



universität
wien

MASTERARBEIT / MASTER'S THESIS

Titel der Masterarbeit / Title of the Master's Thesis

„Maschinelles Dolmetschen – Zukünftige Entwicklung
und deren Bedeutung für den Berufsstand
„Dolmetschende““

verfasst von / submitted by

Maria Cecilia Radwanski MA

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

Wien, 2017 / Vienna 2017

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

A 065 348 375

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

Masterstudium Dolmetschen Italienisch Polnisch

Betreut von / Supervisor:

Univ.-Prof. Mag. Dr. Gerhard Budin

Danksagung

Ein großer Dank gebührt allen voran meinen Eltern, die mir ermöglicht haben, meinen Leidenschaften nachzugehen und mir gleichzeitig die Bedeutung einer für das Wachstum und die Weiterentwicklung unabdingbaren disziplinierten Einstellung stets vor Augen geführt haben.

Genauso möchte ich mich bei meiner Tante, Maria Ausilia Bandino für Ihre entschiedene, fürsorgliche und richtungsweisende Art bedanken. Stets ist sie mir eine große Stütze und hat sie mich gelehrt, dass die Ausdauer früher oder später immer reiche Früchte trägt.

Meine besondere Wertschätzung möchte ich Monika Lerch zum Ausdruck bringen: dafür, dass sie jegliche in mir aufkommenden Zweifel beseitigen kann, für die Kraft, den Ehrgeiz und das Vertrauen, die sie mir bedingungslos erweist und dank denen sie nicht zuletzt eine große Motivationsquelle für das Studium, das Privat- und Berufsleben ist.

Schließlich gilt meinem Betreuer, Herrn Univ-Prof. Mag. Dr. Gerhard Budin mein herzlicher Dank. Trotz der vielen Verpflichtungen, zahlreichen anderen Studenten und StudentInnen gegenüber, stand er mir vor sowie während des Verfassens der vorliegenden Arbeit bei jeglichen Fragen sofort mit konstruktiver Kritik und aufschlussreichen Antworten stets zur Seite.

Danksagung.....	2
0. Einleitung.....	1
1. Maschinelles Dolmetschen (MD).....	2
1.1. Begriffserklärung.....	3
1.2. Geschichte.....	5
1.3. Anwendungsbereiche.....	10
1.3.1. Multilinguale Dialogsysteme.....	10
1.3.2. Telekommunikation.....	11
1.3.3. Militäreinsätze.....	11
1.3.4. Akademische Zwecke.....	12
2. Das MD-System.....	12
2.1. Technik – Wie funktioniert ein MD-System?.....	12
2.1.1. Automatische Spracherkennung (ASR/ <i>Speech-to-Text</i>).....	14
2.1.2. Maschinelle Übersetzung (MÜ/ <i>Text-to-Text</i>).....	16
2.1.3. Sprachsynthese (<i>Text-to-Speech/TTS</i>).....	23
2.2. Technische Entwicklung und Vergleich.....	24
2.2.1. Verbmobil.....	28
2.2.2. Jibbiggo.....	30
2.2.3. Skype Translator.....	31
2.2.4. Lecture Translator (LT).....	32
2.3. Voraussichtliche zukünftige (technische) Entwicklung und Erwartungen.....	37
2.3.1. Paradigmenwechsel – SMT zu NMT.....	37
2.3.2. Technischer Ausblick vs. Limitationen.....	39
2.3.3. MT/MD-Marktanalyse: Ein Überblick.....	43
3. Qualitätsmessung von Dolmetschleistungen.....	44
3.1. Begriffsbestimmung Qualität.....	45
3.2. Qualitätsparameter.....	47
3.3. Methodologien der Qualitätsmessung.....	53
4. MD vs. Humandolmetschen.....	55
4.1. Dolmetschleistungen im Vergleich.....	56
4.1.1. Lecture Translator vs. Humandolmetscher.....	56
4.1.2. TC-Star-System vs. Humandolmetscher.....	60
5. Berufsstand „DolmetscherIn“ in Österreich.....	62
5.1. Exkurs: Darstellung in österreichischen (Online-)Presseartikeln.....	68
5.1.1. Imagearten.....	71
5.1.2. Dominierende Thematiken.....	73
5.2. Bedeutung des MD für den Berufsstand „DolmetscherIn“.....	74
5.2.1. Voraussichtliche Erwartungen und Entwicklung.....	76
6. Zusammenfassung.....	77
7. Bibliographie.....	79
Anhang.....	89
Verbmobil – Technische Ziele.....	89
Fragebogen Fünfer Studie.....	90
Abstract.....	92

Abbildung 1: Darstellung der Dolmetschtätigkeit aus dem Alten Ägypten	4
Abbildung 2: Haremhab Relief.....	4
Abbildung 3: Artsrounis „maschinelles Gehirn" (vgl. Hutchins 2004).....	7
Abbildung 4: Sprach-zu-Sprach-Übersetzungssystem (Waibel 2015b:104).....	14
Abbildung 5: Ein typisches Spracherkennungssystem (Waibel 2015b:105).....	15
Abbildung 6: MÜ-Dreieck nach Vauquois (vgl. Carstensen et al. 2001:516).....	17
Abbildung 7: Maschinelle Übersetzung (Text-to-Text)	18
Abbildung 8: Direkte Übersetzung (Carstensen et al. 2001:518).....	19
Abbildung 9: Transferansatz (Carstensen et al. 2001:518).....	19
Abbildung 10: Interlingua-Ansatz (Carstensen et al. 2001:518)	20
Abbildung 11: Beispiel für ein multilinguales Interlingua-System (vgl. Weinstein 2002:4)..	21
Abbildung 12: Simultandolmetschen und -übersetzen über Richtungs-lautsprecher und Heads-Up-Display-Brillen	27
Abbildung 13: Benutzeroberfläche des Forschungsprototypen von Verbmobil.....	29
Abbildung 14: EU-Bridge-Architektur: MD als cloudbasierter Dienst (Waibel 2015:116)....	33
Abbildung 15: Übersetzungsprozess des LT	35
Abbildung 16: Paradigmenwechsel in der MT (Vogel 2016).....	39
Abbildung 17: MÜ vs. HÜ (Vogel 2016: 16:03).....	44
Abbildung 18: Qualität von Dolmetschleistung aus verschiedenen Perspektiven (Pöchhacker 2002:98)	46
Abbildung 19: Qualitätsparameter des Dolmetschens als Produkt oder Dienstleistung (Pöchhacker 2002:97).....	48
Abbildung 20: durchschnittliche Beurteilung der von Fünfer (2013) abgefragten Kriterien durch die ZuhörerInnen der MD und der HD auf einer Skala von 1 (positivster Wert) bis 6 (negativster Wert)	58
Abbildung 21: HD vs. MD im Rahmen des TC-Star-Projektes (Waibel 2011: 17:06).....	61
Abbildung 21: Beziehung 1 von Auto- und Heterobild.....	67
Abbildung 22: Beziehung 2 von Auto- und Heterobild.....	67
Abbildung 24: Abhängigkeit eines Image von eigener/fremder Information (Merten 2007:249)	68
Tabelle 1: Entwicklung sprachübersetzender und -dolmetschender Systeme	25
Tabelle 2: zukünftige Herausforderungen der MT- und MD-Forschung	43
Tabelle 3: Unique User/Monat (vgl. ÖWA 2017)	69
Tabelle 4: Gliederung nach Themen- und Imageart	71

0. Einleitung

Was vor 10 Jahren noch Science-Fiction war, scheint heute mittlerweile eine Selbstverständlichkeit zu sein. Technologische Fortschritte sind aufgrund ihrer rapiden Entwicklungen kaum noch mitzuverfolgen, sodass man hier und da, als LaiIn auf diesem Gebiet nun doch mit einer Neuigkeit überrascht werden kann. So geschah es wahrscheinlich, als der Skype Translator 2014 auf den Markt kam und zahlreiche Schlagzeilen machte. Nicht jeder hätte vermutlich damit gerechnet. Ähnlich hatte es sich Jahre zuvor (2009) mit Jibbigot verhalten. Allein diese zwei Beispiele sind als Beweis für den kontinuierlichen, unvermeidbaren Fortschritt im Bereich (Sprach)Technologie und in dem Falle auch Computerlinguistik zu sehen. Vorliegende Arbeit soll diese Thematik näher beleuchten und auch Aspekte berücksichtigen, die neue Fragen aufwerfen. Zweifelsohne setzt der sprachtechnologische Fortschritt voraus, dass sich Betroffene – in dem Fall HumandolmetscherInnen und –übersetzerInnen – auf die sich dadurch ergebenden Situationen umstellen können. Gewollt oder ungewollt wird sich der Bereich des maschinellen Dolmetschens, im Folgenden auch MD, weiterentwickeln und hier und da auch die eine oder andere Hürde bewältigen. Sollte sich der/die HumandolmetscherIn nun dadurch bedroht fühlen? Und schließlich die große Frage: Wird die Maschine, den Menschen ersetzen? Es ist zweifelsohne eine Frage, die manchmal mit einem „Vielleicht“, oft mit einem „Nein“ und meistens mit einem „Jein“ beantwortet werden kann. In vorliegender Arbeit wird versucht, das derzeitige Geschehen auf sprachtechnologischer Ebene näher zu bringen und etwaige gefundene Antworten zu kommentieren und argumentieren.

Die gesamte Arbeit ist als ein Zusammenspiel von Theorie und Praxis zu sehen, indem anhand konkreter Beispiele, theoretische Entwicklungen bestätigt oder erläutert werden sollen. Im Groben befasst sich die Arbeit mit zwei Hauptthematiken. Zum einen mit dem maschinellen Dolmetschen und zum anderen mit dem Berufsstand Dolmetschender. Beide Themen können hier und da miteinander verschmelzen.

Allen voran wird in Kapitel 1 das maschinelle Dolmetschen definiert, indem auch ein Bezug zu dessen geschichtlichen Entwicklung dargelegt wird. Dem folgt – in Form von Unterkapiteln – eine kurze Aufzählung der Einsatzbereiche des maschinellen Dolmetschens.

Der zweite Teil der Arbeit in Kapitel 2 befasst sich mit dem technischen Aspekt des maschinellen Dolmetschens. Dabei sollen unter anderem einige Meilensteine des MD vorgestellt werden. Kapitel 3 soll schließlich als eine Art Überleitung zum vierten Abschnitt dienen, in dem auf den Vergleich zwischen Mensch und Maschine eingegangen werden soll, nachdem davor auf den Qualitätsbegriff in Bezug auf die Bewertung von Dolmetschleistungen Bezug genommen wurde.

Die weiteren Kapitel stellen vielmehr den Berufsstand Dolmetschender (am Beispiel Österreich) in den Vordergrund. Dabei soll berücksichtigt werden, wie dieser Beruf von außen wahrgenommen wird sowie, auf welche Art und Weise professionelle DolmetscherInnen und ÜbersetzerInnen nach außen auftreten.

Abschließend können in Kapitel 5 alle im Vorhinein dargestellten Erkenntnisse in einer Zusammenfassung, miteinander verschmelzen und können auf Basis des zuvor Dargelegten, voraussichtliche Szenarien vorgestellt werden.

1. Maschinelles Dolmetschen (MD)

In einer Welt, in der die Vernetzung auf globaler Ebene zu einer unabdingbaren Realität geworden ist, muss die Kommunikation in verschiedenen Sprachen und somit auch für verschiedene Kulturen als selbstverständlich angesehen werden. Bis 2020 werden weltweit um die sechs Milliarden (Mobil)Telephone und rund eine Milliarde Personal Computer in Umlauf (vgl. Waibel 2016: 1:13) sein. Nicht zuletzt durch den Einzug von mittlerweile leistbaren Smartphones, konnte die digitale Barriere auch in ärmeren Ländern in großem Ausmaß reduziert werden. Somit sind nun auch große Teile ärmerer Bevölkerungen zu so genannten „Digirati“ (vgl. Carbonell 2016: 4:13), also „vernetzten“ Personen geworden. Während der Vernetzungsprozess eine mittlerweile größtenteils überwundene Hürde ist, steht nun die Herausforderung in Zusammenhang mit der sprachlichen Barriere umso mehr im Vordergrund. Obwohl weltweit rund 6.900 Sprachen¹ (Stand 2012)² verzeichnet sind, wird nur 1% dieser Sprachen seitens multilingualer Unternehmen wie z.B. Google, in Systemen maschinellen Übersetzens und/oder maschinellen Dolmetschens eingesetzt. Somit werden auch alle Informationen einzig in dieser Anzahl an Sprachen (1%) verfügbar gemacht. Wie

¹ Nur 1,2% davon (77 Sprachen) werden von über zehn Millionen Personen beherrscht (vgl. Carbonell 2016: 07:43) während über 50% der Sprachen (ca. 3.000) vom Aussterben bedroht sind

² Waibel (2016: 1:15) spricht von 6.000 Sprachen

können nun die sprachlichen Barrieren überwunden werden, ohne dabei auf das Englische als Lingua franca auszuweichen? Englisch stellt zwar in vielen Fällen eine Stütze dar, doch muss dem hinzugefügt werden, dass nur 30-40% all jener, die der Sprache europaweit mächtig sind, diese auch verhandlungssicher beherrschen (Waibel 2015a: 1:30). Von dieser Tatsache absehend würde die Lösung des Problems einzig mittels Einsatz des Englischen als Lingua franca eine Bedrohung für die Wahrung sprachlicher und kultureller Identitäten darstellen (vgl. Waibel 2015: 103). Das Ziel ist also eine multilinguale Vernetzung, durch den Einsatz möglichst vieler Sprachen. Neben der Herausforderung in Zusammenhang mit der Multilingualität, darf auch die Multimodalität von Sprache nicht unterschätzt werden und muss bei der Entwicklung innovativer Sprachtechnologien immer öfter berücksichtigt werden. Sprache kann sowohl auf verbaler Ebene (Sprechstil, Internet- oder SMS-Sprache bzw. verzerrte Sprache usw.), als auch auf nonverbaler Ebene (Gestik, Mimik, Bilder usw.) stattfinden. Praktische Beispiele für letztere Kategorie wären der *Road Sign Translator* aus dem Jahr 2001, ein System zur Übersetzung chinesischer Straßenschilder (vgl. Kapitel 2.2) oder die Übersetzung von Facebook-Postings über das Jibbiggo-System (vgl. Kapitel 2.2.2).

1.1. Begriffserklärung

Eine nähere Auseinandersetzung mit dem Thema des maschinellen Dolmetschens bedarf allen voran einer Definition beziehungsweise Präzisierung des Begriffes *Dolmetschen*. Praxisbezogen ist hier von einer Tätigkeit die Rede, deren erste Nachweise uns in das Jahr um 1340 v. Chr., also in das Alte Ägypten³ führen. Aus dieser Zeit lassen einerseits Hieroglyphen (vgl. Abb.1), als auch das sogenannte Haremhab-Relief (vgl. Abb. 2) darauf deuten, dass das Zurückgreifen auf sprachliche Mittler⁴ schon seit frühester Zeit notwendig, gar unabdingbar war. Die Hieroglyphen in Abbildung 1 sind überwiegend lautliche Schreibungen, mit dem Mann am Ende als Determinativ zur weiteren Einordnung. Der sitzende Mann deutet an, dass eine Personenbezeichnung vorliegt. Die Hand am Mund im unteren Teil der Abbildung 1 präzisiert, dass ein Bezug zu einer Handlung der geistigen Sphäre, so wie das Sprechen, Denken etc. besteht. Die Darstellung links vom sitzenden Mann stellt jeweils die Grundbedeutung „lösen“ dar, die im übertragenen Sinne mit „ein Problem lösen, ausdeuten, interpretieren“ übersetzt werden kann. Ein *nomen agentis*, das gerne auf unsere Darstellung schließen ließe, würde demnach „Interpret, Exeget“ o.Ä. bedeuten. Die Wörterbücher

³ die Zeit um 1340 v.Chr. fällt in das Zeitalter des Neuen Reiches bzw. unter die 18. Dynastie (vgl. Küffer 2017)

⁴ hier wird bewusst auf die weibliche Form verzichtet, da die Dolmetschtätigkeit in der Antike ausschließlich von Männern ausgeübt wurde

verzeichnen zwar das Wort „Interpret“ nicht, die Existenz dieses Begriffes kann dadurch jedoch keineswegs ausgeschlossen werden. Was als nicht anzuzweifelnde Tatsache gilt, ist, dass die altägyptischen „Interpreten“ nicht als „Dolmetscher“ oder „Übersetzer“ angesprochen werden können. Vielmehr verstand man die Dolmetschenden zu der Zeit als jene, die sich mit den „Unverständlich-Redenden“ abgaben (Gundacker 2016).

Auch in Abbildung 2 lässt sich so ein „Sich-Mit-Unverständlich-Redenden-Abgebender“ erkennen. Bei dieser Darstellung handelt es sich um das Grabmal des Stadthalters und späteren Pharaos Haremhab. Die Szene zeigt auf der rechten Seite eine Menschenmenge, bei der es sich um eine Delegation von Abgesandten aus Vorderasien, Libyen und Nubien handelt, die die Gnade des Pharaos erfleht. Die Figur im Mittelpunkt stellt einen Sprachmittler dar, der sich umdreht und dem Pharaos die Bitte weiterleitet.



