bilingualen Wörterbüchern und Regelsammlungen arbeiten, setzen daten-basierte Systeme (auch korpus-basierte Systeme genannt) auf umfassende Textkorpora. Hier kann ebenfalls zwischen zwei Varianten unterschieden werden, der beispieldbasierten und der statistischen maschinellen Übersetzung.

Das Prinzip der beispielbasierten MÜ beruht darauf, dass Übersetzungen wie in einer TM gespeichert werden. Muss später ein Satz übersetzt werden, der nur Elemente enthält, die im System bereits erfasst sind, dann wird sofort eine Übersetzung generiert. Wenn jedoch im Quelltext keine bekannten Satzteile enthalten sind, kann ein beispielbasiertes System keinen Zieltext zusammensetzen (diese Ergebnisse versucht man oft durch das Hinzufügen von Regelsammlungen zu verbessern). Der Nutzen ist somit recht gering und der Einsatz nur in einer eingeschränkten Textumgebung sinnvoll. Mit dieser Erkenntnis scheidet dieser Ansatz für die DGT und die Kommission aus (vgl. Ramlow 2009: 79-80).

Die seit den 1990er Jahren entwickelten Systeme zur statistischen maschinellen Übersetzung weisen in ihrer Funktionsweise viele Ähnlichkeiten zu TMs auf. Ihre Leistung hängt maßgeblich davon ab, wie viele und welche Art von Texten ihnen zuvor eingespeist wurden. Als Grundlage dient nämlich ein Korpus, das aus Humanübersetzungen und den dazugehörigen Quelltexten besteht. Diese studiert der Computer und lernt daraus, die jeweiligen Sprachen in der Übersetzung neuer Texte anzuwenden. Soll die Maschine Übersetzungen hauptsächlich in einem bestimmten Fachbereich liefern, dann empfiehlt es sich, das Korpus gezielt mit einer großen Menge an Texten aus diesem Bereich zusammenzusetzen, d.h. im Hinblick auf zufriedenstellende Ergebnisse, gilt es eine überlegte Auswahl zu treffen. Im Falle eines internationalen Unternehmens könnten die eingegebenen Texte etwa aus Jahres- und Quartalsberichten und Ähnlichem bestehen (vgl. Jekat/Volk 2010³: 648). Analog bedeutet das in Anbetracht der hauptsächlichsten EU-Textsorten (siehe unten Kapitel 3.2.1.), dass ein idealer SMÜ-Korpus auf politischer und rechtlicher Thematik basiert.

Das SMÜ-System analysiert anschließend das gesamte Korpus und erstellt auf dessen Basis automatisch sprachliche Entsprechungen sowie Wahrscheinlichkeiten (z.B. für bestimmte Wortfolgen und die Verwendung von Ausdrücken). Wie bei einem TM werden die jeweils zusammengehörenden Sätze der Ausgangs- und der Zielsprache miteinander verbunden. Dieser Vorgang geschieht mittels eines Alignierungsprogrammes, das mithilfe von Anhaltspunkten (z.B. Eigennamen oder Zahlen) eine Zuordnung durchführt.

Als nächstes bearbeitet das Programm die Sätze auf ihrer Wortebene, um eindeutig festzustellen, welche beiden Wörter einander entsprechen. Dieser Schritt erweist sich bereits als eine größere Herausforderung, da wie schon zuvor erwähnt, nicht alle Ausdrücke eine deckungsgleiche Bezeichnung in einer anderen Sprache aufweisen und dies den Computer irreführen könnte. Die Satzelemente „*weiterhin nachfrageschwach*“ können z.B. in einer englischen Übersetzung als „*demand continued to be weak*“ vorkommen (Jekat/Volk 2010³: 649). Eine gewisse Fehlerrate muss auf dieser Ebene somit einberechnet werden. Aus den einander zugeordneten Worten erstellt das Programm abschließend ein „aligniertes Wortfolgenlexikon“ (Jekat/Volk 2010³: 649). Mithilfe eines Decoders setzt das Programm bei der Übersetzung die zielsprachlichen Texte zusammen, in dem es die aus dem Trai-

ningskorpus erlernten Wortfolgen nach deren Wahrscheinlichkeit beurteilt und die demnach beste Satzzusammensetzung aussucht. Dieser Teil des System bildet das Übersetzungsmodell (vgl. Jekat/Volk 2010³: 648-650).

Um diese Funktion zu optimieren, wird zusätzlich ein Sprachmodell der Zielsprache erstellt. Dabei untersucht die Maschine unabhängig von der Ausgangssprache ein größtmögliches Korpus der Zielsprache und merkt sich typische Merkmale, z.B. gängige Wortfolgen. Die Texte dieses zweiten Korpus können sowohl aus den vorhin erwähnten Humanübersetzungen bestehen als auch aus anderen Texten in der Zielsprache (allerdings sollten sie den intendierten Fachbereich betreffen). Die daraus gewonnen Erkenntnisse über die Zielsprache sollen dazu beitragen, dass sich die vom SMÜ-System generierte Übersetzung flüssig liest, während die aus dem Trainingskorpus abgeleiteten Informationen der Übersetzungstreue dienen (vgl. Jekat/Volk 2010³: 650).

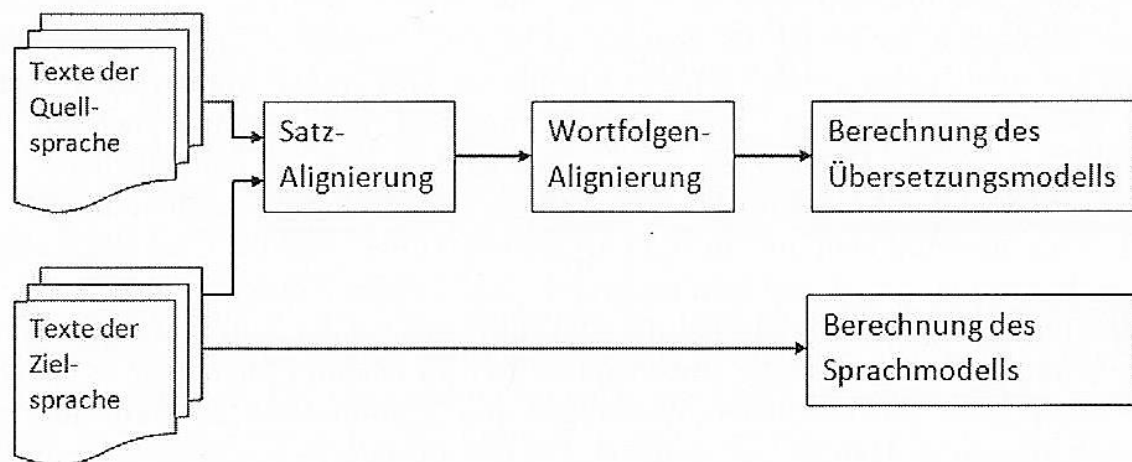


Abbildung 2. Training eines SMÜ-Systems (Jekat/Volk 2010³: 649)

Ein zu erwähnender Nachteil besteht darin, dass in den klassischen SMÜ-Systemen nur statistische Informationen und somit keinerlei sprachlichen Regeln herangezogen werden. Das Programm erkennt deshalb beispielsweise nicht, dass *sage* und *sagst* Konjugationsformen desselben Verbes sind. Gerade für das Deutsche, das zu den Sprachen mit hoher Flexion und vielen Komposita gehört, erweist sich die Anwendung von reinen SMÜ-Systemen somit nicht als ideal (vgl. Ramlow 2009: 82 und Jekat/Volk 2010³: 652). Da die Ergebnisse dementsprechend nicht zufriedenstellend sind, laufen in der DGT bereits Bemühungen eine Kombination aus statistischen und regel-basierten Systemen zu erstellen und in der Praxis anwendbar zu gestalten. Dies ist nicht zuletzt für das Gelingen der neuen Projekte in sämtlichen EU-Sprachen eine notwendige Entwicklung (vgl. <https://scic.ec.europa.eu>; zu Projekten siehe Kapitel 5).

Während anhand der Beschreibung der Ablauf und die Komponenten der SMÜ-Systeme komplexer und aufwändiger erscheinen, als die der zuvor erläuterten MÜ-Programme, erweist sich die Situation in der Realität anders. Hier kommen die Vorteile dieser Herangehensweise zum Vorschein: SMÜ-Systeme werden als robust, weniger Zeit beanspruchend und wartbarer beschrieben, da sie nach der Eingabe der Korpora den Rest der Arbeit vollautomatisch erledigen (vgl. Jekat/Volk 2010³: 647). Nicht zuletzt haben wohl diese Gründe dazu geführt, dass die Kommission sich 2010 dazu entschlossen hat, auf das SMÜ basierte MT@EC umzusteigen und seine Vorzüge für den Übersetzungsdienst zu nutzen.

Bezüglich der Erstellung eines Korpus, der für eine sinnvolle Nutzung umfassend genug ist, hat die EU einen Vorteil, da sie über eine sehr große Anzahl an Texten verfügt, die sie in das SMÜ-System eingeben kann. Dies bezieht sich auf sämtliche Dokumente, die während dem jahrzehntelangen Bestehen der Union entstanden sind und unter diesen vor allem auf Rechtstexte. Ein Großteil von diesen ist auch für Personen außerhalb der EU-Institutionen frei zugänglich, etwa über das *Acquis Communautaire*. Zudem bieten die von professionellen ÜbersetzerInnen eingespeicherten Daten, die im DGT Translation Memory seit Jahren fortlaufend gesammelt werden, einen weiteren ausgezeichneten Ausgangspunkt für ein gutes Training des MT@EC.

Die Grundsteine sind somit gelegt, doch wie sich zeigen sollte, wird eine darüber noch weit hinausgehende Datenmenge benötigt, um eine konsistent gute Qualität gewährleisten zu können. Es handelt sich dabei ebenso um ein Problem, dessen Lösung sich die Kommission im Rahmen ihrer zahlreichen aktuellen Projekte (u.a. ELRC, siehe Kapitel 5.3.3.) annimmt.

2.2.3.3. Das BLEU Verfahren

In dem Kontext der Europäischen Union spielt eine verlässliche Qualität selbstverständlich eine gewichtige Rolle. Für diese Zwecke ist es interessant, einen Blick auf die Evaluierungsmethoden der soeben beschriebenen MÜ-Maschinen zu werfen.

Für die Feststellung der Übersetzungsqualität von maschinellen Systemen wurden mehrere automatische Bewertungen entwickelt. Einerseits wird damit eine subjektive Beurteilung durch Menschen ausgeschlossen und andererseits gestaltet sich der Prozess wesentlich schneller, einfacher und günstiger. Die Evaluierungssysteme teilen grundsätzlich dasselbe Auswertungsprinzip: Die maschinell erzeugte Übersetzung wird mit mindestens einer Humanübersetzung verglichen. Welche Elemente dieser Vergleich beinhaltet, variiert allerdings von Methode zu Methode und so können sich teilweise auch die Noten, die von den einzelnen Systemen für die Übersetzung vergeben werden, durchaus voneinander unterscheiden.

Eines der am häufigsten angewendeten Evaluierungssysteme ist das von Papini entwickelte BLEU Verfahren (*BiLingial Evaluation Understudy*). Dabei werden aus zwei bis vier Wörtern bestehende Wortfolgen – so genannte N-Gramme – einem Vergleich mit ausgewählten Referenztexten unterzogen. Je mehr Referenzen dem Evaluationssystem zur Verfügung stehen, desto aussagekräftiger ist das gelieferte Urteil und soll somit einer menschlichen Überprüfung gleichkommen (vgl. Jekat/Volk 2010³: 653-654 und Sproat 2009: 238). Stellt BLEU fest, dass eine große Anzahl an Wortfolgen in der SMÜ-Übersetzung mit der hochwertigen Humanübersetzung übereinstimmt, wird der Maschine ein gutes Zeugnis ausgestellt. Das grundlegende Beurteilungsmotto von BLEU lautet somit: Je mehr das Produkt der Maschine dem Produkt des Menschen gleicht, desto besser fällt die vergebene Note aus (vgl. Kit/Tak-ming 2015: 226).

Sämtliche Probleme bei der Evaluation beseitigt BLEU jedoch nicht. So hängt das Endergebnis stark von den ausgewählten „humanen“ Referenztexten ab (Nach welchen Kriterien soll deren Auswahl getroffen werden? Wann gilt eine Referenz als gute Übersetzung?). Ein Kritikpunkt ist ebenso, dass es eine Abweichung von der Humanübersetzung nicht zwangsläufig mit einem Fehler gleichbedeutend sein muss – die von der Maschine generierte Formulierung könnte durchaus den gleichen Stellenwert haben, wie die Wortwahl eines Menschen. Trotz dieser Schwachstellen gehört diese Methode zu den beliebtesten und zurzeit den besten ihrer Art (vgl. Hutter/Jekat 2010³: 676).

2.2.3. CAT Tools

CAT Tools werden oftmals als die bedeutendsten Hilfsmittel für ÜbersetzerInnen bezeichnet (z.B. Drugan 2013: 91) und sind aus dem professionellen Übersetzungsprozess nicht mehr wegzudenken. Sie können als eine Software-Plattform beschrieben werden, die mehrere Funktionen in sich vereint, so etwa mehrsprachige Datenbanken, Textbearbeitung, maschinelle Übersetzung, die Optionen zur Analyse, zur Textextraktion sowie zum Projektmanagement und Übersetzung in einem Editor. Das Kernstück bildet die Translation Memory, das als Gedächtnis der ÜbersetzerInnen bezeichnet werden kann (vgl. Melby/Wright 2015: 662-663).

Im Vergleich zu der MÜ gelten die heute gängigen CAT Tools allerdings als eine Entdeckung der letzten dreißig Jahre, die sich im Kreise der professionellen ÜbersetzerInnen erst vor kurzem etabliert hat. Während die kommerzielle Verwendung dieser Programme in den 90er Jahren begann, lässt sich der Ursprung der CAT Tools in der Geschichte der Übersetzungstechnologien viel weiter zurückverfolgen. Die Vorgänger der heutigen CAT Tool Technologien können nämlich bereits auf die 1960er Jahre zurückgeführt werden, bis hin zu den Vorgängerinnen der heutigen EU, nämlich der Europäischen Gemeinschaft für

Kohle und Stahl. Auch wenn Zweck und Einsatzmöglichkeiten stark von den aktuellen Systemen abwichen, hatte man bereits vor fünf Jahrzehnten den enormen Nutzen der Wiederverwendbarkeit von gespeicherten Humanübersetzungen erkannt. Die Umsetzung der Ideen erfolgte somit Schritt für Schritt und die technologischen Möglichkeiten waren in den 90er Jahren an einem Punkt angelangt, an dem eine uneingeschränkte Realisation schließlich durchgeführt werden konnte (vgl. Reinke 2004: 36 und Hutchins 1998: 287).

Eines der ersten Systeme (das von der EGKS verwendete) wurde im ALPAC Bericht besonders gelobt und die Vorteile der maschinengestützten Humanübersetzung gegenüber der rein maschinellen Übersetzung hervorgehoben. Der Bericht beschrieb es wie folgt:

The translator indicates, by underlining, the words with which he desires help. The entire sentence is then keypunched and fed into a computer. The computer goes through a search routine and prints out the sentence or sentences that most nearly match (in lexical items) the sentences in question. The translator then receives the desired items printed out with their context and in the order in which they occur in the source. (ALPAC 1966: 27)

Während somit die Grundidee der CAT Tools bereits damals vorhanden war, wies das System der EGKS einen deutlichen Unterschied zu den aktuellen CAT Tools der DGT auf. Bereits der Name des Organs, das für die Entwicklung und Instandhaltung verantwortlich war, machte dies deutlich, wurde das System doch von der *Terminological Bureau of the European Coal and Steel Community (CECA)* entwickelt. Die TM diente in diesem Fall somit primär als terminologisches Hilfsmittel, denn die Speicherung der Daten war zunächst als Information zum Kontext bestimmter Termini gedacht (ähnlich den Angaben in einer Terminologiedatenbank) und weniger zur Auffindung bzw. Verarbeitung der alten Übersetzungselemente in neuen Übersetzungen mit ähnlichen Textteilen. Die grundlegende Idee der TM Systeme wurde folglich von dem Bestreben ausgelöst, ÜbersetzerInnen die terminologische Arbeit zu erleichtern und nicht zuletzt gleichzeitig die für aufwändige Recherche notwendige Zeit zu kürzen (vgl. Reinke 2004: 36 und Hutchins 1998: 290).

Diese Basis wurde in den 70er Jahren weiterentwickelt, indem es zu dem Vorschlag kam, ein Programm zu erstellen, das nicht bloß mono- und bilinguale Wörterbucheinträge und Kontextsätze zur leichteren Recherche liefert, sondern miteinander übereinstimmende Segmente gleich als Übersetzungsvorschlag anbietet. Diese würde das Programm aus Archiven heraussuchen, in dem alte Übersetzungen (auf Magnetbändern) gespeichert waren. Für eine erfolgreiche Umsetzung war es aufgrund dem Stand der Technologie jedoch noch zu früh (vgl. Hutchins 1998: 291).

Der Gedanke eines Übersetzungsarchives wurde später auch von Peter Arthern aufgegriffen und weiterentwickelt. Während sich der zuvor erläuterte Vorschlag für solch einen Speicher bloß auf jene Sätze bezog, die eine 100%ige Deckung aufweisen, ging Arthern einen Schritt weiter, denn das Programm sollte auch bei bloß ähnlichen Sätzen reagieren. Wie ähnlich sich Sätze sein müssen, um von dem Computer einen Vorschlag aus den gespeicherten

Übersetzungen zu erhalten, erläuterte Arthern damals jedoch noch nicht genau. Er betonte allerdings die zahlreichen Vorteile dieser Vorgehensweise und sah die Anwendung vor allem für die ÜbersetzerInnen der Europäischen Kommission als überaus ökonomisch, da er aus eigener Erfahrung wusste, dass Texte vielfach Passagen enthielten, die sich wiederholten und bereits in vorangehenden Dokumenten vorhanden waren. Auch Artherns Idee musste noch rund ein Jahrzehnt ruhen, bevor die Technologie weit genug war, um ein derartiges TM System auch in Wirklichkeit ausprobieren zu können. Aus heutiger Sicht kann jedoch gesagt werden, dass er anhand der EU-Apparate die richtigen Schlussfolgerungen getroffen hat (vgl. Reinke 2004: 37).

An dieser Stelle ist auch Alan Melby zu erwähnen, der 1981 seine Vorstellungen von einem zweisprachigen Konkordanzprogramm beschrieb. Er sah dieses als ein Hilfsmittel mit viel Potenzial nicht nur während dem Übersetzungsvorgang an sich, sondern auch zur Sicherstellung einer konsistent verwendeten Terminologie.

Das System baute bereits auf der Idee der Segmentierung auf. Zu den einzelnen Satzsegmenten sollten mögliche passende Elemente der Zielsprache im jeweils relevanten, passenden Kontext bereitgestellt werden (vgl. Hutchins 1998: 196).

[T]he source and translated texts are divided into units...‘translation segments’. For a given translation segment in the source text, the corresponding target language translation segment contains the translation of everything in the source segment, but there need be no decision as to the translation of a particular word. (Melby 1981: 545)

Interessant ist zudem, dass die Unterteilung des Textes in die einzelnen Segmente noch von der übersetzenden Person manuell vorgenommen werden musste. Während das Prinzip des Konkordanzprogrammes an sich nicht ganz neu war, war Melby dennoch der erste, der es als Werkzeug für ÜbersetzerInnen gestaltete (vgl. Reinke 2004: 38 und Hutchins 1998: 196).

Diese Überlegungen bildeten die Grundsteine für die derzeitigen Technologien, aber bis zu dem Zeitpunkt der tatsächlichen Umsetzung im Sinne eines CAT Tools wie etwa SDL Trados sollte es noch an die zehn Jahre dauern (vgl. Reinke 2004: 38). Ende der 80er Jahre konnte mit dem Erscheinen der ersten privaten Rechner jedoch ein großer Fortschritt verzeichnet werden. Zuvor konnte auf derartige Hilfsmittel nur über Netzwerke zugegriffen werden, die in diesem Umfang aufgrund des Aufwandes, der Menge an Daten und der Kosten hauptsächlich nur in größeren Unternehmen und Organisationen vorhanden waren. Die PCs ermöglichten es den ÜbersetzerInnen nun schrittweise, derartige TM Datenbanken auch persönlich anzulegen, sie auf eigenen Geräten lokal zu verwalten und mit Textverarbeitungsprogrammen und Druckern zu arbeiten (vgl. Hutchins 1998: 197-198). Die Forschung und Entwicklung der Tools wurde weiter fortgesetzt und ebenso von der EU unterstützt, etwa 1995 mit der Förderung des internationalen Translator’s Workbench Projektes (siehe Kapitel 4.2.5.).

Die heute zum Einsatz kommenden CAT Tools sind bereits aus mehreren Bausteinen zusammengesetzt und ermöglichen es ÜbersetzerInnen jeglicher Art (ob nun bei der DGT, in einem Übersetzungsbüro oder in einem freiberuflichen Umfeld) komplexe Aufgaben zu erledigen. Die Bearbeitung von einer Vielzahl an Formaten, die Angabe von Metadaten, benutzerdefinierte Segmentierung und Planer zur Auftragsabwicklung (Übersicht über Verlauf der Übersetzung, Analyseberichte über Wort- und Zeichenanzahl etc.) gehören mittlerweile zum Standard.

Eine Hilfestellung bieten auch Angaben zu den Übersetzungsvorschlägen. Je nach Wunsch können ÜbersetzerInnen eine Hemmschwelle einstellen: Sind in einem zu übersetzenden Dokument Segmente vorhanden, die bereits übersetzten Segmenten ähnlich oder gar ganz gleich sind, bietet das CAT Tool Vorschläge aus seinem Speicher an. Wie hoch die Ähnlichkeit sein muss (sollen nur Übereinstimmungen mit 100% angezeigt werden oder sind diese auch mit 60% noch sinnvoll?), ist dem Ermessen der ÜbersetzerInnen überlassen. Das Tool zeigt bei seinen Vorschlägen jeweils den Prozentsatz an und arbeitet somit nach dem Match Prinzip. Die meisten CAT Tool Systeme verwenden zumindest drei Match Varianten, nämlich „*exact*, *fuzzy* und *term matches*“ (Bowker 2008: 96). Wie die Bezeichnungen schon andeuten, handelt es sich bei *exact matches* um Segmente, die miteinander vollkommen übereinstimmen; bei *fuzzy matches* um Satzteile, die sich mit den neuen Segmenten teilweise überschneiden und bei *term matches* um bereits bekannte Termini, die das Tool aus der meist integrierten Terminologiedatenbank hervorruft und anzeigt (vgl. Bowker 2008: 96-101).

Dieser Überblick über einige der für die Übersetzung wesentlichen Sprachtechnologien hat bereits verdeutlicht, wie vielfältig sich die Einsatzbereiche der Sprachtechnologien gestalten. Von Relevanz waren in der Geschichte der EU-ÜbersetzerInnen auch Spracherkennungsprogramme. Die herkömmlichen Diktiergeräte der 50er und 60er entwickelten sich mit der Zeit zu komplexeren Softwareprogrammen, die in dieser neuen Form auch heute im Einsatz sind. So können Spracherkennungssoftwares eingesetzt werden, mit derer Hilfe ein fließendes Diktieren möglich ist. Dies bedeutet einerseits eine enorme Zeitersparnis und andererseits eine geringere gesundheitliche Belastung (besonders für Hand und Handgelenke). Spracherkennung ist jedoch nur in einigen wenigen Amtssprachen vorhanden, weshalb die Nutzung dieses Hilfsmittels derzeit noch recht begrenzt ist (vgl. Europäische Kommission 2013: 13).

Zudem haben sich in den letzten Jahren Werkzeuge für den Bereich des Projektmanagements als nützlich erwiesen. Im Falle der riesigen Mengen an Übersetzungsaufträgen verwaltenden DGT sind sie nicht bloß nützlich, sondern mittlerweile notwendig geworden (derzeit: Suivi, Poetry). Dies bringt mit sich, dass Tools für Terminologie- und Übersetzungsmanagement ebenso eine tragende Rolle spielen und die Arbeit maßgeblich prägen.

Zurzeit laufen in diesen Bereichen jedoch keine umfangreichen und staatenübergreifenden Projekte, weshalb auf diese in der vorliegenden Masterarbeit nicht genauer eingegangen werden kann.

3. Europäische Kommission

Die Geschichte der EU ist gleichzeitig auch die Geschichte der Amtssprachen. Mit dem Anwachsen derer Anzahl änderten sich die Anforderungen an die Werkzeuge für die Übersetzung und die technischen Möglichkeiten der Umsetzung. Aus diesem Grund werden in diesem Kapitel die wesentlichsten Meilensteine und Züge der EU und der Kommission aufgezeigt.

3.1. Überblick Geschichte der Europäischen Union

Der Grundstein der heutigen Europäischen Union wurde 1950 gelegt und war eine Konsequenz der zwei Weltkriege, die zuvor ganz Europa erschüttert hatten. Mit einer gemeinsamen Union – einem vereinten Europa – sollten in der Zukunft Kriege verhindert, ein dauerhafter Frieden gesichert und durch enge Beziehungen ein wirtschaftlicher Neuanfang vorangetrieben werden. Um dieses Ziel zu erreichen, strebten die führenden Nationen eine wirtschaftliche und politische Einigung der Staaten an. Unter ihnen war Robert Schuman, zu jener Zeit der französische Außenminister, der 1950 im Rahmen einer Rede schließlich von der Gründung einer westeuropäischen Kohle- und Stahlindustrie sprach. Zwei Jahre später wurde sein Plan in die Tat umgesetzt und die *Europäische Gemeinschaft für Kohle und Stahl* (EGKS) mit sechs Mitgliedern gegründet: Frankreich, Deutschland, Belgien, die Niederlande, Luxemburg und Italien. In dieser Phase entstand auch die Vorgängerin der Europäischen Kommission in Form der Hohen Behörde. Die Behörde agierte als ein unabhängiges Gremium, in dem alle damaligen Mitgliedstaaten vertreten waren. Sie hielt die Befugnis inne, Regelungen bezüglich der Stahl- und Kohleindustrie der sechs Mitgliedsländer aufzustellen und über den neu erschaffenen gemeinsamen Markt eine gewisse Kontrolle auszuüben (vgl. Lengauer 2012: 11, 13 und <http://www.zukunfteuropa.at>).

In dieser Form wurde mit der EGKS die erste supranationale Organisation geschaffen. Neben der Kommission zeichneten sich in der neu gegründeten Organisation auch die anderen Organe der heutigen Union ab: neben einem Gerichts- und Rechnungshof, waren bereits ein Besonderer Ministerrat (Vorläufer des Rates der EU) und eine Gemeinsame Versammlung (Vorläufer des Europäischen Parlamentes) vorhanden (vgl. Schmale 2000: 236).

Der Beginn der Union war somit zunächst stark auf einen wirtschaftlichen Bund ausgerichtet, was sich im Rahmen der Römischen Verträge erneut zeigen sollte. Die im Jahre 1957

unterzeichneten Dokumente führten neben die EGKS die *Europäische Wirtschaftsgemeinschaft* (EWG) ein und begründeten gleichzeitig die *Europäische Atomgemeinschaft* (EURATOM). In den ersten Jahren ihres Bestehens hatte sich die EGKS als erfolgreich herausgestellt und mit der Einführung der EWG sollten deshalb die von der Organisation umfassten wirtschaftlichen Einflussbereiche erweitert werden. Über die Kohle- und Stahlindustrie hinaus stand nun die Errichtung eines einheitlichen Binnenmarktes im Mittelpunkt, der durch geringere Zölle und eine vorteilhaftere Handelspolitik gefördert wurde. Zu diesem Zeitpunkt wählt die Organisation Brüssel als Hauptquartier ihrer Arbeit (vgl. Lengauer 2012: 13 und <http://www.zukunfteuropa.at>). Mit der Gründung der EWG ging auch ein Aufruf an die restlichen (west)europäischen Staaten einher, sich der Vereinigung anzuschließen – damit war allen voran Großbritannien gemeint, das sich jedoch erst 1973 zu einem Beitritt entschließen würde (vgl. Schmale 2000: 239).

Acht Jahre später erfolgte mit der Unterzeichnung der so genannten Fusionsverträge der nächste Schritt in Richtung Europäische Union. Die Arbeitsweise der Organisation erwies sich mit den drei europäischen Gemeinschaften als unvorteilhaft, weshalb eine Reform beschlossen wurde. Im Zuge einer Fusion im Jahre 1967 wurden die EGKS, EWG und EURATOM nun zu der *Europäischen Gemeinschaft* (EG) vereint und erhielten gemeinsame Organe. Mit der Zentralisierung geht auch der neue Status der Kommission einher: sie hat nun die Aufgaben der Hohen Behörde inne und ist für alle drei Gemeinschaften zuständig (vgl. Lengauer 2012: 14 und <http://www.zukunfteuropa.at>).

1979 kam es mit der Errichtung des *Europäischen Währungssystems* (EWS) zu einem nächsten Meilenstein in der Geschichte der Europäischen Union. Der EWS bedeutete die Einführung des *European Currency Unit* (ECU), einer europäischen Währungseinheit (vgl. Bergmann 2015: 1020-1021 und <http://www.zukunfteuropa.at>).

Der nächste Vertrag mit tragender Bedeutung ist jener von Maastricht aus dem Jahre 1992. Aus den Europäischen Gemeinschaften entsteht mit dessen Unterzeichnung die *Europäische Union*. Dies bringt viele Veränderungen mit sich, mitunter die UnionsbürgerInnen-schaft sowie eine gemeinsame Außen- und Sicherheitspolitik. Weitere wichtige Ereignisse der 90er Jahren betreffen das Inkrafttreten des Schengener Abkommens, die unionsrechtliche Festlegung der vier Grundfreiheiten des freien Waren-, Dienstleistung-, Personen- und Kapitalverkehrs und die Gründung der *Europäischen Zentralbank* (vgl. Lengauer 2012: 81 und <http://www.zukunfteuropa.at>).

Erst 1990 wurde mit der Schaffung einer *Wirtschafts- und Währungsunion* (WWU) der Weg in Richtung einer gemeinsamen europäischen Währung eingeleitet. 2002 tauschten die ersten zwölf Mitgliedsstaaten ihre nationalen Währungen auf den Euro aus, darunter auch Österreich. Mit der WWU wird weiterhin das Ziel verfolgt, die Wirtschaftspolitiken der

einzelnen Mitgliedsstaaten so gut wie möglich aufeinander abzustimmen (vgl. Lengauer 2012: 78-79).

2002 traten die Römischen Verträge außer Kraft und dies bedeutete gleichzeitig das offizielle Ende der EGKS. Nur sieben Jahre später kam es mit dem Inkrafttreten des Vertrages von Lissabon zu der Auflösung der EWG – ihre Aufgaben sind damit vollständig auf die Europäische Union übergegangen.

In ihrer derzeitigen Form kann die EU als eine Besonderheit bezeichnet werden: Als supranationale Organisation agiert sie weder als Bundesstaat, noch als Konföderation (wobei es vereinzelt immer wieder Bestreben nach „Vereinigten Staaten von Europa“ wie in den USA gab und teilweise noch immer gibt), mittlerweile wird sie meist als Staatenverbund bezeichnet (vgl. Schmale 2000: 245).

Aktuelle Themen und Krisen prägen die Zusammenarbeit der Staaten und das Geschehen in der Europäischen Union. In Hinblick auf diese sind neue Partnerschaften, Verträge und die Notwendigkeit von Reformen auch in Zukunft sehr wahrscheinlich. Schließlich gehören auch Verhandlungen um neue Mitgliedschaften und in manchen Fällen Austritte ebenso zu einem der stets wiederkehrenden Themen – die Union wird ohne Zweifel auch weiterhin von Veränderungen geprägt sein.

Im Laufe der europäischen Geschichte hat sich die Zusammensetzung der Mitgliedstaaten im Zuge mehrerer Beitrittswellen geändert: von der ursprüngliche Zahl von sechs Gründerstaaten ist sie auf mittlerweile 28 angewachsen.



Abbildung 3. Verlauf der Beitritte und mögliche Kandidaten (Bergmann 2015: 17)

Die Abbildung stellt einerseits die politische Entwicklung Europas dar und gibt andererseits auch Aufschluss über die sprachlichen Veränderungen – mit der Erweiterung der Mitgliedstaaten, kam es auch zu einer Erweiterung der Amtssprachen und mit dem Abschluss der Verträge wiederum zu neuen Sprachregelungen.

Die Aufgabe diese Sprachen zu koordinieren und die europäische Sprachenpolitik umzusetzen, stellte die Union und im Besonderen den Übersetzungsdienst der Kommission mehrmals vor große Herausforderungen. Zuletzt war dies etwa 2004 der Fall, als insgesamt zehn neue Mitgliedstaaten beitraten (so genannte Osterweiterung). Der Beitritt eines neuen Landes bedeutet in der Regel die Einführung einer neuen Amtssprache (abgesehen von einigen wenigen Ausnahmen wie dem Beitritt Österreichs und zunächst Irlands oder der Wiedervereinigung Deutschlands). Eine neue Amtssprache wiederum bedeutet u.a. einen immensen organisatorischen Aufwand, Umstrukturierungen in den Direktionen sowie die

Aufgabe, neues Personal für die „Beitrittssprachen“ zu finden und Übersetzungstools vorzubereiten, damit etwa litauische ÜbersetzerInnen genauso gut arbeiten können wie bis zu jenem Zeitpunkt italienische ÜbersetzerInnen (Europäische Kommission 2009: 20, 36).

Welche Rolle die Kommission im Detail einnimmt, welche Prinzipien und Aspekte die Sprachenpolitik umfasst, was diese für die DGT und das Übersetzen bedeuten und welchen Einfluss auf die Bewältigung der Herausforderungen Sprachtechnologien haben können, soll in den nächsten Abschnitten erläutert werden.

3.2. Die Kommission

In diesem Abschnitt soll die Kommission samt ihrem Aufbau und ihren Aufgaben skizziert werden. Die Arbeit und die Funktionsweise dieses Organs hat eine enorme Bedeutung für die *Generaldirektion Übersetzen* (*Directorate General of Translation*, DGT), da die Kommission ihre primäre Auftraggeberin und somit eng mit ihr verbunden ist. Von den einzelnen Themen, die in den Sitzungen der Kommission aufgegriffen werden oder den Rechtstexten, die auf europäischer Ebene erlassen werden bis hin zu den Kommissionsprojekten, in denen Übersetzung bzw. Übersetzungstechnologien involviert sind – all diese Aspekte üben Einfluss auf das Aufgabenprofil der DGT aus. Die Tätigkeitsfelder und die Befugnisse der Kommission bestimmen unmittelbar die Art der Texte, die zur Übersetzung an die DGT weitergeleitet werden und so ergeben sich aus diesen Umständen auch die besonderen Ansprüche des Übersetzungsdienstes bezüglich der technologischen Hilfsmittel, die für die Übersetzung der Kommissionstexte benötigt werden.

Die Hauptaufgabe der Europäischen Kommission ist es, die „allgemeinen Interessen der Union“ zu fördern und dementsprechend „geeignete Initiativen“ zu ergreifen (Artikel 17 Absatz 1 EUV). Diese zwei Aspekte stellen wichtige Unterscheidungsmerkmale gegenüber den anderen EU-Institutionen dar: die Kommission sollte ihrem Zweck nach allein das Gesamtinteresse der Union verfolgen – die Interessen der einzelnen Mitgliedstaaten stehen nicht im Mittelpunkt.

Das im Vertrag über die Europäische Union festgelegte Initiativrecht der Kommission besagt, dass sie über das exklusive Recht verfügt, Gesetzesanträge zu stellen, die sodann dem Rat und dem Parlament vorgelegt, unter Umständen abgeändert und von diesen beschlossen werden. Dieses Recht beruht auf der Grundidee, dass die Kommission aufgrund der ihr zugewiesenen Aufgaben die europäischen Gesetze und damit die europäische Staatengemeinschaft am besten lenken kann (vgl. Hartmann 2009: 55-56).

Der Rat und das Parlament sind aus dem Vorgang der Gesetzgebung nicht vollkommen ausgeschlossen, denn beiden Organen steht außerdem die Möglichkeit zu, die Kommission zu einer Initiative aufzufordern – ohne diesem Schritt ist es den anderen Organen nicht

möglich, Beschlüsse zu fassen. In speziell geregelten Ausnahmefällen können auch Institutionen wie die Europäische Zentralbank einen Rechtsakt fordern. Mit der Umsetzung der verabschiedeten Gesetze können grundsätzlich der Rat, die Kommission oder die Mitgliedstaaten selbst beauftragt werden (vgl. Hartmann 2009: 57 und <http://www.europarl.europa.eu>).

Artikel 17 EUV enthält weitere grundlegende Bestimmungen, die die Kommission betreffen, so auch konkrete Angaben zu den Aufgabenbereichen:

[...] Sie sorgt für die Anwendung der Verträge sowie der von den Organen kraft der Verträge erlassenen Maßnahmen. Sie überwacht die Anwendung des Unionsrechts unter der Kontrolle des Gerichtshofs der Europäischen Union. Sie führt den Haushaltsplan aus und verwaltet die Programme. Sie übt nach Maßgabe der Verträge Koordinierungs-, Exekutiv- und Verwaltungsfunktionen aus. Außer in der Gemeinsamen Außen- und Sicherheitspolitik und den übrigen in den Verträgen vorgesehenen Fällen nimmt sie die Vertretung der Union nach außen wahr. Sie leitet die jährliche und die mehrjährige Programmplanung der Union mit dem Ziel ein, interinstitutionelle Vereinbarungen zu erreichen (Art. 17 Abs. 1 EUV).

Die Kommission verfügt somit über ein weiteres Spektrum an Rechten und Pflichten, die sie zu einem einflussreichen Organ machen. Dieser Umstand führt dazu, dass die Kommission als eine Art mitregierendes Organ der EU beschrieben werden kann. Die alleinige „Regierung“ ist ihr jedoch nicht übertragen, denn – wie bereits angesprochen – ist sie in vielen Fällen auf die Zusammenarbeit mit dem Parlament, dem Europäischen Rat und den Mitgliedstaaten angewiesen. Etwaige Verhandlungen, die vor einer Entscheidungsfindung ablaufen und der Einfluss nationaler Regierungen dürfen in der tatsächlichen Praxis jedoch nicht außer Acht gelassen werden (vgl. Hartmann 2009: 58-60).

Das Organ setzt sich derzeit aus 28 Mitgliedern zusammen, d.h. jedes der aktuell 28 Mitgliedstaaten entsendet eine Person nach Brüssel (diese agieren vor Ort laut der Gründungsverträge dennoch nicht als nationale VertreterInnen). Unter den 28 Mitgliedern befindet sich auch der bzw. die PräsidentIn der Kommission, der bzw. die von dem Parlament auf Vorschlag des Europäischen Rates ernannt wird (vgl. <http://ec.europa.eu>^b).

Erst darauf folgt die Ernennung und Entsendung der restlichen Mitglieder, indem der bzw. die neue PräsidentIn gemeinsam mit den Oberhäuptern der Regierungen und dem Europäischen Rat die KommissarInnen auswählt. Entsprechend dem Prinzip, dass die Kommission frei von mitgliedstaatlichen Interessen handelt, müssen sich folglich alle Mitglieder des Kollegiums zu Unabhängigkeit verpflichten (vgl. Lengauer 2012: 32).

In dieser Form stellt sich die Kommission als Kollegium dem Europäischen Parlament, das nun darüber abstimmt, ob die Kommission in dieser Zusammensetzung ernannt werden soll. Um eine Entscheidung treffen zu können, finden Anhörungen der KommissarInnen statt, in Zuge derer die wichtigsten Punkte ihrer bevorstehenden Amtszeit im Mittelpunkt

stehen. Nach erfolgreicher Abstimmung wird die Kommission vom Rat offiziell für eine Amtszeit von fünf Jahren ernannt (vgl. Hartmann 2009: 62 und Lengauer 2012: 32-33).

Der bzw. die PräsidentIn verfügt über die Kompetenz die Leitlinien der Kommissionsarbeit festzulegen und hat darüber hinaus einen großen Einfluss auf die interne Organisation, da es zu seinen bzw. ihren Aufgaben gehört, die Ressortzuständigkeit unter den restlichen KommissarInnen zuzuweisen (vgl. Lengauer 2012: 32 und Hartmann 2009: 71).

Der bzw. die Vizepräsident/in der Kommission wird von dem Europäischen Rat und unter der Zustimmung des bzw. der Kommissionspräsidenten/in gewählt. Zugleich hat sie eine zweite wesentliche Position inne, da sie bzw. er damit automatisch zusätzlich zum Hohen Vertreter für Außen- und Sicherheitspolitik wird und den Vorsitz für Auswärtigen Angelegenheiten innehat (vgl. Pollak/Slominski 2012: 91-92).

Die Kommission hat ihren Hauptsitz in Brüssel, sie arbeitet teilweise aber auch in Luxemburg. An diesen Orten halten die Mitglieder der Kommission üblicherweise wöchentlich Sitzungen ab, können aber bei wichtigen Angelegenheiten auch zu Dringlichkeitssitzungen zusammenkommen. Ihre Entscheidungen trifft die Kommission als Kollegium und benötigt für einen Beschluss eine einfache Mehrheit – derartige förmliche Abstimmungen gehören jedoch eher zur Seltenheit (vgl. <http://ec.europa.eu>^c und Lengauer 2012: 32).

Ähnlich einer nationalen Regierung sind die Aufgaben unter den Kommissionsmitgliedern nach einzelnen Ressorts aufgeteilt und können somit mit Ministerien verglichen werden. Zurzeit gibt es mehr als 30 Generaldirektionen und Dienststellen, die für jeweils unterschiedliche Bereiche zuständig sind. Aus organisatorischer Hinsicht werden die Direktionen in zwei Unterbereiche eingeteilt: Politikfelder (*policies*) und Außenbeziehungen (*external relations*), während interne Dienste (*internal services*) sowie allgemeine Dienste (*general services*) zu den Dienststellen zählen (vgl. Bergmann 2015: 456-457). Zu den von den GDs abgedeckten Themengebieten gehören etwa Bereiche wie Gesundheit, Bildung und Kultur, Justiz und Verbraucher oder Umwelt. Sprachbezogene Generaldirektionen sind die SCIC (Dolmetschen) sowie die Generaldirektion Übersetzen und sind organisatorisch gesehen Teil der Politikfelder (vgl. <http://ec.europa.eu>^a).

Die Generaldirektionen werden jeweils von einem/r GeneraldirektorIn geleitet, die den Weisungen ihres/r Fachkommissar/in folgen muss und somit direkt der Kommission unterstellt sind. Einigen GDs kommt in der politischen Praxis meist größere Bedeutung zu als anderen. An dieser Stelle sind vor allem „die Generaldirektionen für Außenbeziehung, Handel, Wirtschaft und Finanzen, Industrie und Unternehmen, Wettbewerb, Binnenmarkt und Landwirtschaft zu nennen“ (Hartmann 2009: 70-71).

Im Rahmen ihrer politischen Tätigkeit setzt sich die Kommission mit einer Vielfalt an Themen auseinander. Betrachtet man einige wichtige Konferenzen und Verhandlungen der Kommission, so umfasst das Spektrum Bereiche die von Urheberrecht, Menschenrechten

bis hin zu Hygienebestimmungen und handelsrechtlichen Fragen reichen. Sie tritt in diesem Sinne nicht zuletzt als Kontrollinstanz auf. Die Koordinierung des europäischen Binnenmarktes ist dabei stets ein Schwerpunkt. Dazu gehört es, dem Rat Empfehlungen über Leitlinien der Wirtschaftspolitik zu geben, die vom Rat ergänzt und bearbeitet werden, um so die Koordinierung der EU Wirtschaftspolitik sicherzustellen (vgl. Hartmann 2009: 81-82).

In vielen Fällen ist eine enge Zusammenarbeit mit dem Rat und den einzelnen Regierungen notwendig. Da die Kommission häufig in ihrer Rolle als europäische Regulierungsbehörde auftritt (Abgasquote, Verbraucher, Lebensmittel etc.), muss sie sich mit den nationalen Behörden der Mitgliedstaaten absprechen. Denn während die Kommission die gesetzlichen Rahmenbedingungen vorgibt, sind die nationalen Staaten für die Umsetzung meist selbst verantwortlich und so müssen die Ziele regelmäßig abgesteckt werden. Die Umsetzung neuer Projekte jeglicher Art kann zudem meist nur mit der Unterstützung der einzelnen Mitgliedstaaten stattfinden (vgl. Hartmann 2009: 83).

3.2.1. Ordentliche Gesetzgebung

Einen Großteil der Übersetzungsarbeit machen Texte aus, die das rechtliche Geschehen in der Union betreffen. Nicht nur die fertigen, verabschiedeten Gesetzestexte müssen übersetzt werden, sondern auch viele der Dokumente, die im Vorfeld eingereicht, abgeändert und diskutiert werden.

Das ordentliche Gesetzgebungsverfahren der Union wird seit dem Vertrag von Lissabon als Mitentscheidungsverfahren bezeichnet und kommt auf dem Weg zur Entstehung neuer europäischer Rechtsakte am häufigsten in Anwendung. In dem Verfahren nimmt das Europäische Parlament die führende Rolle ein, doch als mit dem Initiativrecht ausgestattetes Organ hält die Kommission einen wichtigen Teil inne. In dem Falle, dass die Kommission von ihrem Initiativrecht Gebrauch macht, hat sie ihren schriftlichen Gesetzesvorschlag gleichzeitig an den Rat, das Parlament und die nationalen Parlamente zu übermitteln. Zur Erstellung des Vorschlages wird mit den unterschiedlichen Kommissionsdienststellen konsultiert und mitunter werden Sachverständige oder internationale Organisationen herangezogen, um möglichst alle Aspekte abzudecken. Hat sich das Kollegium nach den Beratungen auf einen fertigen Entwurf geeinigt, der an Rat und Co. weitergeschickt werden soll, wird der Vorschlag meist in allen Amtssprachen im Amtsblatt der Europäischen Union veröffentlicht. In der siebten Legislaturperiode (sprich 2009-2014) machte die Kommission insgesamt 584 Mal von ihrem Initiativrecht Gebrauch (vgl. <http://www.europarl.europa.eu>).

Das Amtsblatt gibt einen Einblick in die Arbeit der DGT ÜbersetzerInnen, da dieses zu einem Großteil aus den Übersetzungen der Generaldirektion besteht. Die aktualisierte

Version wird jeden Werktag veröffentlicht und ist über EUR-Lex der Öffentlichkeit zugänglich. Damit soll nicht zuletzt das Prinzip der Demokratie und der freie Zugang zum Recht für alle BürgerInnen gewährleistet werden (vgl. Europäische Kommission 2009: 52)

Wird nicht die Kommission selbst tätig und kommt der Anstoß für ein neues Gesetz auch nicht von Rat und Parlament, besteht seit 2012 europaweit die Möglichkeit einer BürgerInneninitiative. Damit EU-BürgerInnen diese einleiten können, benötigen sie eine Millionen Unterschriften, die innerhalb eines Jahres und mindestens in sieben Mitgliedstaaten gesammelt werden. Weiters muss die Initiative ein Thema betreffen, dass innerhalb der Kommissionskompetenzen angesiedelt ist – aber da das Spektrum der Kommissionszuständigkeit recht weit ist (dazu gehören insgesamt 85 Politikbereiche), bestehen diesbezüglich eigentlich nur minimale Einschränkungen. Sind diese formalen Kriterien erfüllt, muss sich die Kommission mit der BürgerInneninitiative zumindest ernsthaft befassen und dazu eine Entscheidung abgeben. Diese Stellungnahme muss in allen Amtssprachen veröffentlicht werden (vgl. <http://www.europarl.europa.eu> und <http://ec.europa.eu>^d).

Je nachdem, ob ein Gesetzesvorschlag von den einzelnen Institutionen und den nationalen Parlamenten positiv angenommen wird oder abgelehnt wird, finden mehrere Sitzungen und Lesungen des Textes statt. Die Kommission muss den Text mitunter mehrere Male prüfen und abändern, bis dann im Idealfall eine Einigung und letztendlich ein neuer Gesetzesakt entsteht. Die DGT ist dabei in jeder Phase der Gesetzgebung mit eingebunden (vgl. <http://www.europarl.europa.eu>).

Dieser Abschnitt illustriert, dass Kommunikation und somit Übersetzung auf vielen verschiedenen Ebenen abläuft und das Feld an möglichen RezipientInnen weit ist. Die von der DGT erstellten Dokumente dienen der internen Kommunikation zwischen EU-PolitikerInnen und BeamtInnen als auch zwischen nationalen Behörden und EU-Behörden. Sie dienen der Verständigung zwischen der Union als Einheit und der breiten Öffentlichkeit, genauso wie auch der Verständigung zwischen einzelnen BürgerInnen und ausgewählten EU-Organisationen.

Aus der Vielfalt der Themen und der Kommunikationssituationen heraus ergibt sich eine Vielfalt an Textsorten und Funktionen, mit denen die ÜbersetzerInnen konfrontiert sind. Die jeweilige Textsorte übt für die Herangehensweise an die Übersetzung und den Einsatz von Sprachtechnologien einen Einfluss aus. Wie die Darstellungen des vorangehenden Kapitels gezeigt haben, trägt die Art des zu übersetzenden Textes in großem Maße zur Nützlichkeit einzelner Tools bei (siehe auch Mai 2015). Betrachtet man die Aufteilung der Textsorten in Abbildung 4, können beispielsweise Rückschlüsse darauf gezogen werden, mit welchen Texten MÜ-Systeme trainiert werden sollen oder inwieweit ÜbersetzerInnen bei Matches mit hoher Ähnlichkeit rechnen dürfen sowie für welche Domänen Terminologiedatenbanken hilfreich sein werden. Nach Betrachtung der Kommissionsaufgaben ist es

wohl wenig überraschend, dass der Schwerpunkt dabei vor allem auf rechtliche Texte zu setzen ist.

Document types at the European Commission

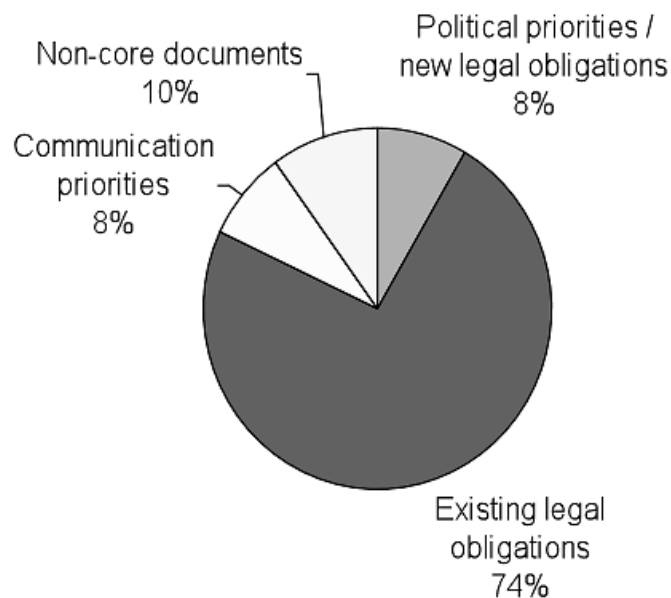


Abbildung 4. Die häufigsten Dokumenttypen in der Europäischen Kommission (<http://www.dfki.de>)

3.2.2. Besonderheiten der EU-Sprache

Eine erfolgreiche Kommunikation ist nicht bloß nach außen hin zwischen der Kommission und den nationalen Regierungen bzw. EU BürgerInnen erforderlich, sondern natürlich auch zwischen den EU-BeamtInnen selbst. Für die alltägliche interne Kommunikation gibt es keine rechtlich verbindlichen Vorschriften, in der Praxis haben sich jedoch drei dominante Sprachen durchgesetzt. Seit 2001 gelten in der Kommission deshalb Deutsch, Englisch und Französisch als Verfahrenssprachen (*procedural languages*), wobei die beiden letzteren häufiger im Gebrauch sind als das Deutsche (vgl. Wagner et al. 2012: 10 und Gerhards 2010: 130).

Für Verhandlungen gilt jedoch stets, dass die Dokumente zumindest in diesen drei Sprachen vorliegen müssen, die Übersetzung in die restlichen 21 Amtssprachen muss „erst“ innerhalb von 24 Stunden vorhanden sein. Seit der Erweiterung der Union um mehr als 15 Mitgliedsstaaten seit den 90er Jahren, ist das Englische noch stärker vertreten – getreu dem Motto „*the more languages, the more English*“ (Gerhards 2010: 134). Das Englisch, das sich dabei entwickelt hat, weist jedoch mittlerweile für die EU typische Merkmale auf. Es ist „durch

ein eigenes Vokabular gekennzeichnet“ und dadurch mitunter auch für MuttersprachlerInnen gewöhnungsbedürftig (Reithofer 2014: 54).

Vereinzelte ist auch noch die traditionelle Verfahrenssprache Französisch anzutreffen, doch ähnlich wie das Englische in einer meist vereinfachten Form, die von den SprecherInnen auch als „Franglais“ eingesetzt wird (Hartmann 2009: 80).

In dieser internationalen Umgebung hat sich der oft zitierte *Eurospeak* entwickelt, der die für die Europäische Union typische Terminologie enthält. Bekannte Beispiele hierfür wären die Begriffe *acquis*, *subsidiarity*, *comitology*, *additionality* oder *democratic deficit* (vgl. <http://ec.europa.eu>¹). Der Ursprung dieser Termini und Formulierungen stammt einerseits aus der vorhin beschriebenen Sprachsituation bzw. internen Kommunikation und andererseits aus dem Bestreben, EU-Recht und nationales Recht eindeutig voneinander zu differenzieren. Die vorgesehene Umsetzung von Verordnungen, Richtlinien etc. soll nicht durch missverständliche Termini gefährdet werden (vgl. Wagner et al. 2012: 64).

Die Eigenheiten des Eurospeak zu beachten, erweist sich bei der Verwendung von Übersetzungstools als unerlässlich. Einerseits gilt dies für die Terminologearbeit, jedoch auch bei Eingaben in Translation Memory Speicher und der Speisung von MÜ-Systemen mit Paralleltexten, in denen der *richtige* Eurospeak enthalten sein sollte, anstatt von Ausdrücken in einem allgemeinsprachlichen Kontext. Der Einfluss der besonderen Sprachgewohnheiten reicht weiter als nur die mündliche Kommunikation: Dieser Umstand hat Auswirkungen auf die Arbeit der ÜbersetzerInnen und kann durchaus die Anwendung der Tools beeinflussen.

Wie soeben angesprochen, ist ein Teil des Eurospeak bzw. der Sprachkultur der Europäischen Institutionen die Herausforderung, dass die meisten BeamtenInnen und PolitikerInnen nicht in ihrer Muttersprache kommunizieren. Die Verfahrenssprachen Englisch und Französisch (primär jedoch Englisch) bilden zu einem hohen Prozentanteil die alltäglichen Kommunikation und zugleich die Ausgangssprache, in der Dokumente jeglicher Art verfasst werden (z.B. sind im Bereich Außenbeziehungen schätzungsweise 90% aller Texte auf Englisch, vgl. Mai 2015). Die beiden Sprachen sind zum Großteil die erste oder zweite Fremdsprache der EU-MitarbeiterInnen, aber unter Umständen auch erst die dritte oder vierte. Es ist unvermeidlich, dass die nicht in der Muttersprache verfassten Texte unklare oder fehlerhafte Stellen enthalten. Dies ergibt sich beispielsweise durch grammatikalische Fehler, umständliche Formulierungen und die Verwendung von so genannten falschen Freunden, falschen Idiomen und Wortfolgen. Die Qualität der Ausgangstexte leidet darunter und dies wird durch einen allzu langen Textumfang zusätzlich verstärkt. Durch Zusammenarbeit mit anderen Sprachabteilungen und durch Kontaktaufnahme mit der/m AutorIn kann Missverständnissen entgegengewirkt werden, doch handelt es sich dabei um eine zeitaufwendige Vorgangsweise (vgl. Wagner et al. 2012: 70-71).

Zudem muss angemerkt werden, dass in vielen Fällen nicht bloß ein/e AutorIn für ein Dokument verantwortlich ist, sondern wie für internationale Organisationen üblich, gleich mehrere VerfasserInnen. Wie das im Überblick geschilderte Gesetzgebungsverfahren zeigt, betrifft dies vor allem Gesetzestexte (im Falle von Berichten, Briefen, Broschüren etc. gestaltet sich die Situation besser): diese enthalten Input von einzelnen Komitees, EU-PolitikerInnen und nationalen Behörden. Ein Umstand, der die Einheitlichkeit des Textes in der Regel nicht unbedingt positiv beeinflusst und auch die Ermittlung der AutorInnen für eventuelle Nachfragen erschwert (vgl. Wagner et al. 2012: 71-72).

Die Kommission hat diese Schwierigkeiten bereits in den 90er Jahren erkannt und im Rahmen des Vertrages von Amsterdam versucht die Lage zu ändern. Ungenügend formulierte Gesetzestexte würden vor allem für die richtige Umsetzung des EU Rechtes Stolpersteine darstellen und müssten durch die Befolgung von Leitlinien vermieden werden. In Erklärung Nr. 39 hält der Vertrag jene Leitlinien fest, die „zur Verbesserung der redaktionellen Qualität der gemeinschaftlichen Rechtsvorschriften“ eingesetzt werden sollten (ABl. Nr. 340 1997). In einer interinstitutionellen Vereinbarung aus dem Jahre 1999 kam es zu der Konkretisierung der Leitlinien (es handelt sich somit nur um Empfehlungen ohne Rechtswirkung). Unter anderem wird in dem Abschnitt festgehalten, dass es das Ziel sei, „die gemeinschaftlichen Rechtsakte klar, einfach und genau“ abzufassen, „die verwendete Terminologie kohärent“ zu gestalten und dass die Texte letztendlich inhaltlich „möglichst homogen sein“ sowie „allzu lange Artikel und Sätze, unnötig komplizierte Formulierungen und der übermäßige Gebrauch von Abkürzungen vermieden werden“ sollten (ABl. Nr. C 073 1999).

Während in den Verträgen die Übersetzungsdienste selbst nicht erwähnt werden, bedeutet die Umsetzung dieser Leitlinien bessere Ausgangstexte mit weniger Rechercheaufwand, weniger Fehlerquellen und indirekt eine erfolgreichere Anwendung von Sprachtechnologien. Zusammenfassend kann festgestellt werden: es gilt der Grundsatz *Garbage In, Garbage Out*. Klare und auf den Punkt gebrachte Texte erleichtern in jeder Hinsicht die Arbeit der DGT, ob nun mit oder ohne Tool.

Die DGT unterstützt dieses Vorhaben durch mehrere Projekte, u.a. durch die Herausgabe von *Clear Writing* Broschüren, Kampagnen, Vorträge für MitarbeiterInnen (vgl. Wagner et al. 2012: 72-73). Zudem hat die DGT ihren Service erweitert und bietet die „redaktionelle Bearbeitung von Originalen“ an, wobei diese Dienstleistung bisher noch wenig genutzt wird – zusammen mit der redaktionellen Bearbeitung von Webtexten machte dieses Angebot 2014 nur ca. 5% der gesamten DGT Arbeit aus (<http://ec.europa.eu>⁸).

Im Zusammenhang mit ihrer Sprachenpolitik trifft die Europäischen Institutionen oftmals die Kritik, dass zu viele Sprachen und aus diesem Grund zu viel Übersetzung und Dolmetschen im Spiel seien. Es wird die Notwendigkeit hinterfragt, jedem Mitgliedsstaat seine eigene Sprache zu gewähren, da damit viel Aufwand, Personal und letztendlich hohe Kosten

einhergehen würden. Folglich wird die Lösung in der Umsetzung von einer Beschränkung der offiziellen EU-Sprachen gesehen. Zudem würde es dann reichen, eine Übersetzung nur in einzelnen Fällen in Auftrag zu geben, was noch dazu auch für die Sprachdienstleistungen weniger Druck bedeuten würde. Übrig bleiben sollten die größeren Sprachen, allen voran Englisch. Dieses Argument wird unter anderem dadurch gestützt, dass die EU von vornherein die Mehrsprachigkeit ihrer BürgerInnen fördert (BürgerInnen sollten neben ihrer Muttersprache noch zumindest zwei andere Sprachen beherrschen; dies wird im Rahmen von Programmen wie *Erasmus* unterstützt; vgl. Gerhards 2010: 147), weshalb es keine besondere Herausforderung darstellen sollte, einen Text in einer Fremdsprache zu lesen.

Der von vielen vorgebrachte Vorschlag, Englisch als umfassende Sprache heranzuziehen (siehe z.B. Gerhards 2010: 213ff), mag zwar aus theoretischen Gesichtspunkten eine Lösung darstellen, beruht jedoch auf der idealisierten Vorstellung, dass die europäische Bevölkerung tatsächlich überall gleichermaßen gute und vor allem ausreichende Englischkenntnisse aufweist. Die Verbreitung von Englisch als *Lingua Franca* besteht mit Sicherheit, aber bei Weitem nicht in dem Ausmaß, dass sie innerhalb des EU-Rahmens die einzelnen Amtssprachen tatsächlich erfolgreich ersetzen könnte (vgl. Reithofer 2014: 53-54). Dies bestätigen auch jene Daten, die im Kontext von Englisch als Sprache des europäischen Handels gesammelt wurden und auch in Kapitel 5.2.2. erläutert werden.

Alleine die Feststellung, dass die Mehrsprachigkeit über einen rechtlichen, politischen, geschichtlichen, kulturellen und symbolischen Grund verfügt, zeigt wie schwierig eine formale Reduktion der Sprachen wäre – eine „Abschaffung“ aller Amtssprachen bis auf das Englische würde auf einen noch größeren Widerstand stoßen. Welche Bedeutung diesem Thema zukommt, verdeutlicht zudem, dass Änderungen in der Sprachpolitik von den Nationalstaaten seit Jahren abgelehnt werden (vgl. Gerhards 2010: 137). Die Durchführung der derzeitigen Sprachtechnologie Projekte zeigt ebenso, dass ein Vorgehen mit einer Handvoll von Sprachen auf europäischer Ebene nicht die erwünschten Ergebnisse und Erfolge auslösen wird. Statt solche Ideen weiterzuverfolgen, setzt die DGT deshalb auf ein sinnvolles, selektives Übersetzen und sucht gemeinsam mit der Kommission mithilfe von Sprachtechnologien nach effizienten Lösungen für die Zukunft.

4. Generaldirektion Übersetzen

Ein Herzstück der Kommissionsarbeit bildet die Generaldirektion Übersetzen, die während ihres Bestehens zwar einige Umwälzungen erfahren hat, jedoch von Anbeginn die Geschichte der EU begleitet hat. Wie stark die Organe und die Übersetzung miteinander verknüpft sind, zeigt sich unter anderem auch daran, dass bei einer Umwälzung der Union eine Veränderung der Sprachdienste meist Hand in Hand ging.

4.1. Geschichte der DGT

Im Jahre 1951 wurde mit der Gründung der Europäischen Gemeinschaft für Kohle und Stahl ein Grundstein für die Europäische Union gelegt und der weitere Werdegang Europas in die Wege geleitet. Während bei diesem historischen Ereignis in der Regel nur die politische und wirtschaftliche Bedeutung hervorgehoben wird, ist ihre sprachliche Bedeutung nicht zu unterschätzen. Die damals gewählte Lösung, um die Verständigung in dem neu gegründeten Bündnis zu ermöglichen, legte den Leitgedanken fest, der auch heute noch mit 24 EU-Amtssprachen gilt: alle Sprachen sind gleichberechtigt. Dieser Grundsatz verfügte damals über einen starken symbolischen Wert – nach dem Krieg sollte Europa nun wieder enger und vor allem friedlich zusammenarbeiten und dazu gehörte auch, dass in der Organisation zumindest offiziell keine Sprache über die andere dominierte. Der zweite wesentliche Grundgedanke, der bis zum heutigen Zeitpunkt gilt, ist die Auffassung, dass es zu den grundlegenden Rechten der BürgerInnen gehört, die von der Gemeinschaft beschlossenen Rechte und Pflichten in ihrer Sprache lesen und nachvollziehen zu können (vgl. Europäische Kommission 2009: 10).

In die Regelungen der EGKS flossen neben den soeben genannten Prinzipien auch die Sprachenpolitik anderer internationaler Organisationen wie UNO, OECD und NATO ein. All diese Überlegungen haben dazu geführt, dass die ersten „europäischen“ Sprachendienste mit insgesamt vier Amtssprachen (Französisch, Deutsch, Niederländisch und Italienisch) ihre Arbeit aufnahmen. Obwohl sich zunächst Großbritannien nicht an dem Bündnis beteiligen wollte, bestand dennoch von Anfang an eine englische Sektion, da das Englische in den wichtigen Industrien und in der Forschung vorherrschend war (vgl. Europäische Kommission 2009: 11-12).

Auch mit einer – aus heutiger Sicht – sehr kleinen Zahl an Amtssprachen wurde jedoch nicht wahllos jedes Dokument übersetzt. Diesbezüglich wurden in einem Protokoll Regeln festgelegt, die sich auch im heutigen DGT Grundsatz des selektiven Übersetzens widerfinden.

Beschlüsse, Empfehlungen und einzelne Stellungnahmen sowie Schreiben an Unternehmen werden in der Sprache derer verfasst, an die sie sich richten. Schreiben an die Gemeinschaftsorgane werden in einer vom Absender frei gewählten Amtssprache der Gemeinschaft verfasst; die Antwort ergeht in derselben Sprache. Die Versammlung regelt die praktischen Fragen des Gebrauchs der Sprachen selbst, wobei es den Delegierten freisteht, welche der vier Amtssprachen der Gemeinschaft sie verwenden. Das Amtsblatt (ABl.) erscheint in vier Sprachfassungen (Europäische Kommission 2009: 12).

Jedes Organ hatte einen eigenen Sprachendienst, der mit professionellen DolmetscherInnen und ÜbersetzerInnen besetzt war. Der Dienst der Hohen Behörde verfügte 1953 über ein überschaubares Personal: in einem Dokument werden für alle fünf Sprachsektionen insgesamt 35 ÜbersetzerInnen verzeichnet (davon waren bloß zwei MitarbeiterInnen zuständig

für Englisch). Dem sei hinzuzufügen, dass bei Bedarf auch damals schon ÜbersetzerInnen von außerhalb herangezogen wurden.

Eine technologische Hilfestellung erfolgte durch angestellte StenographInnen bzw. Schreibkräften, die die von den ÜbersetzerInnen diktierten Texte aufnahmen und schließlich mit der Schreibmaschine abtippten. Im Späteren kamen auch vermehrt (große, aber dennoch praktische) Diktiergeräte zum Einsatz. An der Arbeit waren zudem ein Planungsbüro, das die Abwicklung der Aufträge koordinierte sowie ein eigener Vervielfältigungsdienst beteiligt, der unter anderem für die Reinschrift der fertigen Übersetzungen verantwortlich war (vgl. Europäische Kommission 2009: 12, 17).

Bei diesem relativ kleinem Sprachpersonal sollte es allerdings nicht lange bleiben. Bereits einige Jahre nach der Gründung wuchs die Menge an zu übersetzenden Dokumenten in der Hohen Behörde so stark an, dass sich die von dem Sprachendienst bearbeitete Seitenanzahl in einem Zeitraum von 24 Monaten fast verdoppelt hatte (auf 61.568 Seiten). Diese Anzahl überforderte die ÜbersetzerInnen, weshalb letztendlich bis 1957 zahlreiche Änderungen durchgeführt wurden (u.a. Lieferfristen, Begrenzung der Textlänge, mehr Personal). Zusätzlich wurde eine Rangordnung eingeführt, laut derer gewisse Texte Priorität hatten, z.B. mussten Übersetzungen, die im Rat, im Beratenden Ausschuss und in der Versammlung benötigt wurden, innerhalb von 48 Stunden fertig sein.

Als besonderer Aspekt, der zu der Zeit erstmals in Erwägung gezogen wurde, war die Schaffung einer zentralen Terminologiedatenbank (Eurodicautom), die von einem eigenen Terminologie- und Recherchedienst betreut wurde (vgl. Europäische Kommission 2009: 13-14).

Mit dem Inkrafttreten der Römischen Verträge erfolgte die rechtliche Regelung der Sprachenfrage in der allerersten Verordnung der EWG und EURATOM (zuvor war diese bloß in allgemeinen Protokollen vermerkt gewesen). Damit ging in den 60er Jahren auch eine verstärkte Standardisierung des Übersetzungsvorganges einher, im Zuge dessen weiter differenziert wurde, welche Texte in welcher Form zu übersetzen bzw. zu veröffentlichen sind und auf welchem Wege die Aufträge koordiniert werden (vgl. Europäische Kommission 2009: 15-16).

Das Resultat der Fusionsverträge war nicht nur eine Zentralisierung der Gemeinschaften an sich, sondern auch die Entstehung eines einheitlichen Übersetzungsdienstes der Kommission, der von da an in Brüssel residierte. Mit den neuen Amtssprachen und dem anwachsenden Übersetzungsvolumen erhalten die Sprachdienste eine immer wichtigere Rolle und das Bewusstsein für die Arbeit steigt. Das Übersetzen wurde nun nicht mehr als eine Nebentätigkeit innerhalb der Generaldirektion von Personal und Verwaltung betrachtet. Mit einem Personal von 1000 ÜbersetzerInnen und einem wachsenden Aufgabenbereich, wurde

sie Ende der 1980er Jahre zu einer selbstständigen und unabhängigen Dienststelle umstrukturiert, zunächst mit dem Namen SdT (*service de traduction*). Diese Entwicklungen haben für den Status der ÜbersetzerInnen eine positive Veränderung bewirkt (vgl. Europäische Kommission 2009: 24-25,30).

Die 90er Jahre brachten mit der Unterzeichnung des Maastrichter Vertrages erneut umfassende Änderungen mit sich. Mit dem Übergang der EG zu der EU änderte der SdT unter anderem seine Strategie, anhand derer Aufträge angenommen bzw. an externe ÜbersetzerInnen weitergeleitet wurden. Zudem wurden eine effiziente Nachfragenregelung und ein erweitertes Dienstleistungsangebot erarbeitet. Nun bot der SdT über die klassischen Übersetzungsaufgaben hinaus auch „Zusammenfassungen, mündliche Übersetzungen, Nachbearbeitung maschinell übersetzter Texte usw.“ an (Europäische Kommission 2009: 33).

Ab diesem Zeitpunkt entfaltet auch die Technologie ihren Einfluss: Mit den Fortschritten im Bereich von Computer und EDV übernehmen ÜbersetzerInnen nun vermehrt jene Aufgaben, die in den Jahrzehnten zuvor von Schreibkräften und SekretärInnen erledigt wurden. Dieser Umstand macht es möglich, dass ÜbersetzerInnen viel autonomer arbeiten und langsam auf mehrere technologische Hilfsmittel zurückgreifen können – ÜbersetzerInnen fertigen das Endprodukt nun von Anfang bis zum Ende selbst an (vgl. Europäische Kommission 2009: 33, 38). Bis 1998 erfolgte die Verteilung von PCs in sämtlichen Bereichen der Dienststellen, was eine Flut an Sprachtechnologien ermöglichte. Neben dem Versuch MÜ-Systeme stärker in die Arbeit zu integrieren, wird der Ausbau der elektronischen Datenbanken, Übersetzungsspeicher, Terminologiedatenbanken und Textverarbeitungsprogramme weiterverfolgt (vgl. Europäische Kommission 2009: 37-38).

Die technische Entwicklung bringt nicht bloß für die Arbeitsweise Vorteile, sondern auch in der Administration. Mit der Einführung von Poetry und Suivi werden die Abwicklung der Übersetzungsaufträge und der Zugriff auf diesbezügliche Daten um ein Mehrfaches erleichtert und verkürzt. Die beiden Programme werden zusammen mit SdT Vista (im Späteren DGT Vista) oder etwa TraDesk für die Verwaltung und Dokumentation der Ausgangstexte, Textinformationen, Referenztexte und die Vergabe der einkommenden Aufträge verwendet (vgl. Europäische Kommission 2009: 38). Innerhalb einer kurzen Zeit hat sich der Übersetzungsdienst der Kommission vollkommen verändert.

Die unten folgende Abbildung verdeutlicht, wie viele Möglichkeiten den ÜbersetzerInnen aus der Entwicklung der Sprachtechnologien entstammen. Während manche Hilfsmittel, wie die Spracherkennung in der Praxis weniger als andere benutzt werden, bildet die Schaltfläche ein vollständiges Netzwerk, das für jede Arbeitsmethode die notwendige Technologie bereithält.

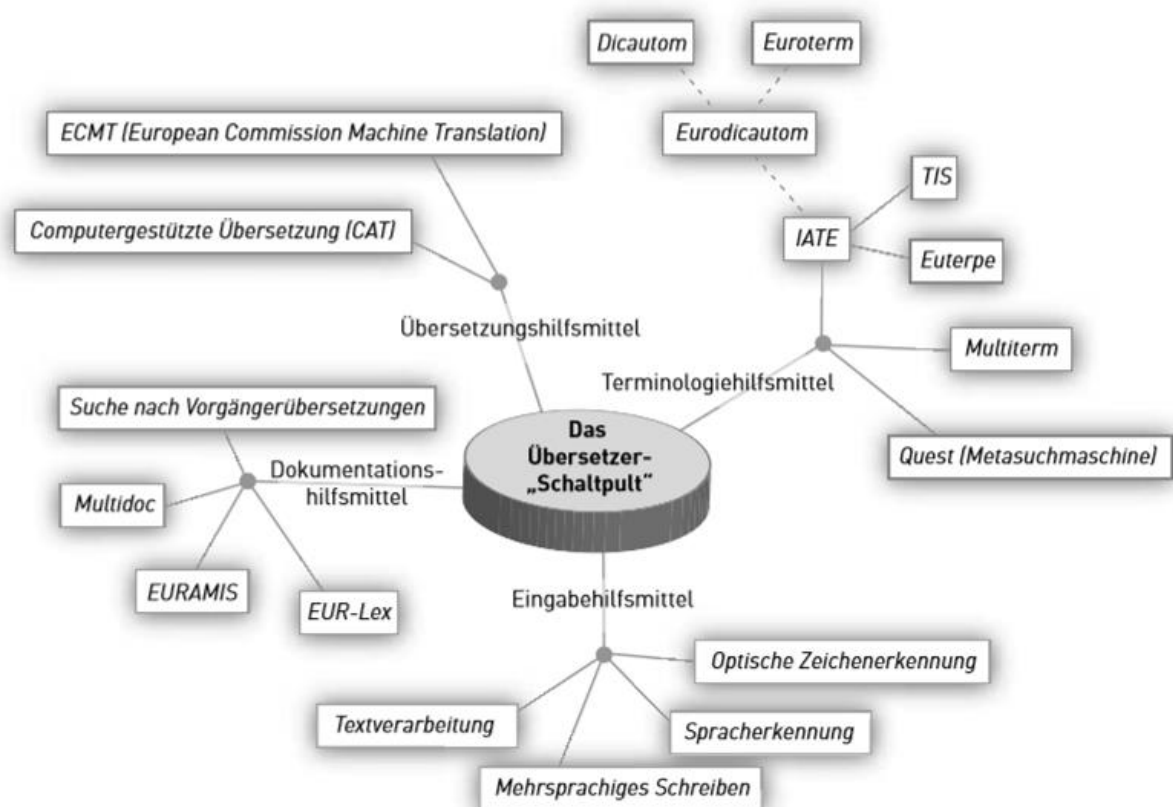


Abbildung 5. Plattform der (sprach)technologischen Hilfsmittel für ÜbersetzerInnen bis ca. 2009 (Europäische Kommission 2009: 37)

Über die nächsten Jahre hinweg ist die Kommission zusammen mit der SdT durchgehend an der Optimierung der Sprachendienste und deren Abläufe involviert. 2002 wird aus dem *Service de Traduction* schließlich die *Generaldirektion Übersetzung* bzw. *Directorate General for Translation*. Nach einer Phase mit einer thematischen Aufteilung der Generaldirektion kehrt man wieder zu den Sprachabteilungen zurück und erhöht den Einsatz von freiberuflichen ÜbersetzerInnen, um ein großes Chaos zu vermeiden (dieses hätten 2004 vor allem die zehn neuen Amtssprachen auslösen können). Der Anfang der 2000er Jahre war somit stark durch Krisenmanagement geprägt. Darüber hinaus wird das Kollegium dazu angehalten, die DGT mehr in den Entstehungsprozess ihrer Dokumente einzubinden, sowie ihre Texte kürzer zu halten (vgl. Europäische Kommission 2009: 43-45).