Abbildung 1: Darstellung der Dolmetschtätigkeit aus dem Alten Ägypten



Abbildung 2: Haremhab Relief

Im Laufe der Geschichte gewann die Tätigkeit des Dolmetschens immer wieder an Bedeutung, bis man anfangs, sie im 20. Jahrhundert zu erforschen und sich folglich die Disziplin der Dolmetschwissenschaft⁵ etablierte. Noch vor dieser Einordnung als eigene Disziplin verfolgte man das Ziel, konkrete Begriffsbestimmungen festzulegen. Die wohl bekannteste Definition des Begriffes Dolmetschen lautet nach Kade (1968:35): „Dolmetschen ist die Übertragung eines einmalig (in der Regel mündlich) dargebotenen Textes der AS in einen nur bedingt kontrollierbaren und infolge Zeitmangels kaum korrigierbaren Text der ZS.“

Weiters ist zum Thema der Begriffsbestimmung in Bezug auf das MD, auf Kades Definition aufbauend, eine Präzisierung vorzunehmen, die es erlaubt, den technologischen beziehungsweise maschinellen Aspekt, der hier hinzukommt, miteinzubeziehen. In Anlehnung an die hauptsächlich auf Englisch verfasste Literatur auf computerwissenschaftlichem Gebiet, kann festgestellt werden, dass noch keine konkrete Definition des Begriffes „Maschinelles Dolmetschen“ vorhanden ist. Gleichzeitig könnte man aber annehmen, dass die englischsprachigen Benennungen dieses Phänomens bereits einen definitionsgebenden Charakter besitzen. So finden beispielsweise folgende Benennungen Anwendung: „Speech-to-Speech-Translation“ (*SST/S2ST*), „Automatic Simultaneous Translation“, „Spoken Language Translation“, „Simultaneous Speech Translation“ (vgl. Fünfer 2013:13).

1.2. Geschichte

Aufgrund der technischen Zusammensetzung von MD-Systemen, war deren Umsetzung allen voran von der Entwicklung zweier Komponenten abhängig: der maschinellen Übersetzung und der automatischen Spracherkennung (vgl. Kapitel 2.).

Als Urväter und Ausgangspunkt für die spätere Entwicklung maschineller Translationssysteme kann man neben Descartes und Leibniz auch Johann Joachim Becher und John Wilkins nennen. Jeder von ihnen strebte nach der Entwicklung einer Universalsprache. Bei Descartes' und Leibnizens *lingua rationalis* beziehungsweise *lingua universalis* handelte es sich um die Idee künstlich geschaffener Sprache, die auf dem Eindeutigkeitsideal basieren sollte. Höchstes Ziel sollte dabei die Rationalisierung des

⁵ seit dem Kongress in Wien (1992) unter dem Titel „Translation Studies – An Interdiscipline?“ kann man von einer Einordnung der Dolmetschwissenschaft als eigene Disziplin der Translationswissenschaft sprechen

Denkens sein. Jeder Vorstellung⁶ könne genau ein Zeichen zugeordnet werden, wodurch jedem Gegenstand der Realität eine exakte Bezeichnung entspräche. Folglich ergeben sich zwei Schlussfolgerungen: Zum einen, dass jede Aussage in eine mathematische Gleichung⁷ überführt, zum anderen, dass die Wahrheit dieser Aussagen durch Nachrechnen überprüft werden könne (vgl. Gardt 1999:135ff.). Dies hätte beispielsweise bei Meinungsunterschieden und Streitigkeiten – auch auf internationaler Ebene – zu einer Lösung führen können, da man durch Nachrechnen, zur Wahrheit hätte gelangen können. Dass eine praktische Umsetzung dieser Ansätze unmöglich gewesen wäre, gab bereits Descartes selbst zu (vgl. Mauthner 1923). Weiters lieferte auch Johann Joachim Becher mit seinem Werk *Character, pro notitia linguarum universali* (1661) einen weiteren Ausgangspunkt für die später weiterentwickelte Idee der Sprachtranskodierung und Sprachtranslation durch die Maschine, indem auch er den Versuch der Konstruktion einer Universalschrift unternahm. Als Basis dafür dienten ihm das Lateinische, Ziffernkombinationen und grammatische Symbole. Es handelte sich also um nichts anderes als um ein Latein-Wörterbuch mit numerischer Notation. Nennenswert ist in diesem Zusammenhang auch das von John Wilkins⁸ (1668) publizierte Werk *An Essay towards a Real Character and a Philosophical Language*. Auch Wilkins war bemüht, eine Kunstsprache anhand lullistischer⁹ und grammatischer Elemente zu schaffen (vgl. Schmidt-Biggemann 1983:193). Genauso wie alle anderen o.g. künstlichen Sprachsysteme, ist auch Wilkins' Versuch als ein Vorschlag zu einem interlingualen Ansatz, jedoch noch lange nicht als eine Form von maschineller Translation zu sehen.

Erst im Laufe des 20. Jahrhunderts ging man in der Geschichte des Maschinellen Dolmetschens beziehungsweise der maschinellen Translation (im Folgenden MT) einen Schritt weiter. So wurden 1933 zwei unabhängig voneinander entwickelte Systeme vorgestellt. Dabei handelte es sich zum einen um das sogenannte „cerveau mécanique“ bzw. „maschinelle Gehirn“, für das der Franzose Georges Artsruni ein Patent erhielt, zum anderen um das vom Russen Petr Petrovich Smirnov-Trojanskij entwickelte MT-System.

Artsrunis Maschine war im Stande, Äquivalente von Wörtern oder Wortkombinationen aus einer Ausgangs- in eine Zielsprache zu generieren. Ermöglicht wurde die Äquivalenzfindung anhand eines eingefügten Wörterbuches in der Zielsprache. Im

⁶ *idées simples* nach Descartes oder *termini primi* nach Leibniz (vgl. Gardt 1999:138)

⁷ z.B.: *animal* = 2, *rationalis* = 3 → *animal* in Verbindung mit *rationalis* ergibt *homo* → *homo* = 6

⁸ Mitbegründer der Royal Society

⁹ *Lullismus*: eine Bewegung, die ab dem 14.Jh. bis in die Gegenwart präsent war und auf Ramon Llull, einen mallorquinischen Philosophen und Logiker des 11/12.Jh. zurückzuführen ist. Dieser bezeichnete die Logik als Kunst und Wissenschaft, mit Hilfe des Verstandes, Wahrheit und Lüge zu unterscheiden, erstere anzunehmen und letztere von sich zu weisen. Auf Basis dessen zeichnete er eine „logische Maschine“, die auf einem mechanischen Kombinieren von Begriffen beruhte. (vgl. Universität de Barcelona 2017)

Großen und Ganzen stellte Artsrounis System, ein maschinelles multilinguales Wörterbuch dar. Das System war in der Lage, gespeicherte Informationen abzurufen und auszugeben. So ein System hätte man in Zukunft in verschiedenen Bereichen einsetzen können: bei der Erstellung von Fahr- bzw. Zeitplänen, Telefonbüchern, telegraphischer Kommunikation, Kryptographie usw.

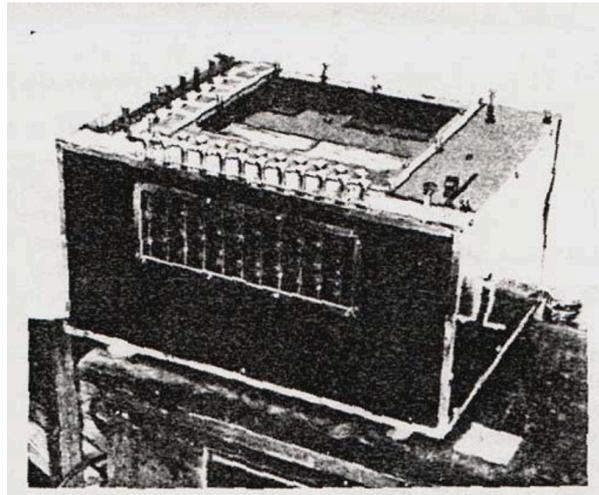


Abbildung 3: Artsrounis „maschinelles Gehirn“ (vgl. Hutchins 2004)

Im Falle von Trojanskis MT-System handelte es sich nicht nur um eine Maschine zum Zwecke von Translationsprozessen auf multilingualer Ebene, sondern enthielt seine Arbeit auch Vorschläge für eine Kodierung und das Erkennen grammatischer Elemente anhand universaler Symbole, was den gegenwärtigen technischen Eigenschaften und dem heutigen Verständnis eines MT-Systems schon etwas näher kommt. Die Entstehung der Parsing¹⁰-Komponente dürfte folglich auf Trojanski zurückzuführen sein. Der von ihm entwickelte Parser führte eine Analyse auf Wort- und Syntaxebene durch und konvertierte die erkannten Grundformen in einem späteren Schritt in Äquivalente der Zielsprache.

Obwohl Trojanski mit dieser Erfindung, seiner Zeit weit voraus war, blieb er außerhalb Russlands lange Zeit unbekannt.

Die Idee von der Entwicklung eines MT-Systems wurde erst Jahre später wieder aufgegriffen. Diesmal war es Warren Weaver, Vizepräsident der Rockefeller Foundation, der auf diesem Gebiet einen Meilenstein setzte als er 1949 das Weaver Memorandum verfasste.

¹⁰ ein Parser erzeugt eine Folge von Symbolen anhand von Grammatikstrukturen der AS; dies soll u.a. ermöglichen, (Syntax)Fehler zu erkennen; eignet sich v.a. für Dialogsysteme (vgl. Kapitel 2)

Dieser Schritt setzte primär in den USA viele Forschungsarbeiten in Gange, bis im Jahr 1954 das erste System für die maschinelle Übersetzung (MÜ) an der Georgetown University vorgestellt wurde.

Der Anstieg an Forschungsaktivität im Bereich der MT, der sich in den USA der 50er und 60er Jahre beobachten ließ, erfuhr 1966 seinen beinahe gänzlichen Stillstand. In diesem Jahr erschien der *ALPAC*-Bericht (*Automatic Language Processing Advisory Committee*). Dabei handelte es sich um ein von einer Gruppe von WissenschaftlerInnen verfasstes Dokument zur Bewertung des Forschungsstandes und der Sinnhaftigkeit weiterer Aktivität auf dem Gebiet der MÜ. Die Schlussfolgerung, dass MÜ-Systeme zu langsam und zu teuer seien, war Grund genug, um jegliche Förderprogramme und Forschungsvorhaben in den USA aufzugeben. Der wohl bekannteste MÜ-Entwickler, der es schaffte, sich trotz *ALPAC*-Berichtes durchzusetzen, war das Unternehmen SYSTRAN, dessen gleichnamiges MÜ-System 1968 entstand und später international eingesetzt wurde – u.a. von der Air Force sowie ab 1976 von der Kommission der Europäischen Gemeinschaften (später EU) und später u.a. auch von Google und Microsoft. Während des in den USA herrschenden, *ALPAC*-bedingten Forschungsstillstandes im Bereich der MÜ, ließ sich in anderen Ländern ein steigendes Interesse für MÜ-Projekte beobachten. Dies war u.a. in Europa, Kanada, und der Sowjetunion der Fall. In Frankreich beispielsweise entstand im Jahre 1970 das TITUS-Projekt am ITF (Institut Textile de France). Das Akronym des Projektes bezog sich sowohl auf das entwickelte MT-System, als auch auf die zu Austauschzwecken gegründete Netzwerkplattform und steht für „Textil-Informationstechnik und –service“ (Zingel 1978). Im selben Jahr entstand weiters SUSY, auch „Saarbrücker Übersetzungssystem“ genannt. Nennenswert sind auch andere Projekte wie TAUM (Traduction Automatique de l'Université de Montréal) in Kanada, GETA-Ariane in Frankreich oder Mu in Japan, um nur einige zu nennen (vgl. Hutchins 2015:123).

Erst mit der Gründung des *Centre for Machine Translation* (CMT) an der Carnegie Mellon University, dank der die wissenschaftliche Aktivität im MÜ-Bereich wieder anstieg, wurden auch in den USA wieder und vermehrt in Japan Forschungsarbeiten initiiert, bis gegen Ende der 80er Jahre mehrere Systeme entwickelt waren und erste Systeme für das maschinelle Dolmetschen entstanden. Der damalige Entwicklungsstand war jedoch unzureichend und hätte in der Praxis nur in den wenigsten Anwendungsgebieten seine Funktion erfüllt. Folglich forschte man im Bewusstsein, dass der Einsatz vom MD nur in bestimmten Anwendungsbereichen erfolgt wäre. Dieses Ziel wurde u.a. vom in den 90er Jahren gegründeten C-STAR (*Consortium for Speech Translation Advanced Research*)

verfolgt. So entstanden vorerst Dialogsysteme wie C-Star, VERBMOBIL, NESPOLE, BABYLON oder TRANSTAC (vgl. Fünfer 2013:20). Erst später erreichte man einen Wendepunkt – wie beispielsweise 2004, als von der EU-Kommission Projekte wie TC-STAR ins Leben gerufen wurden oder 2005 mit der Entstehung vom LECTURE TRANSLATOR. Ab diesem Zeitpunkt sollte die maschinelle Translation um zahlreiche Anwendungsgebiete erweitert werden.

Grundsätzlich lassen sich in der Geschichte des MD-Systems in Bezug auf die Komponente des Maschinellen Übersetzens Entwicklungsphasen in vier Bereichen erkennen:

1. Erste Computer-Phase
2. Zweite Computer-Phase
3. Computerlinguistische Phase
4. Phase der künstlichen Intelligenz

Phase 1 der maschinellen Übersetzung fällt nach Hutchins (vgl. Schwanke 1991:65) auf die Jahre 1946-1954 und lässt sich mit der Auseinandersetzung auf ingenieurwissenschaftlicher Ebene zusammenfassen. Es ist also von einer experimentellen Phase die Rede, in der man mit dem Gedanken der Entwicklung eines Systems zur „zeitsparenden vollautomatischen Übersetzung in hoher Qualität“ (Schwanke 1991:65) spielte. Infolgedessen entstanden MT-Systeme der ersten Generation, die mit dem Ansatz der direkten Übersetzung arbeiteten. Eine Durchführung der empirischen Studien auf diesem Gebiet war jedoch, nicht zuletzt aufgrund der komplexen Programmierung, nur ansatzweise und lokal begrenzt möglich. Die Systeme waren anfangs nicht im Stande, eine Trennung zwischen den sprachlichen Daten i.S. von Grammatik oder Vokabular und den Verarbeitungsalgorithmen vorzunehmen.

Erst in der nächsten Entwicklungsphase ging man einen Schritt weiter. Ziel war es, das direkte MT-System (vgl. Kapitel 2.1.2) möglichst zeitnah zu optimieren. Diese Periode begann 1954 mit der Vorstellung des o.g. ersten MT-Systems an der Georgetown University und endete mit dem ALPAC-Bericht von 1966.

Die dritte Entwicklungsphase lässt sich in den Jahren zwischen 1966 und 1975 beobachten. Man fing nun an, sich mehr auf den linguistischen Aspekt zu konzentrieren und Komponenten wie Parser zu erforschen. Diese Umorientierung mit Schwerpunkt auf

semantischen und syntaktischen Formen dürfte als Folge des ALPAC-Berichtes zu sehen sein, da man nun gezwungen war, von der Idee einer qualitativ hochwertigen vollautomatisierten Übersetzung abzusehen. Dies führte zur Entwicklung von Transfer- und Interlingua-Übersetzungssystemen und von MT-Systemen der zweiten und dritten Generation (vgl. Kapitel 2.1.2).

Die historisch vierte Phase ist bis heute noch aktuell. Sie ist besonders stark durch die Erforschung der künstlichen Intelligenz gekennzeichnet und resultiert in der Praxis einerseits in Hybridformen der o.g. MT-Modelle (direkte Übersetzung, Transfer- oder Interlingua-Übersetzung) im MD, andererseits in einem Zusammenspiel mehrerer Disziplinen wie Informatik, Computerlinguistik und Translationswissenschaft (vgl. Schwanke 1991:66).

1.3. Anwendungsbereiche

Je nach Anwendungsbereich eines MD-Systems, weist dieses bestimmte Eigenschaften in Bezug auf die technische Zusammensetzung auf. Wie bereits erwähnt und auch in folgenden Kapiteln im Detail beschrieben, gibt es beispielsweise verschiedene Übersetzungsansätze, die in die MT-Komponente einer MD-Maschine implementiert werden können. Dabei sind Hybridformen nicht auszuschließen.

Egal, ob als reines MD-System oder als Dialogsystem: der Anwendungsbereich von Speech-to-Speech-Translation-Maschinen ist mittlerweile immer breiter gefächert. Im Folgenden soll dies in einem allgemeinen Überblick beleuchtet werden.

1.3.1. Multilinguale Dialogsysteme

Das wohl bekannteste Projekt, das einen Meilenstein in der MD-Entwicklungsgeschichte setzte, war VERBMOBIL (vgl. Kapitel 2.2.1), das als (Telefon-)Dialogsystem konzipiert wurde. Themenschwerpunkt waren dabei face-to-face-Buchungen, -Reservierungen, -Terminvereinbarungen etc. Dieses Projekt der 90er Jahre resultierte zu der Zeit in einem der innovativsten MD-Systeme weltweit, zumal es verschiedene Komponenten wie jene zur Dialogklärung (Ziliani 2005), zur Nebenläufigkeit oder zur Generierung komplexerer Dialoge vereinte (vgl. Bub et al. 1997:74).

Bi- oder multilinguale Dialogsysteme finden heute vermehrt ihren Einsatz, wobei sie im Vergleich zum VERBMOBIL-System mittlerweile technologisch fortgeschrittener sind. So beispielsweise die Yochina-App, ein mobiler Reiseratgeber (vgl. Xu et al. 2014:51-57).

Besonders charakteristisch bei Dialogmaschinen ist, dass sie meistens bis immer nur domänenorientiert programmiert werden können.

1.3.2. Telekommunikation

Ein Bereich, der inzwischen als eine Selbstverständlichkeit ins Alltagsleben einfließt, ist jener der (Tele)Kommunikation über Telefon- oder Videokonferenzen. So ist das MD auf diesem Gebiet zu etwas durchaus Praktischem, wenn nicht gar Unabdingbarem geworden. Im Jahr 2014 stellte Microsoft den Skype Translator vor, der Chat-Gespräche und Videokonferenzen simultan übersetzt und dolmetscht. Zusätzlich erscheint das Gespräch in Form von Untertiteln. Als weiteres ähnliches MD-Produkt lässt sich iTranslate nennen, das als App für mobile Geräte erhältlich ist und als eine Art persönliche/r DolmetscherIn fungieren soll. Nach mündlicher Spracheingabe wird der Input in die gewünschte Sprache übertragen und als mündlicher Output wiedergegeben.