Die zu übersetzende Menge geht um ca. 30% zurück und der Schreibstil wird von vielen als prägnanter empfunden – aufgrund der Begrenzungen enthalten Dokumente nun nur mehr das Wesentliche. Im Allgemeinen legt die Kommission Wert darauf, dass die Qualität der Originaltexte verbessert wird. In Diesem Sinn erweiterte die DGT ihre Aufgaben und bot „die sprachliche und stilistische Überarbeitung“ von englisch- und französischsprachigen Ausgangstexten an (Europäische Kommission 2009: 45-48, 50).

Dadurch, dass die Kommission und ihr Sprachendienst die Anforderungen erkannt, sich dementsprechend angepasst und sich auch vor großen Umstrukturierungen nicht gescheut haben, ist es gelungen, dass die DGT heute 24 Amtssprachen meistert. Im Laufe der letzten Jahre hat sich dafür ein komplexes organisatorisches System entwickelt, das über klare Regeln und Abläufe verfügt. Die eingesetzten Hilfsmittel variierten, wie dieser kurze Überblick bereits zeigt – eine detaillierte Beschreibung folgt im nächsten Unterkapitel.

4.2. Meilensteine der EU-Sprachtechnologien

Die enorme Veränderung, die Ende der 90er Jahre beobachtet werden kann, ist der Vielzahl an neuen Systemen zu verdanken, die ihren Weg in die DGT fanden und den Arbeitsablauf stark umgeformt haben. Doch die EU wagte bereits zwei Jahrzehnte zuvor den Versuch, das wachsende Sprachproblem mithilfe von Sprachtechnologien zu lösen. Die bedeutendsten dieser Technologien werden auf den folgenden Seiten erläutert.

4.2.1. Systran

Schon früh beschäftigte sich die EU mit der maschinellen Übersetzung und so fand der erste Anlauf in Form des Übersetzungsprogrammes Systran statt. Seinen Anfang nahm das System 1976, als die Kommission zunächst eine Version mit der Sprachrichtung Englisch-Französisch erwarb. Aufgrund der recht zuversichtlichen Rückmeldung wurde in den darauffolgenden Jahren das Repertoire langsam um Deutsch sowie Italienisch und um andere Kombinationen (FR-EN, EN-DE etc.) erweitert. Der damalige Übersetzungsservice entwickelte die Systeme in Luxemburg und musste dabei vor allem auf die zweisprachigen Wörterbücher, die ein ausschlaggebendes Element von Systran bildeten, besonders achten. Nach zahlreichen Testphasen, war das System erstmals 1981 voll einsatzfähig und durchlief auch danach noch mehrere Überarbeitungen. Die Maschine funktionierte anhand des Transfer-Prinzips, weshalb die in den Wörterbüchern enthaltenen Angaben von großer Bedeutung für die Qualität waren – Änderungen und neue Einträge mussten mit Bedacht gewählt werden. Unter anderem versuchte man, der Maschine Informationen über die sprachspezifische Zeitenfolge und die Domäne eines Wortes zu vermitteln, indem jedem Terminus eine von ca. zwanzig Abkürzungen zugeteilt wurde (z.B. PRCH für Chemie). Die Arbeit der Kommission ist jedoch vielfältig, sodass eine solch komplizierte Zuordnung eines Begriffes zu einer einzelnen Domäne sich bald als wenig sinnvoll herausstellte und allgemeinere lexikalische Angaben eingeführt wurden. Zudem versuchte man semantische Codes zu entwickeln, mit deren Hilfe das System die richtige Satzstellung ermitteln würde (indem z.B. Adjektive korrekt mit „ihren“ Nomen verbunden werden) (vgl. Hutchins 1986).

Die von der Maschine aus dem Französischen ins Englische generierten informativen Texte konnten auf Wunsch der NutzerInnen auf dem Wege des *rapid post-editing* überarbeitet werden. Dabei handelte es sich um kein gründliches Korrekturlesen, sondern um eine grobe minimale Nachbesserung, die dafür wenig Zeitaufwand bedeutete (vgl. Hutchins 1986).

Das Systran System machte viele Fortschritte, profitierte jedoch nur bedingt von der Vermischung der lexikalischen und semantischen Regeln sowie seinem komplexen Wörterbuch. Neue technologische oder linguistische Ansätze konnten schwer integriert werden und die Erweiterung der Sprachenpaare war zwar grundsätzlich möglich, aber ab einer gewissen Anzahl an Sprachen nicht mehr für die an sie gerichteten Aufgaben geeignet (vgl. Hutchins 1986).

4.2.2. Eurotra

Diese Überlegungen führten in der Kommission schlussendlich zu der Entscheidung, auf ein System umzusteigen, das das Prinzip der Sprachenvielfalt besser umsetzen konnte und für computerlinguistische Innovationen empfänglicher für war. Als Reaktion darauf wurde 1982 Eurotra offiziell zu einem von der Kommission finanzierten Projekt erklärt, das über einen Zeitraum von fünf Jahren laufen und beträchtliche Ergebnisse mit sich bringen sollte. Es sollte ein fortgeschrittenes System sein, das sich dazu eignet, alle Amtssprachen zu meistern (nach dem Beitritt von Portugal und Spanien waren es zu diesem Zeitpunkt bereits neun Sprachen, an denen die SoftwareentwicklerInnen gleichzeitig arbeiteten) und die modernsten technologischen und sprachwissenschaftlichen Methoden in sich vereint. Auf diesem Wege sollte es zu einem praktischen Hilfsmittel für sämtliche Übersetzungsdienste der Europäischen Gemeinschaften werden. Eine weitere Anforderung war, dass es anpassungsfähig ist und somit Neuerungen einfacher durchgeführt werden können als zuvor mit Systran (gemeint sind damit beispielsweise zusätzliche Grammatikregeln und das Hinzufügen oder Abändern von Wörterbüchern etc.).

In der Anfangsphase sollte Eurotra in begrenzten Bereichen und für bestimmte Textsorten einsetzbar sein, doch darüber hinaus erwartete man sich, dass es nach erfolgreicher Entwicklung auch auf kommerziellerer Ebene Verwendung finden könnte (vgl. Hutchins 1986: 5-6). Nicht Teil der Erwartungen waren jedoch „perfekte“ Übersetzungen – das Ziel war es, den von der Maschine produzierten Text zur Gewinnung von Information zu nutzen und für alle anderen Bedürfnisse eine Nachbearbeitung zu veranlassen. Dementsprechend wären für die Aufgabe des Korrekturlesens passende Tools notwendig (vgl. Hutchins 1986). Eine Besonderheit in der Entwicklung dieses Transfersystems stellte der Projektablauf dar, da bis zu dem Zeitpunkt selten so viele europäische ForscherInnen ihr Wissen zusammentrugen. In diesem Sinne war Eurotra zwar langfristig gesehen nicht die ideale Lösung für die Europäische Union und ihre Übersetzung, doch sie brachte viele Fortschritte mit sich, die nicht zu unterschätzen sind (vgl. Hutchins 1986).

4.2.3. European Commission Machine Translation

Bevor das neueste System MT@EC seinen Platz einnahm, war ECMT im Einsatz. Aufbauend auf den seit 1976 bestehenden Programmen (siehe oben), stand dieses System den Bediensteten der EU-Institutionen und der DGT ab 1998 zur Verfügung. Am meisten wurde sie von der Kommission in Anspruch genommen, insbesondere jedoch von der DGT, die es bereits zu diesem Zeitpunkt in ihre CAT Tools integriert verwendete. Zuletzt hatte die Generaldirektion noch allein im Jahre 2008 über 1,5 Millionen Seiten mithilfe des MÜ Systems übersetzt (vgl. Europäische Kommission 2009b: 12).

Aufgrund seiner Vorgänger handelte sich somit um ein regelbasiertes Übersetzungssystem, dessen Wörterbücher, sprachlichen Regeln und neue Sprachkombinationen manuell eingegeben werden mussten und genau auf die Tätigkeitsfelder der Kommission abgestimmt waren (vgl. <http://ec.europa.eu>^m und Europäische Kommission 2009b: 13).

Die Verbreitung von Internet und der Kommunikation per E-Mails, brachte eine vereinfachte Nutzung mit sich (da auf MÜ über Mails zugegriffen werden konnte) und wurde auch von BeamtInnen verstärkt eingesetzt, um schnell global den Inhalt eines Dokumentes feststellen zu können. Es wurde jedoch davor gewarnt, in eine Fremdsprache übersetzen zu lassen, da man in der Muttersprache den Output der Maschine wesentlich besser beurteilen kann. 2009 bestand das System aus 18 Sprachrichtungen, von denen die Hälfte „zufriedenstellende Ergebnisse“ lieferte (Europäische Kommission 2009: 67).

Doch auch dieses regelbasierte System eignete sich nicht für die Umstände der Union – man bedenke das plötzliche Anwachsen der Sprachenanzahl im Jahre 2004. Während das System auch außerhalb der EU in den öffentlichen Verwaltungen der Mitgliedstaaten verwendet werden konnte, erachtete man die regelbasierte Funktionsweise als veraltet. Zudem war es fraglich, inwiefern die Qualität angehoben werden könnte. Ein weiterer Aspekt, der in die Beurteilung von ECMT einfluss, war die Anforderungen an das Personal. Für die Instandhaltung des Systems, d.h. Aktualisierung und Verbesserung der Regelsammlung und Wörterbücher, wurde ein großes Personal an SpezialistInnen benötigt (vgl. <http://ec.europa.eu>^m). Ein Ausbau der Sprachenpaare erwies sich somit aus mehreren Gründen als ein allzu schwieriges Unterfangen (hoher Aufwand, Kosten etc.), weshalb die Ausweitung des Angebotes auf sämtliche Sprachrichtungen nicht in Betracht gezogen werden konnte. Folglich wurde die Entscheidung getroffen, stattdessen auf die Entwicklung eines vielversprechenderen Systems umzusatteln. Im Mai 2010 wurde die DGT schließlich mit der Betreuung und Ausführung von MT@EC beauftragt (vgl. Europäische Kommission 2013: 12).

4.2.4. Inter-Active Terminology for Europe

Das umfassendste technische Hilfsmittel für europäische Terminologie ist seit 2004 IATE. Zu jenem Zeitpunkt wurden die bis dahin separat bestehenden Terminologiedatenbanken

der einzelnen Institutionen vereint, darunter u.a. Euradicator, die anfängliche Datenbank der Kommission (vgl. <http://iate.europa.eu>^a). Ihr Speicher enthält an die eineinhalb Millionen mehrsprachige Einträge, die mit Informationen über die AutorInnen, Institutionen, Fachbereiche und der Zuverlässigkeit (1-4 Sterne) der Termini versehen sind (vgl. Europäische Kommission 2013: 9).

Der Inhalt von IATE wird durchgehend aktualisiert, denn neue Themengebiete und dazugehörige Texte bringen immer wieder neue Termini hervor. In diesem Sinne steht es den TerminologInnen und ÜbersetzerInnen der DGT jederzeit frei, die Einträge zu ergänzen. Darüber hinaus existieren Projekte, in deren Rahmen Termini in das System eingespeist werden, die voraussichtlich in naher Zukunft in EU-Texten vorkommen werden – auf diese Weise kann den ÜbersetzerInnen zuvorgekommen werden (vgl. <http://iate.europa.eu>^b).

Um aus der Datenbank etwaige falsche Informationen zu eliminieren, soll zudem die Zuverlässigkeit der Angaben überarbeitet werden, um so u.a. alte, doppelt vorhandene oder noch nicht ausreichend überprüfte Termini zu vermeiden (vgl. <http://iate.europa.eu>^b).

4.2.5. SDL Translator's Workbench

Von 1989 bis 1994 wurde die Erstellung eines Translator's Workbench in zwei Phasen von der Union in Kooperation mit Unternehmen und Hochschulen gefördert (Projekte TWB I und TWB II). Es wurde hierbei ein Arbeitsplatz für ÜbersetzerInnen entwickelt, der sich bereits aus umfangreichen Werkzeugen zusammensetzte, wie sie auch zu Beginn dieses Kapitels aufgelistet wurden und für den gesamten Übersetzungsvorgang von Nutzen waren. In einem weiteren Schritt war die Generaldirektion Übersetzen eingebunden und es sollte das in der ersten Etappe entstandene TWB auch für PCs weiterentwickelt werden, doch die endgültige Umsetzung des Vorhabens scheiterte dann aufgrund von budgetären Kürzungen (vgl. Reinke 2004: 45-46).

Zurzeit ist das in der DGT meist verwendete CAT Tool das von SDL entwickelte Translator's Workbench. Der große Vorteil, den das computergestützte Übersetzen bietet, ist der Zugriff auf verschiedene Programme von einer Plattform aus. Über diesen lokalen Übersetzungsspeicher können ÜbersetzerInnen somit auf die in Euramis vorhandenen Daten zugreifen und ähnliche, bereits übersetzte Segmente übernehmen (System der Matches), ihre Arbeit automatisch abspeichern sowie auf Terminologiebestände zugreifen (ebenso über das von SDL stammende MultiTerm). Den Segmenten ist dabei stets eine Reihe von Attributen zuzuordnen, um auf einen Blick die ÜbersetzerInnen, die Dokumentennummer, das Jahr und die Auftrag gebende EU-Institution ausfindig machen zu können (vgl. Europäische Kommission 2009b: 11).

Wie die Bezeichnung des Programmes vermuten lässt, handelt es sich bei dem TWB nicht um ein von der Union selbst entwickeltes Tool. Das Produkt der Firma SDL wurde in einer Ausschreibung ausgewählt und schließlich an die Bedürfnisse der DGT ÜbersetzerInnen spezifisch angepasst. Die Wahl fiel deshalb auf dieses TWB, da es alle EU-Amtssprachen abdeckt und genauso wie auch Euramis mit dem tmx-Format (*Translation Memory eXchange*) arbeitet (vgl. Mai 2015 und Europäische Kommission 2009b: 11).

Obwohl es den Status des Standard CAT Tools hat und an den EU-Arbeitsablauf angepasst ist, besteht weiterhin die Option, andere Softwares heranzuziehen. So gibt es vereinzelt EU-ÜbersetzerInnen, die dennoch zu anderen Hilfsmitteln wie etwa OmegaT greifen (vgl. Mai 2015). CAT Tools haben sich besonders in der Kommission aus einem einfach nachvollziehbaren Grund durchgesetzt. Da sich die Mehrheit der zu übersetzenden Texte aus Rechtstexten wie etwa Gesetzesvorschlägen oder Rechtsakten zusammensetzt, kann auf unzählige Vorgängerübersetzungen und folglich passende Segmente zurückgegriffen werden. Zudem helfen integrierte Terminologiespeicher die Einheitlichkeit der EU-Termini zu wahren (vgl. Europäische Kommission 2009b: 11).

4.3. Qualitätssicherung und Sprachtechnologien

Der Einsatz von Sprachtechnologie verfolgt das Ziel, bei der Übersetzung eine Hilfestellung zu bieten, um einfacher und schneller den fertigen Text verfassen zu können. Dabei stellt sich die Frage danach, wie das Endergebnis beschaffen sein soll – kurz: welche Qualitätsansprüche bestehen gegenüber der fertigen Übersetzung? Die in der Arbeit beschriebenen Arten und Systeme an Sprachtechnologien bieten unterschiedliche Produkte an und somit hängt der Erfolg ihrer Verwendung zu einem beachtlichen Teil auch von den Erwartungen der NutzerInnen ab. Die DGT hat konkrete Anforderungen, die sie an ihre Übersetzungen stellt (siehe *fit-for-purpose*) sowie bestimmte Kriterien und Verfahren, anhand derer sie diese Qualität bewertet. Im nachfolgenden Abschnitt soll daher die Qualitätssicherung in der DGT dargestellt werden.

4.3.1. Ausgangslage

Die Merkmale von guter Qualität umfassend zu definieren, erweist sich oftmals schwierig und so verhält es sich auch im Falle von Übersetzungen nicht anders. Es bestehen unterschiedliche Ansprüche und folglich auch verschiedene Beurteilungskriterien und Standards. Andere Kriterien führen deshalb zwangsläufig zu anderen Qualitätsurteilen. Die Frage danach, was eine „gute“ Übersetzung ist, hängt grundsätzlich davon ab, wer die Zielgruppe ist, welche Funktion der Text erfüllt etc. (vgl. Hutter/Jekat 2010³: 675). Dasselbe trifft für

die Beurteilung vom „Können“ der Sprachtechnologien zu. Wie schon das 2. Kapitel gezeigt hat, hängt die Zufriedenheit der gelieferten Ergebnisse unvermeidbar mit den an sie geknüpften Erwartungen und vorschwebenden Zielen zusammen.

In diesem Zusammenhang ist es das Ziel der DGT, mit dem Einsatz von Sprachtechnologien für EU-Zwecke geeignete qualitative Texte zu erstellen, gleichzeitig aber auch Tools anzubieten, die primär der Verständigung dienen sollen. Somit stellt sich stets die Frage: Welche Technologie hält den qualitativen Erwartungen stand und erfüllt tatsächlich ihre Aufgaben?

Die Qualitätssicherung findet in allen Stadien des Übersetzungsprozesses statt – während der Wahl der geeigneten Hilfsmittel, der adäquaten Aufbereitung und dem effizienten Management des Ablaufs. Ein Großteil findet jedoch noch nach der Beendigung und Abgabe der eigentlichen Arbeit statt und zwar in Form von Evaluation und Feedback. Zukünftige Fehler sollen vermieden und die Systeme ständig erneuert werden (vgl. Drugan 2013:105). Daraus entsteht die logische Schlussfolgerung: andere Tools, anderer Ablauf, andere (idealerweise bessere) Qualität. Durch den gegenseitigen Einfluss aufeinander kann davon ausgegangen werden, dass dies zu einer „anhaltenden Evolution“ führt (Drugan 2013:107).

Sprachtechnologien ersetzen natürlich nicht vollkommen die Aufmerksamkeit und Überprüfung der ÜbersetzerInnen und durch Tools entstehende Fehlerquellen sind sehr wohl weiterhin vorhanden. Denn die Qualität des Endproduktes hängt wie bereits festgestellt von mehreren Faktoren ab, so etwa dem Ausgangstext, dem Tool und den ÜbersetzerInnen. Sind in einer TM beispielsweise fehlerhafte Übersetzungen eingespeichert oder enthalten Terminologiedatenbanken falsche bzw. unvollständige Angaben, dann kann die Qualität unabhängig von den an sich guten Tools nicht zufriedenstellend sein. Dasselbe trifft auch auf maschinelle Übersetzungstechnologien zu: werden einem SMÜ-System unpassende oder zu wenige Parallelkorpora eingespeist, werden die Ergebnisse als enttäuschend einzustufen sein. Es gilt also wiederum das GIGO (*Garbage In, Garbage Out*) Prinzip zu beachten. Aus diesem Grund ist es für ÜbersetzerInnen eine Notwendigkeit, alte Einträge (z.B. insbesondere *fuzzy matches*) nicht ohne Weiteres zu übernehmen, sondern stets im Kontext der neuen Übersetzung zu prüfen und zu aktualisieren.

Diesbezüglich hält die Kommission zusätzliche Hürden bereit, da in den Datenbanken europäische und nationalstaatliche Terminologie aufeinander treffen. Richtige Formulierungen können unter Umständen von den Tools als falsch markiert werden, obwohl sie in dem jeweiligen Kontext der richtigen Wahl entsprechen. Die Achtsamkeit und Vorsicht der ÜbersetzerInnen ist somit in allen Stadien gefordert, sei es bei der Vorbereitung, der Übersetzung oder der nachträglichen Datenpflege. Vereinfacht ausgedrückt, bedeutet dies: das Zusammenwirken guter Tools und guter ÜbersetzerInnen ebnen den Weg zu guter Qualität (vgl. Drugan 2013: 108-109).

4.3.2. Qualitätssicherung in der DGT

Während die Qualität eines Tools wie der maschinellen Übersetzung mit verhältnismäßig einfachen Methoden definiert werden kann, stellt sich die Situation bei von Menschen verfassten Übersetzungen bereits anders dar. Mithilfe von CAT Tools übersetzte Texte beinhalten stets das menschliche Element. Innerhalb der DGT besteht deshalb ein eigenes Vorgehen, um die Qualitätsstandards zu sichern.

In diesem Sinne ist eine durchdachte und effektive Qualitätssicherung der DGT Dienstleistungsprodukte unerlässlich. Mit der richtigen Vorgangsweise können hochwertige Übersetzungen geliefert werden und damit einhergehend die Kosten für fehlerhafte oder schlechte Übersetzungen um ein Vielfaches gesenkt werden.

Ein richtiger Umgang mit Qualitätssicherung ist ausschlaggebend, da fehlerhafte oder unpassende Übersetzungen auf mehreren Ebenen Konsequenzen nach sich ziehen können. Einerseits können dadurch falsche Informationen vermittelt und Missverständnisse verursacht werden, die auf politischer und wirtschaftlicher Ebene Auswirkungen haben. Ebenso darf nicht vergessen werden, dass dies darüber hinaus juristische Folgen haben kann und wegen Fehlübersetzungen durchaus in kostspieligen Gerichtsverfahren entschieden werden muss. Zudem kann dadurch für die EU bzw. die Kommission im Speziellen und nicht zuletzt auch die DGT eine Schädigung des Images einhergehen. Dieser Schaden ist finanziell zwar nicht konkret messbar, aber dennoch nicht zu unterschätzen, denn es gehört zu den grundlegenden Zielen der EU, eine transparente, nachvollziehbare und gute Kommunikation mit den BürgerInnen zu führen. Fehlerhafte, schwer verständliche und missverständliche Texte und Übersetzungen bieten dafür jedoch keine vielversprechende Ausgangslage (vgl. Europäische Kommission 2012b: 46).

Aus diesem Grund verfolgt die DGT das Ziel, den Ablauf der Arbeitsprozesse und die Kommunikation zwischen den ÜbersetzerInnen so gut wie möglich zu gestalten. Ein reibungslos ablaufendes System ermöglicht nicht nur, dass bei der Übersetzung selbst Zeit (und somit Kosten) gespart werden können, sondern auch, dass es zu weniger Fehlern und Nachbesserungen in den Übersetzungen kommt und somit weniger Nachbearbeitungskosten verursacht werden. Genau an diesem Punkt übernehmen Sprachtechnologien eine tragende Rolle.

Im Bereich des Qualitätsmanagements versteht man unter dem Begriff „*cost of poor quality*“, die Kosten, die durch jenen Aufwand entstehen, der benötigt wird, um die schlechte Qualität eines Produktes auszubessern (Europäische Kommission 2012b: 2). Als Definition von guter Qualität werden dabei vor allem die Anforderungen und Erwartungen der KundInnen herangezogen bzw. für welchen Zweck die Übersetzung von diesen im Weiteren angewendet werden soll – dies wird in der DGT auch als „*fit-for-purpose* Übersetzung“ bezeichnet (Europäische Kommission 2012b: 11 und vgl. *ibid.*: 4). Konkret umfasst dies

die präzise Anwendung der jeweiligen Zielsprache sowie fachspezifischer und EU-interner Terminologie, ebenso wie die inhaltliche Übereinstimmung des Originals mit der Übersetzung und die Berücksichtigung der spezifischen Textsortenkonventionen (vgl. Europäische Kommission 2012b: 13).

Studien belegen, dass es sich als kostspieliger erweist, Fehler im Nachhinein auszubessern. Demzufolge ist es sinnvoller, in den Übersetzungsvorgang selbst zu investieren und diesen zu optimieren, damit in der nächsten Runde keine Verbesserungskosten und keine Fristenverschiebungen anfallen (vgl. Europäische Kommission 2012b: 5).

Zu dem Übersetzungsvorgang zählt nicht allein das Übersetzen selbst, sondern bereits die Vorbereitung ebenso wie auch die Nachbereitung. Zudem ist es ausschlaggebend, wer den jeweiligen Text übersetzt (wenn es sich beispielsweise um ein bestimmtes Fachgebiet handelt, dass einen darin erfahrene/n ÜbersetzerIn braucht) und welche Tools eingesetzt werden. Das entsprechende Lektorat und Feedback dürfen ebenso wenig vergessen werden. Das Gesagte gilt sowohl für die angestellten ÜbersetzerInnen als auch für die freiberuflichen, von der EU-anerkannten ÜbersetzerInnen, die ebenso an die strengen Qualitätsstandards gebunden sind. Externe Übersetzungen – dies betrifft ungefähr 28% der jährlich übersetzten Seiten –, werden zusätzlich benotet („*very good, good, acceptable, below standard, unacceptable*“) und Freelancer mit durchgängig schlechten Ergebnissen werden von der Liste gestrichen (Europäische Kommission 2012b: 17).

Eine Übersetzung kann auch als mangelhaft gelten, wenn die Stilistik nicht elegant und klar ist (vgl. Europäische Kommission 2012b: 46). In einigen Fällen werden deshalb auch stilistische Änderungen vorgenommen, um den Wünschen der KundInnen gerecht zu werden (vgl. Europäische Kommission 2012b: 30-31). Daraus kann gefolgert werden, dass ein großer Anteil dieser Fehler bereits im Vorhinein ausgemerzt werden könnte – unter anderem mithilfe von Sprachtechnologien, die die ÜbersetzerInnen beispielsweise automatisch auf unübersetzte Textstellen hinweisen und den konstanten Gebrauch von bestimmten Begriffen erleichtern würden. Ob jedoch stilistische und terminologische Falschangaben dadurch beseitigt werden können, ist fraglich, da hier in erster Stelle die Aufmerksamkeit und die Kenntnisse der ÜbersetzerInnen gefragt sind.

Es kostet zahlreiche Arbeitsstunden, Unklarheiten, mehrdeutige Formulierungen, inkonsequente Formulierungen zu eruieren und klarzustellen – sei es durch eigene Recherche, durch Zusammenarbeit mit KollegInnen und der Erstellung so genannter *Notes* oder durch Kontaktierung der AutorInnen. In diesem Zusammenhang stellt das vorherige Verbessern eine große Hilfe dar, mit der bereits vor dem Übersetzen Unstimmigkeiten entfernt und der Ausgangstext im Allgemeinen verbessert werden kann. Je mehr Qualitätsaspekte in sämtliche Teile des Übersetzungsprozesses eingebunden werden, desto besser das Ergebnis (vgl. Europäische Kommission 2012b: 20-23).

5. Aktuelle Situation und Projekte der Kommission

In den letzten fünf Jahren haben sich im Bereich Sprachtechnologien viele Initiativen entwickelt, aus denen langjährige Programme hervorgegangen sind. Ihre Ausführung ist derzeit im Gange und somit ist ihr Erfolg und ihre langfristige Auswirkung noch nicht eindeutig feststellbar. Die Projekte werden jedoch von Studien und Analysen begleitet und so kann bereits nach einer kurzen Laufzeit ein erstes Bild über ihren Einfluss gegeben werden.

Die Projekte sind allesamt stark miteinander vernetzt und bilden das Gerüst für eine Reihe von Initiativen, die die Wettbewerbsfähigkeit der EU im Visier haben, wie etwa Horizont 2020. Im Mittelpunkt steht ohne Zweifel die maschinelle Übersetzung, mit der erneut große Hoffnungen und Pläne verbunden sind und neuer Aufschwung ausgelöst werden soll. Neben der MÜ sind Schlüsselbegriffe wie *digitale Vernetzung*, *Einheitlichkeit* und *sofortiger Informationsfluss* in allen Projekten allgegenwärtig.

Möchte man einen ersten Überblick über neue Entwicklungen auf dem Gebiet der europäischen Sprachtechnologien bekommen, ist ein Blick auf die Themen und Vorträge renommierter Konferenzen, die unlängst stattgefunden haben, ein ausgezeichneter Ausgangspunkt. Zu nennen wäre etwa die Riga Konferenz 2015, bei welcher sowohl die Idee des mDSM als auch das ELRC der breiten Öffentlichkeit vorgestellt wurden. Was sich genau hinter diesen und ähnlichen Abkürzungen verbirgt, ist Thema dieses Kapitels.

5.1. Machine Translation at the European Commission (MT@EC)

Das Projekt, das auf die DGT den größten Einfluss hat, ist die neue maschinelle Übersetzung. Nachdem der Vorgänger ECMT eingestellt wurde, erfolgte im Juni 2013 die Einführung des neuen Systems, an dem weiterhin weiterentwickelt wird. Der Umstieg von einem regel-basierten auf ein statistisches System sollte die bestehenden Nachteile ausgleichen und die Handhabung der stark angewachsenen Amtssprachen erleichtern. Von dem neuen Programm werden sämtliche Amtssprachen in jeder erdenklichen Sprachkombination abgedeckt – die Gesamtzahl an möglichen Kombinationen beläuft sich somit auf unglaubliche 552, wobei davon 62 auf unmittelbarem Wege bearbeitet werden können, d.h. ohne die Verwendung einer Art Mittlersprache (Europäische Kommission 2014c: 1). Die nachfolgenden Seiten beschreiben, wie das MT@EC System benutzt werden kann, sowie dessen Aufbau und die Möglichkeiten des Zugriffs.

Nach dem Einstellen von ECMT bestanden grundlegende Anforderungen an das neue System, das sich in Form des MT@EC durchsetzte: Es sollte mindestens das Niveau von seinem Vorgänger erreichen und den NutzerInnen weiterhin mindestens dieselben Leistungen bieten können. Natürlich muss es an die Anforderungen der Kommission und deren Aufgabenbereich angepasst sein (z.B. den reibungslosen Informationsaustausch zwischen der

Kommission und diversen Mitgliedstaaten ermöglichen) und zumindest alle Amtssprachen abdecken. Zudem sollte im Sinne eines optimalen Ergebnisses auf die Nutzung der zur Verfügung stehenden Ressourcen Wert gelegt werden (von bestehenden Terminologiedatenbanken bis hin zu Sammlungen an Textkorpora). Erst wenn diese Punkte erfüllt werden könnten, würde es sich aus finanzieller und organisatorischer Sicht lohnen, das maschinelle System in solch einem Umfang umzuwälzen. Um den Voraussetzungen gerecht zu werden, sollte überdies die Vertraulichkeit des Systems gesichert werden, um die mitunter heiklen Daten ausreichend schützen zu können (vgl. <http://ec.europa.eu>^m).

Diese Kriterien wurden von einer im Jahre 2009 in einer vorbereitenden Phase gegründeten ExpertInnengruppe aufgestellt. Im Anschluss wurde sie von der Kommission beauftragt, sich über die Beschaffenheit des neuen Systems Gedanken zu machen. Im Zuge dessen fiel die Entscheidung auf folgenden Aufbau: Der Zusammensetzung des neuen SMÜ-Systems sollten zwei Säulen zugrunde liegen, die so genannten „*data management hub*“ und „*MT engines hub*“ (<http://ec.europa.eu>^m).

Ersteres dient der Sammlung und der Verarbeitung der für das SMÜ unerlässlichen Datenmengen und Sprachtechnologien (siehe z.B. Training, wie in Kapitel 2.2.3.2. beschrieben), während letzteres sämtliche Aspekte des Umgangs mit MÜ-Maschinen umfasst, im Speziellen die Handhabung von MÜ-Anfragen.

Das zweite *hub* ist durch zwei Komponenten gekennzeichnet, den Übersetzungsmaschinen selbst und dem so genannten *Dispatcher*, der für die Bearbeitung der Anfragen zuständig ist. Der *Dispatcher* empfängt die neuen Aufträge, analysiert diese und teilt das zu übersetzende Dokument der zuständigen Maschine zu. Die Wahl fällt er je nachdem, welche Sprache gewünscht wurde und ob der Text inhaltlich einer speziellen Domäne zugeordnet werden kann. Ist die Übersetzung fertig, wird sie auf demselben Wege von der Maschine über den Dispatcher an die Auftraggeber zurückgeleitet. Ihm kommt auch die Aufgabe zu, die Übersetzung wieder in der adäquaten Form (meist im Originalformat) anzuzeigen (vgl. <http://ec.europa.eu>^m).

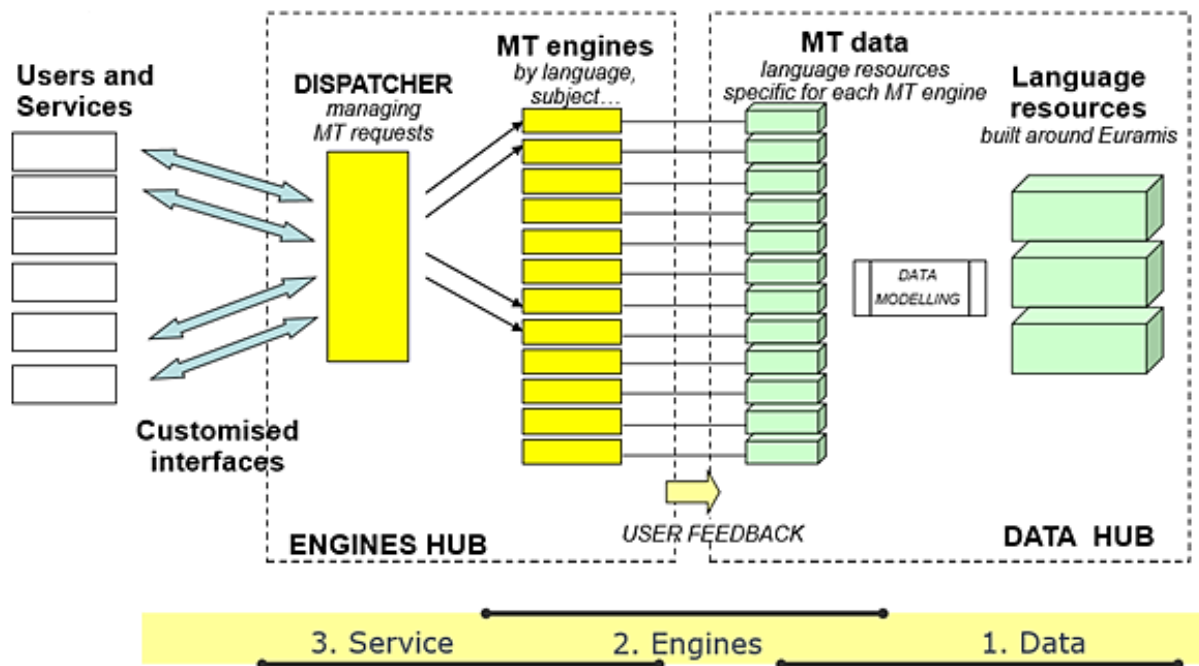


Abbildung 6. Struktur des MT@EC Systems (<https://termcoord.files.wordpress.com>)

Wie die zwei Säulen bereits erkennen lassen, bilden die sprachlichen Ressourcen (sprich die Daten) und die IT Ressourcen die Kernkomponenten. Die Entwicklung und Organisation dieser zwei Bereiche wurde aufgeteilt: Während das ISA (*Interoperability Solutions for European Public Administrations*) Programm die mit IT und der Organisation zusammenhängenden Aufgaben übernimmt, ist die DGT seit Juni 2010 vor allem für die sprachlichen Ressourcen zuständig. Der dafür entwickelte Plan sieht u.a. vor, dass die DGT die von der ISA vorgesehenen Pläne ausführt, Anfragen für kundenspezifische Systeme handhabt und nach dem Ende der ISA Projekte den Service und die Instandhaltung des MT@EC übernimmt (vgl. <http://ec.europa.eu>^m).

Seit 2010 wurde das System Schritt für Schritt eingeführt. In den ersten Phasen lag der Fokus u.a. auf der Evaluierung der Risiken und der Anforderungen der NutzerInnen sowie die Erstellung von Prognosen für die zukünftige Einführung und die Festigung des gesamten Konzeptes (z.B. Festlegung der Grenzen und Möglichkeiten des Systems). Mittlerweile lief bis Ende 2015 die erste Testphase (vgl. <http://ec.europa.eu>^m).

5.1.2. Moses

Für die Entwicklung von MT@EC wurde das SMÜ-Programm Moses verwendet. Es handelt sich um eine seit 2005 bestehende Open-Source Software, die derzeit zu den am verbreitetsten Softwares dieser Art gehört und deshalb oft als Vergleichsbasis herangezogen wird, wenn es um die Bewertung neuer Systeme geht. Neben der EU wird es ebenso in an-

deren internationalen Organisationen verwendet, so ist es unter anderem auch in eLUNa, dem CAT Tool, das speziell für die Übersetzung in den Vereinten Nationen entwickelt wurde, integriert (vgl. Forcada 2015: 161-162 und <https://cms.unov.org>).

Die Zusammenarbeit zwischen Moses und der Europäischen Union begann bereits 2006 im Rahmen des Euromatrix Projektes, das die Förderung der maschinellen Übersetzung für alle Sprachenpaare zu seinem Ziel erklärt hatte. Das System kommt in den Bereichen der Forschung, der Lehre und vermehrt nun auch im kommerziellen Umfeld (gesehen bei TAUS) zum Einsatz. Das SMÜ-System von Moses ist hauptsächlich für zwei Zwecke gedacht: Einerseits als MÜ, die zusätzlich durch Nachbearbeitung von Menschenhand korrigiert wird und andererseits integriert in die computergestützte Übersetzung, z.B. in einem CAT Tool (vgl. <http://www.statmt.org/moses>).

5.1.3. MT@EC außerhalb der DGT

Im Folgenden soll die Verwendung von MT@EC außerhalb der DGT (sprich nicht als ein Tool von deren ÜbersetzerInnen), sondern als ein Service für NutzerInnen wie den restlichen Organen und Behörden, dargestellt werden.

2015 startete eine erste Pilotphase, im Rahmen derer MitarbeiterInnen und Mitgliedsstaaten den Dienst kostenlos bis Ende 2015 nutzen konnten. Seitdem ist die Übersetzung in und aus allen Amtssprachen möglich und seit kurzem wurde das Angebot um Norwegisch erweitert. Nach Wunsch kann ein Ausgangstext mit einmaliger Abgabe in mehrere Sprachen übersetzt werden. Der Ausgangstext kann entweder über das Eingabefenster der MT@EC Internetplattform als freier Text eingegeben oder als eigenständiges Dokument hochgeladen werden. Es wird versucht, das Originalformat und die Struktur des Originaltextes so gut wie nur möglich beizubehalten (vgl. <http://ec.europa.eu>ⁿ).

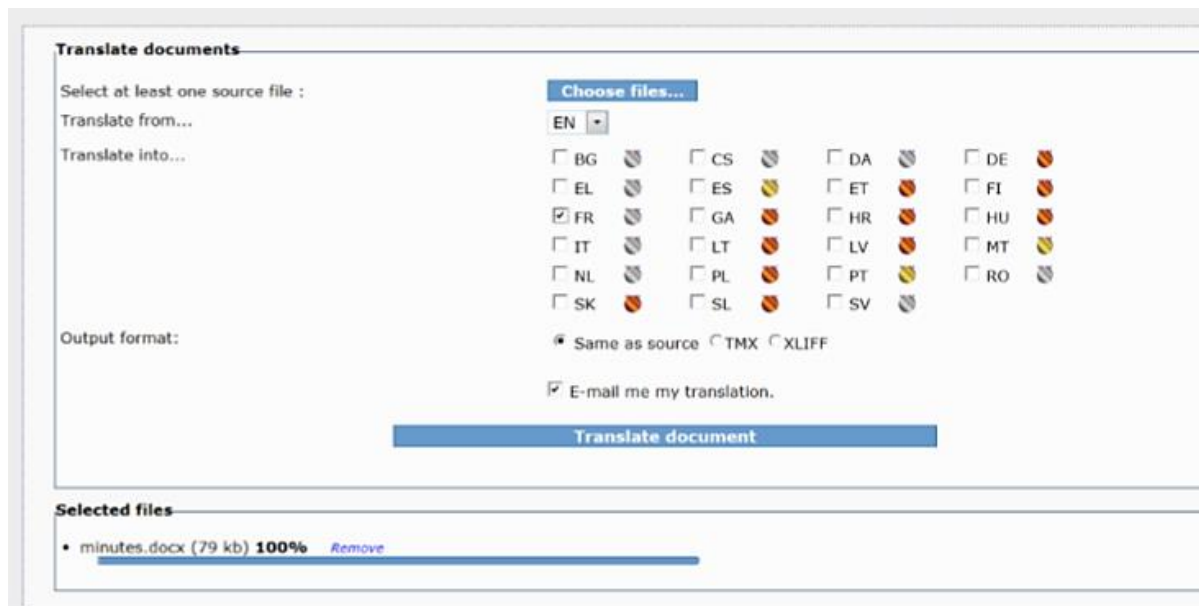


Abbildung 7. Ausschnitt aus dem MT@EC API (Europäische Kommission 2014c: 2)

Die Anleitung listet Microsoft Office Dokumente („doc, docx, xls, xlsx, ppt, pptx“), Open-Office Dokumente sowie „html, odt, pdf, rtf, txt, tmx, xliiff“, als Formate auf, die von dem Programm akzeptiert werden (Europäische Kommission 2014c: 2). Die NutzerInnen erhalten ihre Zieldokumente automatisch in der ursprünglichen Formatierung zurück, wobei dieser Service bei PDF Dokumenten nicht besteht, denn die Übersetzung wird in diesem Fall nur als txt angezeigt (vgl. Europäische Kommission 2014c: 2). Da PDF allerdings sehr praktisch und daher beliebt ist und die neuerliche Erstellung des Layouts mit viel Aufwand einhergeht, haben bereits mehrere EU-Institutionen gemeldet, dass die Beibehaltung des Layouts auch bei diesem Format erwünscht wäre (vgl. Europäische Kommission 2014a: 94).

5.1.3.1. Qualität und Ergebnisse

Der MT@EC Service ist jedoch mit Vorsicht zu genießen: Auch wenn das System auf alle Sprachen angewendet werden kann, müssen sich NutzerInnen bewusst sein, dass die Qualität je nach Sprachenpaar schwankend ist. Während z.B. die Kombination von Italienisch und Englisch gute Ergebnisse liefert, werden NutzerInnen von Ungarisch und Englisch oder sogar Deutsch und Englisch andere Erfahrungen machen. Um die Qualität der einzelnen Sprachkombinationen festzustellen, wird in den EU-Institutionen das BLEU Verfahren angewendet (siehe 2.2.3.3.).

Um auf die soeben angesprochenen Qualitätsunterschiede hinzuweisen, wurde symbolisch jeder Sprache eine Medaille verliehen, die Auskunft über ihren Rang auf dem Qualitätspodest gibt. Folgerichtig ist bei dem Symbol der goldenen Medaille von einer qualitativ hochwertigeren maschinellen Übersetzung auszugehen, während von Silber eine Übersetzung

zu erwarten ist, die zu Verständniszwecken geeignet ist (dies ist meist so zu deuten, dass das Endergebnis stilistisch und grammatikalisch zu wünschen übrig lässt, sich der Inhalt aber ohne größere Probleme ermitteln lässt). Die Bronzemedaille wird an jene Sprachenpaare vergeben, bei denen MT@EC eher nur dazu genutzt werden sollte, sich einen groben Eindruck von dem Inhalt eines Textes zu verschaffen (vgl. <http://ec.europa.eu>ⁿ).

Die Medaillesymbole werden im API neben den Sprachen angezeigt (siehe dazu Abb. 7) und sollen dabei helfen, einschätzen zu können, welche Erwartung grundsätzlich an die angefertigte Übersetzung gestellt werden kann. Klarerweise hängt sehr viel von den konkreten Ausgangstexten ab und so kann die Qualität der tatsächlich gelieferten Zieltexte durchaus besser oder schlechter ausfallen, als das von der Webseite verliehene Qualitätssiegel (vgl. <http://ec.europa.eu>ⁿ). Dennoch ist das Medaillesymbol eine gute Methode, um den NutzerInnen schnell und unkompliziert Informationen über ihre gewünschten Sprachenpaare im MT@EC System zu vermitteln.

To From	BG	CS	DA	DE	EL	EN	ES	ET	FI	FR	GA	HR	HU	IT	LT	LV	MT	NL	PL	PT	RO	SK	SL	SV
BG		B	B	B	B	G	S	B	B	S	B	B	B	B	B	B	S	B	B	S	S	B	B	B
CS	B		B	B	B	S	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B
DA	B	B		B	B	G	S	B	B	B	B	B	B	B	B	B	S	B	B	B	B	B	B	B
DE	B	B	B		B	S	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B
EL	S	B	B	B		G	S	B	B	S	B	B	B	B	B	B	S	B	B	S	B	B	B	B
EN	S	S	S	B	S		G	B	B	S	B	B	B	S	B	B	G	S	B	G	S	B	B	S
ES	B	B	B	B	B	G		B	B	S	B	S	B	S	B	B	S	B	B	S	S	B	B	B
ET	B	B	B	B	B	B	B		B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B
FI	B	B	B	B	B	B	B	B		B	B	B	B	B	B	B	B	B	B	B	B	B	B	B
FR	B	B	B	B	B	G	G	B	B		B	S	B	S	B	B	S	B	B	S	S	B	B	B
GA	B	B	B	B	B	B	B	B	B	B		B	B	B	B	B	B	B	B	B	B	B	B	B
HR	B	B	B	B	B	B	B	B	B	B	B		B	B	B	B	B	B	B	B	B	B	B	B
HU	B	B	B	B	B	B	B	B	B	B	B	B		B	B	B	B	B	B	B	B	B	B	B
IT	B	B	B	B	B	G	S	B	B	S	B	B	B		B	B	S	B	B	S	S	B	B	B
LT	B	B	B	B	B	B	B	B	B	B	B	B	B	B		B	B	B	B	B	B	B	B	B
LV	B	B	B	B	B	S	B	B	B	B	B	B	B	B	B		B	B	B	B	B	B	B	B
MT	S	B	B	B	B	G	S	B	B	S	B	S	B	S	B	B		B	B	S	S	B	B	B
NL	B	B	B	B	B	S	B	B	B	B	B	B	B	B	B	B	B		B	B	B	B	B	B
PL	B	B	B	B	B	G	S	B	B	B	B	B	B	B	B	B	S	B		B	B	B	B	B
PT	S	B	B	B	B	G	S	B	B	B	B	S	B	S	B	B	S	B	B		S	B	B	B
RO	S	B	B	B	B	G	S	B	B	S	B	B	B	S	B	B	S	B	B	S		B	B	B
SK	B	B	B	B	B	S	S	B	B	B	B	B	B	B	B	B	S	B	B	B	B		B	B
SL	B	B	B	B	B	S	S	B	B	B	B	B	B	B	B	B	S	B	B	B	B	B		B
SV	B	B	B	B	B	G	S	B	B	B	B	B	B	B	B	B	S	B	B	B	B	B	B	

G = gold, S = silver and B = bronze

Tabelle 1. Qualität nach Sprachenpaaren, ermittelt durch BLEU Tests
(<http://ec.europa.eu>ⁿ)

Die zum Großteil bronzefarbte Tabelle lässt keine Zweifel daran, dass das neue SMÜ noch reichlich Training und eine umfangreichere Datenmenge benötigt. Es verwundert kaum, dass Sprachenpaare, die Englisch als Zielsprache in ihrer Kombination enthalten, am öftesten den Goldstatus erhalten haben und auch Spanisch über zahlreiche Silbermedaillen verfügt. Im Gegensatz dazu gibt es eine große Anzahl an Sprachen, die nur in wenigen Fällen Silber oder gar Gold verliehen bekommen hat. Betrachtet man die Spalten von Ungarisch, Finnisch, Estnisch, Kroatisch oder Deutsch (nur ein einziges Silber), dann ist die Qualität des maschinellen Outputs als eher bescheiden zu bezeichnen.

Die Gründe für die Unterschiede wurden in dieser Masterarbeit bereits angesprochen. Statistiken von April 2015 sprechen für sich und zeigen deutlich die starke Variation zwischen den Datenmengen. Die Zahl der monolingualen englischen Texte in Euramis beläuft sich auf nahezu 97 Millionen, während im Vergleich dazu weniger als drei Millionen monolinguale kroatische Texte vorhanden sind. Die Tatsache, dass Kroatien ein noch relativ junger Mitgliedstaat ist, bietet zwar eine Erklärung für diese Diskrepanz, doch ein Blick auf die knapp 58 Millionen gespeicherten monolingualen französischen Texte verdeutlicht, dass dieser Unterschied auch im Vergleich zu großen, traditionellen EU-Sprachen besteht. Laut Angaben befindet sich die Datenmenge anderer Amtssprachen zwischen 22 und 38 Millionen (vgl. <http://www.rigasummit2015.eu>). In diesem Sinne ist es klar, dass das Sammeln von nützlichen Daten, vor allem von parallelen Textkorpora, ein wichtiges Ziel darstellt, um das System auf den bestmöglichen Stand zu bringen.

An dieser Stelle muss angemerkt werden, dass die Anzahl der Dokumente natürlich nicht der einzige Faktor ist, der über die Qualität eines MÜ-Systems entscheidet. Wie Vielfältig die Gründe für das Endergebnis sein können, wird nun unter die Lupe genommen.

Die für Mitgliedsstaaten und MitarbeiterInnen zur Verfügung gestellte *Service Description* der Kommission nennt neben nützlichen Informationen zur Nutzung auch drei Auslöser für eventuelle Qualitätsschwankungen (vgl. <http://ec.europa.eu>ⁿ).

Erstens seien dies die involvierten Sprachen. Dies hängt mit den grammatikalischen Merkmalen zusammen (je komplexer die Regeln einer Sprache – vor allem der Zielsprache – desto wahrscheinlicher sind schwerwiegende grammatikalische Fehler).

Zweitens eignet sich die MÜ-Methode nicht für alle Stile. Sobald der Ausgangstext von dem Formulierungsstil abweicht, der für EU-Dokumente charakteristisch ist, sinkt die Qualität des Endproduktes. Bedenkt man wie SMÜ-Systeme funktionieren, ist dies leicht nachvollziehbar: Das System wurde anhand der offiziellen EU-Texte trainiert – an umgangssprachliche oder etwa literarische Formulierungen ist es nicht gewohnt, findet in seinem Speicher keine passenden Ergebnisse und kann mit dem Ausgangstext dementsprechend schlechter umgehen (vgl. <http://ec.europa.eu>ⁿ). Zudem variiert die Qualität der Formulierungen in den Ausgangstexten je nach VerfasserIn (ein Phänomen das bereits in Ka-

pitel 3.2.2. beschrieben wurde). Als Schlussfolgerung ist es bei solchen Texten empfehlenswert, sich über den Zweck der Übersetzung Gedanken zu machen: soll bloß der Inhalt ermittelt werden oder soll die Übersetzung beispielsweise veröffentlicht werden? Wird die Entscheidung getroffen, dass der Stil ebenso wichtig ist und auch in der Zielsprache dieselbe Wirkung erzielt werden soll, bietet die Humanübersetzung durch die DGT (oder zumindest eine letzte Überarbeitung) eindeutig die bessere Lösung. So werden AuftraggeberInnen auch auf der offiziellen Homepage der DGT daran erinnert: „*Remember, if you need an accurate, high-quality translation, the text must be revised by a skilled professional translator*“ (<http://ec.europa.eu>^e).

Drittens ist der Inhalt selbst zu nennen und hängt wiederum von dem Trainingskorpus des Systems ab. Handelt der Text von Gegenständen, die einen EU untypischen Themenbereich betreffen (und damit keine vorherigen Texte der Europäischen Kommission dazu verfügbar sind), erstellt die maschinelle Übersetzung nur unzureichende Sätze. Unbekannte Terminologie kann in diesem Fall nicht verarbeitet werden und wird im Zieltext in Ausgangssprachlicher Form beibehalten – ebenso ein Problem, das bei der Übersetzung durch DGT ÜbersetzerInnen nicht auftreten würde (vgl. <http://ec.europa.eu>ⁿ).

Während diese drei Punkte für alle Sprachen zutreffen, geben sie keine detaillierte Auskunft über die sprachspezifischen Probleme. Das MT@EC Team ist bedacht darauf, diese Unterschiede zu beheben, doch in den letzten Jahren hat sich dies als schwierige Aufgabe herausgestellt. Es finden regelmäßig Treffen zwischen den VertreterInnen der einzelnen Sprachabteilungen und den EntwicklerInnen der MÜ statt, doch die Ergebnisse sind durchwachsen. So beschreibt Markus Foti im Rahmen der diesjährigen MT@EC Konferenz, dass einfache Probleme zwar schnell gelöst werden, doch sobald Grammatik der Auslöser ist, oftmals kaum Fortschritte gemacht werden können. In diesem Zusammenhang durchgeführte Experimente beschäftigten sich bereits damit, wie die richtige Satzordnung im Deutschen hergestellt und wie im Ungarischen mit den Suffixen (das „Anhängen“ von z.B. Präposition an das Wortende) umgegangen werden kann. Im Allgemeinen können SMÜ-Systeme mit flektierenden und komplexeren Sprachen nicht zufriedenstellend umgehen. Die Erstellung von Statistiken anhand von eingespeicherten Texten reicht meist nicht, um die grammatikalischen Eigenheiten und Variationen zu erfassen und sie schließlich im Übersetzungsprozess wiedergeben zu können (besonders wenn die herausfordernde Sprache die Zielsprache ist) (vgl. Foti, 04.12.2015, <https://scic.ec.europa.eu>).

Als mögliche Lösung wird ein Hybridsystem gesehen, das SMÜ und regel-basierte Ansätze miteinander vereint. Erste Versuche mit MT@EC haben jedoch keine guten Ergebnisse mit sich gebracht, weshalb eine Umsetzung dieser Idee in die Wirklichkeit bisher gescheitert ist. Mit dem CEF Programm erhofft man sich eine Veränderungen der derzeitigen Situation. Dadurch, dass dieses große Projekt (Details zu CEF siehe unten) die Vergrößerung des MT@EC Teams und neuen Input zum Resultat hatte, soll nun erneut verstärkt

die Arbeit an einem Hybrid begonnen werden (vgl. Foti, 04.12.2015, <https://scic.ec.europa.eu>).

Über MT@EC kann auch ein bestimmtes Themengebiet (Domäne) angegeben werden. Diese Option ist nicht für alle NutzerInnen zugänglich, sondern nur für EU-Organisationen, für die solche Einstellungen eigens erarbeitet wurden. Benutzerdefinierte SMÜ-Systeme wurden anhand von im Vorfeld gesammelten Material speziell für diese Organe erstellt. Der Zugriff auf die angepassten Systeme steht jedoch nicht automatisch allen anderen Institutionen zur Verfügung und muss zuerst von jenem Organ bewilligt werden, für das es ursprünglich entwickelt wurde (vgl. <http://ec.europa.eu>ⁿ). Dieser besondere Service trägt zu einer verbesserten Qualität des Outputs bei, da Korpus und Terminologie für die Zwecke eines ausgewählten Organs maßgeschneidert werden können und das SMÜ nun auf ein spezialisiertes Wissen zugreift.

Laut Angaben der *Service Description* bearbeitet MT@EC Dokumente innerhalb von durchschnittlich zwei Stunden (mehr als 95% der Aufträge gehen innerhalb dieses Zeitrahmens an die AuftraggeberInnen zurück). Für kürzere Textausschnitte („*snippet characters*“ mit bis zu maximal 4000 Zeichen) beträgt die Wartedauer an die zwei Minuten (dieser Ablauf gilt hier ebenso für mindestens 95% der Aufträge) (<http://ec.europa.eu>ⁿ). Das System ist diesbezüglich auf einem sehr guten Entwicklungsstand, da der reibungslose Ablauf in beinahe 100% der Fälle erreicht werden kann (vgl. <http://ec.europa.eu>ⁿ).

MT@EC ist mit Sicherheitsvorkehrungen ausgestattet, um unter Umständen heikle Dokumente bestmöglich zu schützen. Für den sensiblen Datenaustausch zwischen nationalen und EU-Behörden wird sTESTA genutzt („*Secured Trans European Services for Telematics between Administrations*“) (<http://ec.europa.eu>^j). Es handelt sich dabei um ein eigenes, von anderen Telekommunikationsnetzen unabhängiges EU-weites Netzwerk, das zu diesem Zweck u.a. von der Kommission, dem Rat und Europol gemeinsam entwickelt wurde. Der Service bietet NutzerInnen eine zuverlässige Alternative zum Versenden des Ausgangs- und Zieltextes per E-Mail (vgl. <http://ec.europa.eu>^j).

Zudem speichert MT@EC Originale sowie Übersetzungen 24 Stunden lang, nachdem der Auftrag beendet wurde. Sind NutzerInnen damit nicht einverstanden, kann auch die Option gewählt werden, dass die Dokumente nach dem Herunterladen der Übersetzung sofort aus dem System gelöscht werden (vgl. <http://ec.europa.eu>ⁿ).

Um Lücken und Probleme des Systems schließen und beseitigen zu können, werden NutzerInnen nach dem Erhalt einer Übersetzung um Feedback gebeten. Man erhofft sich dadurch auch mitunter einen Anstoß für zukünftige Anwendungen zu erhalten. Bis Ende 2015 machten jedoch leider nur wenige von der Feedback-Möglichkeit Gebrauch: über einem Zeitraum von etwa zwei Jahren erhielt die DGT an die 500 Rückmeldungen (vgl. <http://ec.europa.eu>ⁿ und Foti, 04.12.2015, <https://scic.ec.europa.eu>).

5.1.3.2. Pilotprojekte für die öffentliche Verwaltung

Öffentliche Verwaltungen in den Mitgliedsstaaten bzw. deren MitarbeiterInnen hatten bis Ende 2015 die Möglichkeit, den neuen MÜ-Service kostenlos zu nutzen. Dafür bestanden keine besonderen Voraussetzungen, die NutzerInnen mussten sich lediglich über einen E-CAS (*European Commission Authentication Service*) Account registrieren und konnten über diesen mit MT@EC Übersetzungen erstellen lassen (vgl. Europäische Kommission 2014c: 3).

Um das System zu testen und weiterentwickeln zu können, hat die Kommission in Kooperation mit der DGT auch nationalen Behörden die Teilnahme an Pilotprojekten angeboten. Interessierte öffentliche Verwaltungen der Mitgliedstaaten haben die Wahl zwischen fünf verschiedenen Optionen, um mit der Übersetzungsabteilung bilateral zusammenzuarbeiten und MT@EC auf eine auf sie zugeschnittene Art anzuwenden (vgl. Europäische Kommission 2014c: 4).

Die zur Wahl stehenden Optionen gestalten sich wie folgt:

Pilot A bietet die Möglichkeit, das MÜ-System mit einem Informationssystem der jeweiligen Verwaltung zu verbinden, damit von dort aus auf das Angebot der DGT direkt zugegriffen werden kann. Die nationalen Behörden müssen die Verbindung zu der MT@EC Plattform jedoch selbst herstellen, können dabei aber mit der Hilfe der DGT ExpertInnen rechnen.

In Pilot B und Pilot C sollen die Voraussetzung geschaffen werden, eine für die Bedürfnisse der Verwaltungen trainierte Version der MT@EC zu erschaffen. Dafür werden die nationalen Einrichtungen aufgefordert, eigenes Material zur Verfügung zu stellen, das in der DGT zum Training verwendet wird. Diese Daten dürfen mit keinen Urheberrechten mehr behaftet sein, ansonsten sind sie für die Bearbeitung nicht geeignet. Die beiden Projekte unterscheiden sich bloß in jener Hinsicht, dass Pilot B das Ergebnis auch allen anderen NutzerInnen zugänglich macht, während das in Pilot C entstandene System die Informationen weiterhin nur für die InhaberInnen der zur Entwicklung übermittelten Daten einsehbar ist (vgl. Europäische Kommission 2014c: 4).

Die letzten beiden Projektvariationen sind sich in ihren angestrebten Zielen sehr ähnlich. Pilot D bietet öffentlichen Behörden die Gelegenheit, eine eigene MÜ-Maschine zu verwenden. Diese würde von der DGT gebaut und anhand von zur Verfügung gestellten Dokumenten fertiggestellt werden. Der Service ist dann nicht mehr in das MT@EC Projekt eingebunden, sondern ein eigenständiges System, das von den Verwaltungen noch dazu lokal verwaltet wird (der Zugang geschieht nicht über die EU-API und sonstige weitere Aufgaben bezüglich der Instandhaltung und Modernisierung passieren ohne Zutun der DGT). Das Endergebnis von Pilot E ist ebenso eine eigene MÜ-Maschine, doch die Arbeit übernimmt gänzlich die jeweilige Behörde, während die DGT bloß als Stütze involviert ist.

Sie bietet Hilfe bei der technischen Umsetzung und macht nach der Abhaltung von Konsultationen mit den Behörden, Vorschläge für das Training und die Installation der neuen Maschine (vgl. Europäische Kommission 2014c: 5).

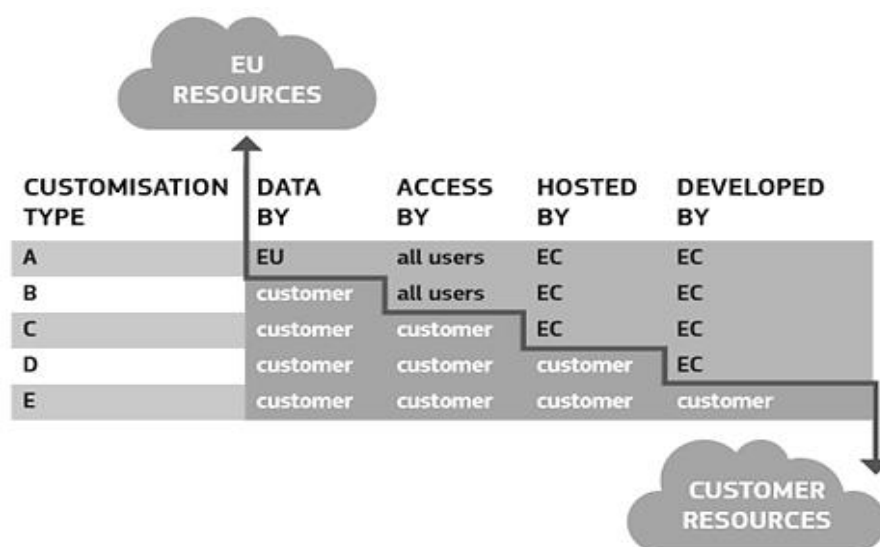


Abbildung 8. Vergleich der Pilotprojekte (Europäische Kommission 2014c: 6)

Durch die Durchführung der Pilotprojekte erhofft man sich schlussendlich, anregende Erkenntnisse für die Verbesserung der MT@EC Methode und der benötigten Daten zu gewinnen. Aus diesem Grund sollten die Verwaltungen, die sich an die DGT wenden, von vornherein große Mengen an typischen Originaltexten samt ihren Übersetzungen sammeln und an die Generaldirektion schicken, damit diese außer dem Training der SMÜ auch genauer analysiert werden können. Auf diesem Wege können typische Fehlerquellen, regelmäßig auftretende Probleme sowie Charakteristika der Texte und der Terminologie ermittelt, in der Entwicklung von kundenspezifischen Anforderungen bedacht und wenn notwendig, für zukünftige Anwendungen behoben bzw. angepasst werden (vgl. Europäische Kommission 2014: 5).

Die DGT bietet im Rahmen der Testphase somit drei Möglichkeiten dafür, wie die von den öffentlichen Verwaltungen zugesendeten Dokumente bearbeitet werden können: zur Evaluation, zur Adaptation und zum Training. Um repräsentative Ergebnisse zu ermitteln, muss eine gewisse Anzahl an Texten zur Analyse gegeben sein. Reichen für die erste Evaluation bereits 20 bis 100 Texte aus, steigt die Menge für weitere Schritte an. Eine erfolgreiche Anpassung des MÜ-Systems an behördenspezifische Umstände benötigt bereits mehrere hundert Texte, während für das sinnvolle Training eines neuen SMÜ-Systems – wie in Kapitel 2.2.3.2. erklärt – mehrere tausende Dokumente gebraucht werden (vgl. Europäische Kommission 2014: 5-6).

Die Erprobung auf nationaler Ebene soll für 2016 neue Anstöße für MT@EC auf EU-weiter Ebene bringen.

5.1.4. MT@EC innerhalb der DGT

Die DGT ist in vieler Hinsicht in dem Projekt involviert, da sie mit dem Projektmanagement betraut ist und ihre MitarbeiterInnen wie schon zuvor erläutert nicht zuletzt MÜ bereits stark in ihre Arbeit einbinden und von ihr profitieren. Die ÜbersetzerInnen der DGT sind in der Verbesserung des MT@EC Output stark involviert. Die bis 2011 entwickelten Übersetzungsmodule werden innerhalb der Kommission natürlich ebenso getestet und in die alltägliche Arbeit eingeführt.

MT@EC wird dabei hauptsächlich als ein in das TWB integriertes Tool verwendet. Es steht jedoch allen ÜbersetzerInnen frei, persönlich für sich zu entscheiden, ob sie bzw. in welchem Ausmaß sie MÜ einbinden möchten oder nicht – im CAT Tool lässt sich die MÜ Option mit einer einfachen Auswahl aktivieren. Als Entscheidungsgrundlage für die Verwendung einer maschinellen Vorübersetzung dienen in der Regel die Textsorte und das Sprachenpaar, anhand derer erfahrene ÜbersetzerInnen schnell erkennen können, welche Tools sich am besten eignen (z.B. Verordnung vs. Broschüre) (vgl. Mai 2015).

In Kontakt mit MT@EC Texten kommen die ÜbersetzerInnen auch bei deren Nachbearbeitung. Laut Statistiken aus dem Jahre 2010 erhielt die DGT über das System an einzelnen Tagen an die 10.000 maschinell übersetzte Seiten, die schließlich nachbearbeitet werden mussten (vgl. Europäische Kommission 2013: 12 und Mai 2015). Darüber hinaus kommen den MitarbeiterInnen (vor allem dem IT Team) als MT@EC ExpertInnen in Rahmen des CEF zusätzlich unter anderem Beratungsfunktionen hinzu (siehe unten).

5.1.5. MT@EC Konferenz

Seit 2012 werden jährlich Konferenzen zu dem neuen maschinellen Übersetzungssystem abgehalten. Die erste legte ihren Fokus noch auf die MitarbeiterInnen der DGT und ihre Erfahrungen, nachdem ECMT abgeschafft wurde. In den darauffolgenden Jahren standen die EU-Institutionen (2013) sowie die öffentlichen Verwaltungen der Nationalstaaten im Mittelpunkt (2014). Die letzte Konferenz fand am 4. Dezember 2015 in Brüssel statt und sollte nun die ersten Schlussfolgerungen und Berichte der unterschiedlichen TeilnehmerInnen vergleichen und die ersten Jahre rekapitulieren. Unter den Anwesenden waren VertreterInnen der Kommission (bzw. der DGT) sowie des Europäischen Parlamentes und von Universitäten sowie ForscherInnen und MitarbeiterInnen von SDL Trados. Das letzte Treffen kann per Video On Demand abgerufen werden und gibt interessante Einblicke in den neuesten Stand und Pläne der verschiedenen Akteure (vgl. <https://scic.ec.europa.eu>).

Im Laufe der Konferenz sticht eine Botschaft deutlich hervor: auch wenn es um die

maschinelle und computergestützte Übersetzung geht, ist der Mensch derjenige, von dem am Ende alles abhängt. Ein weiteres Thema, das angeschnitten wurde, betrifft die Frage, inwiefern Studierende auf neue Aufgaben wie die Nachbearbeitung von maschinellen Übersetzungen vorbereitet werden können und in welcher Form dies in die Studienpläne eingebaut werden kann. Mehrere TeilnehmerInnen plädierten für ein bewusstes Training von qualifizierten *post-editors*. Zuweilen wird eine Unterscheidung zwischen ÜbersetzerInnen und *post-editors* gemacht und die zwei Aufgaben folglich voneinander getrennt gesehen.

MT@EC entwickelt sich unaufhaltsam weiter. Im Dezember 2015, zum Zeitpunkt der letzten Konferenz, war bereits die dritte Version des Systems veröffentlicht worden und verspricht nutzungsfreundlicher zu sein. Als nächster Schritt sollen mehrere NutzerInnen Zugriff auf MT@EC haben: Nebst der DGT sollen nun auch eine größere Zahl an EU-Institutionen damit arbeiten und die Mitgliedstaaten ebenso stärker eingebunden werden (vgl. <https://scic.ec.europa.eu>).

5.2. Europäischer Binnenmarkt und Übersetzung

Das mDSM Projekt (multilingual Digital Single Market) wurde von der europäischen Sprachtechnologiegemeinschaft ELTC ausgearbeitet und der Öffentlichkeit erstmals im Rahmen des *Riga Summit* im April 2015 vorgestellt. Es stellt eine Erweiterung des DSM (Digital Single Market) Programmes der Europäischen Kommission dar, das erst kurz davor im März präsentiert wurde. Bevor das mDSM im Detail erläutert wird, folgen zunächst die wesentlichen Grundlagen des digitalen Binnenmarktes.

5.2.1. Digital Single Market

Der digitale Binnenmarkt ist einer der wirtschaftlichen Prioritäten der aktuellen Kommission und verfolgt das Ziel, die im Internet bestehenden Hürden zwischen den einzelnen Nationalstaaten weiter abzubauen. Zwar besteht seit den 90er Jahren der gemeinsame europäische Binnenmarkt, digital existieren allerdings noch immer sehr viele Hindernisse und Unterschiede zwischen den einzelnen nationalen Märkten. DSM zielt genau darauf ab, diese Probleme zu beseitigen und einen tatsächlichen homogenen Binnenmarkt auch auf digitaler Ebene in die Wirklichkeit umzusetzen, auf dem die vier Grundfreiheiten „frei zirkulieren können“ (<http://ec.europa.eu>^p).

Der DSM basiert auf drei Säulen: „Besserer Zugang zu digitalen Waren und Dienstleistungen“, „Optimale Rahmenbedingungen für digitale Netze und Dienstleistungen“ und „Digitale Wirtschaft als Wachstumsmotor“ (<http://ec.europa.eu>^p).

Die erste Säule umfasst vor allem Maßnahmen, die die elektronische Wirtschaft erleichtern sollen. Dazu gehört beispielsweise die Bekämpfung des so genannten *Geoblocking*. Darunter ist die Preisdiskriminierung am Binnenmarkt zu verstehen (Produkte der gleichen AnbieterInnen haben in einigen Mitgliedsstaaten einen höheren Preis). Weitere Punkte sind die Unterstützung von kleinen und mittelständischen Betrieben im Online-Handel, damit sie einen größeren KundInnenkreis außerhalb der nationalen Grenzen erreichen können sowie die Schaffung günstigerer Versandbedingungen. Derzeit liegt der Prozentsatz jener KMUs, die auch Online-Handel betreiben bei niedrigen 15%. Davon bieten ihre Produkte und Dienstleistungen nur etwa ca. 7% an internationale InteressentInnen an – Daten, die laut der Kommission im Sinne eines europäischen Wachstums zu ändern sind (vgl. <http://ec.europa.eu>^o und Rehm 2015: 1).

Die zweite Säule legt den Fokus auf die Entwicklung geeigneter Technologien und Infrastruktur, mit deren Hilfe die notwendigen Rahmenbedingungen geschaffen werden sollen, um die Ziele des DSM erreichen zu können. Betroffen sind die Bereiche der Telekommunikation, Mediengesetzgebung und Cybersicherheit (vgl. <http://ec.europa.eu>^r).

Die letzte Säule des Planes zielt darauf ab, in die Entwicklung von IT Anwendungen wie Cloud Computing, Big Data u.ä. zu investieren. Europas Kompetenzen im Bereich IT sollen erweitert und im Sinne der Einheitlichkeit gemeinsame Normen festgelegt werden. Die neuen Kenntnisse sind auch für die Nutzung durch öffentliche Behörden angedacht, um in dieser Form Ausgaben senken zu können (vgl. <http://ec.europa.eu>^q).

Der Weg zu einem digitalen Binnenmarkt ist noch nicht fertig beschritten. Die Kommission hat im Rahmen ihrer Digitalpolitik für 2016 für alle drei Säulen weitere Schritte geplant – bis Ende des Jahres sollen insgesamt sechzehn Initiativen durchgeführt werden. Einige Eckpunkte der zukünftige Maßnahmen bilden die Überprüfung und Modernisierung von EU-Recht, das in seiner alten Form den Ausbau des DSM hindert (z.B. Verbraucherschutz, Datenschutzrichtlinien, Telekommunikation), sowie die Verbesserung des Datenflusses und Erhöhung der Personenanzahl, die über Internetzugriff verfügt (vgl. <http://ec.europa.eu>^p und <https://ec.europa.eu>ⁱ).

5.2.2. Multilingual Digital Single Market

Der mDSM folgt dem Aufbau des DSM und verfolgt im Grunde genommen dieselben Ziele. Das Besondere an diesem Projekt ist jedoch, dass deren VerfechterInnen den Faktor Sprache bzw. Mehrsprachigkeit als den Schlüssel zur Verwirklichung des einheitlichen digitalen Marktes sehen. Inwiefern Übersetzung und Sprachtechnologien darin eine Rolle einnehmen, zeigt die Präsentation der mDSM Strategie.

Der Ausgangspunkt der Strategie ist die Beobachtung, dass es trotz anhaltender Bemühungen weiterhin keinen gemeinsamen *europäischen Markt*, sondern aufgrund der Landessprachen voneinander getrennt agierende *europäische Märkte* bestehen. Der wirtschaftliche Kontakt wird dadurch eingeschränkt und die Möglichkeiten eines internationalen Handels nicht ausreichend ausgenutzt. Auch an dieser Stelle wird die kulturelle und sprachliche Vielfalt Europas als schützenswert gesehen, aber gleichzeitig sei dieses Charakteristikum eine wirtschaftliche Bremse. Denn: „*Language barriers constitute market barriers*“ (Rehm 2015: 5).

Englisch als bindendes Glied zwischen den einzelnen Sprachen wird auch von den ForscherInnen des mDSM abgelehnt, da die Zahl an Englisch-SprecherInnen relativ gering ist: 50% der EU-Bevölkerung beherrscht Englisch, allerdings verfügen davon nur an die 10% über genügend Kenntnisse, um sich beispielsweise im Online-Handel zurecht zu finden. Diese und ähnliche Daten waren das Ergebnis einer umfassenden Eurobarometer-Umfrage, die die Kommission 2012 durchführen ließ (vgl. <http://ec.europa.eu>^s und Rehm 2015: 1). Dass eine *Lingua Franca* keine zufriedenstellende Lösung bietet, unterstreicht eine weitere Statistik, gemäß derer an die 60% jener Personen, die Englisch nicht als ihre Muttersprache haben, kaum oder gar keine Einkäufe auf rein englischsprachigen Seiten tätigen (vgl. Rehm 2015: 1).

Zahlreiche EU-BürgerInnen und Unternehmen werden dadurch benachteiligt und haben auf erhebliche Mengen an Online-Daten keinen Zugriff – dieser wird durch ihre Sprache blockiert („*language-blocking*“) (Rehm 2015: 4). Doch *language-blocking* bezieht sich nicht allein auf wirtschaftliche Transaktionen. Die ProjektleiterInnen heben hervor, dass davon auch Regierungsseiten und EU-Daten betroffen sind. Der Internetauftritt von lokalen bzw. nationalen Regierungsbehörden und Ministerien ist oftmals nur in ihrer Landessprache zugänglich. Auf EU-Ebene ist die Situation auch nicht gänzlich anders: Zahlreiche Ressourcen der Europäischen Kommission werden nur in Englisch oder in den drei Verfahrenssprachen (EN, FR, DE) angeboten, weshalb wiederum für an die Hälfte der BürgerInnen die Aufnahme von Informationen erheblich erschwert wird (vgl. Rehm 2015: 6).

Diese Anhaltspunkte sprechen allesamt für den Einsatz von Übersetzungen, um die separaten Märkte mit Brücken zu verbinden. Die Arbeit von HumanübersetzerInnen wirft jedoch erneut Bedenken auf, hinter jenen primär finanzielle und zeitliche Überlegungen stehen. Es wird argumentiert, dass etwa KMUs und kleine, regionale Organisationen ihr volles Potenzial nicht ausschöpfen könnten, weil sie beispielsweise die Übersetzungskosten nicht tragen können und ihnen infolgedessen trotz des Binnenmarktes weiterhin zu viele Türen verschlossen blieben. Für die ProjektleiterInnen ist zudem fraglich, wie ÜbersetzerInnen die großen Datenmengen meistern sollen, weshalb auf menschliche Übersetzung im Rahmen von mDSM bloß in ausgewählten Fällen gesetzt wird. Die geförderten Sprachtechnologien

sollen zudem zu einer effizienteren Arbeitsweise von HumanübersetzerInnen beitragen (vgl. <http://www.multilingualeurope.eu>).