1.3.3. Militäreinsätze

Aufgrund der weltweit zahlreichen militärischen Einsätze, ist der Bedarf an maschinellen Dolmetschsystemen für den Einsatz zu militärischen Zwecken im Laufe der Jahre angestiegen. Dies dürfte gemäß Weinstein (2002:11) nicht zuletzt auf den Mangel¹¹ an HumandolmetscherInnen für solche Einsätze zurückzuführen sein. In der Praxis haben sich für diese Domäne programmierte MD-Systeme aufgrund begrenzter technologischer Kapazitäten¹² noch nicht mit Erfolg durchgesetzt. Die bisherigen Entwicklungen auf diesem Gebiet liefern jedoch gute Voraussetzungen für einen zeitnahen zukünftigen Fortschritt. Bereits 1994 stellte das MIT-Lincoln Laboratory ein potentiell MD-System für die Nutzung bei militärischen Einsätzen vor. Der Entwickler stand aber vor einer Herausforderung, zumal eine praktische Realisierung solch einer Maschine nur in Zusammenarbeit mit Nutzern zu Stande kommen konnte. Letztere sollten helfen, die Domäne durch Vorgabe von Vokabular und Grammatik näher zu definieren, um das System im Nachhinein mit diesen Daten trainieren zu können. Dies in die Praxis umzusetzen erwies sich als schwieriger als gedacht.

Trotzdem entstand 2001 das DARPA-Projekt. Ziel war es, ein Übersetzungs- bzw. Dolmetschsystem für das Militär zu entwickeln, das schnell und domänenorientiert gearbeitet

¹¹ hier: in Bezug auf die USA

¹² u.a. Inkompatibilität mit militärischer Software

hätte. Durch den implementierten *one-way*-Ansatz (Weinstein 2002:12) sollte eine umso adäquatere und genauere Übertragung garantiert werden.

Konkrete Szenarien, in denen MD-Systeme eingesetzt werden sollen, umfassen u.a. militärische Operationen (Kommunikation mit Alliierten, Kommunikation mit der lokalen Bevölkerung, Befragungen, logistische sowie geographische Informationen) sowie humanitäre Hilfe (Katastrophenhilfe, ärztliche Diagnose und Behandlung).

1.3.4. Akademische Zwecke

Dieser Anwendungsbereich ergab sich aufgrund zunehmend international besetzter Studentenschaften. Der 2005 erstmals vorgestellte, am KIT (Karlsruher Institut für Technologie) entwickelte *Lecture Translator* (LT) zählt als erstes, simultanes maschinelles Dolmetschsystem für Vorträge (vgl. Kapitel 2.2.4). Trotz der Tatsache, dass man in international besetztem Umfeld in der heutigen Lingua franca, also Englisch zu kommunizieren pflegt, muss in einigen Fällen von dieser Lösung abgesehen werden. Dies ist beispielsweise im Rahmen universitärer Vorlesungen der Fall. Studenten, die der lokalen Sprache nicht mächtig sind, soll nun aber die Möglichkeit geboten werden, die Vorträge in ihrer Muttersprache oder einer anderen, ihnen verständlichen Sprache mitzuverfolgen. Dies wird dank einem MD-System wie dem *Lecture Translator* ermöglicht.

2. Das MD-System

2.1. Technik – Wie funktioniert ein MD-System?

Eine Dolmetschmaschine muss grundsätzlich drei Haupt-Teilaufgaben lösen. Die ersten zwei, für die maschinelle Übersetzung notwendigen Schritte, werden beim MD um den zusätzlichen Aspekt der Sprachsynthese ergänzt, wodurch sich schließlich folgender Aufbau ergibt:

1. ASR (Automatic Speech Recognition): Komponente für die automatische Spracherkennung
2. MÜ (Maschinelle Übersetzung): Übersetzungskomponente

3. Sprachsynthese

Wie anhand Abbildung 4 ersichtlich, verwandelt die ASR-Komponente über ein Mikrofon gesendete gesprochene Signale in Sprache (L_a). In einem zweiten Schritt erfolgt die Ausgabe der inzwischen verarbeiteten Signale als geschriebener Text (*Speech-to-Text*).

Im Rahmen der maschinellen Übersetzung wird der Text der Ausgangssprache (L_a) in die Zielsprache (L_b) übersetzt (*Text-to-Text*). Die letzte Komponente umfasst schließlich den Prozess der mündlichen Textausgabe (*Text-to-Speech*) in der Zielsprache (L_b).

Da ein maschinelles Dolmetschsystem in jeweils zwei Sprachrichtungen arbeiten muss, das heißt, sowohl von L_a nach L_b , als auch umgekehrt, verfügt dieses über analoge Subsysteme für die jeweils andere Sprache. Somit bestehen MD-Systeme aus jeweils insgesamt sechs Komponenten, die zwar grundsätzlich sprachunabhängig sind, doch gleichzeitig die Integrierung sprachabhängiger Modelle erfordern. Dabei darf nicht außer Acht gelassen werden, dass jede Sprache, eigene Charakteristika aufweist, die bei der Programmierung sprachabhängiger Modelle berücksichtigt werden müssen. Bei der Entwicklung solcher Systeme gewinnt zumeist auch der menschliche Aspekt in Bezug auf die Maschinenleistung (je nach Einsatzgebiet) an Bedeutung (vgl. Waibel 2011: 6:10).

Ein MD-System kann je nach Anwendungsbereich weiters noch um andere Aspekte ergänzt werden. Darunter fiele beispielsweise ein Teil für die Aufarbeitung von Informationen, mit dem Ziel diese wiederum bestmöglich weiterzuverarbeiten.

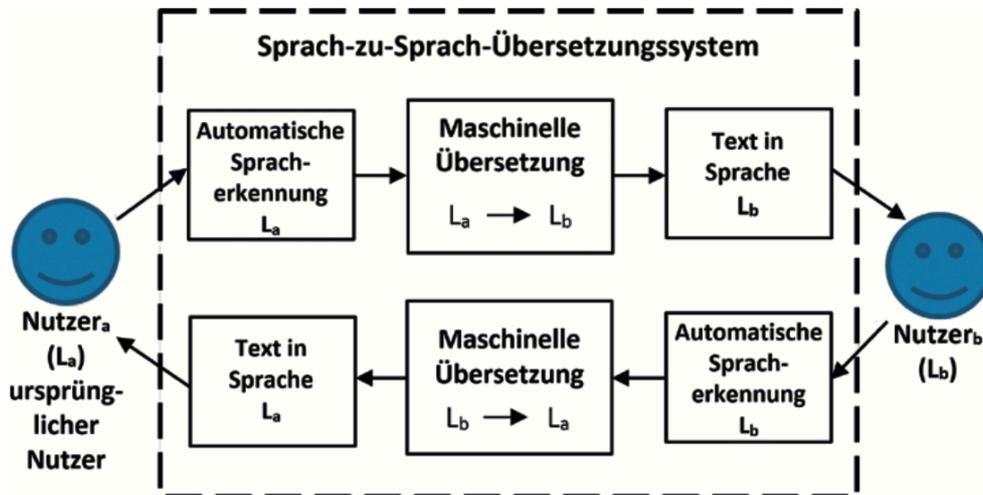


Abbildung 4: Sprach-zu-Sprach-Übersetzungssystem (Waibel 2015b:104)

Da alle drei Haupt-Teilaufgaben als eigene Forschungsgebiete einzuordnen sind (vgl. Waibel 2015b:105), sollen diese im Folgenden näher beleuchtet werden.

2.1.1. Automatische Spracherkennung (ASR/Speech-to-Text)

Die Leistung von MD-Systemen hängt sehr stark von der Komponente für die automatische Spracherkennung ab, zumal diese, wie oben bereits erwähnt, als erste Anlaufstelle für die eingehenden akustischen Signale zu sehen ist. Besonders wichtig für die ASR, sind das Mikrofon, die Laufzeit und die Latenzzeit (vgl. Müller et al. 2016).

Die Hauptaufgabe der ASR-Komponente besteht darin, die einzelnen Wörter als solche zu erkennen und sie in weiterer Folge in digitale Signale umzuwandeln. Ein Problem, das sich daraus ergibt, sind mögliche Störfaktoren oder Mängel, die das Erkennen der eingegebenen Informationen beträchtlich erschweren oder gar unmöglich machen können. Dabei kann es sich um Störgeräusche, Pausen, Disfluenzen, Umfang des Vokabulars, nonverbale Elemente, Dialekte etc. handeln. Auch eine zu hohe Perplexität¹³ kann die Qualität der maschinellen Dolmetschung um einiges verschlechtern. Im Gegensatz zu MÜ-Systemen muss man sich beim maschinellen Dolmetschen auf die sogenannten *language model statistics*, d.h. auf Sprachmodellen basierende statistische Werte stützen. Auf Basis der erhaltenen Informationen, bildet die ASR-Komponente Worthypothesengraphen, errechnet einen

¹³ unter *Perplexität* versteht man die durchschnittliche Anzahl möglicher Fortsetzungen

Wahrscheinlichkeitswert und leitet diese Daten an die MÜ-Komponente weiter. Dabei sollte die Fehlerrate (*word error rate* oder *WER*) nicht mehr als 10% betragen.

Eine besondere Herausforderung stellt bei der ASR-Komponente die Ambiguität von Sprache dar. Diese Mehrdeutigkeit kann sowohl auf akustischer, als auch auf semantischer Ebene stattfinden. Ersteres Phänomen betrifft beispielsweise dieselbe akustische Sequenz von Lauten in zwei verschiedenen Wörtern oder Aussagen. So kann der englische Satz „Give me a new display“ vom ASR-System fälschlicherweise als „Give me a nudist play“ (Waibel 2016: 2:13) erkannt werden. Genauso verhält es sich mit der akustischen Ambiguität der Begriffe „euthanasia“ vs. „Youth in Asia“ oder „This machine can recognize speech“ vs. „This machine can wreck a nice beach.“ (Waibel 2015:105).

Weiters stellen auch Zweideutigkeiten auf semantischer Ebene oft eine Herausforderung dar. So wäre es bei Sätzen wie „If the baby does not like the milk, boil it.“ nicht unwahrscheinlich, dass der Bezug unerkannt bliebe und zur Erhöhung der WER-Rate führe (Waibel 2015a: 3:17). Um dieser Problematik entgegenzuwirken, werden zusätzliche Komponenten (vgl. Abb. 5) in das Spracherkennungssystem eingebaut: ein Aussprachewörterbuch, das jedem Wort eine Aussprache zuordnet, ein akustisches und ein Sprachmodell. In beiden letzteren Fällen beruht der Prozess jeweils auf einer Wahrscheinlichkeitsberechnung in Bezug auf Laute einerseits sowie auf Wortsequenzen (W_1, W_2, \dots) in einem Satz andererseits.

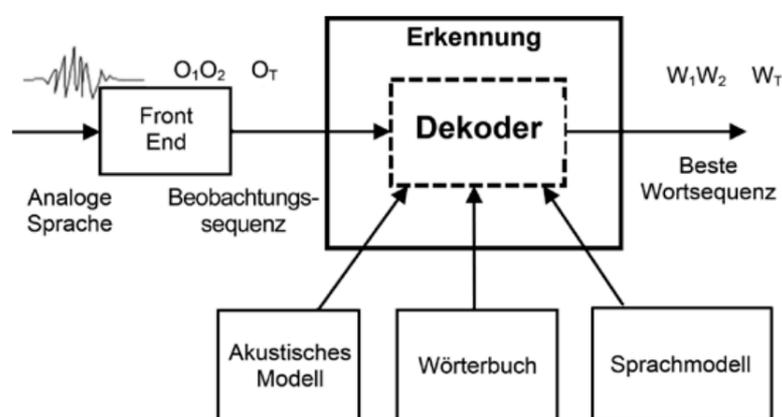


Abbildung 5: Ein typisches Spracherkennungssystem (Waibel 2015b:105)

Die im ASR-System eingebauten Modelle werden anhand von Such- und Optimierungsparametern entwickelt, die wiederum anhand von Lernalgorithmen funktionieren. Zu diesem Zwecke kommen entweder statistische Methoden oder, wie im Falle modernerer Modelle, neuronale Netze zum Einsatz. Dabei erlernt das System anhand bekannter Beispieldaten die jeweils beste Lösung in Bezug auf die Zuordnung von Symbolen zu Signalen. Wichtig ist in diesem Zusammenhang vor allem die Berücksichtigung der Kontextabhängigkeit von Wörtern und Phonemen. Je kontextabhängiger das System, desto niedriger die WER-Rate in Bezug auf oben bereits angeführte mögliche akustische oder semantische Fehler. Folglich kann dank der erhöhten Lernfähigkeit zwar eine Optimierung der ASR-Komponente beobachtet werden, doch kann man gleichzeitig noch keineswegs von einer 0-prozentigen WER-Rate ausgehen. Abgesehen von den im ASR-System eingebetteten Modellen, ist auch die externe Komponente des Signal-Inputs für die Signalaufnahme von großer Bedeutung, zumal auch diese domänenabhängig ist. So wäre es beispielsweise störend, in einem Besprechungs-Setting, ein Nahbesprechungsmikrofon einzusetzen. Bei Vorträgen hingegen macht dieses durchaus Sinn. Folglich wird die WER-Rate je nach Einsatz der externen Input-Komponente, positiv oder negativ beeinflusst und somit zu Unterschieden zwischen 10% und 40% führen (vgl. Waibel 2011: 33:04).

2.1.2. Maschinelle Übersetzung (MÜ/*Text-to-Text*)

Da MÜ-Systeme ursprünglich für geschriebene Sprache konzipiert sind, können sie beim Verarbeiten von Spontansprache Probleme aufweisen. So unterscheidet man grundsätzlich zwischen drei Ansätzen, die über die Vorgehensweise der MÜ-Komponente bestimmen:

1. Direkte Übersetzung
2. Transferansatz (*transfer approach*)
3. Interlingua-Ansatz (*interlingua approach*)

Alle drei Ansätze sind als regelbasierte Übersetzungsmethoden (*Rule Based Machine Translation*) zu sehen, die den erkannten Ausgangstext auf Basis von Regeln in die Zielsprache übertragen. Die verschiedenen Ansätze weisen grundsätzlich Unterschiede in folgenden Bereichen auf (vgl. Werthmann & Witt 2014:87):

- Anzahl der Sprachpaare
- Art der Informationsgewinnung und Analyse der ausgangssprachlichen Eingabe
- Art der abstrakten Repräsentationen für die Übersetzung.

Der Einsatz von Übersetzungsansätzen hängt von den jeweiligen Anwendungsbereichen und Eigenschaften der MÜ- oder MD-Systeme ab. Die Maschinen können bilingual oder multilingual arbeiten. In letzterem Fall kann der Prozess uni- oder bidirektional erfolgen. Auch die Benutzerinteraktion während der Übersetzung lässt sich zu den soeben genannten Faktoren hinzufügen. Interaktive Systeme können Benutzern die Möglichkeit bieten, der Maschine Feedback zu geben – z.B: in Form von Fragen und Antworten. Dies soll den Datenbestand der Maschine in Bezug auf deren Weltwissen optimieren. Je nach gewählter Übersetzungsstrategie, ist auch der Analyseaufwand größer oder geringer und wird man bei jedem Ansatz auf Vor- und Nachteile stoßen (vgl. Abb. 6).

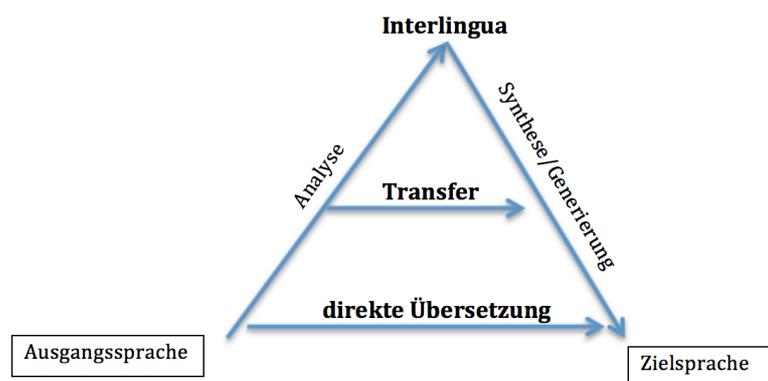


Abbildung 6: MÜ-Dreieck nach Vauquois (vgl. Carstensen et al. 2001:516)

Die direkte Übersetzung wurde im Laufe der MÜ- bzw. MD-Geschichte als erste Übersetzungsstrategie angewendet. Dabei handelt es sich, wie bereits festgestellt, um die Übersetzung auf Grundlage bestimmter Regeln. Um dies zu ermöglichen, muss eine sehr große Datenmenge vorhanden sein. Nicht zuletzt aus diesem Grund ist diese Methode vor allen für Systeme gedacht, die thematisch einzugrenzen sind (*domain limited systems*). Der

Nachteil einer solchen Übersetzungsvorgehensweise ist, dass die MÜ-Komponente jeweils für jedes Sprachenpaar trainiert werden muss. Dies erfolgt mit Hilfe von Paralleltextrn, indem das System, ähnlich wie die ASR-Komponente, auf statistischer Häufigkeit von Ausdrücken samt deren Übersetzung basierend lernt (vgl. Abb. 7). Dabei handelt es sich grundsätzlich um eine Wort-für-Wort-Übersetzung, bei der auf semantische oder morphologische Analysen verzichtet wird. Das Endergebnis ist von der Struktur her mit einem Wörterbuch vergleichbar. Somit wird die Ausgangstextanalyse sehr oberflächlich durchgeführt und erfolgt die darauf aufbauende Übersetzung in die Zielsprache ohne weitere Zwischenschritte, sondern nur durch Einsatz des bilingualen Wörterbuches. Infolgedessen divergiert der Zieltext der direkten Übersetzung nicht allzu sehr vom Ausgangstext.