Der Fokus liegt somit auf dem Einsatz von automatisierten Sprachtechnologien. Mit deren Hilfe sollen Übersetzung und Lokalisation auch für kleinere Unternehmen erschlossen werden, da die Kosten mit maschineller Übersetzung verhältnismäßig gering gehalten und am Markt ihr vielfacher Nutzen eingeholt werden kann. Produkte und Dienstleistungen würden einem internationalen, wenn nicht sogar vollkommen globalen Markt zugänglich sein (vgl. Rehm 2015: 2-3).

Die Entwicklung von Sprachtechnologien und Systemen, die den in der Agenda formulierten Anforderungen gerecht werden, ist das Ziel von CRACKER und LT_Observatory. Dieses Vorhaben erweist sich jedoch in jenem Sinne schwierig, dass auf diesem Gebiet je nach Sprache andere Umstände herrschen. Es ist ein ähnliches Problem anzutreffen, wie auch im Falle von MT@EC: die Systeme liefern qualitativ äußerst abweichende Ergebnisse. Besonders kleinere Sprachen müssen verstärkt unterstützt werden (darunter ist vor allem das Zusammentragen von mehr Sprachressourcen zu verstehen), um ihr digitales Aussterben zu verhindern. Denn im Gegensatz zu Englisch oder Französisch sind Sprachen wie Ungarisch, Maltesisch oder Litauisch im Internet deutlich unterrepräsentiert (vgl. Rehm 2015: 3).

Durch die Erforschung neuer Technologien erhofft man sich bereits innerhalb von zehn Jahren große Fortschritte zu machen und diese in den unterschiedlichsten Sektoren einzusetzen. Vorteile erhofft man sich nicht nur für die Wirtschaft, sondern auch in Bereichen wie Gesundheit und Bildung. Die VerfasserInnen des Berichtes sehen dafür gute Chancen, da Europa in diesem Feld eine attraktive Basis bietet: Mehr als 2500 europäische Unternehmen befassen sich mit Sprache und IT, zudem wächst die Zahl der InternetnutzerInnen stetig an und wirkt dem digitalen Aussterben entgegen (vgl. Rehm 2015: 3-4).

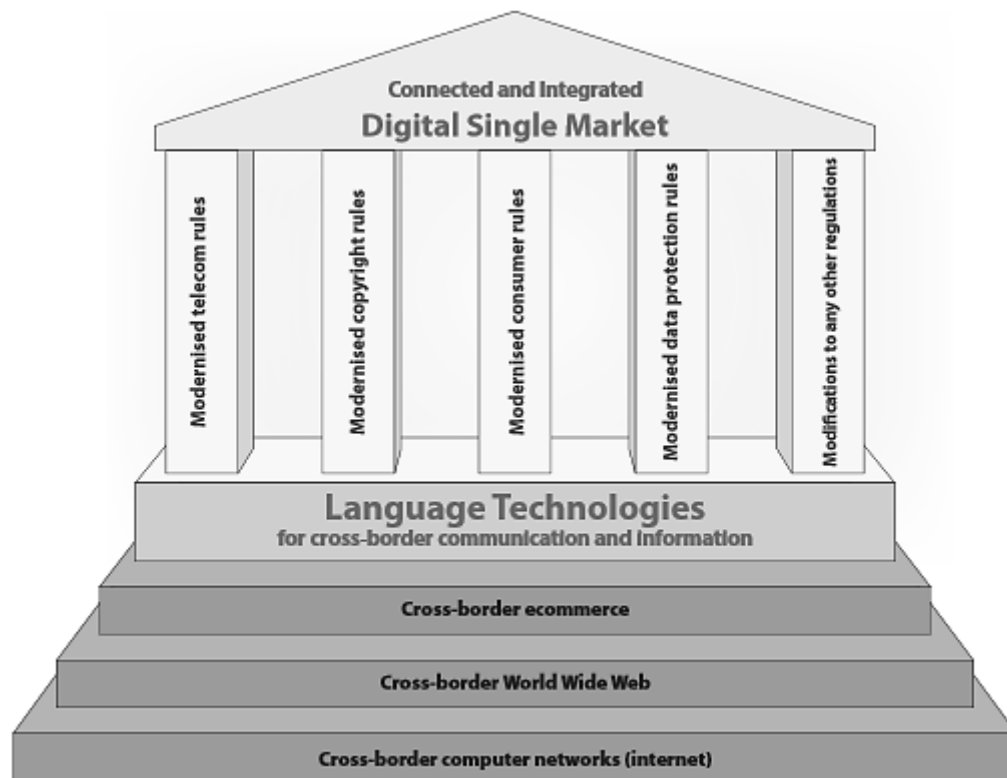


Abbildung 9. Sprachtechnologien als Grundlage des Digital Single Market (Rehm 2015: 11)

Neben der Sprache kommt dem Umgang mit Daten ein weiterer wichtiger Aspekt zu. Durch eine durchdachte Vorgangsweise sollen Daten in verschiedensten Medien (primär Text, aber genauso auch Video, Ton etc.) und verschiedensten Sprachen Europas gesammelt und geordnet werden, um im Späteren eine sinnvolle Nutzung einer gesamteuropäischen Datenmenge gewährleisten zu können. Denn es ist nicht ausreichend die Daten zusammenzutragen, als nächster Schritt müssen sie mithilfe von innovativen Ideen aus dem Bereich der (Sprach)Technologie zugänglich und praktisch einsetzbar gemacht werden. Für die erfolgreiche Umsetzung dieses Planes entwickelten die ExpertInnen des ELTC (*European Language Technology Community*) eine aus drei Ebenen bestehende Strategie, die bis 2020 laufen soll (vgl. Rehm 2015: 7, 10).

Die erste Ebene besteht aus Lösungsansätzen, die für Unternehmen und öffentliche Einrichtungen vorgesehen sind und den Weg für neue Sprachtechnologien ebnen sollen. Die Ansätze sollen gemeinsam mit der Europäischen Kommission ausgearbeitet, verfeinert und vereinbart werden. Von dieser Phase hängen die weiteren Schritte ab: Die konkret beschlossene Form dieser Lösungsansätze bestimmt die Entwicklung der nächsten beiden Ebenen, da diese von dem Ausmaß und den Ansprüchen der ersten Ebene abhängen (vgl. Rehm 2015 10, 12).

Die Lösungen umfassen u.a. Marktforschungstechnologien, sprachenübergreifenden KundInnenservice im Bereich des *e-commerce* sowie Systeme zur Erstellung neuen, vernetzten Inhaltes. Sämtliche dieser Vorschläge sollen durch die Zusammenarbeit von europäischen Projekten entwickelt werden. Ein Beispiel hierfür ist das *Digital Translation Center* (DTC). Dieses ist als eine Cloud angelegt und soll rund um die Uhr sofort vertrauenswürdige Übersetzungen und Dolmetschungen in den europäischen und größeren nicht-europäischen Sprachen liefern. Grundsätzlich sollen NutzerInnen die Wahl zwischen mehreren Angeboten haben: Leistungen reichen von dem „simplen“ automatischen Übersetzen einzelner Wörter und Termini bis hin zu vollständigen Texten sowie zu von professionellen ÜbersetzerInnen angefertigten Arbeiten, die gebührenpflichtig sind, und schließlich kostenlose Humanübersetzungen oder Nachbearbeitungen, die von NGOs wie *Translators without Borders* oder *The Rosetta Foundation* bereitgestellt werden. Es ist darüber hinaus aber auch vorgesehen, dass Terminologie betreffende Fragen an die DTC gerichtet werden können und eine Datenbankpflege von TMs angefragt werden kann. Damit soll allen VerwenderInnen, seien es BürgerInnen, Unternehmen, WissenschaftlerInnen oder öffentliche Organisationen, eine Möglichkeit geboten werden, zu den Services des mDSM und folglich zu den von ihnen gewünschten Inhalten unabhängig von ihrer Landes- bzw. Muttersprache Zugang zu haben. Für Unternehmen beispielsweise könnte diese Funktion auch auf die Produktionsabläufe einen Einfluss haben: Produktinformationen, Warnhinweise oder Garantien könnten in kürzester Zeit für mehrere Regionen und Sprachen gleichzeitig übersetzt und lokalisiert werden (vgl. Rehm 2015: 19).

Während die restlichen Lösungsansätze der ersten Ebene nicht explizit die Übersetzungsdienstleistungen in den Mittelpunkt stellen bzw. andere Arten von Sprachtechnologien verwenden, kommen diese dennoch nicht ohne MÜ oder menschlicher Übersetzungshilfe aus. Als Beispiel kann *Voice of the Customer* genannt werden. Hierbei werden mithilfe von Sprachtechnologie zunächst Daten über die KundInnenzufriedenheit in mehreren Staaten gesammelt (z.B. über Kommentare auf Social Media oder Blogs) werden. In der Auswertungsphase ist der Einsatz von Übersetzungstechnologien vorgesehen, um auf diese Weise die Analyse und weitere Erforschung möglich zu machen. Dieses Konzept soll auch für die EU-Institutionen von Nutzen sein. Mit derselben Auswertungsmethode kann die Meinung und grundlegende Einstellung der EU-BürgerInnen zu aktuellen politischen Themen gemessen werden. Solch eine Analyse lässt das Erfassen eines gesamteuropäischen Bildes zu und geht somit über separat geführte Diskurse in einzelnen Mitgliedstaaten hinaus (vgl. Rehm 2015: 24, 21).

Für Unternehmen von Interesse sind weiters das *Unified Customer Experience* (UCE), bei der automatische Übersetzungstools für Online-Shops und ähnlichen Internetseiten angewendet werden. In Verbindung damit kommen auch halbautomatische Tools für die Lo-

kalisierung und Internationalisierung der Inhalte zum Einsatz. Diese Funktion soll der Anforderung der heutigen Zeit, sofort auf Produkte und Produktinformationen u.ä. zugreifen zu können und mit den KundInnen in Echtzeit und in der Landessprache in Kontakt treten zu können, gerecht werden. Vor allem für KMUs stellt das eine besondere Möglichkeit dar, um europaweit wettbewerbsfähig zu sein (vgl. Rehm 2015: 18-19).

Die zweite Ebene besteht aus den technischen Werkzeugen (Softwareprogrammen, Clouds, Internet etc.) mit denen aus den theoretischen Lösungsansätzen der ersten Ebene nun Praxis werden soll – sprich es geht um ihre Umsetzung. Zusammen sollen die Technologien eine Plattform bilden, deren Elemente miteinander vernetzt, nutzerfreundlich und jederzeit erreichbar sind sowie stetig erweitert bzw. miteinander kombiniert werden können. Nutzt jemand beispielsweise eine Spracherkennungssoftware, wäre damit automatisch auch die Übersetzung des vorgeschprochenen Textes möglich. Aus den einzelnen Anwendungen soll somit eine Technologie-Infrastruktur zusammengesetzt werden, die für NutzerInnen eine einfache Bedienung sowie für EntwicklerInnen und AnbieterInnen leichtere Entwicklungsmöglichkeiten bietet (vgl. Rehm 2015: 29).

Auf dieser Ebene kann auch die Arbeit von HumanübersetzerInnen (stärker) einfließen. Die Angebote der Plattform sollen nämlich bestimmte Kriterien aufweisen, darunter hohe Qualität, Vertraulichkeit und Datenschutz. Idealerweise sollen darunter auch hochwertige Open Source Programme sein, um finanziell benachteiligte NutzerInnen nicht auszugrenzen (vgl. Rehm 2015: 12, 34).

Um genügend mehrsprachige Ressourcen als Grundlage zur Verfügung zu haben, wird vorgeschlagen auch Quellen wie Wikipedia oder BabelNet heranzuziehen. Im Rahmen von Projekten wie z.B. Meta-Share wird versucht, diese Technologien auf schnellem und kostengünstigem Wege umzusetzen. Um das festgelegte Ziel zu erreichen, ist noch viel Arbeit notwendig – dies gilt vor allem für die Zahl an verfügbaren Ressourcen und die Frage nach der Integration von bestimmten Leistungen (vgl. Rehm 2015: 12, 30).

Die dritte und zugleich letzte Ebene ist der Forschung gewidmet, die die soeben beschriebenen Prozesse befähigt. Diese soll mehrere Themengebiete untersuchen, einige Beispiel dafür sind eine hochwertige MÜ und Humanübersetzung. Die Agenda empfiehlt es, sich umfassende Daten über Algorithmen, Ergebnisse statistischer Studien und neue Ansätze anzueignen, damit Rückschlüsse für die Übersetzung zwischen allen europäischen und nicht-europäischen Sprachen abgeleitet werden können. Zudem könnte auf dem Gebiet der Sprachtechnologie die Erforschung von natürlichsprachlichen Dialogsystemen sowie Tools für die Textextraktion und für das automatische Zusammenfassen und Ordnen von Material genau untersucht werden. Die Strategie fordert dazu auf, derzeit noch bestehende Lücken im Wissen über die möglichen Einsatzbereiche von Sprachtechnologien zu füllen und eifrig zu erforschen. Zu diesem Zweck sollen Forschungszentren und EntwicklerInnen europaweit zusammenarbeiten und somit nicht nur die Grundlage für die zweite und erste Ebene

schaffen, sondern das Potenzial eines mDSM auch in Zukunft aufrechterhalten (vgl. Rehm 2015: 16). Eine solche internationale Zusammenarbeit erweist sich auch in jener Hinsicht als ausschlaggebend, dass eine Standardisierung der entstehenden Systeme notwendig ist, damit sie problemlos miteinander kombiniert werden können (vgl. Rehm 2015: 34).

Weitere interessante Herausforderungen sind der Umgang mit der enormen Sprachenvielfalt Europas sowie dem Weltwissen und der Kultur, die sich hinter Sprache verbergen kann. Doch laut den AutorInnen der Agenda können bereits die kleinsten Fortschritte zu einem Durchbruch verhelfen, um den Weg in Richtung neuer Technologien und kommerziell einsetzbarer Systeme einzuschlagen (vgl. Rehm 2015: 31).

In einem nächsten Schritt argumentieren die AutorInnen, dass die positive Wirkung des mDSM über die sprachlichen und wirtschaftlichen Aspekte weit hinausgeht. So sollen auch gesellschaftliche Herausforderungen, die im Rahmen des Horizon 2020 Programmes in Angriff genommen werden, ebenso von Sprachtechnologien profitieren. Diese Bereiche sollen etwa die nachhaltige Land- und Forstwirtschaft, Lebensmittelsicherheit, Klimaschutzaktionen oder die nachhaltige Verwendung von Ressourcen sein. Die Schaffung neuer Arbeitsplätze würde mit diesen Neuerungen ebenso einhergehen und einen Aufschwung für die Sprach- und Technologieindustrie mit sich bringen (vgl. Rehm 2015: 14). Von einem erfolgreichen mDSM erwartet sich CRACKER außerdem, dass ein Aufschwung der Sprachindustrie Hand in Hand mit einem immensen Aufschwung Europas geht. Die Ziele des DSM könnten um ein vielfaches höher gesteckt werden und die 250 Milliarden Marke für den Wirtschaftswachstum deutlich übersteigen – all das durch das intensive Einbinden der Faktoren Sprache und Technologie. In der Agenda wird sogar beschrieben, dass es vorstellbar wäre, die ausgearbeiteten Technologien und Dienstleistungen an andere mehrsprachige Staaten weiterzugeben und ein globales Wirtschaftswachstum voranzutreiben (Indien, Südafrika etc.) (vgl. Rehm 2015: 35).

Neben der Verwirklichung der drei Ebenen sieht es das Projekt als unerlässlich an, parallel zu der Umsetzung Gesetze zu erlassen, die primär der Förderung der Mehrsprachigkeit dienen. Dies würde einerseits für die mDSM selbst förderlich sein, da die Entwicklung der Sprachtechnologien durch zusätzliche rechtliche Grundlagen angetrieben werden muss, andererseits aber auch der Rolle der Europäischen Kommission als ein Organ, das Fortschritt unterstützt und den Bedürfnissen seiner EU-BürgerInnen dient, entsprechen. Veränderungen in der Rechtslage fordern die AutorInnen etwa für die veralteten Regelungen betreffend der Verwendung von für Forschungszwecke wertvollen Datenmengen wie z.B. nationalen Korpora. Auf diese kann momentan nicht zugegriffen werden, obwohl sie für die Optimierung einer Reihe von Sprachtechnologien (allen voran der MÜ) wesentlich wären. Neue nationale und europaweite Gesetze sollen helfen, die Situation zu ändern, denn entgegen der Befürchtungen verletze die Auswertung von Daten zu wissenschaftlichen Zwecken nicht die wirtschaftlichen Interessen von UrheberInnen und HerausgeberInnen der jeweiligen

Texte. Ansonsten bestehe die Gefahr, dass die Sprachtechnologieforschung an der geringen Menge an Ressourcen scheitert und von Europa nach Amerika oder China weiterzieht (vgl. Rehm 2015: 33-34).

Die Erschaffung des mDSM würde nur gelingen, wenn das Projekt über die tatkräftige Unterstützung auf nationaler Ebene (z.B. den einzelnen EU-Mitgliedstaaten) und internationaler Ebene (Kommission, EU Institutionen im Allgemeinen), sowie von Unternehmen jeglicher Größe verfügt. Der entschlossene Wille zu solchen umfassenden Reformen würde jedoch zu diesem Zeitpunkt noch fehlen, wie die AutorInnen zusammenfassend feststellen. Dabei wäre es aus ihrer Sicht für die Zukunft Europas dringend notwendig:

Language Technology made for Europe in Europe will significantly contribute to future European crossborder and cross-language communication, economic growth, and social stability while establishing for Europe a worldwide, leading position in technology innovation, securing Europe's future as a worldwide trader and exporter of goods, services, and information. Only a large, coordinated push of this magnitude will be able to unlock a truly multilingual Digital Single Market (Rehm 2015: 17).

5.3. Connecting Europe Facility

Zu den aktuellen Projekten, in die derzeit investiert wird, gehört ebenso *Connecting Europe Facility* (CEF). Dessen Ziele sind sehr weit gefächert, der Schwerpunkt liegt jedoch auf der Förderung von Europas Infrastruktur und digitaler Vernetzung. Insgesamt umfasst es drei große Themenbereiche: Verkehr, Energie und Telekommunikation. Das Projekt nahm 2014 seinen Anfang und soll planmäßig 2020 zu Ende gehen. In den Projektzielen lassen sich Ähnlichkeiten mit dem DSM erkennen und so liefert auch CEF einen Beitrag zum Gelingen des einheitlichen digitalen Binnenmarktes (vgl. <https://ec.europa.eu>^h).

Aus Sicht der Übersetzung ist der Bereich der Telekommunikation interessant. Dieser Sektor verfügt über die gesamte Projektdauer mit einem Budget von 1,14 Milliarden Euro. Ein Großteil dessen ist den so genannten *Digital Service Infrastructures* (DSIs) gewidmet, die innerhalb des CEF Programms die Verstärkung von grenzüberschreitenden digitalen Dienstleistungen, Unternehmen und Verwaltung vorantreiben sollen (vgl. <http://ec.europa.eu>^h).

Die DSIs können mit Bausteinen verglichen werden (das Projektlogo selbst besteht aus zusammengesetzten Legosteinen und suggeriert bereits diesen Vergleich), die mehrfach eingesetzt und miteinander kombiniert werden sollen. Alle Bausteine erfüllen bestimmte Standards, sodass sie von den VerwenderInnen weiterentwickelt und für ihre individuellen Projekte eingesetzt werden können. Sie sollen als Werkzeuge fungieren, mit denen der reibungslose Fluss der Dienste und der Handel über nationale Grenzen hinweg angetrieben werden kann (vgl. <https://joinup.ec.europa.eu>^a).

Jedes DSI zielt auf die Bewältigung bestimmter Aufgaben ab. Daher rührt auch, dass sich ihr Aufbau und ihr Zweck voneinander unterscheiden. Von dieser Vorgangsweise werden hohe Effizienz, Vernetzung und niedrigere Kosten auf dem europäischen Markt erwartet. In Arbeitsgruppen werden die Fortschritte aufmerksam beobachtet und unter deren Berücksichtigung jährlich weitere Schritte geplant. Nur wenn die nötige Reife erreicht wird, ist der Erhalt des CEF Budgets erlaubt (vgl. <http://ec.europa.eu>^h).

Das Projekt beinhaltet zwei Arten an DSIs: zum einen gibt es die so genannten „*building block DSIs*“, welche solch grundlegende Funktionen erfüllen, dass sie bei mehreren Elementen eingesetzt werden können oder gar als Grundlage für deren Entwicklung dienen (Europäische Kommission 2014: 74). Andere DSIs sind für die Anwendung in einem bestimmten Bereich gestaltet und eine Funktion außerhalb dessen ist an sich nicht vorgesehen („*sector specific DSI*“) (Europäische Kommission 2014: 74).

Den Unterschied zeigt etwa ein Vergleich zwischen der *building block AT* (*Automatic Translation*, CEF.AT) und dem *sector specific eHealth*. Während die maschinelle Übersetzung in jeder Anwendung eingesetzt werden kann, um sie mehrsprachig zu gestalten, sind die Funktionen des eHealth konkret an die Bedürfnisse des europäischen Gesundheitssektors angepasst. Sie haben mit ihren Projektzielen (u.a. ein digitales Netzwerk mit grenzüberschreitenden PatientInnenregistern sowie elektronische PatientInnendossiers und Rezepten) für andere Bereiche – beispielsweise eJustice – kaum bis gar keine Relevanz (vgl. Europäische Kommission 2014: 74 und <http://www.epsos.eu>). Neben „*Automated Translation*“ (auch als „*eTranslation*“ bezeichnet) gelten noch „*eDelivery*“, „*eSignature*“, „*eInvoicing*“ und „*eID*“ zu den fünf grundlegenden Bausteinen des CEF Programmes (<https://joinup.ec.europa.eu>^b).

CEF Building block DSIs

1. eDelivery
2. eID
3. eInvoicing
4. eSignature
5. eTranslation

} ...können eingesetzt werden für...

CEF Sector Specific DSIs

e-Health
Cybersecurity
European e-Justice Portal
Public Open Data
etc.

Tabelle 2. Anwendung CEF.AT (erstellt anhand von <https://joinup.ec.europa.eu>^b)

Aufgrund der Reichweite und der Berührungspunkte des Projektes, ist die DGT bei Weitem nicht die einzige Generaldirektion, die an CEF teilnimmt, denn auch DG Connect und DG Digit übernehmen weitere Rollen in der Ausführung. Im Sinne der guten Zusammenarbeit

stehen die involvierten Generaldirektionen in Kontakt, halten Gesprächsrunden ab und beraten sich gegenseitig. Prinzipiell unterstützen die MÜ-ExpertInnen der DGT die LeiterInnen der DSIs bei der Einführung des MÜ-Systems in ihre eigenen Projekte, sowie auch bei den nachfolgenden praktischen Schritten (z.B. Beurteilung und Verbesserung der Qualität, Hilfe in der Pilotphase) (vgl. Europäische Kommission 2014: 13).

Des Weiteren haben die in CEF involvierten Mitgliedstaaten die Möglichkeit, mit der DGT über mehrere Kanäle Kontakt aufzunehmen, z.B. über ein ISA-Komitee oder den jeweiligen Sprachbeauftragten in den einzelnen Staaten.

Das primäre Ziel der Generaldirektionen war es in diesem Anfangsstadium (somit bis Ende 2015) die Staaten und andere NutzerInnen auf die laufende Pilotphase aufmerksam zu machen und Feedback zu erhalten. Zudem ist es angedacht, die Situation jener Sprachen, die bisher von der MÜ benachteiligt waren, in Angriff zu nehmen (vgl. Europäische Kommission 2014: 13-14). Welche Einsatzbereiche für die CEF.AT genau vorgesehen sind und inwiefern sich die DGT an diesem CEF Baustein beteiligt, wird nun genauer untersucht.

5.3.1. CEF Automated Translation

Der *building block* AT dient zur Ergänzung der anderen vier vielfältig einsetzbaren Bausteine, um diese mehrsprachig zu machen. Nationale und internationale Behörden können sie einsetzen, um auf fremdsprachige Ressourcen zuzugreifen, eigene Informationen in anderen Sprachen verfügbar zu machen und grenzüberschreitend in ihren Landessprachen zu kommunizieren. CEF.AT ist so gestaltet, dass es in bereits vorhandene Systeme eingebaut, aber genauso auch alleine genutzt werden kann – diese flexible Anwendung entspricht dem Grundgedanken des Projektes.

Das Übersetzungssystem baut im Kern auf das DGT MT@EC auf und es wird betont, dass es sich in mehreren Punkten erheblich von MÜ-Systemen wie Google unterscheidet. Zum einen wird verstärkt auf Datenschutz und Vertraulichkeit der zu übersetzenden Texte gesetzt (man bedenke, es sollen gesundheitliche und persönliche Daten von Personen ausgetauscht werden), zum anderen kann der Korpus, aus dem das MÜ-System seine Informationen schöpft, für spezifische Textsorten und Terminologien eingestellt werden (vgl. [https://joinup.ec.europa.eu^{b\)}](https://joinup.ec.europa.eu^{b)})).

Die grundlegenden Funktionen und Kapazitäten des MT@EC müssen jedoch ergänzt werden, um den Anforderungen gerecht zu werden. Dies bedeutet, dass das System deutlich verbessert und um neue Anwendungen erweitert werden muss. Es werden gezielt Aspekte wie die Anzahl der Sprachen, die Qualität des Outputs, die Schnelligkeit der Bearbeitung und die spezifischen Einsatzgebiete des CEF einkalkuliert. Zusätzlich sollen für komplexere und datenärmere Sprachen weitere Programme entwickelt werden und diese mit

MT@EC gekoppelt zum Einsatz kommen, um so die Effizienz zu steigern (vgl. Kimmo Rossi, 04.12.2015; <https://scic.ec.europa.eu>).

Neben der MÜ setzt sich CEF.AT aus einer Bandbreite an Sprechressourcen zusammen wie Textkorpora, TMs und Nomenklaturen. Die Übersetzung erfolgt auf zwei Arten: 1) für die Übersetzung bestimmter Sprachenpaare generiert die AT unmittelbar eine Übersetzung, 2) ansonsten müssen NutzerInnen auf eine Art Relaismethode zurückgreifen, bei der zuerst in eine Sprache wie etwa Englisch und aus dieser erst in die gewünschte Zielsprache übersetzt wird. Der Grund dafür hängt mit der SMÜ und seinen bisher trainierten Sprachrichtungen zusammen (vgl. Europäische Kommission 2014a: 11).

Im Rahmen des CEF Programmes wurde MT@EC das erste Mal in einer realen Auftrags-situation getestet. Wie Kimmo Rossi berichtet, wurde das MÜ-System für die Übersetzung der im Open Data Portal gesammelten Dokumente eingesetzt. Für den enormen Datensatz benötigte MT@EC beinahe ein Monat Zeit, weshalb ersichtlich wird, dass es zwar eine gute Basis darstellt, allerdings Komponenten wie die Kapazitäten des Systems, die Computerressourcen und die Qualität zukünftig noch verbessert werden müssen. So soll für CEF das SMÜ-System von Moses um regelbasierte Elemente ergänzt werden, damit zusammengesetztere grammatikalische Konstruktionen von der Maschine besser gehandhabt werden können (vgl. Rossi, 04.12.2015, <https://scic.ec.europa.eu>). In der Pilotphase haben bereits zahlreiche mitgliedstaatliche Behörden und ihre MitarbeiterInnen über sTesta Zugriff auf den Service und machen von diesem auch Gebrauch: laut den ersten Zahlen ist MT@EC bereits für die Übersetzung von mehr als 7 Millionen Seiten verantwortlich (vgl. Europäische Kommission 2014a: 13).

5.3.2. Der DSI Bericht

Einen ersten Eindruck über die Fortschritte des Projektes gibt der von Deloitte durchgeführte Bericht *CEF DSI Maturity Study*. Er wurde von der Kommission in Auftrag gegeben und von einem unabhängigen ExpertInnenteam erstellt. Er sollte vor allem die Frage danach beantworten, wie ausgereift die bisher entwickelten Bausteine sind (Gegenstand ist nicht nur der CEF.AT) und anhand dessen Prognosen für den weiteren Verlauf erstellen. Der Bericht dient der Kommission folglich als Hilfsmittel, um fundierte Entscheidungen bezüglich der zukünftigen Gesamtstrategie fällen zu können – besonders für den Arbeitsplan von 2015 und 2016 sollen Rückschlüsse gezogen werden.

Im Rahmen von Interviews mit den Organisatoren der DSIs und der Analyse von Fortschritten und der dafür gesammelten Beweise und Informationen, verteilte der Bericht Noten und gibt Auskunft darüber, welche Aspekte noch verbesserungswürdig sind. Als letztes evaluierten die LeiterInnen der Studie, ob die Bausteine für den im Jahre 2015 intendierten

Gebrauch bereit waren und welche Prognosen zu erwarten wären (Europäische Kommission 2014a: i-ii).

Die Erwartungen an die AT gehen davon aus, dass sie zunächst in zwei bis drei der anderen DSI Elemente integriert werden kann. Verläuft der Start erfolgreich, wird in einer nachfolgenden Phase die Zahl an DSIs um weitere zwei bis drei erhöht (vgl. Europäische Kommission 2014a: ii). Wie Abbildung 10. zu entnehmen ist, wurde die erste Phase 2015 durchgeführt.

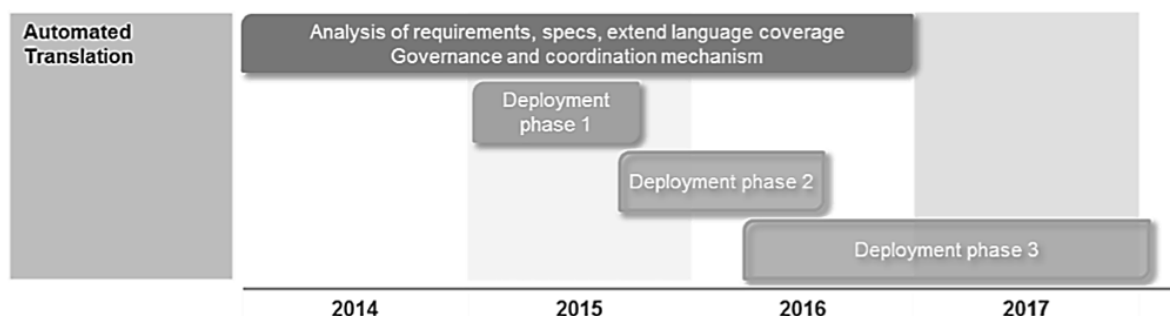


Abbildung 10. Geplanter Ablauf der CEF.AT Einführung (Europäische Kommission 2014a: 17)

Am Beginn stehen vorbereitende Maßnahmen und die Anwendung des MT@EC im Rahmen von erst wenigen DSIs. Darunter ist zu verstehen, dass bereits weiter fortgeschrittene DSIs nach ihren Ansprüchen auf Mehrsprachigkeit analysiert werden (Welche Sprachen werden am häufigsten nachgefragt und in welchem Umfang? etc.). Folglich ist die erste Phase von der Durchführung von zahlreichen Studien, den Analysen von Pilotprojekten und der Kontaktaufnahme mit den EntwicklerInnen und NutzerInnen geprägt, um auf diese Weise so viele Informationen wie nur möglich zur Verfügung zu haben. Anhand dieser werden die nächsten Maßnahmen abgeleitet und der Ablauf der zweiten Phase konkretisiert (vgl. Europäische Kommission 2014a: 14).

In der mittleren Phase steht die Erweiterung und Festigung der in der ersten Phase erreichten Meilensteine im Mittelpunkt. Die AT und MT@EC Systeme sollen weiter ausgebaut und zu diesem Zweck zusätzliche MÜ-Programme und Laufwerke erworben werden. Mit deren Hilfe können mehr Sprachressourcen gespeichert, aufgearbeitet und die Qualität der bestehenden Systeme verbessert werden. Der Erwerb ist nicht zuletzt aber auch deshalb notwendig, um das Angebot für mehr als nur drei DSIs zur Verfügung stellen zu können, denn das Ziel ist es, in der zweiten Phase bereits an die sechs DSIs mit automatischen Übersetzungen ausstatten zu können (vgl. Europäische Kommission 2014a: 14).

Schließlich soll in den kommenden Jahren die dritte und letzte Phase umgesetzt werden. Diese beinhaltet die Einführung einer komplexen Plattform, die mehrere Funktionen in sich vereint. Sie umfasst einen MÜ-Service, der fähig ist, die verschiedensten DSIs mit einer

großen Menge an Übersetzungen zu versorgen. Gleichzeitig dient sie als Speicher mit Ressourcen aus sämtlichen Sprachen. Die Intention ist, alle Sprachen und nicht allein die üblichen Verdächtigen für die MÜ zu trainieren. Daneben soll es gemäß dem Plan möglich sein, dass die Plattform als Produzent für neue MÜ-Systeme, unabhängig von Themengebiet und Sprachenpaar, fungiert (vgl. Europäische Kommission 2014a: 14-15).

Innerhalb der Studie wurden für die einzelnen ausgewerteten Aspekte Noten vergeben, die von der niedrigsten Stufe „*not mature*“ bis hin zu der besten Stufe „*fully mature*“ reichen und anhand derer ersichtlich wird, welche Teilgebiete bis zu der ersten Einsatzphase noch intensiver Entwicklung und Vorbereitung bedürfen (Europäische Kommission 2014a: 6-7). Aus dem Bericht und seinen Ergebnissen erschließt sich, dass das CEF Projekt vielversprechend voranzugehen scheint. Die Deloitte-BegutachterInnen bescheinigen in dem Bericht dem AT Block einen guten Fortschritt (*advanced maturity*) und unterstützen den weiteren Erhalt von Förderungen, damit die Kapazitäten weiter ausgebaut werden können (Europäische Kommission 2014a: ii). Neben DSIs wie eHealth und eJustice gehöre AT zu den am weitesten entwickelten Bausteinen. Zu diesem Entschluss gelangten die ExpertInnen durch die Feststellung, dass es in der Durchführung und den Testphasen bloß geringe Lücken aufweist – diese könne man in den folgenden Umsetzungsphasen füllen. Vorgeschlagen werden die Ausarbeitung weiterer noch ungeklärter Details, z.B. die Frage danach, für welche DSIs der Dienst zuerst angewendet werden soll und wie Kontrollen bezüglich der Organisation des Projekts und der Technologie gestaltet werden sollen (vgl. Europäische Kommission 2014a: 15, v-vi).

Es sind Pilotphasen angedacht, in denen die Funktionsfähigkeit weiter getestet werden kann (die Durchführung erfolgt in einzelnen Mitgliedstaaten). Diese Projekte sind die ersten wesentlichen Indikatoren dafür, dass die Dienste den operativen und technischen Ansprüchen gerecht werden und ihnen die finanziellen Mittel auch in Zukunft zustehen (vgl. Europäische Kommission 2014a: 13). Im Zuge dessen wurde festgestellt, dass es bei der technischen Umsetzung kaum Probleme gibt. Die AutorInnen zeigten sich auch mit der Leitung des Projektes zufrieden und gaben ihr die Note *advanced maturity*. Dieser Teil der Analyse bewertete u.a. ob und in welcher Art und Weise die Mitgliedstaaten in die Entwicklung mit einbezogen sind bzw. waren und als aktive Mitgestalterinnen agieren. Weitere Aspekte betrafen Organisation und das Management des Projektes. Während die Beurteilung gut ausfiel, sei für das Gelingen letztendlich noch eine stärkere Einbindung der Mitgliedstaaten und anderer wichtiger Akteure sowie die gründliche Erarbeitung von Datenschutzbestimmungen erforderlich (vgl. Europäische Kommission 2014a: 15-16).

Eine Umfrage ergab, dass gerade EU-Institutionen einen Service, wie jenen des CEF.AT, sehr schätzen und eindeutig einen Bedarf dafür zu verzeichnen ist. Der Nutzen wird vor

allem im Bereich der Übersetzung EU-Rechts gesehen, genannt werden jedoch auch offizielle Webseiten von nationalen Regierungen und dem Amt für Veröffentlichungen der Union. Man erhofft sich Vorteile aus der Echtzeit-Übersetzung von Inhalten in und aus allen Amtssprachen. Weitere geäußerte Wünsche waren beispielsweise, dass eine weite Bandbreite an Formattypen, allen voran auch PDFs, zur MÜ-Abgabe möglich sein soll und die Übersetzung in jedem Fall mit dem ursprünglichen Layout des ausgangssprachlichen Dokumentes an die AuftraggeberInnen zurückgelangt.

Trotz der grundsätzlich positiven Rückmeldungen, berichteten Institutionen dennoch über Qualitätsmängel. Dies scheint der Grund dafür zu sein, warum AT theoretisch begrüßt wird, aber die befragten EU-Institutionen dem Projekt ihre Unterstützung nicht ohne Vorbehalte ausgesprochen haben (vgl. Europäische Kommission 2014a: 17). Nichtsdestotrotz bestärken die aktuellen Ergebnisse den bestehenden Plan und die Implementierung der AT wird weiterverlaufen (vgl. Europäische Kommission 2014a: 18).

Als Beispiel für zwei wichtigen CEF.AT NutzerInnen innerhalb der Europäischen Kommission gehören IMI und SOLVIT. Diese haben bereits ECMT genutzt und werden das neue SMÜ-System auch in Zukunft stark in ihre Arbeit einbinden. Beide Organisationen haben bereits Zugriff auf die neue Version und sind in die Testphase eingebunden (Europäische Kommission 2014a: 13).