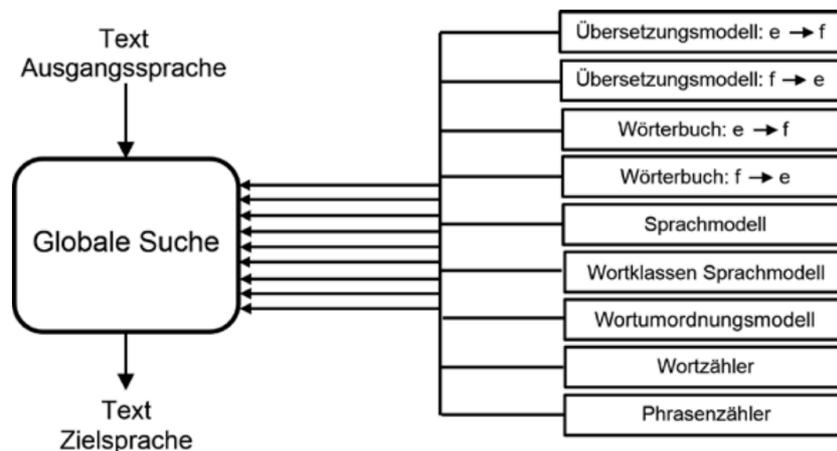


Abbildung 7: Maschinelle Übersetzung (Text-to-Text)

Als vorteilhaft erweist sich bei dieser Methode der geringe Zeitaufwand sowie die Robustheit. Dies geht jedoch vor allem dann auf Kosten der Qualität, wenn Ausgangs- und Zielsprache auf lexikalischer und struktureller Ebene große Unterschiede aufweisen. Um Verbesserungen auf Qualitätsebene zu erzielen, bietet sich die Implementierung eines Moduls für die morphologische Analyse an (vgl. Abb.8), das u.a. erlaubt, die Menge der Einträge zu reduzieren.

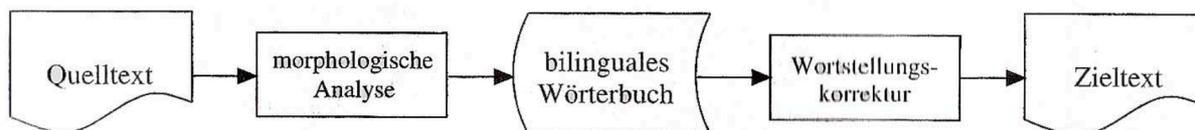


Abbildung 8: Direkte Übersetzung (Carstensen et al. 2001:518)

Nicht zuletzt, um den mit dem direkten Ansatz einherkommenden Limitationen entgegenzuwirken, wurden mit dem Transfer- und Interlingua-Ansatz, Systeme der zweiten Generation entwickelt, dank denen der Schwerpunkt auf syntaktische und semantische Analysen gesetzt wurde.

Der Transferansatz (siehe Abb. 9) sieht vor der Generierung des zielsprachlichen Endergebnisses zwei weitere Arbeitsschritte vor. Hierbei wird der Ausgangstext in einem ersten Schritt semantisch analysiert. Das bedeutet, dass dieser indirekte Übersetzungsansatz nicht auf Wort-, sondern auf Satzebene arbeitet. Die zweite Verarbeitungsphase umfasst eine zielsprachliche Repräsentation anhand von Transferregeln. Dabei sind letztere als Übersetzungsvorschläge zu sehen, die kontextabhängig generiert werden. Diese Berücksichtigung der Kontextabhängigkeit ist schließlich jener Aspekt, der sich positiv auf die Qualität des Endergebnisses auswirkt (vgl. Carstensen et al. 2001:518).

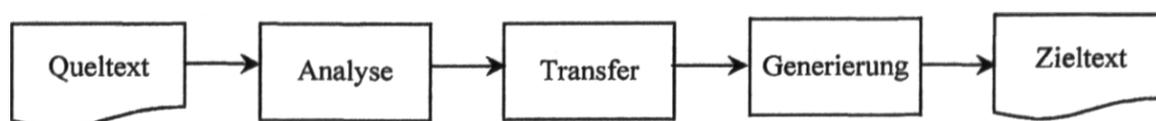


Abbildung 9: Transferansatz (Carstensen et al. 2001:518)

Ähnliches passiert beim Interlingua-Ansatz, bei dem aber zusätzliche, sprachenunabhängige semantische Zwischenrepräsentationen ermittelt werden (vgl. Abb. 10). Hierfür kommen Parser¹⁴ zum Einsatz, die jedoch manuell entwickelt werden müssen. Der Interlingua-Ansatz eignet sich besonders gut für multilinguale Systeme und erfordert einzig Analyse- und Generierungskomponenten für jede einzelne Sprache. Auf bilinguale Komponenten wie zum Beispiel Wörterbücher wird dabei verzichtet. Der dadurch entstehende Aufwand ist umso

¹⁴ ein Parser erzeugt anhand von Grammatikstrukturen der AS, eine Folge von Symbolen – was u.a. ermöglichen soll, (Syntax)Fehler zu erkennen; eignet sich v.a. für Dialogsysteme

größer, da aufgrund der Sprachenunabhängigkeit der Komponente, ein möglichst vollkommenes Sprachverstehen vorausgesetzt wird, was auf einzig maschineller Ebene nicht realisierbar ist (vgl. Carstensen et al. 2001:519). Diese Methode ist im Gegensatz zur direkten Übersetzung für den Einsatz in mehreren Anwendungsgebieten gedacht (*domain unlimited systems*).

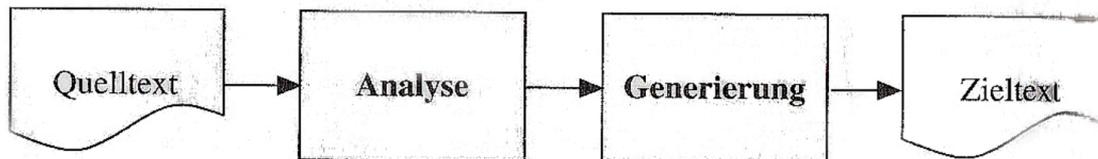


Abbildung 10: Interlingua-Ansatz (Carstensen et al. 2001:518)

Was beim Interlingua-Ansatz nachteilig ausfällt, sind die mit der Ausgangstextanalyse in Verbindung stehenden Herausforderungen in Zusammenhang mit der Erkennung und Auflösung von sprachlichen Mehrdeutigkeiten. Gleichzeitig können die sich ergebenden Repräsentationen, für unterschiedliche Zielsprachen genutzt werden, wodurch der Fokus von Interlingua-Systemen auf multilinguale Lösungen umso nachvollziehbarer erscheint.

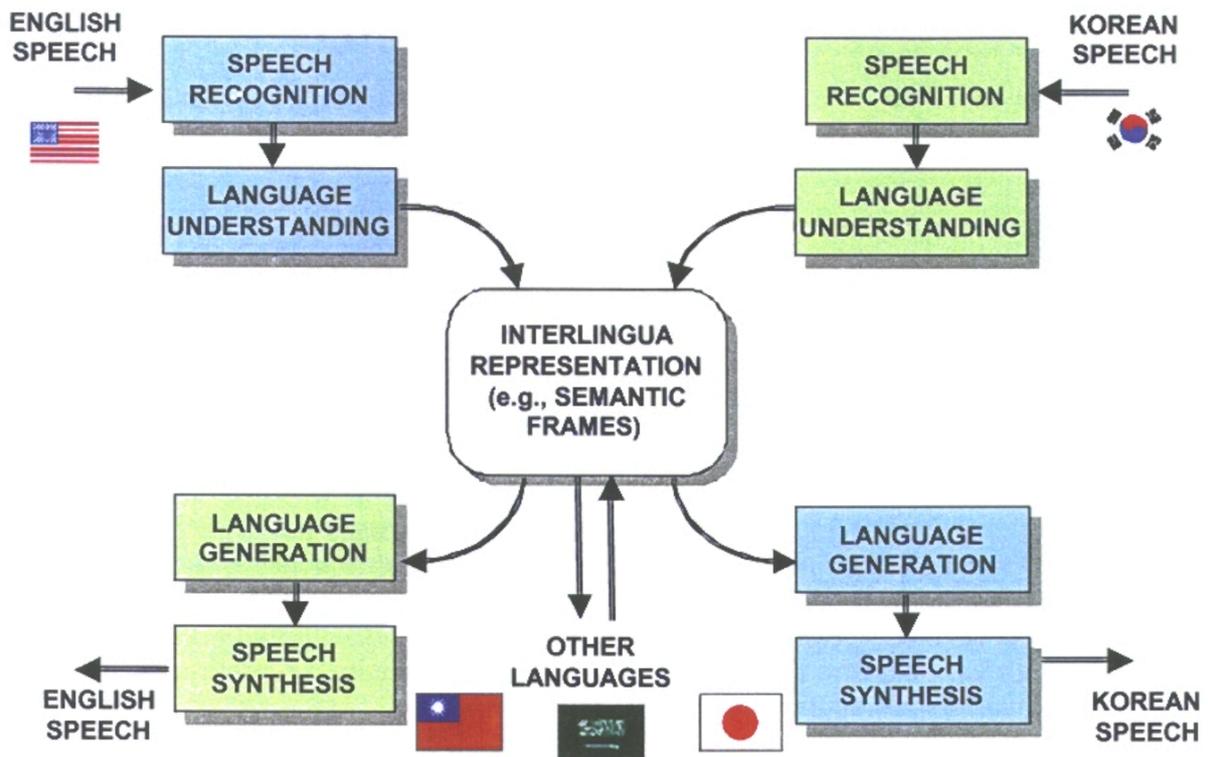


Abbildung 11: Beispiel für ein multilinguales Interlingua-System (vgl. Weinstein 2002:4) ¹⁵

Abgesehen von den drei genannten Übersetzungsansätzen, können auch Hybridformen dieser Strategien in ein MÜ-System eingebaut werden (vgl. Fünfer 2013:16).

Neben der regelbasierten Übersetzungsmethode kommt auch die korpusbasierte, statistische maschinelle Translation zum Einsatz. Dabei werden die Ergebnisse anhand empirischer Auswertungsmethoden erzielt, die bei alignierten ¹⁶ oder Paralleltexten angewendet werden. Dies passiert auf Wort-, Phrasen- und Satzebene (vgl. Werthmann & Witt 2014:91). Auch bei der korpusbasierten Übersetzung unterscheidet man zwischen verschiedenen Ansätzen:

1. Reinstatistischer Ansatz (*Statistical Machine Translation, SMT*): bei dieser Methode kommen weder Wörterbücher, noch Sprachmodelle zum Einsatz. Reinstatistische Systeme funktionieren ausschließlich durch statistische Wahrscheinlichkeitsberechnungen anhand bilingualer Textkorpora.

¹⁵ dabei handelt es sich um ein im Lincoln Laboratory entwickeltes System für die Sprachen Englisch und Koreanisch

¹⁶ dabei handelt es sich um mehrsprachige Korpora mit gleichen Inhalten

2. Beispielbasierter Ansatz (*Example Based Maschine Translation, EBMT*): dieser Ansatz verwendet bereits vorhandene, ähnliche oder identische Segmentübersetzungen als Ausgangspunkt für die Weiterverarbeitung. Grundidee hinter dieser Methode, ist die Möglichkeit der Wiederverwendung abgespeicherter Übersetzungen.
3. Kontextbasierter Ansatz (*Context Based Maschine Translation, CBMT*): hier werden hauptsächlich einsprachige Zieltextkorpora eingesetzt. Weiters funktioniert das System mit Hilfe eines zweisprachigen Vollformenlexikons. Vorteilhaft ist bei diesem Ansatz, dass dank der Kontextabhängigkeit viele sprachliche Mehrdeutigkeiten beseitigt werden können.

Auf alle Übersetzungssysteme bezogen kann gesagt werden, dass deren Leistung sehr stark vom jeweiligen Leistungspotential der einzelnen Komponenten abhängt. Deren Performanz ist wiederum mit weiteren Faktoren wie Sprache, Komplexität der einzelnen morphologischen und lexikalischen Prozesse, jeweilige akustische Bedingungen und das beim Benutzer vorhandene Wissen, in Zusammenhang zu setzen (vgl. Weinstein 2002:4).

Die größte Herausforderung beim maschinellen Übersetzungsprozess stellt, ähnlich wie bei der Spracherkennungskomponente, die Ambiguität der Sprache dar, zumal es schwer ist, dem System das notwendige entsprechende Kontextwissen beizubringen. Dies resultiert in strukturellen Fehlinterpretationen wie im oben genannten Beispiel „If the baby doesn't like the milk, boil it.“ (vgl. Kapitel 2.1.1) oder in semantischen Fehldarstellungen (z.B.: „Zentraleuropa“ wird als „Headquarter Great Grandpa“ übersetzt). Verbesserungen auf diesem Forschungsgebiet wurden durch eine optimale Kombination mehrerer erlernter statistischer Wissenskomponenten wie in Abbildung 7 dargestellt, erreicht (vgl. Waibel 2015:106). Dabei werden im Rahmen einer globalen Suche, Übersetzungs- und Sprachmodelle angewendet. Das Übersetzungsmodell arbeitet nach demselben Prinzip wie die in die Spracherkennungs- und Sprachsynthesekomponenten eingebauten Lernalgorithmen. Dabei kommt die sogenannte Bayes-Formel zur Wahrscheinlichkeitsberechnung zum Einsatz. Anhand der Bayes-Regel findet der statistische Übersetzungsansatz (vgl. Waibel 2011: 7:20) Anwendung. Dabei wählt das Übersetzungsmodell die bestmögliche und statistisch korrekteste Übersetzung aus einem bereits vorhandenen Wörterbuch aus. Die sich daraus ergebende Übersetzung wird an das Sprachmodell weitergeleitet, das wiederum versucht,

anhand weiterer automatischer Algorithmen, die logischste und angemessenste Auswahl zu treffen.

2.1.3. Sprachsynthese (*Text-to-Speech/TTS*)

Die dritte Komponente, die das MÜ- in ein MD-System “umwandelt“, ist jene der Sprachsynthese. In manchen Fällen gibt es die Möglichkeit, die akustische Wiedergabe auch anhand von Untertiteln mitzuverfolgen. Alternativ werden die Untertitel in anderen Maschinen durch eine schriftliche Zusammenfassung ersetzt.

Eine große Hürde bei der Sprachsynthese stellt die Angleichung vom Ausgangstext mit dem Zieltext unter Berücksichtigung von Stimme und Sprechstil dar. Hierzu wird anhand von Versuchen, das Ziel verfolgt, mittels *voice conversion* zu arbeiten, um die Ausgabe auf Basis des Inputs zu generieren.

Ein sehr wichtiger Aspekt der Sprachsynthese ist die Latenzzeit, die beispielsweise beim weiter unten vorgestellten *Lecture Translator (LT)* vor eine große Herausforderung stellte, zumal die MÜ-Komponente aus einem phrasenbasierten Modell bestand. So konnte die Übertragung nur erfolgen, sofern eine ganze Phrase bzw. ein ganzer Satz erkannt wurde, was wiederum die zeitliche Verschiebung erhöhte. Dieses Problem wurde beim LT durch eine Erweiterung des Systems gelöst. Konkret bedeutet dies, dass die Übersetzung in zwei Schritten erfolgt: Vorerst wird nur der erkannte (halbe) Satzteil übersetzt. Sobald der ganze Satz gegeben ist, wird die ursprüngliche Übersetzung aktualisiert. Selbiges Verfahren wird bei der ASR angewendet. Die Bedeutung der Latenzzeit ist bei Systemen wie dem LT besonders groß, da sie im Rahmen interaktiver Szenarien wie zum Beispiel universitäre Vorträge in Einsatz kommen, bei denen auch nonverbale Elemente sowie Präsentationsfolien eine große Rolle spielen.

Die Sprachsynthese gilt als der im Vergleich zur Spracherkennung und Maschinellen Übersetzung einfachere Prozess, zumal hier ein bereits fertiger, zuvor analysierter Text als Signal ausgegeben werden muss. Trotzdem trifft man bei der Programmierung der TTS-Komponente auf Limitationen: Wie läuft die Verwandlung des Textes in phonetische Lautschrift ab? Wie geht man weiters bei der Zuordnung dieser phonetischen Lautschrift zu einer Sprache vor? Dabei muss bedacht werden, dass sich die Ausspracheregeln unterschiedlicher Sprachen deutlich voneinander unterscheiden. Auch in diesem Fall wird die Problematik dank Lernalgorithmen gelöst, die anhand eingespeister, bereits existierender Daten arbeiten. Die Bedeutung des Datenbestandes kann somit nicht oft genug hervorgehoben

werden. Der Prozess der Sprachsynthese ist im Vergleich zur Spracherkennung und zur maschinellen Übersetzung nicht umkehrbar. „Lücken- oder fehlerhafte Verarbeitungen“ können folglich nicht in einer „späteren Komponente [ergänzt] oder [korrigiert]“ (Carstensen et.al. 2001:462) werden, wodurch sich oft Fehlanalysen durch das System fortpflanzen und Folgefehler auslösen können. Umso wichtiger scheinen demzufolge die einzelnen TTS-Komponenten wie beispielsweise zur lexikalischen, prosodischen oder phonetischen Analyse zu sein. Etwaige eingebaute grammatische Sprachmodelle arbeiten kontextspezifisch, mit dem Ziel, eine syntaktische Kongruenz zwischen zusammengehörigen Wörtern zu liefern (vgl. Carstensen et.al. 2001:465).

2.2. Technische Entwicklung und Vergleich

In Kapitel 1.1 der vorliegenden Arbeit wurde im Groben bereits auf die Entwicklungsphasen im Bereich Computerlinguistik/maschinelles Übersetzen/maschinelles Dolmetschen eingegangen. Im Folgenden soll diese Thematik näher beleuchtet werden.