Betrachtet man ihre Funktion, wird deutlich, weshalb in ihrem Fall der MÜ eine bedeutende Rolle zukommt. IMI (*Internal Market Information System*) gilt auf europäischer Ebene dem Austausch von Informationen zwischen Behörden. Sie dient dem Zweck, das Prinzip der Personenfreiheit in der Praxis zu unterstützen, indem beispielsweise Informationen über Personen, Unternehmen und zu der Anerkennung von Qualifikation oder arbeitsrechtlichen Bestimmungen verschiedener Mitgliedstaaten eingeholt werden können. Die staaten- und sprachenübergreifende Kommunikation wird nun durch die Nutzung von CEF.AT sichergestellt (vgl. <http://ec.europa.eu>^k). Eng mit der Tätigkeit von IMI hängt auch die zweite Organisation SOLVIT zusammen, die mittlerweile ebenso auf die Nutzung von MT@EC umgestiegen ist. Die Behörde ist für die Handhabung von Beschwerden wegen Verletzung des EU-Rechtes zuständig. Diese können von sämtlichen BürgerInnen der Union eingereicht werden, die in einem anderen Mitgliedstaat mit solchen Problemen konfrontiert sind und können somit in jeder einzelnen der Amtssprachen verfasst werden. Die für SOLVIT relevanten Themengebiete überschneiden sich teilweise mit jenen von IMI, so sind hier etwa ebenso arbeitsrechtliche Fragen von Bedeutung oder etwa auch rechtliche Fragen zu Visa und Versicherungen (vgl. <http://ec.europa.eu>^l).

5.3.3. European Language Resource Coordination (ELRC)

Das *European Language Resource Coordination* Projekt ist innerhalb der CEF tätig und trägt mit seiner Arbeit zu der AT Funktion bei. ELRC hat sich als Aufgabe gesetzt, aus den Mitgliedstaaten (sowie Norwegen und Island) die größtmögliche Menge an notwendigen Daten zusammenzutragen, die im Anschluss für die automatische Übersetzung als Grundlage dienen. Dies umfasst eine Tournee durch die dreißig teilnehmenden Staaten, bei der Präsentationen und Workshops abgehalten werden. Die unterschiedlichen Organisationen (die Spannweite reicht von lokalen Verwaltungen, nationalen Ministerien bis hin zu internationalen Organisationen, die in Europa tätig sind) und Regierungen sollen hiermit auf das Projekt, rechtliche Hintergründe und die Relevanz eines mehrsprachigen Europas aufmerksam gemacht werden. Die „Reise“ nahm in Athen ihren Anfang, 2016 folgen weitere Stationen u.a. in Portugal, Stockholm und zuletzt auch in Wien. Die davon erhoffte Reaktion ist eindeutig: Die Tournee soll mit der Unterstützung und der Weitergabe von wertvollen Sprachdaten im Rahmen des CEF Projektes zu Ende gehen (vgl. <http://elrc.tilde.com>).

In einem offenen Brief richtet sich die ELRC Gruppe an die Europäische Kommission und macht auf ihre Anliegen aufmerksam. Eines davon, dem durch das Sammeln der Sprachdaten enorm entgegenwirkt werden soll, betrifft kleinere Sprachen, die von einem digitalen Aussterben bedroht sind. Darunter sind jene Sprache zu verstehen, die – wie auch schon im mDSM Bericht angesprochen, im World Wide Web kaum vertreten sind und deren SprecherInnen digital benachteiligt sind. Informationen, die nur in einer Sprache vorhanden sind, bleiben somit einem Großteil der EU BürgerInnen verborgen. Der offene Brief betont deshalb, dass bisher die Auswahl an automatisierten Übersetzungsmöglichkeiten stark von internationalen, außereuropäischen Unternehmen dominiert war, die ihr Hauptaugenmerk folglich auf das Englische gerichtet hatten. Die Aufgabe der Kommission und der Projekte sei es nun, den Markt aktiv mitzugestalten und dies könne nur mit dem Einsatz von Sprachtechnologien geschehen. Die Unterzeichneten plädieren dafür, grundlegende Technologien für alle BürgerInnen, Unternehmen sowie EU-Organisationen bereitzustellen, damit diese in Folge auf dieser Basis weitere Technologien und Anwendungen entwickeln können (vgl. <http://www.multilingualeurope.eu>).

ELRC steht den nationalen Verwaltungen bei der Vorbereitung und Bereitstellung der Daten zur Seite und bietet auch bei Unklarheiten bezüglich technischer Fragen Hilfestellung. Zum einen gibt es zu diesem Zweck auf der offiziellen Webseite ein Helpdesk, zum anderen wurde das *Language Resources Board* (LRB) gegründet, um den Ländern allgemein bei der Teilnahme an ELRC Projekten zu helfen. Auf diese Weise möchte man den Kontakt zu den Staaten und lokalen EntscheidungsträgerInnen aufbauen bzw. aufrechterhalten und sie in den Prozess einbinden (vgl. <http://elrc.tilde.com>).

Natürlich spielt es eine wichtige Rolle, welche Art von Daten die ELRC sammelt. Wie bereits erläutert, funktionieren SMÜ-Systeme im Kontext der EU mit domänen-spezifischen Datenspeichern am besten, doch auch allgemeinere Sprachdaten können zum sprachlichen Wissen der Maschine beitragen. Werden jedoch Ressourcen zu einem bestimmten (Fach)Gebiet weitergegeben, dann sind natürlich vor allem CEF Themenbereiche am hilfreichsten (somit Gesundheit, Kultur, Recht, soziale Sicherheit, Auftragswesen, Verbraucherrechte etc.) (vgl. <http://lr-coordination.eu>). Folglich sind öffentliche Verwaltungen, NGOs, staatliche Organisationen und Co. dazu aufgefordert, folgende Arten an Sprachressourcen in verwertbarem Format (somit z.B. kein PDF) zur Verfügung zu stellen:

- **Translation memories:** linguistic databases that capture translations made by humans. They can be used to facilitate future translations tasks but also for training automated translation systems
- **Translation/language models:** statistic information that assigns a probability to a piece of unseen text, based on some training data
- **Corpora:** monolingual and multilingual corpora, comparable, aligned, parallel documents
- **Lexica:** monolingual and multilingual lists of words, multi-words, sentences, etc. in general or specific subject fields
- **Terminological resources:** structured sets of concepts, with associated linguistic information in a specific subject field
- **Grammars:** sets of rules that formalize a language
(<http://lr-coordination.eu/resources>)

Um die oben aufgelisteten Daten an die ELRC weiterzugeben, wurde die so genannte SHARE Plattform errichtet. Auf dieser müssen sich NutzerInnen mit Angabe ihrer Organisation registrieren und können als nächsten Schritt bereits Dokumente versehen mit einer Beschreibung hochladen (vgl. <http://elrc-share.ilsp.gr/>).

5.3.3.1. Open Data Portal

Ein Beispiel für die Bemühungen, die gemeinsamen europäischen Sprachressourcen aufzustocken, ist das Open Data Portal. Die Erschaffung eines solchen offenen Datenportals, das die Dokumente der EU-Institutionen enthält, ist ganz im Sinne des ELRC, des CEF und mDSM Programmes und wird der Idee, Daten zu sammeln, sie zugänglich und für verschiedenste Zwecke nützlich verwendbar zu machen, gerecht. Die über die Website verfügbaren Datenmengen sind für alle InteressentInnen kostenlos zugänglich – von BürgerInnen und Unternehmen bis hin zu ForscherInnen. Ziel ist auch hier neue, innovative Entwicklungen zu befähigen, durch den freien Informationszugriff die Vernetzung voranzutreiben und infolgedessen nicht zuletzt auch „wirtschaftliches Potenzial freisetzen“ (<https://open-data.europa.eu>^a).

An die 25 Mitgliedstaaten folgen derzeit dem Beispiel der europäischen Plattform: Sie machen ihre nationalen Daten auf eigenen Seiten zugänglich und speisen mit ihren Sammlungen damit gleichzeitig das europäische Kontingent. So bietet beispielsweise die österreichische Regierung Datensätze zu 14 verschiedenen Themengebieten, u.a. Wirtschaft und Tourismus, Arbeit, Verkehr und Technik, Bevölkerung, Land und Forstwirtschaft, Bildung uvm. an. Die VerfasserInnen sind etwa die Stadt Wien oder die Statistik Austria, im Allgemeinen kann jedoch gesagt werden, dass die Datensätze von Gemeinden bis hin zu dem Bund stammen, oder eben einzelnen Organisationen (vgl. <https://www.data.gv.at/>).

Eine kurze Erläuterung auf der Website illustriert, wie sich die Union die praktische Anwendung vorstellt: „Beispielsweise könnte eine Anwendung Daten zu den wichtigsten Wirtschaftszweigen, Schadstoffemissionen und Bildungseinrichtungen kombinieren und Ihnen so bei der Auswahl ihres Wohnortes helfen.“ (<https://open-data.europa.eu>^b) Daten sollen also nicht nur separat eingesetzt werden, sondern Themenbereiche miteinander verknüpft werden (*linked data*), um so Zusammenhänge auf europaweiter Ebene zu erkennen (vgl. <https://open-data.europa.eu>^b). Über das Open Data Portal kann unter anderem auch die Translation Memory der DGT eingesehen und genutzt werden.

5.4. CAT Tool Projekte

Der Fokus der EU-Projekte liegt auf der maschinellen Übersetzung, da man sich durch ihre Optimierung eine effiziente und funktionsfähige Echtzeit-Übersetzung für die digitale Vernetzung erhofft. Andere Tools, die primär zur Unterstützung von professionellen ÜbersetzerInnen dienen, können mit diesen Anwendungen in diesem Sinne nicht dienen, verfügen aber dennoch über praktische Bedeutung. Aus diesem Grund sollen hier einige interessante, von der EU-finanzierte Projekte präsentiert werden. Zurzeit läuft keines der CAT Tool Projekte mehr, deren Realisierung liegt jedoch nicht lange zurück.

5.4.1. Casmacat

Zu erwähnen ist hier CASMACAT, ein Forschungsprojekt, das im Rahmen des *Framework Programme 7* von der Kommission mitunterstützt wurde und bis Oktober 2014 über einen Zeitrahmen von drei Jahren lief. In der ersten Phase wurden ÜbersetzerInnen nach kognitionswissenschaftlichen Aspekten bei der Arbeit mit CAT Tools beobachtet und Aufzeichnungen über deren Augenbewegungen („*eye tracking*“) und Tastennutzung („*key logging*“) gemacht (<http://www.casmacat.eu/>).

Die Auswertung der Ergebnisse wurde im nächsten Schritt dazu verwendet, aufbauend auf diese Analyse neue Tools zu entwickeln, die auf das Verhalten während dem Übersetzungs-

prozess Rücksicht nehmen. Das Endergebnis war ein Open Source Programm eines Translator's Workbench, der über einen integrierten Editor, Server und über Werkzeuge für Analyse und Visualisierung verfügte. Diese Elemente sind als Module gestaltet, sodass sie nach Wunsch eingesetzt werden können, u.a. auch integriert in bereits vorhandenen CAT Tools (vgl. <http://www.casmacat.eu/>).

Eine der Besonderheiten, die CASMACAT bietet, ist die Lernfähigkeit des Systems. Der Workbench merkt sich die von den NutzerInnen getroffenen Entscheidungen und erneuert anhand dieser automatisch sein Wissen. Es handelt sich dabei nicht bloß um einen Übersetzungsspeicher, sondern um viel mehr um ein Training, ähnlich wie das bei SMÜ Systemen.

Neben TMs sind auch hier MÜ-Elemente vorhanden, um den ÜbersetzerInnen interaktiv Vorschläge zu machen, wie sie ihre Sätze beenden könnten. Die Software verwendet die neuesten MÜ-Errungenschaften, assistiert bei dem Lektorat und informiert NutzerInnen über die Vertrauenswürdigkeit seiner Angaben (vgl. <http://www.casmacat.eu/>).

5.4.2. Matecat

Neben CASMACAT wurde ein weiteres CAT Tool namens MATECAT unterstützt. Es handelt sich um ein für sämtliche NutzerInnen kostenloses CAT Tool, das als online Plattform funktioniert. Entwickelt wurde das Projekt ebenso über einen Zeitraum von drei Jahren und endete 2014. Die Website verspricht, dass mittels der Ressourcen, auf die das Programm zugreift, an die 10 bis 20% mehr Matches vorgeschlagen werden, was letztendlich zu einer schnelleren Arbeitsweise führen soll. Dafür werden neben den persönlichen TMs der jeweiligen ÜbersetzerInnen auch maschinelle Übersetzung und das TM von MATECAT verwendet, dessen Speicher mehr als 12 Milliarden Wörter enthält.

Im Bereich des Projektmanagement bietet es eine einfache Bedienung. Ein Ziel des Online Tools ist es, dass gleichzeitige Arbeit von mehreren ÜbersetzerInnen an demselben Projekt mit einem gemeinsamen Glossar und einer gemeinsamer TM ermöglicht werden soll, inklusive der Option, Probleme in einem Chat zu klären. Die NutzerInnenfreundlichkeit steht im Mittelpunkt, so steht zudem ein weitreichender technischer Service zur Verfügung (vgl. <http://www.matecat.com/benefits/>).

6. Interview

Das nachfolgende Kapitel enthält Hintergründe zu dem im Rahmen der Arbeit geführten Expertinneninterview. Zunächst folgen Erläuterungen zu der Wahl der Interviewform und der Fragen, sowie die Vorstellung der Interviewpartnerin Katja Mai. Den Abschluss bildet eine Zusammenfassung der wichtigsten Konklusionen. Das Transkript des gesamten Gespräches befindet sich im Anhang dieser Arbeit.

6.1. Grundlagen

Bei dem Gespräch handelt es sich um ein qualitatives ExpertInneninterview. Es zeichnet sich durch eine offene Fragenstellung aus und ist durch ein Sachinteresse der interviewenden Person geprägt. Es ist so gestaltet, dass das Gespräch zwar gelenkt wird, aber die befragte ExpertIn zu einer freien Rede veranlasst wird (vgl. Mieg & Näf 2005: 4-5, 9). Im Gegensatz zu geschlossenen Fragen (wie sie z.B. in einem Fragenbogen aufzufinden sind) werden der befragten Person nämlich keine festgelegten Antwortmöglichkeiten zur Auswahl vorgelegt und auf diese Weise die Möglichkeit zum selbstständigen formulieren und argumentieren geboten. Diese Methode gehört zu den häufigsten im Rahmen qualitativer Forschung (vgl. Atteslander 2006: 136, 129). Die Wahl fiel deshalb auf den offenen Fragentyp, da dadurch bei der Beantwortung viel Spielraum gelassen werden und die Interviewpartnerin die Aufmerksamkeit auf wichtige, noch unbekannte Aspekte lenken konnte.

Als Gesprächspartnerin wurde zu diesem Zweck Frau Katja Mai ausgewählt. Frau Mai ist Übersetzerin in der Deutschabteilung der DGT in Brüssel und hat sich nach der Kontaktaufnahme sogleich zu einem Interview bereit erklärt. Sie ist bereits seit April 2006 Teil der Abteilung und arbeitet mit den Sprachen Englisch, Französisch und Italienisch (dem Muttersprachenprinzip folgend entstehen die Übersetzungen jeweils ins Deutsche). Vor ihrer Arbeit in der EU war Frau Mai als freiberufliche Übersetzerin tätig.

Mit Frau Mai wurde auf Empfehlung von Herrn Achim Braun, Sprachbeauftragter der österreichischen Außenstelle der DGT, Kontakt aufgenommen. Sie war auf Grund ihrer Position als Übersetzerin in der DGT, ihrer Teilnahme an Arbeitsgruppen zu Übersetzungstools und ihrer langjährigen Erfahrung eine passende Interviewpartnerin für diese Masterarbeit. Im Laufe der letzten Jahre hat sie als Übersetzerin die Entwicklungen in der DGT und den verwendeten Sprachtechnologien aktiv und vor Ort miterlebt und konnte einen Einblick in die ersten Schritte von MT@EC aus dem Blickwinkel einer professionellen Übersetzerin bieten. Damit erfüllt sie in jeder Hinsicht das in der Forschung geltende Kriterium, das Personen als „ExpertInnen“ qualifiziert: darunter ist eine Person zu verstehen, „die aufgrund langjähriger Erfahrung über bereichsspezifisches Wissen/Können verfügt“

(Mieg & Näf 2005: 7). Frau Mai verfügt durch ihren beruflichen Hintergrund über ein weites Fachwissen sowie Praxiserfahrung und ermöglichte es so, hinter die Kulissen und Fakten blicken zu können.

Die Form des Experteninterviews eignete sich somit besonders, da es wichtig war, persönliche Erfahrungen einer Person einfließen zu lassen, die in die fortlaufenden Entwicklungen involviert ist. Zudem zeigte die Interviewpartnerin eine hohe Bereitschaft Einblick in ihre Arbeit zu geben. Diese Aspekte bildeten eine gute Ausgangslage für ein adäquates Leitfadeninterview (vgl. Atteslander 2006: 131).

Für ein bestmögliches Gelingen des Gespräches ist es notwendig, dass sich die interviewende Person gründlich vorbereitet, um mit den ExpertInnen auch interagieren zu können. Denn können InterviewerInnen den Ausführungen nicht folgen oder keine adäquaten Fragen stellen, wirken sie dadurch als Laien – ein Umstand, der sich auf die Antworten und das Gespräch als Ganzes nachteilig auswirkt (vgl. Mieg & Näf 2005: 9-10).

In diesem Sinne wurden in der Vorbereitung diese Aspekte berücksichtigt. Einerseits war das Verfassen der Masterarbeit zum Zeitpunkt des Interviews bereits im Gange, wodurch eine Vertrautheit mit wichtigen Sprachtools, den Termini und der DGT gegeben war. Andererseits wurden zusätzliche Materialien, die im Zuge der Kontaktaufnahme erhalten wurden, durchgelesen und in der Fragenstellung beachtet. Mit diesem Vorgehen sollte sichergestellt werden, dass wesentliche Terminologie und Sachverhalte bekannt sind und dass keine Fragen gestellt werden, die sich auch einfach unter Herannahme von Broschüren und Literatur beantworten ließen.

Vor der Durchführung des Interviews wurde anhand der Interessensgebiete ein Leitfragenkatalog erstellt. Diese Liste bildet den Leitfaden eines ExpertInnengespräches und hat „das mehr oder weniger grob strukturierte schriftliche Frageschema, welches dem Interviewer bei der Interviewführung als Gedächtnisstütze dient“ zu enthalten (Mieg & Näf 2005: 14). Somit wurden die Fragen gesammelt, anschließend thematisch und nach ihrer Priorität geordnet.

Im Vorfeld wurden Frau Mai die Leitfragen des Interviews zugesendet, um ihr bereits einen Überblick über den Verlauf und die Themenschwerpunkte des Gespräches zu ermöglichen. Die im Katalog enthaltenen Fragen wurden mit dem Ziel ausgewählt, neben den dargelegten idealen Abläufen, Zahlen und Fakten einen Einblick in die realen Abläufe und Umstände der Praxis des DGT Übersetzens zu erhalten. Ein Großteil des Materials, das sich mit den MT@EC und Co. Projekten und den innerhalb dessen entwickelten Sprachtechnologien beschäftigt, geht auf diese Aspekte kaum ein. Sie erweisen sich jedoch als bedeutend, da die Verwendung von Tools und eine tägliche Arbeit mit ihnen zwangsläufig zu einem anderen Blickwinkel führt, als dies bei ForscherInnen einer Studie oder bei der Betrachtung der the-

oretischen Funktionsweise der Fall ist. Im Fokus stand bei dem Gespräch somit, die Literatur mit der Realität zu vergleichen, Beispiele aus der Praxis zu sammeln, Sachverhalte zu klären und zu vertiefen sowie neue Erkenntnisse aus der Erfahrung einer professionellen Übersetzerin zu ziehen, die an den (sprach)technologischen Entwicklungen beteiligt ist (somit im wahrsten Sinne über *hands-on experience* verfügt). Folglich wurden an die Interviewpartnerin bewusst auch Fragen gestellt, die sich auf ihre ganz persönliche Einschätzung und Meinung richteten.

Gemäß dem Fokus der Arbeit betrafen die Fragen die Sprachtechnologien und ihre Anwendung, sich daraus ergebende Vor- und Nachteile, sowie Zukunftspläne. In die Fragen flossen neben eigener Recherche auch die von Frau Mai und Herrn Braun im Vorfeld zur Verfügung gestellten Materialien ein. Das im Anhang festgehaltene Interview wurde am 22. Dezember 2015 telefonisch durchgeführt und aufgenommen. Die Dauer des Gespräches betrug ca. 40 Minuten.

6.2. Konklusion und Zusammenfassung

Für die Auswertung des Interviews wurden die Aussagen nach ihrer Thematik gesammelt und zusammengefasst, um so die jeweilige Kernaussage herausfiltern zu können. Anhand dieser Feststellungen konnte ein Vergleich zwischen den zuvor in der Fachliteratur zusammengetragenen Meinungen und der Erfahrung von Frau Mai vorgenommen werden.

Die gegebenen Antworten betreffen mehrere Themengebiete, die in der vorliegenden Arbeit angesprochen werden. Somit wird auf Mais Aussagen bereits in mehreren Kapiteln der Masterarbeit verwiesen (siehe etwa zu Besonderheiten der EU-Sprache, Textsorten in der Kommission, Translator's Workbench oder zu der Rolle der ÜbersetzerInnen). Aus diesem Grund soll an dieser Stelle keine gesonderte ausführliche Analyse erfolgen, sondern bloß die entscheidenden Erkenntnisse hervorgehoben werden.

- Im Hinblick auf die Textsorten geht aus dem Gespräch die Dominanz von Texten mit grundlegend rechtlicher Thematik deutlich hervor. Die Ausprägung der Rechtstexte ist jedoch sehr unterschiedlich, da die darin behandelten Sachverhalte verschiedener Natur sind und jeweils für andere Zwecke bestimmt sind (Kommunikation mit Kommissionsmitgliedern, mitgliedstaatlichen Behörden, Informationen an BürgerInnen, Berichte mit rechtlichen Elementen). Der Eindruck, dass nur klassische Gesetzestexte übersetzt werden, erwies sich als irreführend, denn innerhalb der Kategorie „EU-Rechtstext“ gibt es ein breites Spektrum an Varianten, Zielgruppen und Stilen.
- Bezüglich der Tools überraschte die Bemerkung, dass es zwar ein verbreitetes Standardtool gibt (SDL Trados), deren Verwendung den ÜbersetzerInnen jedoch nicht „vorgeschrieben“ wird und deshalb KollegInnen mitunter mit anderen Systemen (wie OmegaT) arbeiten. Diesbezüglich wurde von Frau Mai allerdings angesprochen, dass diese

Freiheit bei der Wahl des bevorzugten Tools natürlich an unvermeidbare Probleme gekoppelt ist, da eine Einheitlichkeit der Übersetzungen deutlich schwieriger zu erreichen ist. Zu denken ist etwa an eine erschwerte Zusammenarbeit bei umfangreicheren Aufträgen, an denen mehrere ÜbersetzerInnen gleichzeitig arbeiten oder die nachträgliche Alignierung für die ordnungsgemäße Archivierung im Euramis Speicher.

Für die MT@EC Nutzung in der DGT ist dies zusätzlich aus jenem Grund zu beachten, dass die maschinelle Übersetzung nicht nur als ein gesondertes Programm eingesetzt wird, sondern innerhalb der Übersetzungsdirektion vorrangig als ein in ein CAT Tool integriertes Werkzeug verwendet wird. Bei der Einführung und der Anpassung an bereits bestehende Tools könnte dies unter Umständen eine zu beachtende Rolle spielen.

- Im Zusammenhang mit den Sprachtechnologien bestätigte sich die Annahme, dass ein idealer Ablauf nicht immer gewährleistet werden kann. Bei einer Modernisierung müssen etwa bestehende Lizenzen beachtet werden. Auftretende Fehler und Verzögerungen haben für den Ablauf „Nebenwirkungen“: rechtliche Vorgaben schränken das Potenzial einiger Tools ein oder führen zu zeitlichen Verschiebungen, die Implementierung aktueller Versionen bedarf eines großen Aufwandes. Die Umsetzung und erfolgreiche Integration der Tools in den tatsächlichen Arbeitsablauf ist mit vielen Hürden verbunden – dies ist ein Aspekt, der auch in den momentanen und zukünftigen Projekten miteinberechnet werden muss.
- Aufgrund des Interviews konnte weiters der Workflow innerhalb der DGT besser beurteilt werden und daraus zusätzliche Schlüsse in Bezug auf die Qualitätssicherung und den Ablauf der Zusammenarbeit zwischen den einzelnen MitarbeiterInnen gezogen werden.
- Die aus der Literatur erarbeiteten Informationen konnten durch aufschlussreiche Beispiele aus dem tatsächlichen Übersetzungsalltag der DGT ergänzt und unterlegt werden. Die Entscheidung, welches Sprachtool zum Einsatz kommen soll, illustrierte Frau Mai an zwei aktuellen Projekten, der Al-Qaida Verordnung und dem Europäischen Semester. Ihre Beschreibungen hinsichtlich der MT@EC Sprachen und den unterschiedlichen Ergebnissen haben die in der MT@EC Konferenz der Kommission angesprochenen Probleme und die Grundlagen des Medaillen-Qualitätssystems verdeutlicht und konkrete Beispiele geliefert, in denen die maschinelle Übersetzung besser in den Übersetzungsprozess eingebaut werden kann.
- Die im Gespräch geschilderte Situation der ÜbersetzerInnen in der DGT bzw. die Arbeit betreffende Zukunftsprognosen wurden den anderen ExpertInnenberichten gegenübergestellt. Frau Mais Beobachtungen bestärken diese teilweise und spiegeln die in der Literatur beschriebenen Vermutungen wieder: Sie spricht den stetig steigenden Druck sowie die Tendenz zu der noch intensiveren Nutzung von Übersetzungstools an – zwei Aspekte, die ohne Zweifel die Arbeit stark beeinflussen, laut Befürchtungen der Interviewpartnerin aber eher in die negative Richtung gehend. Im Gegensatz dazu konnten

in der Literatur auch positivere Einstellungen gelesen werden. Speziell auf die DGT bezogen stellt Mai ebenso wie Hartmann (2009: 70) fest, dass bestimmte Direktionen mehr Gewicht haben als andere und dies indirekt Auswirkungen auf die DGT hat.

- Das Interview enthält auch nützliche Informationen zu der Realität der Verfahrenssprachen. Wie vermutet, setzt sich in diesem Bereich das Englische extrem stark durch und verdrängt das Französische. Für ÜbersetzerInnen wie auch Frau Mai bedeutet das, dass ihre Ausgangstexte zu ca. 90% auf englischen Ausgangstexten beruhen. Diese geäußerte Prozentzahl bestätigt die Wichtigkeit von stilistisch und grammatikalisch besseren AT Texten und die Nützlichkeit von Kampagnen, die besonders für MitarbeiterInnen der Kommission mit einer nicht-englischen Muttersprache gedacht sind.

Zusammenfassend kann gesagt werden, dass die aus dem Interview gewonnen Informationen zum größten Teil mit den vorherigen Annahmen übereinstimmen, jedoch diese um zahlreiche neue Details und Elemente aus der Praxis erweitert haben.

7. Ausblick und Schlussfolgerungen

Nach den ersten Testphasen zeigen sich die ProjektleiterInnen, wie etwa in Kapitel 5.3. geschildert, positiv gestimmt und von den ersten Studien bestätigt. Auf diesen Auswertungen beruhend, werden die Initiativen in geplanter Weise durchgeführt werden können und laut erster Prognosen die erhofften Ergebnisse erfüllen. Solch großangelegte Projekte haben über die wirtschaftliche Situation hinausgehend auch auf die Sprach- bzw. Übersetzungsindustrie und deren TeilnehmerInnen Einfluss. Auf Überlegungen, welche Auswirkungen dies auf ÜbersetzerInnen im Allgemeinen aber speziell in der EU haben können, gehen die folgenden Seiten ein.

7.1. Ausblick ÜbersetzerInnen und Sprachtechnologien

“I am sure the first question which a translator asks about machine translation is ‘How will it affect my job?’”

Peter J. Arthern 1979: 97

Die vorangehenden Teile der Masterarbeit haben die Entwicklung der Sprachtechnologien allgemein und im Besonderen im Kontext der Europäischen Kommission skizziert. Daraus ist unter anderem deutlich geworden, dass sich mit der Zeit und den neuen Technologien stets auch die Arbeit der ÜbersetzerInnen verändert hat. Bereits der einfache Vergleich der Hilfsmittel der 50er Jahre zu den Hilfsmitteln des 21. Jahrhunderts zeigt den überaus großen Unterschied nicht nur zwischen den Tools selbst, sondern auch wie sie die Arbeitsweise

geprägt haben. Von der Schreibmaschine und den Terminologiekarteien bis hin zu den heutigen CAT Tools und nicht zuletzt der maschinellen Übersetzung.

Somit stellt sich die Frage nach dem Einfluss und den Auswirkungen, die diese neuen Technologien mit sich bringen. Bei den Tools handelt es sich nicht um bloße Werkzeuge, sondern um Technologien, die maßgeblich das Übersetzen auf mehreren Ebenen verändern: die kognitive Aktivität der ÜbersetzerInnen, ihre gesellschaftliche Position und ihren Status als ExpertInnen (vgl. Pym 2011:1). Auch Dickinson et al. stellen fest, dass keineswegs außer Acht gelassen werden darf, welche Auswirkungen Computer und Sprachtechnologien sogar auf die Gesellschaft als Ganzes haben können (Dickinson et al. 2013: 215). Auf die Europäische Union und die DGT bezogen bedeutet dies: Welche Möglichkeiten im Bereich der Übersetzung könnten in der Zukunft vorhanden sein? Inwiefern könnte sich die Rolle der ÜbersetzerInnen (in der DGT) in den nächsten Jahren möglicherweise verändern?

Es geht dabei vor allem um die Frage, ob sich durch den Fokus auf die Erweiterung der Nutzung von Tools und dem Bestreben, diese vermehrt und in so vielen Aufgabenbereichen wie nur möglich einzusetzen – um auf diese Weise kosten- und zeiteffizient zu arbeiten und im Idealfall die notwendige Qualität aufrecht zu erhalten – nicht die Aufgabenbereiche der ÜbersetzerInnen verschieben. Damit ist gemeint, dass sich beispielsweise die Tätigkeit viel stärker in Richtung der *editing* Aufgabenbereiche richten könnte als dies bisher der Fall war. Darunter sind das so genannte *pre-editing* und *post-editing* zu verstehen, bei denen jeweils die Vorbereitung eines Textes zur optimalen maschinellen Bearbeitung und die Nachbearbeitung bzw. Verbesserung der maschinell übersetzten Texte im Vordergrund steht. Bereits ein kurzer Blick ins Internet zeigt, wie sehr diese Auffassung in den Mittelpunkt gerückt ist. Einer der größten Marktanbieter von Übersetzungstechnologien, SDL Trados, wirbt auf seiner offiziellen Website damit, dass maschinelle Übersetzung und *post-editing* „die Zukunft des Übersetzens“ seien und bietet den interessierten KundInnen zu dem Thema zugleich ein kostenloses e-Book an, das in diese Materie einführen soll (<http://www.sdl.com>). Auch in der Ende 2015 abgehaltenen MT@EC Konferenz in Brüssel waren die Formen des *editing* und die dafür benötigten Kompetenzen ein Schlüsselthema.

In der Vergangenheit bestand diesbezüglich, wie diese Masterarbeit bereits aufgezeigt hat, selbstverständlich ein gänzlich anderes Bild. Gratis Versionen von MÜ sind bereits jeder Person mit Computer und Internetverbindung zugänglich. Doch auch aus der professionellen Übersetzung und der öffentlichen Verwaltung ist die Anwendung von maschineller und computergestützter Übersetzung nicht mehr wegzudenken.

Die seit den 50er Jahren bestehende Hoffnung der ComputerlinguistInnen einen derartigen Durchbruch zu erzielen, dass die Übersetzung von bestimmten Textsorten gänzlich ohne menschliche Hilfe zu ermöglichen, ist für professionelle ÜbersetzerInnen hingegen wohl als Befürchtung zu bezeichnen. Diese Befürchtungen bestehen nicht unbedingt nur darin, durch Computersysteme vom Markt verdrängt zu werden, sondern zusätzlich darin, dass

automatisierte Tools die Preise negativ beeinflussen (vgl. Hutchins 1998: 289). (Hutchins bezieht sich hier zwar auf die 1960 Jahre, doch die Situation hat sich nach der Wahrnehmung der Autorin der Masterarbeit nicht allzu verändert. Ähnliche Bedenken schwingen beispielsweise auch im mit Katja Mai geführten Interview mit.). Man wolle darüber hinaus nicht zu einem so genannten „*garbage collector*“ (Hutchins 1998: 10) werden; einer Person, die von einer eigenständigen Arbeit zu der „Bereinigung“ von maschinell erzeugten Texten übergeht.

Grundsätzlich kann die Haltung von professionellen ÜbersetzerInnen gegenüber MÜ somit als eher skeptisch beschrieben werden. Dies steht im Kontrast zu dem recht guten Ruf, über den CAT Tools verfügen. Die Softwares haben sich in Profikreisen als unerlässliche Hilfe durchgesetzt: Ihre Verwendung kann fast schon als ein Muss bezeichnet werden und eine Anforderung, die auch immer mehr AuftraggeberInnen stellen. Diese Einstellung hat letztendlich mit Sicherheit damit zu tun, dass CAT Tools die Funktion haben, die eigene Arbeit zu fördern und zu speichern, während MÜs oftmals schlechte Qualität, wenig Kreativität und Unnützen zugeschrieben wird. Trotz alldem teilt sich die Mehrheit der WissenschaftlerInnen die Meinung, dass diese Befürchtungen nicht begründet sind, da der maschinellen Übersetzung trotz ihrer heutigen immer besser werdenden Qualität deutliche Grenzen gesetzt sind. Auf Veränderungen ist natürlich dennoch einzugehen.

Es ist unbestreitbar, dass die neuen Sprachtechnologien unzählige Vorteile mit sich gebracht haben. Auf europäischer Ebene zeigen die EU-Projekte, wie mit deren Hilfe die Vernetzung Europas und die Bereitstellung von Daten vorangetrieben werden kann. Für ÜbersetzerInnen bedeuten sie eine schnellere und geordnetere Arbeitsweise. In der DGT kann dadurch ein TM mit einem Speicher genutzt werden, der Segmente aus mehreren Jahrzehnten beinhaltet. Anthony Pym macht jedoch auf vorhandene, wenn auch auf den ersten Blick gewisser Maßen versteckte Nachteile aufmerksam. So argumentiert er unter anderem, dass vor allem TMs das Gedächtnis der übersetzenden Personen bzw. dessen Kapazitäten beeinflussen. Mithilfe der TMs ist es möglich, eine große Menge an Texten im CAT Tool zu speichern und somit auf eine ebenso umfangreiche Menge an Übersetzungsvorschlägen zurückzugreifen. Während dies als Vorteil aufgefasst werden kann, sieht Pym darin gleichzeitig auch ein verstecktes Problem: während die ÜbersetzerInnen eine schwierige Stelle leicht erkennen und das Aufrufen von möglichen Lösungen bereits schnell erfolgt, nimmt die Auswahl des adäquaten Wortes aufgrund der hohen Anzahl an Vorschlägen oftmals einen längeren Zeitraum ein. Eine zu große Auswahl erschwert somit eher den Übersetzungsprozess, lenkt laut Befragungen häufig von anderen Elementen ab und das genutzte Tool garantiert in dieser Hinsicht somit keineswegs ein effizientes Arbeiten (vgl. Pym 2011: 2). Pym fährt weiterhin fort, indem er seine Ansicht darlegt, dass die neuen Übersetzungstechnologien die Linearität der zu übersetzenden Texte unterbrechen. Als Beispiel erwähnt er die Segmente, mit denen in den gängigen CAT Tools gearbeitet wird. Die Texte werden

nicht mehr als Einheit gesehen, sondern auf einzelne Sätze aufgespalten und in abgesonderten Stücken bearbeitet. Dies könne sich unter anderem auf die Kohäsion auswirken (vgl. Pym 2011:3). Studien aus dem Jahre 2008 haben diese von Pym beschriebenen Effekt bestätigt und die Zahlen haben gezeigt, dass durch die Aufteilung eines Textes auf einzelne Satzsegmente, der Fokus unvermeidbar auf diese, voneinander unabhängig behandelte Segmente gerichtet wird, was natürlich Einfluss auf das Endprodukt hat. Ob dies nun bedeutet, dass die Übersetzung einen wörtlicheren Charakter hat oder nicht, konnte jedoch noch nicht geklärt werden. Aspekte wie Erfahrung und Fachkenntnisse spielen selbstverständlich beim Endprodukt ebenso eine Rolle und so kann eine TM Segmentierung auch so manch Vorteile mit sich bringen (vgl. Declercq 2015: 482).

Dickinson et al. nähern sich dem Einfluss der Sprachtechnologien aus einer anderen Richtung und beschreiben zunächst die Veränderungen im Arbeitsablauf. Während sie ihre Szenarien nicht spezifisch für den Beruf der ÜbersetzerInnen beschreiben, sondern allgemein den Einfluss von Sprachtechnologien auf den Arbeitsfluss von Menschen auslegen, liegt es dennoch nahe, dass sie in gleichem Maße Prognosen für ÜbersetzerInnen erfassen und sich auf die DGT beziehen lassen. Die Autoren schildern folgende zwei alternative Szenarien, die sie jeweils als *deskilling* und *upskilling* bezeichnen (Dickinson et al. 2013: 215-216).

Unter *deskilling* verstehen die Autoren den Prozess, infolge dessen die Arbeitsabläufe aufgrund der neuen Entwicklungen und technologischen Möglichkeiten in mehrere, einfachere Schritte aufgeteilt werden. Diese vereinfachten Schritte erfordern von den ausführenden ArbeiterInnen im Gegenzug weniger Erfahrung und weniger Wissen. Aus diesem Grund wird ein Teil des Prozesses automatisiert und die Entlohnung gesenkt (vgl. Dickinson et al. 2013: 215, 216).