Erste praktische Umsetzungen von Text-to-Speech-Forschungssystemen wurden Anfang der 90er-Jahre gewagt, als die Maschinen begannen, den dafür notwendigen minimalen Reifegrad zu erreichen (vgl. Waibel 2015b:108). Tabelle 1 nach Waibel (2015b:107) zeigt einen klaren Überblick über die im Laufe der Entwicklungsphasen entstandenen Systeme und den Fortschritt in Richtung Maschinellem Dolmetschen wie man sie heute kennt.

	Jahre	Vokabular	Sprechstil	Domäne	Geschwindigkeit	Plattform	Beispielsysteme
Erste Dialog Demonstrationssysteme	1989–1993	begrenzt	beschränkt	limitiert	2–10 × RT	Workstation	JANUS-1, C-STAR-I
Einfache Phrasenbücher	seit 1997	begrenzt, modifizierbar	beschränkt	limitiert	1–3 × RT	Handheld-Gerät	Phraselator, Ectaco
Spontane Zwei-Weg-Systeme	seit 1993	unbegrenzt	spontan	limitiert	1–5 × RT	PC/Handheld-Gerät	JANUS-III, C-STAR, Verbmobil, Nespole, Babylon, Transtac
Übersetzung von Nachrichtensendungen, politischen Reden	seit 2003	unbegrenzt	Lesen/Vorbereitete Rede	offen	offline	PCs, PC-Clusters	NSF-STRDUST, EC TC-STAR, DARPA GALE
Simultanübersetzen von Vorträgen	seit 2005	unbegrenzt	spontan	offen	Echtzeit	PC, Laptop	Lecture Translator
Kommerzielle Konsekutivübersetzer per Telefon	seit 2009	unbegrenzt	spontan	offen	online und offline	Smartphone	Jibbiggo, Google, Microsoft
Simultandolmetscherleistungen	seit 2012	unbegrenzt	spontan	offen	online	Server, Cloud-basiert	KIT, EU-Bridge, Microsoft

Tabelle 1: Entwicklung sprachübersetzender und -dolmetschender Systeme

Es müssen vor allem zwei Unterscheidungen vorgenommen werden. Allen voran liegt eine klare Gegenüberstellung von domänenbeschränkten vs. domänenunbeschränkten Systemen vor. Gleichzeitig lassen sich auch die Übersetzungs- und/oder Dolmetschmodi der Maschine, in *konsekutiv* vs. *simultan* unterteilen. Dabei gehen domänenbeschränkte Systeme zumeist Hand in Hand mit dem Einsatz mittels konsekutiver Übersetzung oder Dolmetschung, während domänenunlimitierte Maschinen, in Echtzeit arbeiten.

Die erste Hürde, die im Laufe der Entwicklungsphasen überwunden werden musste, stellte stets die Spontansprache mit all ihren Herausforderungen dar. Dazu gehören beispielsweise Wiederholungen, Füllwörter, Pausen usw. Nachdem diese vom System erkannt werden, müssen sie in einem Vorbereitungsprozess für die Übersetzung herausgefiltert und gegebenenfalls korrigiert werden. So entstanden zwischen 1993 und 2000 die ersten domänenbeschränkten spontansprachlichen Systeme. Vor allem VERBMOBIL bildete zu der Zeit einen Meilenstein in der Entwicklung der Sprachübersetzung und bedarf im Folgenden (vgl. Kapitel 2.2.1) einer näheren Auseinandersetzung. Trotz des großen Fortschrittes in der Entwicklungsphase bis zum Jahr 2000, blieben die Systeme für den Einsatz zunächst

unbrauchbar, da zum einen die Domänen nicht ausreichend klar definiert waren und zum anderen auch die Geschwindigkeit der Maschinen unzureichend war. Weitere technische Herausforderungen brachte in der darauf folgenden Phase die Berücksichtigung der Umsetzungsart mit sich, da die Systeme je nach Einsatz, konsekutiv oder simultan arbeiten sollten. Ersterer Modus betrifft Dialogsysteme im mobilen Einsatz bzw. Szenarien, in denen wenige Sätze zeitverschoben in der Zielsprache wiedergegeben werden sollen. Dabei beträgt der gesamte Wörterbestand zumeist nicht mehr als 40.000 Termini. Diese Methode hat sich vor allem für den Feldeinsatz wie beispielsweise bei medizinischen Hilfeinsätzen, in Krisengebieten usw. gut bewährt, nachdem sie im Rahmen humanitärer und logistischer Übungen der US-Regierung vorerst getestet wurde. Dies bestätigt auch die Domänenabhängigkeit des automatischen Konsekutivdolmetschens, was anfänglich auch nur mit einer kleinen Interessenten- und Abnehmergruppe einherging. Erst mit dem Einzug von Smartphones machte es Sinn, die Technologie auf diesem Gebiet weiterzuentwickeln und mobiler, offener und unbeschränkter zu gestalten. Als wahrscheinlich erfolgreichstes System dieser Art ist JIBBIGO (2009), der weltweit erste mobile Speech-to-Speech-Übersetzer (siehe Kapitel 2.2.2.), zu sehen.

Der simultane Modus hingegen war von Anfang an für domänenunbeschränkte Systeme konzipiert. Das erste Projekt dieses Forschungsbereiches stellte das TC-STAR-System dar, das im EU-Parlament im Rahmen des Vergleiches zwischen Mensch und Maschine erprobt wurde (vgl. Kapitel 4). Als Meilenstein auf diesem Gebiet kann zweifelsohne der vom Karlsruher Institut für Technologie und der CMU (Carnegie Mellon University) entwickelte und 2005 erschienene *Lecture Translator* genannt werden, auf den in einem separaten Kapitel näher eingegangen werden soll (vgl. Kapitel 2.2.4.).

Einen weiteren Bestimmungsfaktor in Bezug auf die Entwicklung von automatischen Dolmetschsystemen stellt die Multimodalität von Sprache und Kommunikation dar. Dies bestätigt u.a. Paul Watzlawicks (2011:53) erstes Axiom: „Man kann nicht nicht kommunizieren“. Kommunikation findet auf verschiedene Weisen ihren Ausdruck – durch Sprache, aber auch durch Mimik, Gestik, Bilder oder durch verzerrte (SMS und Social Media-)Sprache, die heute mehr denn je zu Tage kommt. Diese Beobachtungen bewirkten die Entwicklung neuer Forschungssysteme für das Übersetzen und Dolmetschen. Nennenswerte Beispiele sind u.a.:

- Heads-Up-Display-Brillen, sog. *Goggles*: diese übersetzen simultan und wurden für die Sprachen Deutsch und Spanisch entwickelt. Der/die BenutzerIn sitzt dem/der GesprächspartnerIn gegenüber und sieht den Zieltext als Untertitel über die Brille aufscheinen. Heute ist dieses System auch über mobile Geräte wie Smartphones, Smartwatches und dergleichen verfügbar.
- Dolmetschung mit Richtungs-lautsprechern: dabei werden die Lautsprecher auf bestimmte Punkte im Raum fokussiert (vgl. Waibel 2015b:119), sodass die Simultandolmetschung auch ohne Kopfhörer nur von den ZielrezipientInnen akustisch wahrgenommen wird. Dabei können über mehrere Richtungs-lautsprecher in demselben Raum, Dolmetschungen in verschiedenen Sprachen gleichzeitig angeboten werden.
- Handschrifterkennung: dieses System erkennt eine Schrift oder Zeichen wie Straßenschilder und übersetzt diese. Ein Beispiel dafür wäre der 2001 erschienene *Road-Sign-Translator*, der chinesische Schriftzeichen übersetzt (vgl. Yang et al. 2001).
- Stille Sprache (*Silent Speech*): erkennt artikulierte Mundbewegungen als Sprache, auch wenn nichts laut ausgesprochen wurde.

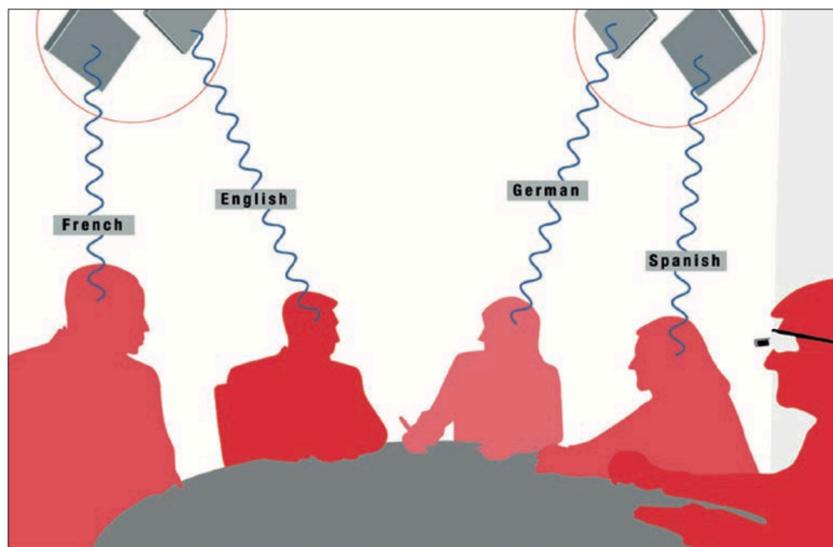


Abbildung 12: Simultandolmetschen und -übersetzen über Richtungs-lautsprecher und Heads-Up-Display-Brillen

Nachdem ein genereller Überblick über die Entwicklungsphasen sprachübersetzender und –dolmetschender Systeme geliefert wurde, sollen im Folgenden einige Systeme im Detail vorgestellt werden.

2.2.1. Verbmobil

Verbmobil ist auf sprachtechnologischem Gebiet als großer Meilenstein in der Entwicklung von Dialogsystemen zu sehen. Das Projekt entsprang einem Leitvorhaben des deutschen Bundesministeriums für Bildung und Forschung, das die Forschungsarbeiten mit rund 60 Millionen Euro (Pöchhacker 2016:189) förderte. Die Durchführung des Projektes dauerte insgesamt 8 Jahre (1993-2000) und wurde vom Deutschen Forschungszentrum für Künstliche Intelligenz (DFKI) in Zusammenarbeit mit anderen Forschungszentren, Universitäten und Unternehmen der Informationstechnologie übernommen (vgl. Wahlster 2000). Ziel war die Entwicklung eines mobilen MD-Systems für Spontansprach- bzw. Dialogübersetzungen in ausgewählten Anwendungsbereichen wie Terminverhandlungen oder Reiseplanungen. Der Fokus lag dabei auf Echtzeitübertragungen zur Unterstützung von Kommunikation in Form von Dialogen. Ziel des Projektes war, dass das System im Falle unzureichender Kenntnisse seitens der GesprächspartnerInnen, einspringen würde. In Zweifelsfällen haben GesprächspartnerInnen die Möglichkeit, in die eigene Muttersprache, also Deutsch oder Japanisch zu switchen während das Dialogsystem die Übersetzung der Passage ins Englische übernimmt. Aus technischer Sicht gleichen die Komponenten prinzipiell jenen eines bidirektionalen sprachübersetzenden Systems. Der Unterschied basiert jedoch auf den festgelegten Zielen und Voraussetzungen.

Das Verbmobil-Projekt ist in zwei Phasen durchgeführt worden. Im Laufe der Phase 1, die von 1993 bis 1996 dauerte, wurde der Schwerpunkt auf Terminverhandlungsdialoge zwischen den Sprachen Deutsch und Japanisch, mit der Zwischensprache Englisch gesetzt. Aufgrund der erfolgreichen Realisierung des Verbmobil-Projektes in Phase 1, wurde die zweite Phase für die Jahre 1997 bis 2000 genehmigt. Im Rahmen letzterer konzentrierten sich die Forschungsarbeiten auf die bidirektionale Übersetzung spontansprachlicher Dialoge aus den Themenbereichen Reiseplanung und Hotelreservierung in den Sprachkombinationen Deutsch-English und Deutsch-Japanisch.

Der nach der ersten Phase erzielte Erfolg des Verbmobil-Projektes lässt sich mit der Erreichung aller der insgesamt zehn technischen Ziele (s. Anhang) nachweisen. Darunter

fallen unter anderen: die „Erkennung fließend gesprochener Spontansprache für Deutsch, Japanisch und Englisch über [ein] Nahbesprechungsmikrofon“, ein Wortschatz von ca. 2500 Wörtern für die Sprachrichtung Deutsch nach Englisch sowie „über 70% approximativ korrekte Übersetzungen bei der End-to-End Evaluation in der Domäne Terminverhandlung“ (Wahlster 2000). Als approximativ korrekt gilt eine Übersetzung dann, „if it preserves the intention of the speaker and the main information of his/her utterance.“ (Tessitore 2000:629). Die WER-Rate konnte im Laufe der ersten Phase von anfänglichen über 50% auf rund 14% reduziert werden. Der erzielte Erfolg war in großem Maße auf die klar strukturierte, aber gleichzeitig flexible und erweiterbare Architektur des Verbmobil-Prototyps zurückzuführen. Wie anhand Abbildung 13¹⁷ erkennbar, umfasst Verbmobil u.a. voneinander unabhängige Module, die untereinander kommunizieren.

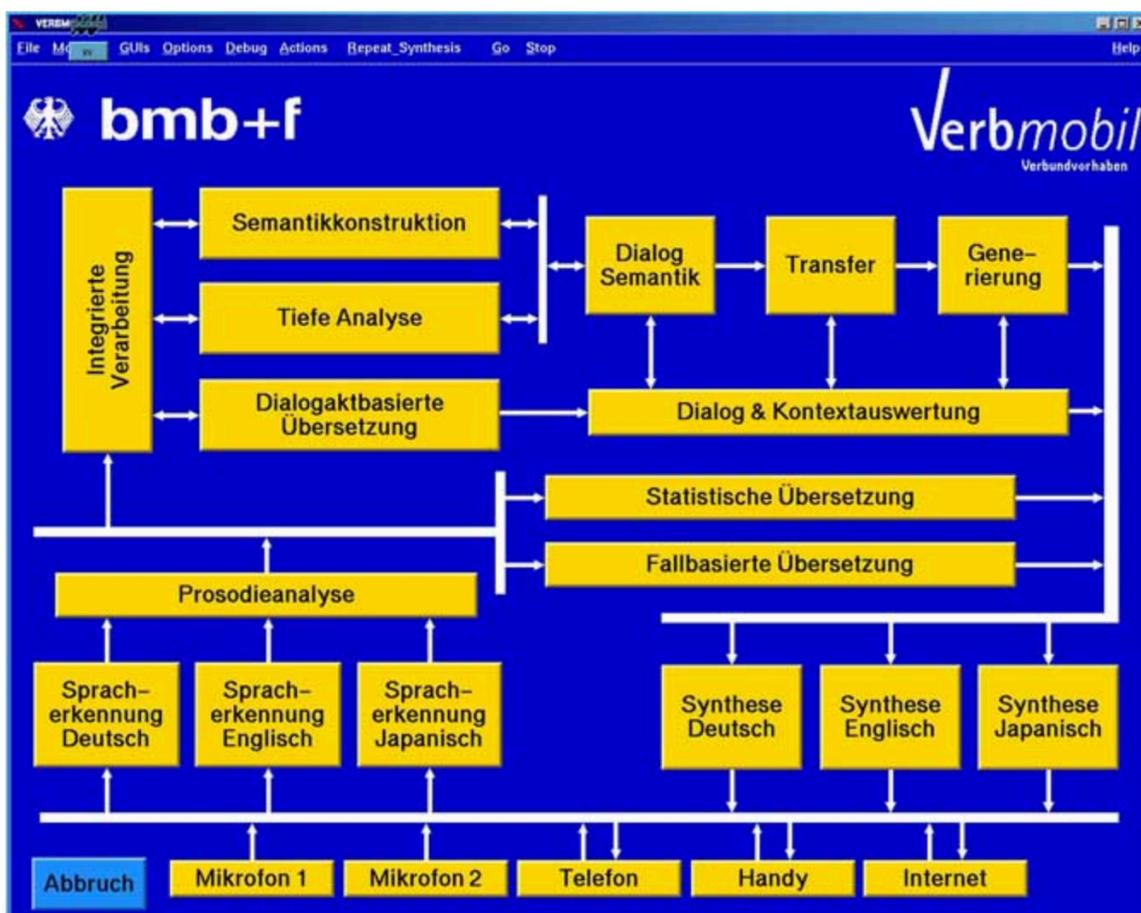


Abbildung 13: Benutzeroberfläche des Forschungsprototypen von Verbmobil

¹⁷ Abb. 13 umfasst nur die Hauptmodule der insgesamt 43 Komponenten

Im Groben besteht das System aus einigen Hauptmodulen wie der Spracherkennungskomponente, den Modulen zur syntaktisch-semantischen Analyse, zum Transfer oder zur Generierung. Die Koordinierung dieser Bereiche erfolgt wiederum über andere Module. Dieser Aufbau bringt den Vorteil mit sich, dass die Software anpassbar und wiederverwendbar ist, was die Weiterentwicklung des Systems vereinfacht. Für den Entwicklungsprozess selbst wirkte es sich insofern positiv aus, als die Software mit ihren Komponenten geographisch unabhängig entwickelt werden konnte.

Eine Besonderheit des Systems ist, dass es als einziges, prosodische Informationen zur Weiterverarbeitung der akustischen Eingabe nutzt, was einen Ersatz für die zur Segmentierung notwendige, jedoch fehlende Interpunktion darstellt und den Verarbeitungsprozess signifikant beschleunigt.

Im Rahmen der zweiten Phase lag der Forschungsfokus (s. Anhang) auf der direkten Übersetzung spontansprachlicher Dialoge für die Sprachenkombinationen Deutsch-Englisch, mit einem Datenbestand von 10.000 Wörtern und Deutsch-Japanisch mit 2.500 Wörtern. Da die Entwicklung eines multilingualen Systems das Ziel war, wurde auf die Implementierung sprachunabhängiger und in hohem Maße reversibler Prozesse und sprachlicher Wissensquellen gesetzt. Letztere können sich automatisch an neue Domänen anpassen, wobei das System selbst auch in der zweiten Phase domänenbeschränkt bleibt. Die Adaptierbarkeit ist insofern wichtig, als der Verbmobil-Prototyp somit als Ausgangspunkt für die Entwicklung späterer Forschungssysteme gilt.

In Phase 2 des Verbmobil-Projektes wurde die Dialogführung ohne Mikrofon, das heißt durch Freisprechen ermöglicht. Getestet wurde das System einerseits für Szenarien wie Hotelreservierungen und Verkehrsroutenanfragen und andererseits für den Einsatz bei mehrsprachigen multimedialen Telekonferenzen zum Thema Reiseplanung.

2.2.2. Jibbig

Jibbig stellt den 2009 erschienenen, weltweit ersten tragbaren Sprachübersetzer für iOS- und Android-Geräte dar, der sowohl als Online-, als auch als Offlinelösung nutzbar ist und somit eine mobile und durch ausfallende Roamingkosten, verhältnismäßig günstige Applikation ist. Entwickelt wurde das System vom Startup-Unternehmen *Mobile Technologies LLC*¹⁸. Ursprünglich war es für den Einsatz im humanitären sowie im touristischen Bereich konzipiert. So wurde es u.a. von der US-Regierung und von karitativen NG-Organisationen in

¹⁸ das Unternehmen wurde 2013 von Facebook übernommen (vgl. Cohen 2013)

Thailand, Kambodscha und den Honduras genutzt, wobei es der Kommunikation zwischen Ärzten und Patienten diene. Eine Evaluierung des Jibbig-Systems erfolgte 2013 im Rahmen des *Medical Civil Action Program* in Thailand. Die Ergebnisse zeigten, dass erfolgreiche Interaktionen zwischen Ärzten und Patienten zu bis zu 95% und ohne Intervention seitens Humandolmetscher, möglich waren (vgl. Waibel 2015b:112).

Weiters hat Jibbig auch im Einsatz als Reise-App sehr große Erfolge erzielt. So besteht für Reisende, abgesehen von der Offline-Funktion noch die Möglichkeit, regionale Übersetzungspakete für Reisen in mehrere verschiedensprachige Länder herunterzuladen (vgl. Businesswire 2012). Insgesamt ist die Jibbig-Applikation in über 15 Sprachen und Dialekten mit einem Wortschatz von insgesamt 40.000 Wörtern verfügbar.

Seit der Übernahme des Entwicklerunternehmens *Mobile Technologies* durch *Facebook*, wird Jibbig auch für die Übersetzung von Facebook-Postings eingesetzt.

2.2.3. Skype Translator

Mit der Ankündigung des Markteintrittes durch den Skype Translator, setzte *Microsoft* einen weiteren Meilenstein im Bereich des Maschinellen Dolmetschens. Die Skype-Applikation ermöglicht Echtzeit-Dolmetschungen von Spontansprache in über 50 Sprachen (vgl. Paul 2015). Somit ist der erste große Schritt in Richtung multilingualer Kommunikation in simultanem Modus gesetzt.

Beim Skype Translator kommen die drei üblichen Speech-to-Speech-Komponenten zum Einsatz, d.h. Spracherkennung, maschinelle Übersetzung und Sprachsynthese. Was den Skype Translator jedoch von anderen sprachübersetzenden Systemen unterscheidet, ist, dass sich dessen Entwickler im Detail mit der Thematik des Sprechstils auseinandergesetzt haben. Je mehr man die Unterschiede zwischen geschriebener und gesprochener Sprache berücksichtigt, desto zufriedenstellendere Ergebnisse wird auch das Speech-to-Speech-System liefern. Um dies zu erreichen, hat Skype die davor eingesetzte ausschließlich phrasenbasierte statistische maschinelle Übersetzung (SMT) durch eine syntaxbasierte Phrasen-SMT ersetzt (vgl. Chong 2014), die es ermöglicht, eventuelle Unterschiede der Sprachstile zu erkennen und Ambiguitäten zu reduzieren oder gar zu beseitigen. Dies erfolgt indem Sätze zuerst in einzelne Wörter segmentiert werden und erst dann in der Zielsprache entsprechend eingeordnet werden.

Der Lernprozess des Skype Translators erfolgt auf Basis von Social Media-Datenbasen. Diese Quelle soll der Maschine dazu verhelfen, die menschlichen

Kommunikationscharakteristika besser zu verstehen. So definiert auch Jurafsky (zit. in Chong 2014) die Lernprozesse anhand von Social Media-Quellen als: „important examples of a new line of research in computational social science, showing that subtle social meaning can be automatically extracted from speech and text in a complex natural task“. Die damit verbundene Herausforderung besteht in Zusammenhang mit den je nach Quelle (Facebook Posting vs. SMS usw.) variierenden Eigenschaften der benutzten Sprache. Dafür wurde eine weitere Komponente entwickelt, die diese Sprache verarbeitet und in entsprechende Segmente umwandelt, die wiederum an die MT-Komponenten weitergeleitet werden. Dank diesem Prozedere kann das System zu Social-Media-Sprachübersetzungen um 6% genauer arbeiten (vgl. Chong 2014).

Die Nutzungsstatistik für das sich noch in der Testphase befindende S2S-System ergab für das Jahr 2015 folgende Ergebnisse:

- Die beliebteste Sprachenkombination ist Französisch zu Englisch.
- Seit Einzug des Skype Translators auf den Markt, konnten 400% mehr Telefonate verzeichnet werden.
- Die meisten Telefonate erfolgen von Deutschland nach Ghana (vgl. Skype Team 2016).

Derzeit umfasst das Angebot für die S2S-Translation die Dolmetschung in sieben Sprachen: Deutsch, Englisch, Französisch, Italienisch, Mandarin, Portugiesisch und Spanisch (vgl. Kannenberg 2016). Weiters bietet Microsofts System die schriftliche Untertitelung bzw. Echtzeit-Übersetzung in 50 Sprachen an.

2.2.4. Lecture Translator (LT)

Das Projekt des Lecture Translators entsprang dem Bedürfnis und der Idee, für Studierende aus dem Ausland herrschende Sprachbarrieren zu durchbrechen, mit dem Ziel der Internationalisierung der in gegebener Ausgangssituation deutschsprachigen Universität. Initiiert wurde das LT-Projekt am Karlsruher Institut für Technologie (KIT), unter der

Leitung von Professor Alexander Waibel und in Zusammenarbeit mit der Carnegie-Mellon-Universität (CMU).

Die Idee des Englischen als Lingua franca und als Vortragssprache im Rahmen universitärer Vorlesungen wurde dabei in den Hintergrund gerückt. Auch die Erwartung und Annahme, dass ausnahmslos alle Studierenden aus dem Ausland Deutsch lernen sollen, um den Unterricht mitverfolgen zu können, wurde als eher unrealistisch erachtet. Stattdessen sollte eine Lösung gefunden werden, die „sprachliche und kulturelle Vielfalt und Toleranz unterstützt“ (Waibel 2015b:112) hätte, ohne auf HumandolmetscherInnen angewiesen zu sein, was auf Dauer mit höheren Kosten verbunden gewesen wäre. Dies war der Ausgangspunkt für die Entwicklung einer modernen Sprachtechnologie zur Überwindung von Sprachbarrieren.

Der LT wurde erstmals im Jahr 2005 vorgestellt. Es handelt sich um ein cloudbasiertes domänenunbeschränktes System, das die Architektur vom EU-Bridge-Projekt¹⁹ nutzt (vgl. Abb. 14).

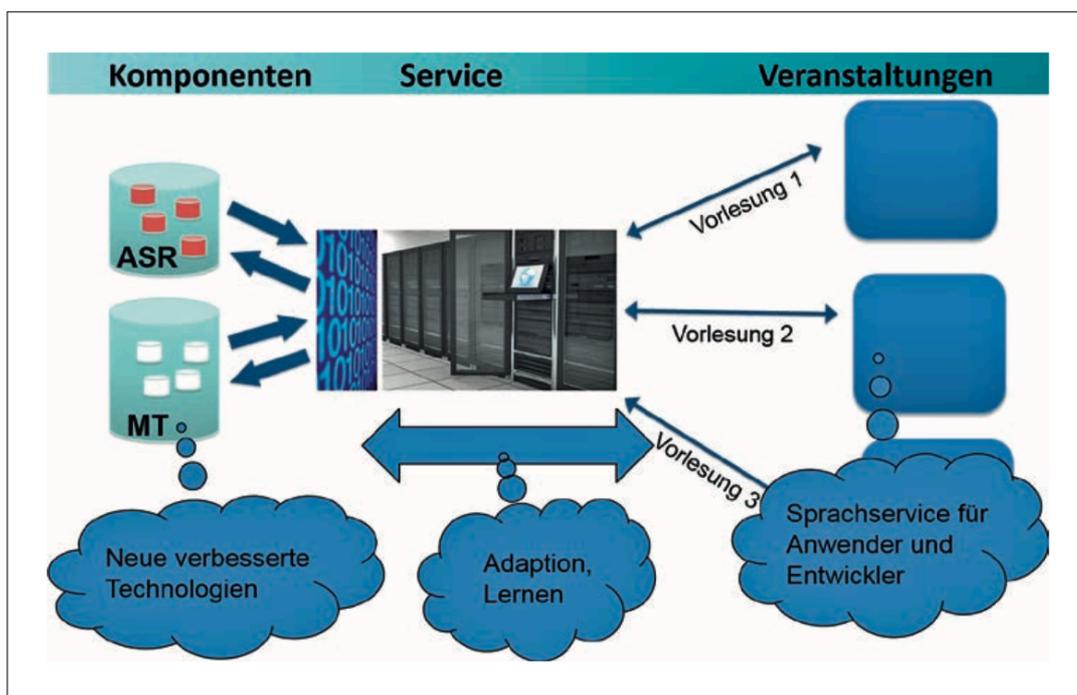


Abbildung 14: EU-Bridge-Architektur: MD als cloudbasierter Dienst (Waibel 2015:116)

Die Architektur des Systems ist von drei Hauptkomponenten abhängig:

¹⁹ EU-Bridge ist ein Multisite-Programm zur Beseitigung von Sprachbarrieren. Im Rahmen des Programmes werden u.a. Sprachtechnologien entwickelt (vgl. EU-Bridge)

1. zentraler Server, der als „Mediator“ fungiert
2. *worker* (ASR, MÜ, Segmentierungskomponente etc.)
3. BenutzerInnen, die eine Dienstleistung erfragen

Der Prozess beginnt bei den *workern*, die dem „Mediator“ Informationen über ihre eigene Funktion übermitteln. Der zentrale Server speichert die zu erfüllende Aufgabe ab und erkennt, in welcher Reihenfolge die *worker* in Einsatz kommen müssen, um den Auftrag der BenutzerInnen erfolgreich erfüllen zu können.

Die Entwicklung der LT-Sprachtechnologie war insofern eine Herausforderung, als es sich hier um einen maschinellen Simultandolmetscher für Monologe handelt, die im Konsekutivmodus aus maschineller Sicht aufgrund der niedrigeren Ansprüche in Bezug auf die Latenzzeit, einfacher zu entwickeln wären. Die Erfordernis einer möglichst geringen Latenz ist besonders im Deutschen problematisch, da hier das Verb meist am Satzende steht. Im Laufe des gesamten Arbeitsprozesses des Systems, lassen sich drei Hauptherausforderungen erkennen:

1. Wie erkennt man die Sprache und konvertiert sie in Text? (*Speech-to-Text*)
2. Wie konvertiert man den Text der Ausgangssprache in Text der Zielsprache? (*Text-to-Text*)
3. Wie wird die Ausgabe so präsentiert, dass sie Studierende auf iPhones, iPads oder Laptops mitverfolgen können?

Die Antwort auf diese Fragen findet sich in den Hauptkomponenten der Maschine (vgl. Kapitel 2.1.): Für den Spracherkennungsprozess wird das ebenso am KIT entwickelte Janus Recognition Toolkit (JRTk) eingesetzt. Das akustische Modell wurde dafür mit Hunderten von Stunden an Aufnahmematerial trainiert. Das Ergebnis der ASR des Lecture Translators wird an die Komponente für die Resegmentierung weitergeleitet, wo Anfang und Ende eines Satzes bestimmt werden und die Interpunktion gesetzt wird. Diesen Schritten folgt der Übersetzungsprozess. Dieser beruht beim LT auf statistischer maschineller Übersetzung anhand implementierter Wörterbücher (*Discriminative Word Lexica*) und ist somit

phrasenbasiert. Abgesehen von den Wörterbüchern kommen auch (bilinguale) Sprachmodelle zum Einsatz. Diese Übersetzungsmethode wurde anhand verschiedener Ausgangstexte sowie themenbezogener Daten trainiert. Insgesamt wurden dafür 1,8 Mio. Sätze in das System eingespeist (vgl. Müller et al. 2016).

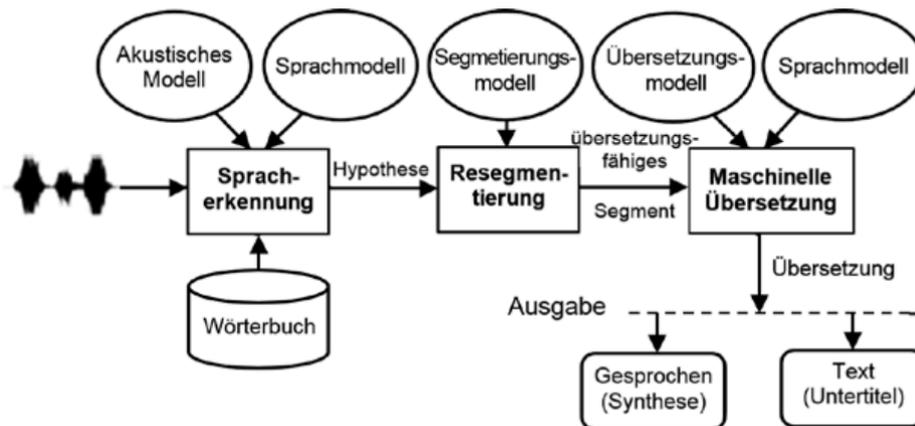


Abbildung 15: Übersetzungsprozess des LT

Zusätzlich kann die Ausgabe in Form von Untertiteln auf dem jeweiligen Ausgabegerät mitverfolgt werden. Die Verbindung zum System erfolgt über einen Webbrowser. Mögliche Sprachkombinationen sind: die Übertragung aus dem Deutschen ins Englische sowie umgekehrt, aus dem Englischen ins Spanische und ins Französische sowie aus dem Deutschen ins Französische.

Neben der Online-Dolmetschung bietet das System auch eine Offline-Lösung, bei der kein Echtzeitmodus notwendig ist. Dabei kann sowohl die Transkription, als auch die Übersetzung unter Berücksichtigung eines längeren Kontexts bestimmt werden. Die Offlinelösung ermöglicht weiters eine Archivierung und Nachbearbeitung der Übersetzung.

Trotz der mittlerweile schon sehr fortschrittlichen Entwicklung des Lecture Translator, gilt es weiterhin, einige mit der Transkription und Übersetzung, vor allem aus der deutschen Sprache, in Verbindung stehende Problematiken zu lösen oder in möglichst hohem Maße zu reduzieren. Neben der bereits erwähnten Verb- bzw. Wortstellung, stellen auch Komposita eine Herausforderung für die Übersetzungskomponente dar, zumal diese auf sinnvolle Art und

Weise zerlegt werden müssen. Die MÜ-Komponente muss z.B. erkennen, ob es Sinn macht, den Terminus „dramatisch“ in „Drama“ und „Tisch“ zu zerlegen. Diesem Problem versucht man mittels Einsatz von Algorithmen entgegenzuwirken. Das mit den Komposita zusammenhängende Problem unterstreicht die Bedeutung der Kontextabhängigkeit, die Fügen (2008:68) mit folgenden Worten zusammenfasst: „The more information about a speaker or topic is known in advance, the better the model can be adapted and the better the system might perform ultimately.“

Ein weiterer Aspekt, der beim LT eine große Rolle spielt, ist die Bedeutung von vorlesungsspezifischen Fachwörtern. Dafür besitzt der LT das oben erwähnte Wörterbuch, das entweder über automatische Algorithmen oder anhand manueller Korrekturen seitens Benutzer aktualisiert wird. Schwierig für das System ist auch das Erkennen von Code-Switching, wenn im Rahmen deutschsprachiger Vorträge, englische Fachwörter eingesetzt und oft sogar dekliniert verwendet werden.

Neben anderen Faktoren wie Lesbarkeit bzw. Interpunktion, Pronomina, Eingabe über ein Nahbesprechungsmikrofon und Spontansprache, ist auch die sprachliche Skalierbarkeit (*Portability*) ein Thema, auf das man sich im Rahmen weiterer Forschungen in Zukunft zweifelsohne konzentrieren wird. Ziel dabei ist es, Übersetzungen zwischen möglichst vielen Sprachen und Kulturen zu erreichen.

Trotz der soeben genannten Limitationen und Herausforderungen, hat sich der LT als ein vielversprechendes MD-System erwiesen. Mittlerweise wird das System am KIT seit Jahren bei zahlreichen, auch parallel laufenden Vorlesungen genutzt und ist es auch sehr wahrscheinlich, dass das System schon bald in anderen Universitäten bzw. anderen Ländern eingesetzt werden wird. Zusammenfassend kann mit Waibels (zit. in KIT:2015) Worten behauptet werden: „The translation is not always perfect [...], but it is part of the language tools, by means of which students are enabled to better follow the lectures in spite of language barriers.“

Anhand von Benutzerstudien konnte bewiesen werden, dass der LT seinen Skopos erfüllt, zumal er den Studierenden zu einem besseren Verständnis der universitären Vorträge verhilft.