Die gegenteilige Alternative dazu wird als *upskilling* bezeichnet. In dieser Version übernimmt die Technologie automatisierte, routinemäßig ablaufende Prozesse und lässt seinen AnwenderInnen mehr Zeit und Energie, um sich den komplexeren und abstrakteren Aufgaben zu widmen, die von der neuen Technologie nicht mehr gemeistert werden können. Der Umgang mit den Tools erfordert zusätzliches Know-how, das sich die jeweiligen ArbeiterInnen aneignen müssen (vgl. Dickinson et al. 2013: 216). Als einfaches Beispiel für das *upskilling* kann hier etwa die Rechtschreibüberprüfung genannt werden, die es den NutzerInnen ermöglichen soll, sich mehr auf den Inhalt und die Formulierung als auf bestimmte Grammatikregeln und Rechtschreibung zu konzentrieren (vgl. Dickinson et al. 2013: 216).

In diesem Sinne kann diese Überlegung auch auf MitarbeiterInnen der DGT analog angewendet werden und argumentiert werden, dass etwa maschinelle Übersetzungstools die grundlegende Struktur der Übersetzung vorgeben, während dies den ÜbersetzerInnen mehr Zeit für die Arbeit an z.B. schwierigen terminologischen Fragen lässt. Im Falle von CAT

und Terminologie Tools der DGT wird auf gleiche Weise der Arbeitsaufwand eingeschränkt: dadurch, dass Paralleltexte und bereits vorhandene ähnliche oder gleiche Übersetzungen automatisch aus der Datenbank herausgesucht und geladen werden, sowie dazugehörige Informationen angeboten und nützliche Vorschläge geliefert werden, bleibt mehr Zeit und Konzentration für die Bearbeitung neuer Textstellen. Diese Arbeit mit den CAT Tools, Terminologiedatenbanken und weiteren Hilfsmitteln erfordert einen professionellen Umgang und entspricht einer komplexen Tätigkeit, während repetitiv auftretenden Aufgaben wegfallen bzw. von den Technologien übernommen werden. Die derzeitigen Projekte und Pläne der DGT beurteilend wird diese Vorgehensweise auch in der Zukunft beibehalten werden.

So stellen auch Dickinson et al. fest, dass ÜbersetzerInnen, die am professionellen Übersetzungsmarkt eintreten wollen, sich ExpertInnenwissen über Tools aneignen werden müssen (bzw. ist dies bereits jetzt der Fall. Wie auch Katja Mai im Interview hervorgehoben hat: vor wenigen Jahren war die Anwendung von Tools noch keine Voraussetzung, heute sind diese jedoch nicht mehr wegzudenken). Translation Memories, elektronische Wörterbücher und Datenbanken, automatisierte Übersetzungsprogramme und CAT Tools sind zentrale Bestandteile dieses Know-hows (vgl. Dickinson et al. 2013: 208).

Der Autor weist aber darauf hin, dass diese Technologien zwar einen Teil der Arbeitsprozesse übernehmen werden, aber die HumanübersetzerInnen dadurch keineswegs ausschließen werden. Dabei handle es sich nicht alleinig um eine Frage der Qualität, sondern auch um Vertrauens. Wenn es sich um heikle, vertrauliche und wichtige Texte handelt – wie sie in einer Organisation wie der Europäischen Union regelmäßig üblich sind – beanspruchen AuftraggeberInnen verantwortungsvolle ÜbersetzerInnen, die für die Genauigkeit, Richtigkeit sowie Adäquatheit des übersetzten Endproduktes und für Vertraulichkeit bürgen können (vgl. Dickinson et al. 2013: 208). In dieser Hinsicht ist auch interessant zu beobachten, wie solch ein Aspekt in einem Bereich wie dem eHealth Baustein des CEF Projektes gehandhabt wird. In diesem Zusammenhang bildet die Frage nach Datenschutzbestimmungen und letztendlich die Vertraulichkeit und Sicherheit der Informationen einen Schwerpunkt.

Die Konsequenzen der *deskilling* und *upskilling* Szenarios können folgendermaßen zusammengefasst werden: Die Ergebnisse und Forschung zeigen, dass eine Kombination von Humanübersetzung und Maschine der ausschlaggebende Faktor sein werden. „Ideale“ professionelle ÜbersetzerInnen bringen somit fachspezifisches Wissen sowie Weltwissen mit und sind gleichzeitig bereit mit den Technologien zu arbeiten und ihre Kompetenz mit der Weiterentwicklung der verschiedensten Tools auf diesem Gebiet fortlaufend zu erweitern. Es steht außer Frage, dass sich die Technologien und ihr Einsatzgebiet bzw. ihr „Können“ verbessern werden, aber dies sollte eher als Vorteil und Erleichterung für die ÜbersetzerInnen aufgefasst werden. Dickinson et al. beschreiben Sprachtechnologien somit viel mehr als

Chance effizienter und schneller arbeiten zu können und keineswegs als Bedrohung (vgl. Dickinson et al. 2013: 208).

Ob die Autoren diese Auffassung vertreten, da sie Vertreter des Bereichs der Computerlinguistik sind und die neuen Entwicklung aus einem anderen Blickwinkel sehen als TranslatorInnen, sei dahingestellt. Allerdings zeigt die Tatsache, dass ÜbersetzerInnen trotz der jahrzehntelangen Entwicklung und der seitdem andauernden Experimenten mit Sprachtechnologien unerlässlich sind, dass der Beruf der „Humanübersetzer“ und der DGT weiterhin nicht gefährdet ist.

Von vielen Seiten wird immer wieder stark betont, dass als Reaktion auf die neuesten Entwicklungen neue Methoden in die Ausbildung einfließen sollten. Gemeint ist damit jedoch nicht nur die Vermittlung von einem weitreichenden Verständnis für einzelne gängige Tools und ihrer spezifischen Funktionen. Viel wichtiger sei es, im Allgemeinen einen sicheren Umgang mit Technologien zu fördern und die Prinzipien solcher Softwares zu verstehen. Denn ein Tool, das heute weit verbreitet ist, mag in einigen Jahren bereits als überholt gelten und schnell abgelöst werden. Somit gehe es nicht ausschließlich um das wie, sondern vordergründig um das wann und warum Sprachtechnologien von ÜbersetzerInnen verwendet werden sollen und in welcher Weise sich ein einzelnes Tool auf den Übersetzungsprozess auswirken (vgl. Bowker 2015: 95). Sich mit Sinn und Zweck eines Hilfsmittels für einen jeweiligen Text auseinanderzusetzen und Veränderungen schnell aufgreifen und meistern zu können, sei ein wesentlicher Aspekt, der aufgegriffen werden sollte, denn diese Fähigkeit trägt zu der Konkurrenzfähigkeit bei und hilft den Ansprüchen der KundInnen gerecht zu werden. Ein weiterer zu beachtender Punkt, der auch im Rahmen der MT@EC Konferenz hervorgehoben wurde, ist die Bedeutung des *post-editing* (vgl. Hannah Riley, 04.12.2015; <https://scic.ec.europa.eu>). Wie ein Blick auf die Projekte und die Arbeit der DGT bereits klar zeigt, ist PE als Teil des Übersetzungsprozesses unumgänglich. Oftmals fehlen professionellen ÜbersetzerInnen jedoch die Übung und die richtige Herangehensweise an diese Aufgabe, denn es handelt sich nicht um die Korrektur eines normalen Textes, sondern die Eigenheiten eines MÜ Systems müssen beachtet werden. Diese Tatsache müsse in der universitären Ausbildung angesprochen, bewusst gemacht und den Studierenden die Bearbeitung von Maschinenoutput antrainiert werden (vgl. Bowker 2015: 88).

Für zukünftige ÜbersetzerInnen der Kommission empfiehlt es sich zusätzlich ein sicherer Umgang mit Management Tools anzueignen, wie es beispielsweise Euramis oder Suivi sind. Ein Vorteil gegenüber freiberuflichen KollegInnen ist jedoch, dass sie über das IT Team des DGT verfügen, sich in einem Notfall jederzeit an dieses wenden können und regelmäßig Workshops abgehalten werden (z.B. Translating Europe Workshops 2016) (vgl. Bowker 2015: 96).

Mit Sicherheit kann festgestellt werden, dass Übersetzungstools bereits ein fester Bestandteil der Übersetzungsarbeit geworden sind und dies auch in Zukunft verstärkt der Fall sein wird. So erklären auch Dickinson et al., dass sich die Technologien stetig verbessern werden. Dies sei anhand von mehreren Indikatoren erkennbar: die zur Verfügung stehenden Materialien zum Trainieren der Programme wachsen mit der Zeit kontinuierlich an, die Hersteller von Übersetzungstechnologien werden weiterhin mit ProgrammierexpertInnen an der Nutzbarkeit für eine größere Anzahl von Sprachen arbeiten (nicht zuletzt, um diese Systeme auch kommerziell einsetzen zu können) und die viel versprechenden Entwicklungen im Bereich der statistischen MÜ lassen ebenso auf einen positiven Werdegang schließen (vgl. Dickinson et al. 2013: 208-209). Die Förderung dieses Vorganges und das Bestreben, Daten zu sammeln und einzusetzen – wie die Tätigkeit des ELRC zeigt – sind Teil dieses Prozesses und treiben die Entwicklung an.

Pym vertritt ebenso diese Meinung und verweist darauf, dass Kritik bzw. Widerstand gegenüber neuen technologischen Entwicklungen sich als wenig Erfolg bringend erweisen wird. Ähnlich den Schilderungen des *upskilling* Szenarios argumentiert er, dass sich das Machtgefüge verschieben wird: es wird sich von jenen Personen, die „bloß übersetzen können“ auf diejenigen Personen verlagern, die die neuen Tools kennen und in ihrer Arbeit anwenden können. Dies können neben professionellen ÜbersetzerInnen unter anderem auch Mitglieder anderer Berufsgruppen sein, etwa ProjektmanagerInnen, MarketingexpertInnen oder ProduktentwicklerInnen. Pym schlägt sogar vor, dass sich Übersetzen als eine der grundlegenden Fähigkeiten, die jede Person mithilfe der Fertigkeit „Sprache“ vollziehen kann, etablieren könnte. Neben Sprechen, Hören, Lesen und Schreiben kann mit den neuen Technologien Übersetzen ebenso nicht mehr an einen besonderen Personenkreis gebunden, sondern allen frei zugänglich und handhabbar sein. Als Folge dessen wird ein Teil des Übersetzungsmarktes von den professionellen, hauptberuflichen ÜbersetzerInnen an Freiwillige übergehen, die mithilfe ihres aus anderen Bereichen mitgebrachten Wissens mit Tools wie MÜ und CAT Tools lernen umzugehen und für ihre Bedürfnisse einzusetzen (vgl. Pym 2011: 5-6).

Pym folgert jedoch, dass dieser Vorgang deshalb nicht zwangsläufig als schlecht betrachtet werden muss. Einerseits ist er der Ansicht, dass es weiterhin bestimmte Übersetzungsnischen geben wird, die nicht an Laien übergehen werden. Andererseits erklärt er, dass der Markt (und das Internet) genug Raum für beide Gruppen bietet und sogar eine Zusammenarbeit bzw. ein Dialog zwischen laienhaften und professionellen ÜbersetzerInnen auf mehreren Ebenen durchaus positive Aspekte in sich birgt (vgl. Pym 2011: 6-7).

Die Situation der Kommission bestätigt seine Aussage: auf einer Schiene laufen die Projekte wie mDSM und CEF primär über maschinelle Übersetzung, doch daneben existiert die Arbeit der DGT, die außerhalb dieser Projekte weiterhin reguläre Aufträge erhält, welche von ÜbersetzerInnen bearbeitet werden.

Einen ähnlichen Ausblick schilderten auch die vortragenden (Sprachtechnologie)ExpertInnen der 5. *European Master's in Translation* (EMT) Konferenz im November 2011. Die in Brüssel stattfindende Konferenz unter dem Motto „Die Zukunft des Übersetzens meistern“ wurde gemeinsam von der Europäischen Kommission und der Generaldirektion Übersetzen organisiert und galt vor allem den Studierenden der Translationswissenschaften (vgl. Europäische Kommission 2011: 1). Der offizielle Bericht der Veranstaltung fasst dabei die wichtigsten Erkenntnisse der TeilnehmerInnen kurz zusammen und spiegelt den Einfluss der Technologien auf die Übersetzungsbranche wieder.

Anwesend war unter anderem Sharon O'Brien (University of Dublin), die neben dem *post-editing* vor allem im *pre-editing* einen zukünftigen Schwerpunkt sieht und auf die Vorbereitung des Textes mehr Fokus lenkt. Im Gegensatz zu Bowker (siehe oben), erläutert sie, dass nicht erst während dem *post-editing* an die Besonderheiten eines Tools und die Erwartungen der KundInnen zu denken sei. Die Arbeit mit MÜ verläuft erfolgreicher, wenn der Text bereits vor der Eingabe in das Programm für die Übersetzung vorbereitet wird. In diesem Zusammenhang erwähnt sie auch das *fit-for-purpose* Prinzip: Was möchten die KundInnen von dem jeweiligen Text? Wofür ist dieser gedacht? Abhängig dieser Beantwortung der Fragen, soll auch die adäquate Vorbereitung vorgenommen und das adäquate Tool gewählt werden (vgl. Europäische Kommission 2011: 2).

Neben zahlreichen Vortragenden sprach auch Jaap van der Meer, der Gründer von TAUS, über die Zukunft des Übersetzens. Seine Ansichten sind sehr im Einklang mit den zuvor geäußerten Ausblicken, vor allem jedoch auch mit Pym. So spricht er von einer „deutlichen Spaltung im Markt“ und beschreibt einen Prozess, bei dem ein stetig wachsender Anteil der Übersetzungen von Maschinen übernommen wird (Europäische Kommission 2011:3). Das Endergebnis weise zwar oftmals keine hohe Qualität auf, habe aber den Vorteil, dass die Übersetzung in Echtzeit entsteht – ein Merkmal, das auch die EU zu ihren Gunsten einsetzen möchte. Demgegenüber wird es weiterhin einen Markt für Humanübersetzer geben, die „für lokalisierte Übersetzungen von hoher Qualität in mehr Sprachen verantwortlich bleiben würden“ (Europäische Kommission 2011: 3).

Der Bericht stellt fest, dass all diese Aspekte in die zukünftigen Projekte der EMT einfließen werden (z.B. in der Zusammenarbeit mit Hochschulen) und auch in der Kommission nicht außer Acht gelassen werden. Schon im Jahre 2011 betonte Rytis Martikonis, Generaldirektor der DGT, dass in diesem Sinne „Qualität, Flexibilität und Präzision“ wichtige Schlüsselbegriffe sein werden und die DGT auch dementsprechend auf die Veränderungen reagieren wird – nicht zuletzt durch den Einsatz von MÜ (Europäische Kommission 2011: 1).

Das im Rahmen dieser Masterarbeit geführte Gespräch mit der DGT Übersetzerin Katja Mai bestätigt ebenso die in der Fachliteratur und von den KonferenzteilnehmerInnen beschriebenen Prozesse. Während die ÜbersetzerInnen der Kommission zwar nicht einer Konkurrenz wie der des freien Marktes ausgesetzt sind, handelt es sich dennoch nicht um ein gänzlich abgeschirmtes System. Innerhalb der Europäischen Union besteht ebenso Druck effizienter und schneller zu arbeiten. Dies zeigt sich in der starken Anwendung von CAT Tools und den neuen Versuchen der maschinellen Übersetzung (vgl. Mai 2015).

Des Weiteren beschreibt Mai, dass in den kommenden Jahrzehnten das Personal der DGT nicht mehr so konsequent nachbesetzt werden soll, da auch die Kommission die Senkung ihrer Kosten plant. Die DGT sei davon stärker betroffen als andere Generaldirektionen, die politische Arbeit leisten. In dieser Hinsicht würden sehr wohl Befürchtungen bestehen, dass die Computerübersetzung im Vergleich zu der Humanübersetzung verstärkt in den Vordergrund rücken wird, nicht zuletzt um die kleinere Zahl an Übersetzungspersonal zu kompensieren. So betont Mai, dass in Zukunft möglicherweise mehr auf maschinelle Übersetzung zurückgegriffen wird, um dem Druck standzuhalten und die Aufträge rechtzeitig erfüllen zu können. In dem Fall würde sich der primäre Aufgabenbereich der DGT auf die Vor- und Nachbearbeitung bzw. der Evaluierung der Texte richten (vgl. Mai 2015).

Fraglich sei bei diesem Szenario jedoch, ob und in welchem Maße sich Quantität und Qualität in der Waage halten werden. Denn nicht zuletzt könnten die Forderung nach schnelleren Übersetzungen sowie eine stetig steigende Menge zu einem niedrigeren Qualitätsanspruch der AuftraggeberInnen führen. Mai fügt jedoch auch hinzu, dass aufgrund der heiklen Themen (z.B. mit Rechtswirkung versehene Rechtstexte) bestimmte Aufgaben komplexe Ansprüche mit sich bringen und bloß durch eine erhöhte Verwendung von Sprachtechnologien nicht gelöst werden können – ein Text müsse u.a. verständlich und lesbar bleiben. Mai stellt fest, dass es sich hierbei nur um Prognosen handle, doch es sei klar, dass sich mit der verstärkten Anwendung von CAT Tool und Co. die Arbeit natürlich ändern würde (vgl. Mai 2015).

Im Laufe der Recherche sind regelmäßig wissenschaftliche Artikel und Konferenzen zu der Frage anzutreffen, ob und inwiefern Computer das Übersetzen übernehmen und revolutionieren könnten. Es ist eine Frage, die viel Interesse weckt und deren Beantwortung sehr facettenreich ausfällt. Selbstverständlich kann auf all die in diesem Kapitel erläuterten Fragen keine eindeutige Antwort gegeben werden, da wie auch Mai oben betont, diese momentan nur Prognosen sind. Welche Ergebnisse sich aus der fortlaufenden Forschung und welche Ideen sich in der Praxis tatsächlich umsetzen lassen, anschließend im Arbeitsalltag Fuß fassen werden und wie sich die derzeitigen Kommissionsprojekte langfristig auf die Übersetzung in der Europäischen Union auswirken werden, werden allein die nächsten Jahre zeigen. Das Fortbestehen der Sprachtechnologien im Allgemeinen hingegen ist kein fraglicher Punkt und in diesem Sinne stellt auch Anthony Pym fest: „*The technology, for*

better or for worse, is here to stay“ (Pym 2011:6). Es ist ohne Zweifel ein brisantes Thema, das bereits seit Mitte des letzten Jahrhunderts beschäftigt und zu Diskussionen anregt und dies wohl auch in der Zukunft tun wird.

7.2. Schlussfolgerungen

Die Europäische Union ist in den vergangenen Jahrzehnten in vielerlei Hinsicht gewachsen. Ihr Aufgabenbereich hat sich erweitert und die von ihr beeinflussten Politikbereiche haben sich um ein Vielfaches vermehrt – sie reichen nun von Wirtschaft und Politik bis hin zu Umwelt, Gesundheit, Mehrsprachigkeit und vielem mehr. Die Zahl ihrer Mitgliedstaaten hat sich beinahe verfünffacht und damit zugleich kletterte die Zahl ihrer Amtssprachen auf insgesamt vierundzwanzig.

Mit neuen Gesetzen, Beitritten und europaweiten Krisen haben sich der Informationsfluss und folglich die Zahl der Dokumente schlagartig erhöht. Daran waren neue Anforderungen an die Generaldirektionen gekoppelt. Sie mussten im Laufe der Jahre mit der Kommission mitwachsen und ihre Struktur mehrmals umgewälzt werden, um einen reibungslosen Ablauf zu garantieren. Die Generaldirektion Übersetzen bzw. ihre Vorgängerinnen bildeten davon keine Ausnahme.

Dabei kommt zum Ausdruck, wie vernetzt die Europäische Kommission und die DGT tatsächlich sind. In ihrer Zusammenarbeit übernehmen sie in der Umsetzung und Vereinheitlichung der EU-Gesetzgebung die führende Rolle und tragen zu der „europäischen Kommunikation“ bei, sei es intern oder mit den Mitgliedstaaten und der Öffentlichkeit. Durch ihre Übersetzungs- und Sprachdienstleistungen gestaltet die DGT das Bild der Öffentlichkeit von der Kommission aktiv mit. Zudem zeigt eine Betrachtung der neuesten Projekte in welchem Ausmaß die DGT die Ziele der Kommission unterstützt, ausführt und möglich macht.

Die Vernetzung und Modernisierung (besonders in den 90er Jahren) haben viele Veränderungen mit sich gebracht. So wie sich die Kommission und ihr Alltag durch neue Technologien verändert haben, so hat sich zugleich automatisch der Arbeitsalltag der DGT verändert. Der Austausch von Informationen verläuft schneller, ist globaler und mehr Akteure sind an mehreren Orten in Echtzeit involviert. Die Ansprüche und Erwartungen an die Übersetzung ändern sich dadurch, weshalb Anpassungsbedarf besteht. Die daraus folgenden Konsequenzen illustrieren die Veränderungen in der Arbeitsweise der DGT ÜbersetzerInnen und der Fokus auf Systeme wie es auch MT@EC ist. Es ist unbestreitbar, dass sich die EU-Praxis (in bestimmten Bereichen) von der Humanübersetzung immer stärker in Richtung Technologie und digitale Welt bewegt.

Die Funktionen und Anwendungsbereiche des Computer Know-hows sind jedoch nicht uneingeschränkt, wie die ersten Pilotphasen erkennen lassen: Sie können nur in ausgewählten Domänen und für bestimmte Aufgabengebiete eingesetzt werden, für die die Maschinen gezielt trainiert werden müssen. Es sind Bereiche, in denen Texte vorhersehbare Elemente und Muster enthalten. Der Einsatzbereich für die Technologien könnte sich in den nächsten Jahren aber durchaus erweitern. Mit dem Ausbau der Systeme und der Sammlung von genügend Übersetzungen, Ausgangstexten, verbesserter Alignierung und Hybriden, deuten die in der Arbeit dargelegten Informationen darauf hin, dass dies sogar sehr wahrscheinlich eintreten wird, denn die EU bildet für die Entwicklung und den Fortschritt von Sprachtechnologien gute Rahmenbedingungen. Sie bietet Möglichkeiten, die außerhalb einer derartigen Organisation nicht bestehen würden. Einerseits ist ein seltenes internationales Umfeld vorhanden, dass die Zusammenarbeit der Mitgliedstaaten und ihrer Universitäten, Unternehmen und WissenschaftlerInnen fördert. Andererseits ist eine rechtliche Verankerung der Amtssprachen und der gewählten Sprachenpolitik vorhanden, die die Übersetzung und ihre Hilfsmittel zu einem Kernstück machen und Sprache bzw. Übersetzen somit zu einem Bereich geworden ist, auf dessen Aktualisierung und Effizienz man bedacht ist.

Zudem können im EU-Rahmen die Vorteile von computergestützter Übersetzung ideal ausgenutzt werden: die Mehrheit der Texte weist starke Textsortenkonventionen auf und es bestehen bestimmte Erwartungen und Standards. Diese Muster können bei der Übersetzung mit Sprachtechnologie sehr gut genutzt werden.

Die Analyse des MT@EC und seinem jetzigen Stand machten bewusst, dass auch SMÜ noch nicht an seinem endgültigen Ziel angekommen ist. Bis zum Projektende im Jahre 2020 müssen von den LeiterInnen eine Handvoll Hürden in Angriff genommen werden. Im Vergleich zu den Vorgängerversionen hat sich allerdings die Art dieser Probleme verlagert. Mit der statistischen Methode haben sich die Schwerpunkte von der manuellen Erstellung grammatikalischer Regeln und fachspezifischer Wörterbücher auf die Erlangung der notwendigen Datenmenge (Trainingskorpora, Übersetzungsspeicher etc.) verschoben. Einen wichtigen Unterschied zeichnet jedoch aus, dass während die Lösung der Probleme der regelbasierten Systeme einen unverhältnismäßigen Aufwand, geringe Verbesserungsmöglichkeiten und überhöhte Kosten verursacht hätte, man sich bei den SMÜ-Systemen der EU durch gezieltes Vorgehen bereits in naher Zukunft erhofft, eine annehmbare Lösung zu finden. Die Vorteile liegen darin, dass durch die Kooperation mit Mitgliedstaaten (siehe ELRC) und der Zusammenabriet mit den DGT ExpertInnen den derzeitig vorhandenen Lücken besser entgegengewirkt werden kann. Im Gegensatz dazu gestaltete sich die manuelle Aktualisierung der regelbasierten Systeme als zeitaufwändig sowie umständlich, ohne den Aussichtschancen eines SMÜ. Nun setzt man stattdessen auf eine starke IT Struktur, die nachdem die gesammelte Daten eingespeichert wurden, voll automatisch von selbst rennt, ohne großes Zutun von außen.

Aktuelle Konferenzen haben allerdings zu der Erkenntnis geführt, dass dies nicht zwangsläufig das Ende der regelbasierten Systeme bedeutet. Nach den ersten Testphasen und einem Blick auf die Medaillenverteilung der Amtssprachen musste man sich eingestehen, welche großen Datenunterschiede zwischen den einzelnen Sprachen existieren. Im Zuge dessen wurde erkannt, dass ein reines SMÜ-System für eine Gruppe von Amtssprachen nicht ausreichend zufriedenstellende Ergebnisse bringen würde. In Zukunft sollen nun auch Hybride entwickelt werden, die die besten Funktionen aus der statistischen und regelbasierten Methode kombinieren werden.

Wenngleich der Fokus auf der SMÜ liegt, konnten auch Neuerungen bei den CAT Tools entdeckt werden. In deren Bereich stellten sich die Verbindung mit dem Internet und die Vernetzung der ÜbersetzerInnen als ein aktuelles Thema heraus. Die zuletzt unterstützten Projekte wie Casmacat und Matecat sind beide kostenlose Online Plattformen und machen somit professionelle Übersetzungstools auch für InteressentInnen außerhalb des Kreises der SprachexpertInnen zugänglich. Damit wird mitunter erneut die von Pym und seinen KollegInnen geäußerte Beobachtung unterlegt, dass Übersetzungstools nicht mehr unbedingt exklusive Werkzeuge von ÜbersetzerInnen sind. Ob die neuen Erkenntnisse von Casmacat und Co. eventuell auch innerhalb der DGT Einfluss auf den TWB haben werden, wird noch zu sehen sein. Aufgrund des langwierigen Vorganges und den Lizenzregelungen scheint eine Modernisierung momentan nicht in Reichweite zu sein.

Der Vergleich von der in Kapitel 7.1. gezeichneten zukünftigen Rolle und dem zu erwartendem Arbeitsumfeld von ÜbersetzerInnen – sei es auf einem freien Markt oder in einer Organisation wie der EU – und den aktuellen Projekten der Kommission unterstreicht die Tatsache, dass die Arbeit immer stärker von Technologien geprägt sein wird. Ohne Kenntnisse dieser wird das Übersetzen nicht mehr möglich sein. Dabei erweist es sich als notwendig, nicht allein über Grundwissen bezüglich der einzelnen Tools zu verfügen, sondern durchaus über ein vertieftes ExpertInnenwissen. Damit gepaartes translationswissenschaftliches Wissen und eine professionelle Herangehensweise sind ebenso unumgänglich. Die „klassischen“ Übersetzungsfähigkeiten müssen erweitert werden und die angebotene Qualität hochwertig sein. In Zukunft wird etwa verstärkt die Vorbereitung und Nachbearbeitung von Texten oder Zusatzaufgaben wie Alignierung von Textsegmenten und der Umgang mit Projektmanagement Tools für die Koordinierung der Übersetzungsaufträge einfließen. Die Kapitel lassen die Schlussfolgerung zu, dass eine Kombination dieser Merkmale einen Vorteil gegenüber Maschinen und aus anderen Fachbereichen kommenden Personen bedeutet, die nun auch Zugriff auf Tools haben können.

In diesem Zusammenhang wird eindeutig, dass es keine Vorteile mit sich bringt, Technologien zu ignorieren oder sie als Bedrohung aufzufassen. Für ÜbersetzerInnen empfiehlt es

sich, die aus dem „Vormarsch“ der Sprachtechnologien folgenden Konsequenzen anzunehmen, ihre positiven Seiten zu nutzen und nicht zuletzt neue Trends zu erkennen, um auf kommende Technologien reagieren zu können.

Übersetzen getrennt von Sprachtechnologien zu sehen ist mittlerweile unmöglich geworden. Sich vor allem mit maschineller Übersetzung vertraut zu machen, bietet einen großen Vorteil. In Anbetracht der derzeitigen Entwicklungen in der Kommission betrifft dies vor allem zukünftige SprachexpertInnen und ÜbersetzerInnen im Umfeld der europäischen Institutionen. Eine Vertrautheit mit den Sprachtechnologien ist nicht allein deshalb wesentlich, um solche Entwicklungen mitverfolgen zu können, sondern diese im Idealfall auch mitgestalten zu können.

In diesem Sinne betont der überwiegende Teil der dazu verfassten Literatur, dass die Aussage *Mensch vs. Maschine* irreführend sei, denn es ginge weiterhin um *Mensch mit Maschine*. So beschreiben z.B. Dickinson et al. Sprachtechnologien viel mehr als Chance effizienter und schneller arbeiten zu können und keineswegs als Bedrohung (vgl. Dickinson 2013: 208). Das Fazit: MT@EC und Co. haben den Zweck mithilfe der Maschine zu **über-**setzen und nicht HumanübersetzerInnen zu **ersetzen**.

Interessant ist, dass sich die Auslöser für die Motivation Sprachtools zu verwenden, seit Anbeginn kaum verändert haben. Wie in den Kapiteln der vorliegenden Arbeit zu lesen ist, sprach die Kommission bereits 1970 von einem Sprachproblem und einem rasant steigenden Übersetzungsvolumen, das auf eine effiziente Art und Weise eingeschränkt werden müsse. Ähnliche Aussagen sind heutzutage, mehr als 40 Jahre später, weiterhin anzutreffen.

Diesbezüglich scheint es grundsätzlich einen Zwiespalt zwischen der symbolischen Bedeutung bzw. Notwendigkeit und den aufgebürdeten Lasten der Amtssprachenzahl zu geben – ein Problem, zu dessen Lösung an jedem Punkt der Geschichte die jeweils modernsten zur Verfügung stehenden Technologien herangezogen wurden, unabhängig davon, ob es sich um insgesamt sechs oder den heutigen vierundzwanzig Sprachen handelte. Mehr als vierzig Jahre später, besteht dieser Aspekt weiterhin, paart sich jedoch mit Projekten, die (zusätzlich) die digitale Vernetzung und Echtzeit-Kommunikation ins Auge fassen.

Es ist allerdings anzumerken, dass sich bei genauer Betrachtung ein leichter Widerspruch zwischen den im Rahmen der Projekte vertretenen Prinzipien und der Realität in anderen Bereichen des Informationsflusses finden lassen. Denn bis hin zu dem idealen Ziel, dass alle Amtssprachen in gleichem Maße vertreten sind und jegliche Dokumente in allen Sprachen und für alle UnionsbürgerInnen zugänglich sind, ist es noch ein weiterer Weg. Der Status von Englisch als *Lingua Franca* und als Bindeglied in internationalen Umgebungen ist mittlerweile fest verankert. Dies spiegelt sich auch in der Kommission und ihrer Kommunikation wieder bzw. die Art und Weise wie sie Prioritäten bei der Übersetzung setzt. Die Situation kann als Gratwanderung bezeichnet werden: Möglichst viele Dokumente sollen unbedingt übersetzt werden, um die Demokratie, die Transparenz und einen

freien Informationsfluss zu ermöglichen, doch gleichzeitig reichen die Mittel nicht, um alle Aspekte dieser Grundidee umzusetzen, da Prioritäten gesetzt werden müssen.

Daraus folgt, dass trotz der Bestreben und Projekte die EU selbst oftmals auf Englisch zurückgreift (siehe z.B. Broschüren im EU-Bookshop). Anhand der in der Arbeit erarbeiteten Aspekte wird auch in dieser Situation wohl ein Kompromiss notwendig sein: Wichtige Dokumente, die etwa die Gesundheit oder Bildung der EU-BürgerInnen betreffen (siehe Inhalt der CEF Bausteine) sollten auch tatsächlich in ihrer Sprache verfügbar sein bzw. mit den Tools die Option bereitgestellt werden, eine Übersetzung erhalten zu können. Für Daten außerhalb dieses Spektrums muss zunächst eine gute Lösung gefunden werden. Angesichts dessen, wie stark die Kommission derzeit in die Entwicklung der Tools involviert ist, könnte eventuell bereits in den nächsten Jahren die Bandweite an übersetzten Texten und auf diese Weise die „Übersetzungskapazität“ Schritt für Schritt erweitert werden.

Die Betrachtung der geschichtlichen Entwicklung machte deutlich, dass die Antwort auf die Frage, wie die europäische Sprachenpolitik getreu ihrer Prinzipien, gleichzeitig aber auch getreu einer effizienten und pragmatischen Ausführung am besten umgesetzt werden kann und welche Tools diesem Zweck dienen, stets zu Diskussionen anregen wird.

Darüber hinaus zeigen die aktuellen Projekte CEF und mDSM auch eine andere Facette der Sprachtechnologien. Die wirtschaftlichen Ausmaße und Möglichkeiten der Sprachtechnologien wurden erkannt, die Potenziale der Übersetzungstechnologie werden ernst genommen und ein eigener Wirtschaftszweig ist entstanden, der u.a. von der Kommission gefördert wird und vielfältig für ihre Zwecke eingesetzt wird. Der Einsatz der Technologien geht weit über das Ideal und die rechtlichen Vorgaben der Gleichstellung aller Sprachen hinaus, denn Sprache ist ein wirtschaftlicher Faktor geworden. Die Tools – selbstverständlich allen voran die MÜ – entwickeln sich von unzugänglichen Übersetzungstools für SpezialistInnen und experimentellen Unterfangen der Computerlinguistik zu ernstgenommenen Technologien, von denen man sich umfassende Lösungen und Motoren für den europäischen Handel und eine umfassende Vernetzung des EU-Binnenmarktes verspricht. Die Kommission vertraut auf die Ergebnisse und investiert somit in ein solch großes und alle Mitgliedstaaten übergreifendes Projekt wie z.B. das CEF, bei dem ein enormer Aufwand betrieben wird und keine Kosten gescheut werden. Aufgrund dessen scheint es sehr wahrscheinlich, dass solche Projekte auch in Zukunft weitergeführt werden.

8. Bibliographie

Anmerkung zu den URLs: sämtliche Links wurden am 24.04.2016 nochmals eingesehen.

ALPAC (Hrsg.). 1966. *Languages and machines: computers in translation and linguistics. A report by the Automatic Language Processing Advisory Committee, Division of Behavioral Sciences, National Academy of Sciences, National Research Council.* Washington, D.C.: Printing and Publishing Office National Academy of Sciences.

Amtrup, Jan W. 2010³. Aspekte der Computerlinguistik. In: Carstensen, Kai-Uwe et al. (Hrsg.), 1-17.

Arthern, Peter. 1979. Machine Translation and Computerized Terminology Systems. A Translator's Viewpoint. In: Snell, Barbara M. (Hrsg.). *Translating and the Computer.* Amsterdam: North-Holland Publishing Company, 77-108.

Atteslander, Peter. 2006¹¹. *Methoden der empirischen Sozialforschung.* Berlin: Erich Schmidt Verlag.

Bergmann, Jan. (Hrsg.) 2015⁵. *Handlexikon der Europäischen Union.* Baden-Baden: Nomos Verlagsgesellschaft.

Bowker, Lynne. 2002. *Computer-Aided Translation Technology. A Practical Introduction.* Ottawa: University Ottawa Press.

Bowker, Lynne. 2015. Computer-Aided Translation. Translator training. In: Chan, Sin-Wai (Hrsg.), 88-104.

Carstensen, Kai-Uwe/Ebert, Christian/Ebert, Cornelia/Jekat, Susanne/Klabunde, Ralf/Langer, Hagen (Hrsg.). 2010³. *Computerlinguistik und Sprachtechnologie. Eine Einführung.* Heidelberg: Spektrum Akademischer Verlag.

Carstensen, Kai-Uwe. 2012. *Sprachtechnologie. Ein Überblick.* In: <http://www.kai-uwe-carstensen.de/Publikationen/Sprachtechnologie.pdf>

Chan, Sin-Wai (Hrsg.). 2015. *Routledge Encyclopedia of Translation Technology.* Oxon/New York: Routledge.

Declercq, Christophe. 2015. Editing in Translation Technology. In: Chan, Sin-Wai (Hrsg.), 480-493.

Dickinson, Markus/Brew, Chris/Meurers, Detmar. 2013. *Language and Computers.* West Sussex: Wiley-Blackwell.

Forcada, Mikel L. 2015. Open-source machine translation technology. In: Chan, Sin-Wai (Hrsg.), 152-166.

- Gerhards, Jürgen. 2010. *Mehrsprachigkeit im Vereinten Europa. Transnationales sprachliches Kapital als Ressource in einer globalisierten Welt*. Berlin: Springer VS Verlag.
- Hansen-Schirra, Silvia. 2012. Nutzbarkeit von Sprachtechnologien für die Translation. In: *trans-kom* Band 5, Nummer 2: 211-226.
- Hartmann, Jürgen. 2009². *Das politische System der Europäischen Union. Eine Einführung*. Frankfurt: Campus Verlag.
- Hutchins, John. 1986. *Machine Translation: past, present, future*. In: <http://www.hutchinsweb.me.uk/PPF-14.pdf>
[Ehemals in Druck erschienen bei: Chichester/New York: Ellis Horwood/Halsted Press.]
- Hutchins, John. 1996. ALPAC: the (in)famous report. In: *MT News International*, Nummer 14: 15-21. Zugriff über: <http://www.mt-archive.info/MTNI-14.pdf>
- Hutchins, John. 1998. The Origins of the Translator's Workstation. In: *Machine Translation*, Vol.13, Nr.4: 287-307.
- Hutter, Hans-Peter/Jekat, Susanne. 2010³. Evaluation von sprachverstehenden und -generierenden Systemen. In: Carstensen, Kai-Uwe et al. (Hrsg.), 659-678.
- Jekat, Susanne/Volk, Martin. 2010³. Maschinelle und computergestützte Übersetzung. In: Carstensen, Kai-Uwe et al. (Hrsg.), 642-658.
- Kit, Chunyu/Tak-ming, Billy Wong. 2015. Evaluation in machine translation and computer-aided translation. In: Chan, Sin-Wai (Hrsg.), 213-236.
- Lengauer, Alina. 2012. *Einführung in das Europarecht. Europäisches Verfassungsrecht*. Wien: Rechtswissenschaftliche Fakultät Universität Wien.
- Melby, Alan K. 1981. A bilingual concordance system and its use in linguistic studies. In: Gutwinski/Jolly (Hrsg.). *The Eighth LACUS Forum 1981*. Columbia, SC: Hornbeam Press, 541-549.
- Melby, Alan K./Wright, Sue Ellen 2015. Translation Memory. In: Chan, Sin-Wai (Hrsg.), 662-677.
- Menzel, Wolfgang. 2010³. Zur Geschichte der Computerlinguistik. In: Carstensen, Kai-Uwe et al. (Hrsg.), 18-25.
- Mieg, Harald/ Näf, Matthias. 2005². *Experteninterviews*. Zürich: Institut für Mensch-Umwelt Systeme (HES).
- Pollak, Johannes/Slominski, Peter. 2012². *Das politische System der EU*. Wien: Facultas.