2.3. Voraussichtliche zukünftige (technische) Entwicklung und Erwartungen

Die bisherige technische Entwicklung sowie die oben näher beschriebenen großen Fortschritte im Bereich der Sprachtechnologie und Computerlinguistik, haben bereits bewiesen, dass Verbesserungen und Erweiterungen dieses Forschungsgebietes als selbstverständlich anzusehen sind. So besteht des Weiteren auch kein Zweifel darüber, dass die Zukunft, innovative, erfolgsversprechende Projekte und Ansätze mit sich bringen wird, dank denen es womöglich auch hier und da gelingen wird, bisherige Hürden zu überwinden und Problematiken zu beseitigen.

2.3.1. Paradigmenwechsel – SMT zu NMT

Heute kann man bereits von einer Phase sprechen, die als Anfang eines Paradigmenwechsels auf sprachtechnologischem Forschungsgebiet zu sehen ist. An dieser Stelle erscheint es relevant, an die Bedeutung des Begriffes Paradigmata zu erinnern, die nach Kuhn (1976²:10) als „allgemein anerkannte wissenschaftliche Leistungen, die für eine gewisse Zeit einer Gemeinschaft von Fachleuten maßgebende Probleme und Lösungen liefern“ definiert wird. Weiter meint er: „Ein Paradigma ist das, was den Mitgliedern einer wissenschaftlichen Gemeinschaft gemeinsam ist, und umgekehrt besteht eine wissenschaftliche Gemeinschaft aus Menschen, die ein Paradigma teilen“ (Kuhn 1976²:187). Beginnt die Aufrechterhaltung der „allgemein anerkannten Leistungen“ zu scheitern, so kommt es zu einem Paradigmenwechsel, im Rahmen dessen eine neue Theorie die alte verdrängt.

In der Computerlinguistik können wir im Groben zwei bedeutende Paradigmenwechsel erkennen (vgl. Abb. 16). Wie bereits in vorhergehenden Kapiteln beschrieben, kann die regelbasierte Übersetzung als Ausgangspunkt aller Entwicklung von maschinellen MT-Systemen verstanden werden. Regelbasierte Forschungssysteme waren sehr stark von der Expertise von Fachleuten abhängig. So wurden etwaige Fortschritte durch den Einsatz und die Zusammenarbeit von Linguisten bzw. Sprachwissenschaftlern, Informatikern und anderen erzielt. Diese Phase der Forschung im Bereich regelbasierter maschineller Übersetzung hielt sich eine verhältnismäßig lange Zeit aufrecht, bis die statistische MT einen ausreichenden Reifegrad erreicht hatte, um ab dem Jahr 2000 in Forschungssystemen eingesetzt zu werden. Zwar wird diese Phase noch weiter bestehen, wie auch Waibel bestätigt (2015a: 6:10), doch ist bereits ein Paradigmenwechsel im Gange, der die Nutzung neuronaler Netze in den Vordergrund stellt. Die Funktionsweise neuronaler Technologien orientiert sich

an den Prozessen im menschlichen Gehirn. Im Falle maschineller Übersetzung, zeichnet sich das u.a. durch den Verarbeitungsprozess der jeweiligen Repräsentationen aus. Dabei werden im Gegensatz zur SMT zusammenpassende Wörter über einen Encoder aufgenommen und erkannt bevor sie folglich als ganze Sätze repräsentiert werden. In einem nächsten Schritt werden die Repräsentationen in die Zielsprache umgeformt und generiert. Die neue Architektur neuronaler Systeme ermöglicht die Sammlung von Wörtern und dazugehörigen Informationen in einem multidimensionalen Kontext (vgl. Vogel 2016: 8:33). Trotz dieses neuen Ansatzes kann noch nicht von einem voraussichtlich rapiden Fortschritt in der NMT die Rede sein, da die in der SMT bereits vorhandene Herausforderung weiterhin bestehen bleibt. In beiden Fällen handelt es sich um datenbasierte Systeme, die auf die Einspeisung einer angemessenen Datenmenge angewiesen sind und die stets nach einer Regel funktionieren: Je mehr Daten, desto höher die Qualität. Der jetzige Forschungsstand verspricht vorerst noch keine baldige Lösung des mit der Datenmengenabhängigkeit zusammenhängenden Problems. Vielmehr wird davon ausgegangen, dass diese Hürde durch Zurückgreifen auf frühe Methoden bewältigt werden kann. Dies erfordert eine erneute Auseinandersetzung mit der Wort-für-Wort Übersetzung, zumal die NMT zur Zeit wortbasiert funktioniert. So stellt sich die Frage, ob und wie eine Implementierung syntaktischer Strukturen möglich sein wird. Zweifelsohne werden die Erkenntnisse aus den letzten 50 Forschungsjahren eine große Stütze für die Entwickler von Systemen der neuen Art sein. Ein Versuch, neuronale Netze in der sprachtechnologischen Forschung einzusetzen, wurde bereits vor rund 30 Jahren gewagt. Das Unterfangen scheiterte zum einen auf Grund der limitierten Hardware und zum anderen wegen unzureichender mathematischer Implementierungen. Da jedoch auf diesem Gebiet mittlerweile große Fortschritte nachgewiesen werden konnten, ist es auch möglich, den Ansatz der NMT wieder aufzugreifen.

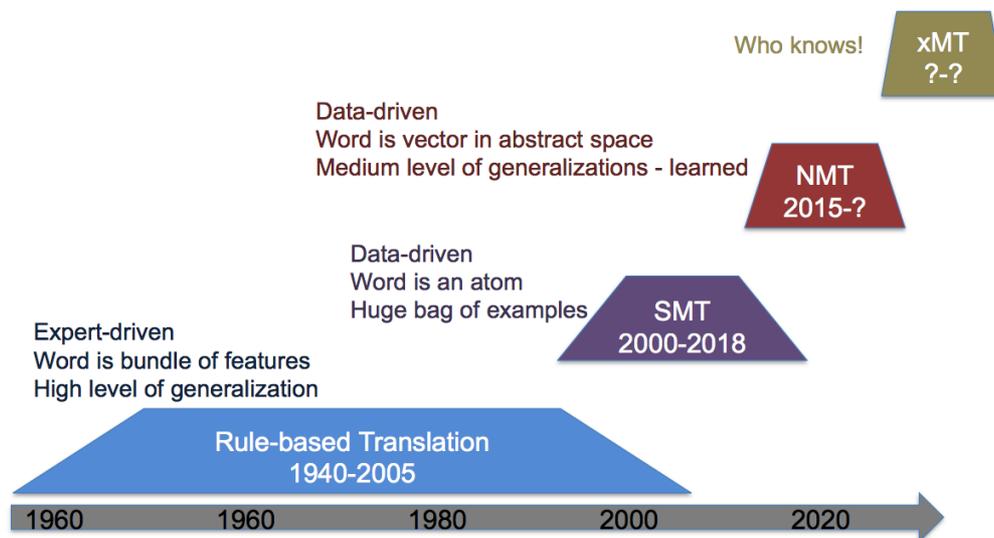


Abbildung 16: Paradigmenwechsel in der MT (Vogel 2016)

Schlussfolgernd wird davon ausgegangen, dass der Trend in Richtung Weiterentwicklung von NMT-Technologien beibehalten wird. Bis dato hat sich der Einsatz von neuronalen Netzen in der MÜ als vorteilhaft erwiesen, was nicht zuletzt anhand verbesserter Ergebnisse der Spracherkennungsprozesse deutlich wird. Es wäre realistisch anzunehmen, dass die statistische maschinelle Übersetzung ab dem Jahr 2018 gänzlich vom NMT-Paradigma ersetzt werden wird. Aufgrund der schnellen technologischen Fortschritte kann man jedoch einen erneuten Paradigmenwechsel nicht ausschließen. Die Frage, ob das Problem der Prozesse einschränkenden und verlangsamenden Abhängigkeit von Daten bis dahin gelöst werden wird können, bleibt vorerst noch unbeantwortet.

2.3.2. Technischer Ausblick vs. Limitationen

In vorliegender Arbeit wurden bereits viele in Zusammenhang mit der zukünftigen Entwicklung stehende technische Herausforderungen hervorgehoben. Zusammenfassend soll im Folgenden näher und auf diese eingegangen werden.

Ein Aspekt, der in Bezug auf die Problematik der sprachlichen Barriere zweifelsohne von Bedeutung ist, stellt die Skalierbarkeit (*Portability*) dar. Diese bestimmt, inwiefern ein maschinelles Übersetzungssystem um zusätzliche Sprachen erweitert werden kann, um die Kommunikation zwischen möglichst vielen Sprachen und Kulturen über ein System zu

ermöglichen. Der Einsatz solcher Systeme erfordert die Entwicklung umso sprachunabhängigerer Komponenten. Dieser Prozess wurde bereits initiiert zeigt erfolgsversprechende Umsetzungen. Dank dieser Entwicklung sollen bessere Ergebnisse in den Bereichen der Adaption, der Inferenz sowie der Abstraktion erzielt werden. Weiters vermag es auch die bessere Nutzung einsprachiger Datenbestände oder von „Crowdsourcing“, was das Herausfiltern vielsprachigen Wissens der Menschheit verbessern und gar kostenlos ermöglichen soll (vgl. Waibel 2015b:115). Eine Methode, die sich zu diesen Zwecken als aussichtsreich erwiesen hat, ist die Verarbeitung mittels neuronaler Netze (vgl. Kapitel 2.3.1), die u.a. am KIT in sprachunabhängigen Systemen bereits eingesetzt werden. Im Allgemeinen sind die Prognosen in Bezug auf die Erlernbarkeit seitens Technologien gut, da letztere anhand von Datenbasen für jede Sprache trainiert werden können. Die Grundvoraussetzung stellt hier eine umfangreiche Datenerhebung dar, die jedoch durch das World Wide Web immer einfacher wird. So werden moderne MÜ-Systeme mit über 1 GWords (vgl. Waibel 2015b:121) trainiert. Dies macht das Doppelte von dem aus, was ein Mensch im Durchschnitt während seiner ganzen Lebenszeit spricht (ca. 0,5GWords).

Weiters kommt der Sprachunabhängigkeit auch der Interlingua-Ansatz zu Gunsten. Dabei besteht insofern wiederum das Problem der Skalierbarkeit, als es schwierig ist, alle möglichen Ausdrucksmöglichkeiten für sprachunabhängige universale Zwischenrepräsentationen in allen zu übersetzenden Sprachenpaaren zu berücksichtigen. Es gälte, eine ideale Repräsentationssprache zu entwickeln, was bisher jedoch nicht realisiert wurde. Aktuelle Systeme sind bis dato hauptsächlich zu Forschungszwecken und nur für bestimmte Domänen oder „für Zielsprachen mit kontrollierbarem Vokabular und stark vereinfachten Satzstrukturen konzipiert“ (Waibel 2015b:110).

Schlussfolgernd kann bestätigt werden, dass die sprachliche Barriere noch lange nicht beseitigt wurde. Diese Hürde ist zusätzlich durch weitere Unzulänglichkeiten der maschinellen Übersetzungs- und Dolmetschsysteme bedingt. So weisen die Komponenten zur Spracherkennung und zur maschinellen Übersetzung eine noch unzureichende Performanz auf, was durch die zu hohe Fehlerrate und Ungenauigkeit bewirkt wird. Dies führt zu einem weiteren Manko, das die Umsetzung von menschlichen Faktoren betrifft. Während Humandolmetscher die Möglichkeit und Fähigkeit der Fehlerausbesserung oder –beseitigung besitzen, bieten Maschinen jeweils immer nur eine zielsprachliche Repräsentation an. Auch die adäquate Erkennung sprachlicher Differenzen und dementsprechend auch die angemessene Übermittlung der Informationen in der Zielsprache, haben noch keinen

ausreichenden Reifegrad erreicht. Außerdem lassen sich auch Schwierigkeiten in Bezug auf die Schaffung eines möglichst natürlichen bidirektionalen Dialogs (bei Dialogsystemen) erkennen.

Idealerweise würde ein MD-System auf Basis von Adaption funktionieren, indem es sich verschiedenen SprecherInnen und Domänen automatisch anpasst. Diese Anforderung erfragt zweifelsohne noch weiterer Forschungsarbeiten, da hierbei auch mehrere Faktoren, wie beispielsweise das Medium zur Signaleingabe (Mikrophon) oder Nebengeräusche, mitberücksichtigt werden müssen.

Anhand dieser wenigen Beispiele erkennt man schon, dass der Weg zu einem „idealen“ Dolmetschsystem aus technischer Sicht noch sehr lange sein kann. Zwar beweisen neueste sprachtechnologische Entwicklungen wie der Lecture Translator oder Skype Translator, dass ein schneller Fortschritt durchaus gegeben ist, doch stellen sich zwei große Fragen:

1. Wie sollen oben genannte Mängel behoben und entsprechende technische Funktionen durch Implementierung in die Systeme, in der Praxis umgesetzt werden?
2. Ist die Umsetzung solcher Projekte aus finanzieller und zeitlicher Sicht realisierbar?

Die Kosten für die Entwicklung eines Speech2Speech-Translation-Systems belaufen sich gemäß Waibel (2011: 24:20) je nach Anforderungen, auf rund 500.000 bis 20 Mio. \$²⁰. Die Kosten gelten für ein System, das mit einem Sprachenpaar und in einer Domäne programmiert wurde. Derzeit werden zwischen 20 und 50 Sprachen zu Zwecken maschinellen Übersetzens und Dolmetschens genutzt, da die Entwicklung solcher Technologien, abgesehen von den Kosten, auch viel Zeit in Anspruch nimmt. So wird im Durchschnitt jeweils über ein Jahr an der Implementierung einer Sprache in ein System gearbeitet, wofür jeweils ca. 5 Mio. \$ investiert werden (vgl. Waibel 2016). Infolgedessen wirken sich die hohen Kosten auf die o.g. Skalierbarkeit negativ aus, zumal auch bedacht werden muss, dass bestimmte Sprachen in den Vordergrund gesetzt werden müssen. Darunter fallen allen voran die gemäß Waibel (2011: 25:50) als „WMD“ (wealthy – many – dangerous) definierten Sprachen. Dabei bezieht

²⁰ Diese große Kostendifferenz kann auf Grund der jeweils eingesetzten Sprache entstehen. Für Sprachen, die nur in Wort vorhanden sind, müssen vorerst Autographien und erst darauf aufbauend ASR- und MÜ-Systeme entwickelt werden, was mit umso höheren Kosten verbunden ist.

sich „wealthy“ auf Sprachen aus reicheren Ländern, „many“ auf Sprachen, die von einem hohen Bevölkerungsanteil gesprochen werden und „dangerous“ auf Sprachen, die als „gefährliches“ Kommunikationsmittel angesehen werden. Im Gegensatz dazu stehen u.a. Sprachen, die nur in Wort, nicht aber in Schrift zirkulieren. Allein in Indien gibt es zum Beispiel rund 300 Sprachen, von denen viele nur gesprochen werden. Diese mit den Kosten in Verbindung stehende Hürde wirkt sich auf die Umsetzung von MD-Projekten für humanitäre Zwecke aus, die zumeist den Einsatz ebensolcher Sprachen erfordern.

Besonders teuer sind Prozesse wie die Datensammlung für die Spracherkennung und die Übersetzung sowie das „Erlernen“ der Daten durch die Systeme, Lernmodelle, die Implementierung der Daten, Testdurchführungen sowie kontinuierliches Nachgehen in Bezug auf weitere Forschungsarbeiten zu Gunsten des technischen Fortschrittes.

Schlussfolgernd beweist der aktuelle technische Entwicklungsstand, dass die Sprachtechnologie auf jeden Fall einen unterstützenden Charakter angenommen hat, was auch durch den vermehrten Einsatz von CAT-Tools bestätigt wird. Sowohl die Leistungen in Bezug auf Datenbestände und Datenverarbeitung, als auch auf die Technologien, zeigen, dass die Umsetzungen durchaus sinnvoll genutzt werden können. Die einzige Gefahr besteht stets darin, die entsprechenden Lösungen zu integrieren um somit Theorie und Praxis in einem System zu vereinen (vgl. Waibel 2015a: 19:30). Folgende Tabelle soll veranschaulichen, in welchen Bereichen sich zukünftige Forschungsarbeiten (weiterhin) konzentrieren werden (vgl. Waibel 2016):

Forschungsgebiet	Schwerpunkte
NMT (neuronale Netze)	- Datenmengenabhängigkeit
Sprachambiguität, Grammatik	- Morphologie, Übersetzung von Pronomen, Wortstellung
Lernalgorithmen	- Inkrementelles Lernen - Croudsourcing
Qualitätsverbesserung von Sprachübersetzung	- Latenzzeit - Fehler und Disfluenzen - Menschliche Faktoren (Intonation,

	Betonungen, Emotionen) - Kontextspezifische MÜ (Gender, formelle/informelle Übersetzung)
Plausibilität in der maschinellen Übersetzung	-

Tabelle 2: zukünftige Herausforderungen der MT- und MD-Forschung

2.3.3. MT/MD-Marktanalyse: Ein Überblick

Um voraussichtliche zukünftige Ansprüche, die der Sprachbranchen-Markt auf nationaler und internationaler Ebene stellen wird, erkennen zu können, ist es vorerst notwendig, einen Überblick über die aktuelle Situation zu verschaffen.