Pym, Anthony. 2011. What technology does to translating. In: *The International Journal for Translation & Interpreting Research*, Vol 3, Nr 1: 1-9.

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

Rehm, Georg (Ltg.). 2015. *Strategic Agenda for the Multilingual Digital Single Market. Technologies for Overcoming Language Barriers towards a truly integrated European Online Market*. (Version 0.5) Zugriff über: <http://www.meta-net.eu/projects/cracker/multimedia/mdsm-sria-draft.pdf>, Stand: 20.10.2015

Reinke, Uwe. 2004. *Translation Memories. System-Konzepte-Linguistische Optimierung*. Frankfurt: Peter Lang Verlag.

Reithofer, Karin. 2014. *English als Lingua Franca und Dolmetschen*. Tübingen: Narr Verlag.

Schmale, Wolfgang. 2000. *Geschichte Europas*. Wien: Böhlau Verlag.

Seewald-Heeg, Uta/Stein, Daniel. 2009. Maschinelle Übersetzung – von der Theorie zur Anwendung. In: *Journal for Language Technology and Computational Linguistics* 24/3: 5-70.

Sproat, Richard. 2010. *Language, Technology and Society*. Oxford: Oxford University Press.

Wagner, Emma/Bech, Svend/Martínez M. Jesús. 2012². *Translating for the European Union institutions*. Manchester: St. Jerome Publishing.

Zimmermann, Harald H. 2004. Maschinelle und Computergestützte Übersetzung. In: *Grundlagen der praktischen Information und Dokumentation*. München: K.G. Saur, 475-480.

Literatur der Europäischen Kommission:

Europäische Kommission. 2009a. *Geschichte des Übersetzungsdienstes der Europäischen Kommission*. Luxemburg: Amt für Veröffentlichungen der Europäischen Union.

Europäische Kommission. 2009b. *Übersetzung: Hilfsmittel und Arbeitsablauf*. Luxemburg: Amt für Veröffentlichungen der Europäischen Gemeinschaften.

Europäische Kommission. 2011. *Bericht über die 5. EMT- Konferenz (European Master's in Translation), Brüssel, 30. November 2011*. Zugriff über: http://ec.europa.eu/dgs/translation/programmes/emt/conferences/emt_conference2011_report_en.pdf

Europäische Kommission. 2012a. *Übersetzung und Mehrsprachigkeit*. Luxemburg: Amt für Veröffentlichungen der Europäischen Union.

Europäische Kommission. 2012b. *Quantifying Quality Costs and the Cost of Poor Quality in Translation. Quality Efforts and the Consequences of Poor Quality in the European Commission's Directorate-General for Translation*. Luxemburg: Amt für Veröffentlichungen der Europäischen Union.

Europäische Kommission. 2013. *Übersetzung – Arbeitsablauf und technische Hilfsmittel*. Luxemburg: Amt für Veröffentlichungen der Europäischen Union.

Europäische Kommission. 2014a. *CEF DSI maturity. Final report. A study prepared for the European Commission DG Communications Networks, Content & Technology*. Luxemburg: Amt für Veröffentlichungen der Europäischen Union.

Europäische Kommission. 2014b. *MT@EC. Secure Machine Translation for the European Union*. Luxemburg: Amt für Veröffentlichungen der Europäischen Union.

Europäische Kommission. 2014c. *Machine translation by the European Commission for public administrations in the member states. Description of pilots*. Zugriff über: http://ec.europa.eu/isa/documents/mt-ec-for-member-states-description-of-pilots_en.pdf

EU-Gesetzesquellen:

Vertrag über die Europäische Union. 2012. Amtsblatt Nr. C 326, 1-390.

Vertrag von Amsterdam. 1997. Amtsblatt Nr. C 340, 1-144.

Mitteilung Interinstitutionelle Vereinbarung vom 22. Dezember 1998 – Gemeinsame Leitlinien für die redaktionelle Qualität der gemeinschaftlichen Rechtsvorschriften. 1998. Amtsblatt Nr. C 073, 1-4.

Interview:

Mai, Katja. 22. Dezember 2015.

Sonstige Quellen:

<http://www.casmacat.eu/>

<https://cms.unov.org/meta/mercury-quick-start>

<https://www.data.gv.at/>

- ^a http://ec.europa.eu/about/ds_de.htm
- ^b http://ec.europa.eu/about/index_de.htm#president
- ^c http://ec.europa.eu/about/index_de.htm#regular
- ^d <http://ec.europa.eu/citizens-initiative/public/basic-facts>
- ^e http://ec.europa.eu/dgs/translation/translationresources/machine_translation/index_en.htm
- ^f http://ec.europa.eu/dgs/translation/whoweare/index_de.htm
- ^g http://ec.europa.eu/dgs/translation/whoweare/translation_figures_de.pdf
- ^h http://ec.europa.eu/newsroom/dae/document.cfm?doc_id=4634
- ⁱ <https://ec.europa.eu/digital-agenda/en/digital-single-market>
- ^j <http://ec.europa.eu/idabc/en/document/2097/5644.html>
- ^k http://ec.europa.eu/internal_market/imi-net/about/index_de.htm
- ^l http://ec.europa.eu/ipg/content/tips/words-style/jargon-alternatives_en.htm
- ^m http://ec.europa.eu/isa/actions/documents/isa_2.8_machine_translation_workprogramme.pdf
- ⁿ http://ec.europa.eu/isa/documents/mt-ec-service-description_en.pdf
- ^o http://ec.europa.eu/priorities/digital-single-market/better-online-access-digital-goods-and-services_de
- ^p http://ec.europa.eu/priorities/digital-single-market_de
- ^q http://ec.europa.eu/priorities/digital-single-market/digital-driver-growth_de
- ^r http://ec.europa.eu/priorities/digital-single-market/environment-where-digital-networks-and-services-can-prosper_de
- ^s http://ec.europa.eu/public_opinion/archives/ebs/ebs_386_en.pdf
- ^t http://ec.europa.eu/solvit/what-is-solvit/index_de.htm
- <http://elrc-share.ilsp.gr/>
- <http://elrc.tilde.com/de/uber-cef>
- <http://www.epsos.eu/oesterreich/start/ueber-epsos.html>
- http://www.europarl.europa.eu/external/appendix/legislativeprocedure/euoparl_ordinarylegislativeprocedure_complete_text_de.pdf
- ^a http://iate.europa.eu/about_IATE.html

^b http://iate.europa.eu/faq/IATE_FAQ_DE.htm

^a <https://joinup.ec.europa.eu/community/cef/description>

^b https://joinup.ec.europa.eu/community/cef/og_page/catalogue-building-blocks

<http://lr-coordination.eu/resources>

<http://www.matecat.com/benefits/>

<http://www.multilingualeurope.eu/de>

^a <https://open-data.europa.eu/de/about>

^b <https://open-data.europa.eu/de/linked-data>

http://www.rigasummit2015.eu/sites/rigasummit2015.eu/files/cef_29_04_2015_spyridon_pilos_building_the_automated_translation_platform.pdf

<http://www.statmt.org/moses>

<https://scic.ec.europa.eu/streaming/machine-translation>

<http://www.sdl.com/de/download/an-introduction-to-machine-translation-and-postediting/71101/>

<https://termcoord.files.wordpress.com/2011/06/machine-translation-for-translators-and-terminology-applications-in-the-eu1.ppt>

<http://www.zukunfteuropa.at/site/4664/default.aspx>

Abbildungs- und Tabellenverzeichnis:

Anmerkung: Ich habe mich bemüht, sämtliche InhaberInnen der Bilderrechte ausfindig zu machen und ihre Zustimmung zur Verwendung der Bilder in dieser Masterarbeit einzuholen. Sollte dennoch eine Urheberrechtsverletzung bekannt werden, ersuche ich um Meldung bei mir.

Abbildung 1. Vauquois Dreieck (Seewald-Heeg/Stein 2009: 9).....	11
Abbildung 2. Training eines SMÜ-Systems (Jekat/Volk 2010 ³ : 649)	14
Abbildung 3. Verlauf der Beitritte und mögliche Kandidaten (Bergmann 2015: 17).....	23
Abbildung 4. Die häufigsten Dokumenttypen in der Europäischen Kommission (http://www.dfki.de)	29
Abbildung 5. Plattform der (sprach)technologischen Hilfsmittel für ÜbersetzerInnen bis ca. 2009 (Europäische Kommission 2009: 37).....	36
Abbildung 6. Struktur des MT@EC Systems (https://termcoord.files.wordpress.com)	47

Abbildung 7. Ausschnitt aus dem MT@EC API (Europäische Kommission 2014c: 2)	49
Abbildung 8. Vergleich der Pilotprojekte (Europäische Kommission 2014c: 6)	55
Abbildung 9. Sprachtechnologien als Grundlage des Digital Single Market (Rehm 2015: 11)	61
Abbildung 10. Geplanter Ablauf der CEF.AT Einführung (Europäische Kommission 2014a: 17).....	69
Tabelle 1. Qualität nach Sprachenpaaren, ermittelt durch BLEU Tests (http://ec.europa.ec ⁿ).....	50
Tabelle 2. Anwendung CEF.AT (erstellt anhand von https://joinup.ec.europa.eu ^b).....	62

Anhang: Interview-Transkript

Anmerkung: Das nachfolgende Transkript umfasst das eigentliche Gespräch“, d.h. Begrüßung, Einleitung, Verabschiedung etc. wurden hier nicht festgehalten, da sie für die Arbeit keine Relevanz haben.

Während an dem Inhalt des Interviews nichts verändert wurde, wurden zu Gunsten der besseren Lesbarkeit und Verständlichkeit der schriftlichen Version einzelne redundante Wörter gestrichen oder Sätze gekürzt bzw. ergänzt, die sich aus der mündlichen Gesprächssituation ergaben, für das Transkript jedoch keine Bedeutung hatten.

Zu Beginn erläuterte Frau Mai als Teil ihrer Vorstellung in einem kurzen Exkurs den Einsatz ihrer Arbeitssprachen und die Dominanz des Englisch:

In der Realität ist es aber wirklich so, dass wir sehr viele englische Originale haben. Ich arbeite vor allem im Bereich Außenbeziehungen, also humanitäre Hilfe, Erweiterung, Entwicklungshilfe etc. und da sind die Originale mindestens zu 90% auf Englisch. In manchen Bereichen ist das unterschiedlich, aber die Tendenz ist ganz massiv [in Richtung des Englischen]. Gerade seit 2004, seitdem die „neuen“ Mitgliedstaaten [oder jetzt auch nicht mehr so neuen MS] dabei sind, hat Englisch wirklich die Überhand gewonnen. Es gibt manchmal noch französische Originale aus dem Entwicklungsbereich, wenn es um französischsprachige Länder geht oder um die ehemaligen französischen Kolonien – da kommt manchmal noch etwas auf Französisch. Aber mittlerweile könnte man sagen, dass die reale Arbeitssprache in der Kommission überwiegend das Englische ist, so dass ich jetzt fast nur noch aus dem Englischen übersetze.

Welche Textsorten übersetzen Sie als Mitarbeiterin der DGT?

Schätzungsweise sind bestimmt $\frac{2}{3}$ Rechtstexte verschiedenster Art, angefangen von ganz einfachen Beschlüssen (z.B. im Außenbereich gibt es viele Finanzierungsbeschlüsse zu bestimmten Projekten [bis hin zu Gesetzen]. Dann gibt es auch politische Texte, also Verordnungsvorschläge, dann offizielle Mitteilungen der Kommission, z.B. Mitteilung der Kommission an den Rat und an das Parlament. Dies gehört natürlich alles zu den Rechtstexten aber dabei handelt es sich natürlich auch um eine relativ weite Spanne. Finanzierungs-vorschläge können beispielsweise auch technische oder allgemein gehalten sein bis hin zu eben Verordnungsvorschlägen, die dann „mehr“ wirkliche Rechtstexte sind [im engeren Sinne]. Und dann gibt es noch einmal einen Teil, bei dem ich sagen würde, dass er vielleicht 20% ausmacht. Da haben wir manchmal Texte zu Videos, z.B. gibt es eine Serie von Videos, die zur Erweiterung gemacht werden. Da haben wir die Untertitel und den Text, der eingeblendet werden soll. Das kommt also vor oder irgendwelche Broschüren. Ein Teil sind auch Fragen von Parlamentariern an die Kommission.

Es gibt auch Fragen von der Kommission, die innerhalb einer bestimmten Zeit beantwortet werden müssen, das sind praktisch politische oder inhaltliche Fragen. Und es gibt auch Bürgerschreiben.

Welche Tools verwenden Sie dabei jeweils? Macht die jeweilige Textsorte bei der Wahl des Tools einen Unterschied? Gibt es Tools, die nur für bestimmte Texte geeignet sind?

Der Großteil der Übersetzer arbeitet eigentlich mit Studio, bis auf ein paar Kollegen der älteren Generation, die weniger mit Studio arbeiten und dann gibt es auch eine kleine Gruppe von Übersetzern, die mit OmegaT arbeiten. Es gibt noch Fälle, z.B. kurze Pressemitteilungen oder eben Bürgerschreiben – also wenn es wirklich ein Text ist, der im Grunde komplett neu übersetzt wird, dann arbeiten wir manchmal auch ohne Studio und dann einfach mal direkt in Word. Aber das ist zum Teil nach persönlicher Vorliebe und nach Text. Aber solchen Fällen wird das dann nachträglich aligniert, damit es auch in unsere Datenbank kommt. Normalerweise arbeiten wir noch mit Studio und dann gibt es eben noch diese Möglichkeit OmegaT zu nutzen und das wird von den meisten Kollegen z.B. für größere Pakete genutzt oder für Texte, wo sich OmegaT besser eignet. Aber es gibt ein paar OmegaT Fans, vor allem in der portugiesischen Abteilung. Das ist aber auch nicht ohne Probleme, weil wir ja auch oft zum Beispiel an größeren Dokumenten arbeiten, die aufgeteilt werden und mehrere daran arbeiten – mit einem gemeinsamen Speicher. Und wenn dann Kollegen mit OmegaT arbeiten, dann wird es schwierig. Aber Standard ist, dass wir mit Studio arbeiten.

Oder auf anderem Wege und normalerweise werden die Ergebnisse dann in diese zentrale Datenbank, unseren Speicher Euramis geschickt, in dem praktisch alles abgelegt wird. Im Grunde genommen unabhängig davon, wie es produziert wurde. Gerade wenn Texte auf anderem Wege erstellt wurden, können wir für diese auch nachträglich ein Alignement anfordern und dann wird es trotzdem an die Datenbank geschickt.

Welche Vorteile und welche Nachteile sehen Sie bei den zurzeit verwendeten Sprachtechnologien der GDU?

Was ich dazu eigentlich als erstes sagen möchte: ich war in einer Arbeitsgruppe, die sich Gedanken darüber gemacht hat, wie es in der nächsten Zeit weitergehen soll und habe dadurch noch einen anderen Blick darauf. Was im Moment im Grunde genommen unsere Arbeit zum Teil behindert, ist die Tatsache, dass wir mit Studio eigentlich mit einer Einzeltextversion arbeiten, also wir arbeiten nicht *server-based*. D.h. wir müssen, wenn wir zusammenarbeiten eigentlich Umwege gehen, die es zum Teil etwas schwerfällig machen. Wir arbeiten z.B. mit einer gemeinsamen Memory, wenn wir zu mehreren arbeiten – aber durch die Art wie wir das machen, gibt es bestimmte Zeitverzögerungen. Dann gibt es Möglichkeiten in Studio, die wir auf Grund der Einwendung in unser gesamtes System eigentlich nicht nutzen können. Denn bei uns ist das System so strukturiert, dass wir praktisch für

jeden Text, also auch für jeden Textteil, eine einzelne Nummer haben und dieser Text wird einzeln bearbeitet. Der normale Workflow ist so, dass man für jeden Text ein eigenes Projekt anlegt. D.h. wenn wir können dann z.B. drei Teile haben für den gleichen Text und in dem Programm, also in unserer Struktur ist es so vorgesehen, dass wir da drei einzelne Projekte anlegen. Somit kann man eigentlich die Vorteile, die man in Studio hat, nämlich dass man in einem Projekt mehrere Texte bearbeiten kann, eigentlich bei uns nicht nutzen oder wiederum nur mit Umwegen. Hier hat man Reibungsverluste. Das ist, was die Nutzung des Potenzials der Tools angeht ein Problem. Dann gibt es natürlich die Probleme mit dem Tool selbst. Wir arbeiten mit Studio 2014 – und da gibt es immer wieder Sachen, die nicht so funktionieren, wie man sich die eigentlich wünschen würde oder wo uns auch nicht ganz klar ist, wie das Programm eigentlich funktioniert bzw. funktionieren sollte. Studio gehört den „Studio-Leuten“, wir haben nur eine Hälfte davon, Leute, die uns da ein Wiki anlegen mit allen Fehlermeldungen und den Lösungen, also wir haben schon ein gutes IT Team, das uns da unterstützt, aber es gibt Sachen an die die auch nicht herankönnen, weil das Programm derart gestaltet ist.

Der Großteil der aktuellen Projekte betrifft überwiegend den Bereich der maschinellen Übersetzung. Gibt es auch neue Entwicklungen im Bereich der CAT Tools oder Bestreben in diesem Bereich etwas zu verändern?

Es ist so, dass es vor zwei, drei Jahren eine Ausschreibung gab. Vorher haben wir mit *Trados Workbench* gearbeitet. Diese Ausschreibung hat SDL Studio gewonnen. Dieser Vertrag läuft 2016 aus. Wir sind an die Regeln der Haushaltsordnung gebunden und können also nicht frei entscheiden „Wir möchten heute dies und morgen das“. Wir können auch nicht sagen, „Ach wir haben uns jetzt so gut an Studio gewöhnt, wir arbeiten da einfach damit weiter.“ Es gibt jedenfalls rechtliche Vorgaben, die wir einhalten müssen und das war auch das Problem, das wir eigentlich vorher hatten. Wir haben sehr lange mit TWB gearbeitet und wenn man das Programm erst einmal kauft, hat man mit der Lizenz die Möglichkeit, das Programm unendlich weiter zu nutzen. Was man aber nicht hat, ist das Recht auf regelmäßige Updates, sodass wir praktisch lange mit TWB gearbeitet haben, ohne eigentlich an die neuen Entwicklungen angebunden zu sein. Deswegen gab es jetzt eine Arbeitsgruppe, die sich darüber Gedanken gemacht hat, wie man so etwas künftig vermeiden kann, dass man so einen großen Bruch dann hat. Da gibt es halt Überlegungen, ob man vielleicht auf Dauer ein internes Tools entwickelt. Zum Beispiel ist da jetzt das Parament dabei, das hat schon für einen Teil der Dokumente mit denen sie arbeiten, also Änderungen oder Änderungsanträge, etwas entwickelt. Es gibt ja auch sehr viele Open Source Programme und auch die Vereinten Nationen haben sich ein eigenes Programm entwickelt. In diesen Arbeitsgruppen haben wir auch die verschiedenen Möglichkeiten uns angesehen und sie sind alle mit Vor- und Nachteilen verbunden. Wir haben den Vorteil, dass wir ein großes System sind, das heißt wir haben Support, wir haben Leute, die sich intern wirklich spezialisieren

können und da Tools auch so anpassen, wie es in unsere Arbeitsumgebung passt. Gleichzeitig ist das aber natürlich auch eine Schwerfälligkeit. Auch wenn wir jetzt z.B. eine neue Ausschreibung machen würden, würde das alles dann Jahre dauern. Studio 2015 ist jetzt auch schon auf dem Markt und in verschiedenen Versionen und das dauert immer eine ganze Weile, bis das bei uns ankommt, bis es dann bei uns getestet wird, bis es eingebunden werden kann, das muss auch alles wieder praktisch an unsere anderen Tools angepasst werden – da gibt es immer eine große zeitliche Verzögerung. Kurz gesagt, sind Überlegungen da, angedacht war zum Beispiel die Möglichkeit mit mehreren Tools zu arbeiten, da vielleicht verschiedene Textarten unterschiedlich gut geeignet sind. Das scheint sich, so wie mein Eindruck ist, nicht wirklich durchzusetzen, weil wir sehr viel Aufwand haben. Wir machen ja auch interne Schulungen, die Leute müssen kompatibel sein. Wie ich vorhin gesagt habe, wenn ich mit einem Programm arbeite und mein Kollege mit dem anderen, dann muss es auch eine Möglichkeit geben zusammenzuarbeiten. Ein gewisses Maß an Flexibilität ist vorgesehen, aber gleichzeitig muss es auch klar sein, dass wir alle miteinander arbeiten können. Die Überlegungen etwas selbst zu entwickeln ist im Gespräch, aber es wird noch dauern. Und es gibt bestimmte Anforderungen bei der Kommission. Wir brauchen ein Programm, das mit vielen verschiedenen Datenformaten arbeiten kann. Wir arbeiten zwar vor allem mit Text, aber wir müssen genauso in der Lage sein PowerPoint Präsentationen oder irgendwelche Excel Texte zu bearbeiten. Und das ist bei den meisten Tools, die zum Beispiel auf Open Source basieren schwer, denn das sind meistens eher Schmalspurprogramme, was den Input angeht. Wir versuchen am Ball zu bleiben, aber wir haben halt die Vor- und Nachteile einer großen Organisation.

Die Umsetzung gestaltet sich demnach schwierig.

Ja, es dauert einfach. Wir müssen zu viele Meinungen berücksichtigen, wir müssen sehr viele technische Vorgaben berücksichtigen, wir [sind an...bestimmte...gebunden] eine Verwaltung. Wir müssen einerseits mit den Auftraggebern kommunizieren können, wir haben ein System, wo die Aufträge reinkommen, die müssen dann an den Übersetzer weitergeleitet werden in den verschiedenen Sprachen und über das System auch wieder zurückgeladen werden und das ist dann alles eben auch an das entsprechende Tool angebunden. Und wir müssen sichern, dass unsere Ergebnisse richtig in die Datenbank eingespeichert werden, sodass wir auch nach Dokumentennummer die Texte wieder auffinden können.

Im Gegensatz dazu sind im Bereich der maschinellen Übersetzung konkrete Pläne vorhanden, deren Umsetzung zurzeit im Gange ist. Laut Angaben ist die Prozentzahl zurzeit eher gering. In welchem Umfang kommt maschinelle Übersetzung in der Praxis zum Einsatz? Wird die Anwendung in der Zukunft steigen?

Es gibt eine besondere Auftragsart, die sich offiziell „post-edition rapide“ nennt. Das sind grob überarbeitete maschinelle Übersetzungen. Der Auftrag ist offiziell nur kurz eine Übersetzung zu Verständniszwecken anzufertigen, wo maschinelle Übersetzung zugrunde liegt

und sie grob, minimal durchbearbeitet. Ich denke darauf bezieht sich dieser Prozentanteil [von unter 1%]. Das ist etwas, was eigentlich noch neu ist und zu ganz speziellen, begrenzten Zwecken eingesetzt wird. Ansonsten ist es jetzt so, dass für bestimmte Sprachkombinationen, auf jeden Fall aus dem Englischen und aus dem Französischem ins Deutsche automatisch, wenn ein Auftrag reinkommt auch eine MT abgefragt wird. Das heißt ich habe die Möglichkeit, das wird mir, wenn ich mein Projekt anlege als Ressource automatisch mitangeboten. Das heißt ich habe ein System, bei dem ich sagen kann, welche Ressourcen ich zu meinem Projekt dazuladen möchte und dann kann ich das ankreuzen oder nicht, und das würde ich textabhängig machen. Ich habe hier jetzt zum Beispiel eine Verordnung, [nämlich] die 240. Veränderung der Al-Qaida Verordnung. Das ist eine Verordnung, die seit 2002 besteht und jedes Jahr kommen x Änderungen, weil Personen aus dieser Liste gestrichen werden – das basiert auf der Liste der Vereinten Nationen – Leute und Unternehmen werden auf diese Liste gesetzt, Einträge werden geändert und gestrichen. Das ist zum Beispiel ein Fall, bei dem ich maschinelle Übersetzung garantiert nicht verwenden werde, weil es in dem Fall überhaupt keinen Sinn macht – weil ich die Vorläufertexte berücksichtigen muss. Es gibt [darin] praktisch keinen neuen Text und wenn ich da maschinelle Übersetzung verwenden würde, dann würde das nur fehlgeleitet werden. Das gleiche gilt für Verordnungen und Texte, bei denen es ganz feste Vorgaben gibt und wir haben oft Verordnungen, die als Rechtsgrundlage dienen und die dann weitergeführt werden oder Dokumente, die darauf basieren. Dadurch, dass unsere MT auf dem gesamten Fundus basiert und ich nicht sagen kann, ich möchte jetzt die MT nur auf der Grundlage des Außenbereichs [gemeint: Außenbeziehungen] oder bestimmter Dokumente, wird sich die Terminologie aus all dem zusammensetzen, was in Euramis ist. Deswegen kann ich nicht sicher sein, dass mir genau das angeboten wird, was ich brauche, also die passende Terminologie. [Es gibt aber auch Texte,] wo es gut funktioniert. Wir haben jetzt z.B. immer Texte zum Europäischen Semester, die sehr einheitlich sind und wo sich die Sätze wiederholen, wo sich die Formulierungen wiederholen und die gleichen Formulierungen auch immer auf die gleiche Weise zu übersetzen sind. Wo es um die wirtschaftliche Entwicklung in den einzelnen Ländern und dann gibt es einen relativ festen Fundus an Termini und relativ viel Material auch dazu, weil alle halben Jahre für den Großteil der Länder entsprechende Bericht erstellt werden, also gibt es auch viel Material in den Daten der maschinellen Übersetzung. Dadurch wird dann die Qualität besser. Bei den Texten, die praktisch wirklich frei zu übersetzen wären, da merkt man immer wieder, da schwimmt die maschinelle Übersetzung – da kommen dann oft sehr lustige Sachen heraus, die uns wirklich sehr erheitern, das macht immer wieder mal Spaß. Aber das hat damit zu tun, dass eben praktisch kaum Material da ist und die maschinelle Übersetzung, die ja eine statistische maschinelle Übersetzung ist, sucht mehr und findet weniger. So, z.B. sind die Texte des Europäischen Semesters beschreibende Berichte, die werden wahrscheinlich auch im Amtsblatt veröffentlicht werden, haben die nicht die glei-

chen Ansprüche und die gleiche Rechtswirkung, wie eine Verordnung, wo jedes Wort stimmen muss und wo ich auch nicht viele Freiheiten habe. Tendenziell, kann man [die Verwendung der MT in der DGT] so grob Einstufen.

Dies beinhaltet dann wohl auch schon im Großen und Ganzen die Antwort auf die nächste Frage: Wie stark muss die maschinelle Übersetzung durchschnittlich noch bearbeitet werden, damit man von einer adäquaten Übersetzung im Sinne des GDU fit-for-purpose Prinzips sprechen kann?

Genau, es hängt von der Textsorte ab und immer wieder von... Also tendenziell ist es so: kurze, einfach strukturierte Sätze im Englischen gut aufgebaut und klar verständlich, die kommen im Deutschen auch gut an. Lange, komplizierte, verschachtelte Sätze – da kommt die maschinelle Übersetzung oft nicht mit. Ich muss wirklich sagen, das schwankt massiv. An sich ist es so, dass sich das System schon weiterentwickelt hat und dass es manche Kollegen tatsächlich systematisch nutzen, sobald es sich eignet und in anderen Fällen weniger. Das ist ja noch nicht sehr lange bei uns eingeführt und auch weiterhin freigestellt, sodass es wirklich unterschiedliche Strategien gibt. Aber wie gesagt, bei manchen Texten ist es wirklich eine deutliche und sehr gute Hilfe.

Haben Sie abgesehen von den aktuellen auch ältere Versionen der GDU Übersetzungstools verwendet? Wie waren diese im Vergleich zu den neuen technologischen Hilfsmitteln?

Nein, mit denen habe ich nicht gearbeitet. Vor allem bei Systran haben sich die portugiesische und besonders die spanische Abteilung mitgearbeitet. Sie hatten auch bessere Ergebnisse. Im Deutschen haben wir jetzt angefangen, seitdem wir dieses statistische System MT@EC haben. Vorher nicht wirklich. Die Ergebnisse sind auch je nach Sprachkombination extrem unterschiedlich. Ich denke, die regelbasierte maschinelle Übersetzung war für die Deutsche wirklich nicht sehr nützlich.

Welchen Einfluss haben die neuen Technologien auf die Arbeit der GDU bzw. ihrer Übersetzer und Übersetzerinnen? Verschieben sich dadurch immer mehr die Aufgabenbereiche? Als Ausblick: inwiefern könnte sich die Arbeit in den nächsten Jahren verändern?

Die Situation hat sich bereits in den letzten zehn Jahren massiv verschoben und stark entwickelt. Vor meiner Anstellung bei der Kommission war ich als freiberufliche Übersetzerin und Dolmetscherin tätig und habe damals noch keine CAT Tools verwendet – es war eine Zeit in der es auch noch ohne ging und die Verwendung der CAT Tools gerade erst anging. Das Thema schürt immer große Ängste. Es gibt zum Teil Widerstände gegen die maschinelle Übersetzung und oft insgesamt gegen neue Tools, was zuerst auch oft eine Schreckreaktion ist. Es gibt Befürchtungen. Es ist sowieso so, dass in nächster Zeit, in den nächsten Jahren, nicht nur in der Kommission sondern speziell auch in der DGT, tendenziell mehr Personal abgebaut wird noch, als in den anderen Generaldirektionen. Die DGT ist eine Generaldirektion in der man mehr einsparen kann als in den Generaldirektionen, die politische

Arbeit leisten. Das heißt, dass dies einerseits ausgeglichen werden soll durch mehr Zusammenarbeit mit Freelancern, aber es wird auch den internen Druck steigern, weil auch das auf die Verteilung der Stellen nicht unbedingt in allen Sprachabteilungen gleichen wird. Je nach Altersstruktur werden Stellen nachbesetzt oder nicht. Also bei uns wird es so sein, dass in der deutschen Abteilung z.B. nach und nach eher ältere Kollegen dabei, die in den nächsten Jahren ausschieden werden und viele von ihnen werden nicht ersetzt. Oder es wird dann durch junge und weniger erfahrende Kollegen übersetzt. Es ist zu erwarten, dass wir noch mehr Druck bekommen, weil wir manche Sachen wiederum auch nicht rausgeben können und wenn wir Sachen maschinell herausgeben, dann ist es für uns ja auch mit Arbeit verbunden – das Vorbereiten, das Evaluieren und zum Teil Nachbereiten. Somit ist die Befürchtung schon, dass wir uns einfach gezwungen sehen werden mehr z.B. auch die maschinelle Übersetzung zu nutzen, um unsere Standards halten zu können. Die Frage ist, wie sich Qualität und Quantität weiterhin in der Waage halten werden. Es gibt Leute, die befürchten, dass die maschinelle Übersetzung praktisch als Druck dafür genutzt wird, dass man damit die Ansprüche an die Qualität im Grunde genommen senkt und man eben nur eine Art Nachbearbeitung von der maschinellen Übersetzung macht. Aber das ist ein Teil unserer Arbeit, wir haben eigentlich wie gesagt selten Texte, die komplett neu sind. Wir haben oft Texte, wo wir ganz viel bestehende Vorgaben schon einerseits nutzen, aber auch beachten müssen, sodass wir meistens schon viele Teile – *100% Matches* oder *Fuzzy Matches* – schon haben und das mit den neuen Teilen ergänzen und dann mit der MT arbeiten. Also man kann das nicht wissen, aber es könnte schon sein, dass sich tendenziell auf Dauer, wie überall auf dem freien Markt auch, durch den Druck, um die Produktion aufrechterhalten zu können, man die Tools noch intensiver einsetzt und dadurch verändert sich natürlich die Arbeit. Das sind halt Prognosen. Das wird vielleicht dazu führen, dass man noch selektiver schaut, welche Texte man wie bearbeiten muss.

Diese Rechtstexte, bei denen es ganz fein darauf ankommt und wo man einerseits bestehende Terminologie berücksichtigen muss, die sich zum Teil auch widerspricht – wir haben auch manchmal [den Fall], dass verschiedene bestehende Verordnungen zusammengeführt werden. Die wurden aber z.B. entweder vor langer Zeit oder von unterschiedlichen Referaten bearbeitet, d.h. sie haben eine unterschiedliche Terminologie. Wenn man die dann zusammenführt, dann ist das immer sehr schwierig, da das zu finden, das rechtlich noch stimmt, was die Rechtssicherheit wahrt und trotzdem lesbar bleibt.

Vielen Dank für das Interview!

Curriculum Vitae

Persönliche Daten:

Vor- und Nachname: Borbála Katalin EKE
Geburtsdatum: 07.08.1992
Geburtsort: Wien, Österreich
Adresse: Khevenhüllerstraße 19/2/4
1180 Wien
E-Mail Adresse: borbala.eke@gmail.com
Sprachkombination: Deutsch – A, Ungarisch – B, Englisch – C

Ausbildung :

seit 2013 Universität Wien – Masterstudium Übersetzen
Schwerpunkt Fachübersetzen
seit 2012 Universität Wien – Diplomstudium Rechtswissenschaften
November 2014 Teilnahme United Nations Shadowing Programme 2014
UNIS Vienna Office, UNO City
2010 – 2013 Universität Wien – Bachelorstudium
Transkulturelle Kommunikation
Juli 2009 EF Sprachkurs; Oxford, Großbritannien
2002 – 2010 Bundesgymnasium Gymnasiumstraße 83, 1190 Wien
1998 – 2002 Volksschule Krottenbachstraße 108, 1190 Wien

Berufserfahrung Übersetzen und Sprachen:

seit März 2016 Volunteer im *Trommons* Projekt von *The Rosetta Foundation*
Übersetzen und Korrekturlesen
2015 Untertitelung Spielfilm *Underdog* (ung. OT: *Fehér Isten*)
mit Mag. Mandana Taban, Worldpool
April – September 2015 Untertitelungspraktikum, Okto Community-TV GmbH
Selbstständiges Erstellen von Untertiteln (DE-EN, EN-DE, HU-DE) mit
Annotation Edit, Verfassen kurzer Inhaltsangaben für Sendungshomepage
seit September 2014 Nachhilfe in Deutsch und Englisch
Vorbereitung von Studierenden auf den Cambridge Certificate Test, Verbesserung der Deutschkenntnisse für Beruf und Alltag, Lektorat von Seminararbeiten

Sprachkenntnisse:

Deutsch auf Muttersprachenniveau (C2)

Ungarisch auf Muttersprachenniveau (C2)

Fließend Englisch in Wort und Schrift (C1)

Basiskenntnisse Französisch (A2)

EDV Kenntnisse:

Microsoft Office (Word, Excel, PowerPoint)

Grundlegendes DTP Know-how (Scribus)

CAT-Tools (SDL Trados + SDL Multiterm)

Untertitelungstools (Annotation Edit, WinCaps)

Abstract

Deutsch:

Die vorliegende Masterarbeit beschäftigt sich mit den Sprachtechnologien, die in der Europäischen Kommission eingesetzt werden. Im Zentrum stehen die aktuellen Pläne, die in enger Zusammenarbeit der Kommission und der Generaldirektion Übersetzen durchgeführt werden. Im Rahmen dieser Projekte steht MT@EC, ein noch relativ junges System zur maschinellen Übersetzung, im Mittelpunkt.

Doch Sprachtechnologien spielen in der Kommission nicht erst seit den letzten Jahren eine Rolle. Um den Kontext der derzeitigen Neuerungen nachvollziehen zu können, wird deshalb ein Blick in die Vergangenheit der DGT geworfen. Zu diesem Zweck werden die Meilensteine in der Entwicklung von Übersetzungstechnologien wie der maschinellen Übersetzung und TM-basierten Tools dargestellt, die den Weg zu den jetzigen Anwendungen geebnet haben.

In das Ergebnis fließt auch ein Interview mit Katja Mai, Übersetzerin in der DGT, sowie Überlegungen zu den Auswirkungen der Sprachtechnologien auf ÜbersetzerInnen und den Übersetzungsprozess in der Generaldirektion mit ein.

Der Aufbau der Arbeit gestaltet sich folgendermaßen: zunächst beinhalten die ersten beiden Kapiteln einen Überblick über die Grundlagen der in der Arbeit behandelten Sprachtechnologien sowie über die Europäische Kommission samt der Generaldirektion Übersetzen. Der dritte Abschnitt widmet sich den derzeitigen Entwicklungen. Abschließend folgt die Beantwortung der zu Beginn gestellten Forschungsfragen samt Ausblick und Schlussfolgerungen zu der Beziehung zwischen Tools und ÜbersetzerInnen.

English:

This master's thesis concerns the language technologies, which are in use at the European Commission. At its center are the current plans, which are being executed in a close cooperation between the Commission and the Directorate General for Translation. Within these projects the focus is on MT@EC, a relatively young machine translation system.

However, language technologies have been a part of the Commission for more than just the recent years. In order to have a good understanding for the ongoing innovations, the history of the DGT is explored as well. Therefore, the thesis illustrates milestones in the evolution of translation technologies, e.g. machine translation and TM based tools, which have paved the way for their present use.

An interview with the DGT translator Katja Mai is also part of the content as well as reflections on the effects that language technologies might have on translators and the Directorate General's translation process.

The thesis has the following structure: the first two chapters consist of an overview of the basics of the examined language technologies and of the European Commission as well as the Directorate General for Translation. The third section details the current developments. The final part answers the research questions and includes an outlook as well as a conclusion concerning the relationship between tools and translators.