Weltweit kann ein 7%iges jährliches Wachstum im Bereich humaner Übersetzungsdienstleistungen verzeichnet werden, was mit einer über 30 Milliarden Euro Industrie gleichzusetzen ist. Die maschinelle Übersetzung wächst mit jährlichen 20-25% noch viel schneller. Dies ist auf verschiedene Faktoren zurückzuführen. Unter anderem darauf, dass immer mehr Unternehmen sowie Institutionen wie das Europäische Parlament, Dienste maschineller Übersetzung in Anspruch nehmen. So geben sich auch Humantranslatoren immer öfter (gezwungenermaßen) mit dieser Entwicklung zufrieden und arbeiten zudem auch vermehrt mit computerunterstützter Übersetzung (CAT-Tools). Weiters bieten viele Translation-Memory-Provider Programmierungsschnittstellen bzw. APIs (*application programming interface*) an, die mit MT-Systemen verknüpft werden. Es ist also davon auszugehen, dass man diese Tendenz der wachsenden MT-Nachfrage und des daraus resultierenden Angebotes auch in Zukunft beobachten wird (vgl. Vogel 2016: 14:01). Die immer größer werdende Bedeutung der Internetpräsenz wie z.B. über Social Media oder Online Shopping-Plattformen, führt in hohem Maße zu einem erhöhten Übersetzungsbedarf, wobei die Übertragung in diesem Bereich in erster Linie und nicht zuletzt auf Grund der enormen, täglich zu übersetzenden Textmengen, durch maschinelle Systeme erfolgt. So beispielsweise bei Google Translate, für das im Durchschnitt täglich bis zu 143 Milliarden Wörter in 100 Sprachkombinationen übersetzt werden müssen (Stand 2016) (vgl. Vashee 2016). Auch auf Facebook wird ein MT-System eingesetzt, das rund 40 Sprachen abdeckt und über zwei Milliarden Textsegmente pro Tag übersetzt. Vogel (2016) nimmt auch auf die Anzahl der Übersetzungen auf ebay Bezug, aus denen hervorgeht, dass das Unternehmen den

täglichen Einsatz von 200.000 ÜbersetzerInnen brauchen würde, um alle Texte ohne MT-System übertragen zu können. Daraus kann abgeleitet werden, dass es sich in den soeben genannten Fällen um keinerlei Aufträge handelt, die HumanübersetzerInnen übernehmen oder auch jeweils zeitnah realisieren könnten. Setzt man die Quantität in den Vordergrund, so resultiert, dass die Maschinen im Vergleich zum Menschen 100 Mal mehr in Bezug auf Webinhalte übersetzen. Abbildung 17 macht diesen beträchtlichen quantitativen Unterschied deutlich.

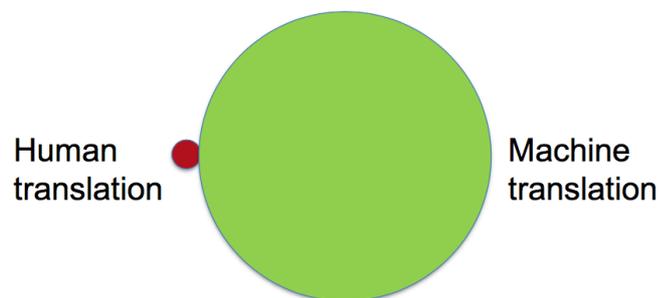


Abbildung 17: MÜ vs. HÜ (Vogel 2016: 16:03)

So wundert auch der massive Anstieg an MT-Startups seit dem Jahr 2005 nicht. Innerhalb von zehn Jahren wurden über 190 Unternehmen mit Spezialisierung im Bereich maschinelle Übersetzung bzw. Sprachtechnologie und Lokalisierung gegründet (vgl. Diño 2015). So unter anderem Mobile Technologies, der Entwickler von Jibbig, Safaba, das von Amazon gekauft wurde oder Apptek, das von ebay übernommen wurde.

3. Qualitätsmessung von Dolmetschleistungen

Um Dolmetschleistungen evaluieren und miteinander vergleichen zu können, ist es vorerst notwendig, zu bestimmen, was eine qualitative Dolmetschung ausmacht bzw. woran diese gemessen wird. Zu diesem Zwecke ist eine Auseinandersetzung mit dem Qualitätsbegriff unabdingbar.

3.1. Begriffsbestimmung Qualität

Es ist sehr schwer bis gar unmöglich, eine allgemein geltende Definition des Qualitätsbegriffes festzusetzen. Nicht zuletzt die hohe Anzahl an verschiedenen Definitionen stellt diese These unter Beweis. Die zahlreichen Qualitätsbegriffe sind nicht einzig auf die verschiedenen Disziplinen zurückzuführen, denen sie entspringen, sondern weisen auch in einzelnen Disziplinen, sich voneinander unterscheidende, manchmal gar kontroverse Definitionen auf. Dies trifft dementsprechend auch auf den translationswissenschaftlichen Ansatz zu. In vorliegender Arbeit soll die Qualität in Bezug auf Dolmetschleistungen im Vordergrund stehen. Um diese aber zu bestimmen, ist es vorerst notwendig, sich die Frage zu stellen, aus welcher Perspektive das Qualitätskonzept betrachtet werden soll und welche Parameter zu diesem Zwecke eingesetzt werden sollen. Die Ermittlung der Perspektive ist insofern von großer Bedeutung, als die Qualitätsbestimmung per se in erster Linie auf subjektiven Wahrnehmungen aufbaut bzw. von diesen stark geprägt wird. Eine möglichst allgemein geltende Definition bedarf folglich der Festlegung von Normen, Zielen, Standards und/oder Konsensformulierungen. Dementsprechend bestätigt die DIN EN ISO 9000, dass es bei der Qualität um die Erfüllung von Anforderungen geht, also um den „Grad, in dem ein Satz inhärenter Merkmale Anforderungen erfüllt.“ (zit. in Piechotta 2008:6). Im Unterschied zum Verständnis dieser Definition im wirtschaftlichen Bereich, sollten die Anforderungen in Bezug auf Übersetzungs- und Dolmetschleistungen nicht einzig oder hauptsächlich vom Kunden bestimmt werden dürfen. Nichtsdestotrotz kann Kurz' (2001:405) Ansatz ein zu berücksichtigendes Kriterium sein: „Quality is evaluated by customers in terms of what they actually receive in relation to what they expected. Clearly, the customer defines quality.“ Ein vorwiegend bei den KundInnen liegendes Recht zur Anforderungsbestimmung in Bezug auf die Qualität einer Dolmetschleistung würde jedoch in der Praxis die Möglichkeit von Manipulationen der Dolmetschleistung nicht ausschließen, was mit einem „Malinche-Dilemma“ vergleichbar wäre. Abgesehen davon stellt sich die Frage, inwieweit KundInnen die für die Festlegung von Qualitätsanforderungen jeweiliger Dolmetschleistung notwendigen Kompetenzen besitzen würden. Der Begriff „inhärent“ impliziert, dass einem Produkt bestimmte messbare Merkmale attribuiert werden können. Bei einem abstrakten Produkt wie einer Dolmetschleistung müssen diese Merkmale jedoch vorab bestimmt werden.

Das soeben Beschriebene soll keineswegs die Miteinbeziehung der KundInnenperspektive ausschließen, sondern vielmehr darauf aufmerksam machen, dass sich der Qualitätsbegriff aus translationswissenschaftlicher Sicht etwas komplexer gestaltet. Eine

Definition, die in diesem Zusammenhang den Nagel auf den Kopf trifft wird vom AIIC (2012) angeboten: „that elusive something which everyone recognizes, but no one can successfully define.“ So müssen neben der KundInnenperspektive auch die Wahrnehmungen seitens der anderen an der Kommunikationshandlung Beteiligten, berücksichtigt werden. Einen Vorschlag für die Konstellation in einer Dolmetschsituation wird von Pöchhacker (2002:98) vorgeschlagen (vgl. Abb. 18).

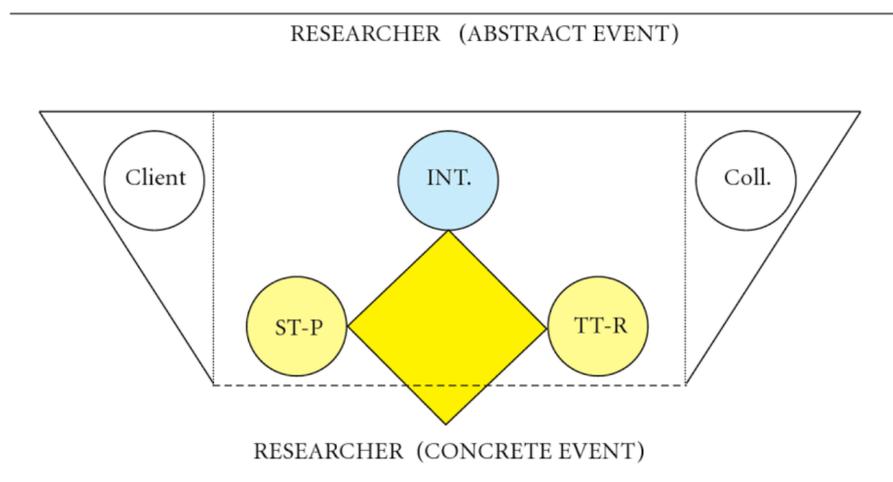


Abbildung 18: Qualität von Dolmetschleistung aus verschiedenen Perspektiven (Pöchhacker 2002:98)

Anhand Abbildung 18 lassen sich die drei an der Kommunikationssituation beteiligten Akteure erkennen. Neben der Perspektive der KundInnen und ZuhörerInnen, darf auch die Position der dolmetschenden Person nicht außer Acht gelassen werden. Letztere fungiert sowohl als TextproduzentIn bzw. SenderIn (*ST-P*, *sender/text producer*), als auch als TextempfängerIn (*TT-R*, *target text receiver*). Weiters zeigt sich auch ein viertes, nur indirekt am Kommunikationsakt beteiligtes Subjekt, das eine wissenschaftliche Perspektive einnimmt und vorrangig objektive Beobachtungen in Bezug auf die Leistung und/oder die direkt Beteiligten anstellt. Man kann sich folglich darüber einig sein, dass das Konzept der Qualität anhand verschiedener Variablen und aus verschiedenen Perspektiven betrachtend, bestimmt werden kann, worauf u.a. Kalina (2005:771) aufmerksam macht: „quality may mean many different things to different persons“.

Die abgebildete Konstellation hebt hervor, dass die Qualitätsbestimmung einer Dolmetschleistung in großem Ausmaß subjektiv ist, was wiederum bestätigt, wie schwer es

ist, „Qualität“ zu definieren. Abgesehen vom AIIC geht auch Pöchhacker (1994a:235) von dieser Annahme aus: „ We seem to know what the product should be like, but less sure about a method for establishing what a particular product is like in a given situation.“ Auch gemäß Grbič (2008:252) bleibt der Dolmetschqualitätsbegriff nicht fassbar: „there is no correct definition of (interpreting) quality, which is to say no one is (entirely) right or (entirely) wrong“. Riccardi (2005) fasst die mit einer Begriffsdefinierung zusammenhängende Herausforderung wie folgt zusammen:

„Für professionelle Dolmetscher war der Begriff „Qualität“ lange Zeit etwas implizites, etwas was man sofort erkennt, aber schwer mit Worten fassen kann; etwas, dass sich einer genauen Erklärung entzieht und zu dessen Definition sehr allgemeine Kriterien verwendet werden, wie etwa Klarheit, Texttreue, Kohäsion und Äquivalenz. Tatsächlich entzieht sich der Begriff Qualität beim Simultandolmetschen einer absoluten Definition, die immer und überall anwendbar ist. Qualität beim SD ist ein relativer Begriff und kann nur unter Berücksichtigung des jeweiligen Kommunikationsereignisses definiert und erklärt werden.“ (Riccardi 2005:99)

Es wurden also mehrere Versuche vorgenommen, genauere Vorstellungen vom Qualitätsbegriff zu übermitteln, wobei wie Riccardi feststellt, die zur Definitionsbestimmung herangeführten Variablen teilweise zu allgemein gehalten werden, im Sinne, dass sie nicht auf jede Situation zutreffen können. So definiert Gile (2003:110) Qualität als ein durch Parameter bestimmtes Konzept: „quality is either explicitly or implicitly seen as a weighted average of quality components, such as informational fidelity, linguistic correctness of the target speech, voice quality, prosodic quality, booth manners, out-of-the-booth behaviour, etc.“ Zweifelsohne stellen etwaige, vorab bestimmte Parameter einen unabdingbaren Aspekt der Qualitätsbestimmung dar und sollen demnach näher erläutert werden.

3.2. Qualitätsparameter

Ähnlich wie beim Definieren des Qualitätsbegriffes, weisen auch die Qualitätsparameter für Dolmetschleistungen mehrere Erwartungen und Anforderungen auf. In diesem Fall scheinen jedoch die in der Literatur vorgeschlagenen Parameter fassbarer zu sein, zumal sie sich wiederholen bzw. die meisten von ihnen immer wieder angeführt werden. Moser-Mercer (1994:44) bedient sich der von ihr festgelegten Parameter, um die „optimale Qualität“ (*optimum quality*) zu beschreiben. Sie führt an:

„Optimum quality in professional interpreting implies that an interpreter provides a complete and accurate rendition of the original that does not distort the original message and tries to capture any and all extralinguistic information that the speaker might have provided subject to the constraints imposed by

certain external conditions. [...] Optimum quality is the quality an interpreter can provide if external conditions are appropriate.”

Die für Moser-Mercer relevanten Qualitätsparameter sind demnach: Genauigkeit (*accurate rendition*) und Vermittlung der intendierten Nachricht, also Äquivalenz. Beide Voraussetzungen werden auch in Pöchhackers unten angeführtem Modell aufgegriffen, wobei dem Analysegegenstand *Dolmetschen*, zwei unterschiedliche Rollen attribuiert werden. Einerseits die Rolle einer Dienstleistung zur Kommunikationsermöglichung, andererseits eine Handlung zur Textproduktion bzw. die Dolmetschhandlung als Produkt. Dabei beziehen sich die Parameter sowohl auf semantisch-lexikalische Aspekte, als auch auf außersprachliche sozio-pragmatische (vgl. Pöchhacker 2002:97) Charakteristika (vgl. Abb.19).

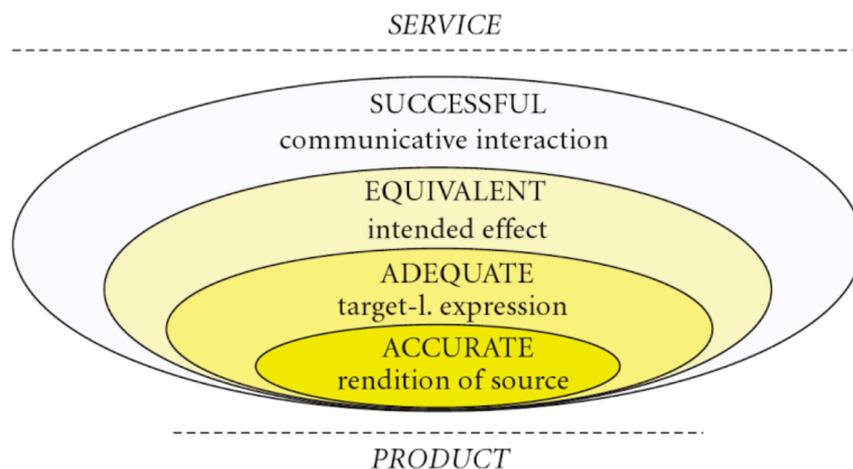


Abbildung 19: Qualitätsparameter des Dolmetschens als Produkt oder Dienstleistung (Pöchhacker 2002:97)

So soll die Dolmetschleistung je nach Kategorisierung als Produkt oder Dienstleistung, in jeweils anderer Gewichtung vorkommende Anforderungen erfüllen:

1. Der Ausgangstext (AT) muss sich durch Genauigkeit und Vollständigkeit auszeichnen. Dabei spielen auch außersprachliche Informationen eine große Rolle. (*accurate*)
2. Der Zieltext (ZT) soll auf sprachlicher und prosodischer Ebene in adäquater Weise wiedergegeben werden. (*adequate*)

3. Die intendierte Botschaft soll treu in die Zielsprache übertragen werden. Hier soll die vom Ausgangstext-Sender gewünschte Wirkung im Mittelpunkt stehen. (*equivalent*)
4. Das Ziel soll, unter Erfüllung der Parameter, die erfolgreiche Kommunikation sein. (*successful*)

Pöchhacker's Modell stellt natürlich nur einen unter vielen Vorschlägen zur Qualitätsbestimmung beim Dolmetschen dar. Diese sind wiederum Perspektiven abhängig, wodurch man wieder zur ursprünglichen Frage zurückgeführt wird: Ist eine allgemein gültige Festlegung von Qualitätsparametern und somit von Qualität beim Dolmetschen in Sicht? Eine Antwort auf diese Frage scheint nach wie vor negativ auszufallen bzw. nicht gegeben zu sein. Die Beurteilung von Dolmetschleistungen dürfte in zu hohem Ausmaß von folgenden Faktoren abhängig sein:

- Perspektive und Subjektivität: Welche Rolle ist der/dem Beurteilenden zuzuordnen? (DolmetscherIn, RezipientIn, KundIn, beobachtende WissenschaftlerIn)
- Kommunikationssituation: In welchem Setting und unter welchen Bedingungen findet die Dolmetschhandlung statt?
- Rolle des Analysegegenstandes (Dienstleistung oder Produkt)

Je nach Gewichtung der o.g. Faktoren werden auch etwaige Studiendurchführungen eine jeweils andere Gestaltung erfahren. Prinzipiell kann man verschiedene Zugänge erkennen. Zum einen ZuhörerInnenbefragungen oder -evaluierungen, zum anderen Textanalysen bzw. der Vergleich vom Ausgangstext mit dem Zieltext.

Jegliche oben angeführte, offene Fragestellungen deuten im Großen und Ganzen darauf hin, dass eine hohe Anzahl an Studien zu derselben Thematik verschiedenste, teils sogar widersprüchliche Ergebnisse mit sich bringen könnte. Letztere können jeweils auf vorab vorgelegten, sich je nach Perspektive oder Kommunikationssituation unterscheidenden Beurteilungskriterien basieren, deren Bestimmung wiederum die Beurteilenden beeinflussen können. So stellt auch Tiselius (2012) fest: „I would posit that features that are not rated as important (such as accent) affect how the user perceives important features (e.g. content or

correct terminology).“ Eine ähnliche Meinung vertreten auch andere Wissenschaftler (vgl. Collados & García 2015:378), die anführen: „In fact, the deficiencies of certain formal parameters not only affect the perception of the quality of an interpretation, but also may have a negative impact on the subjects’ understanding.“

Die Komplexität in Bezug auf die Qualitätsbestimmung des Dolmetschens bleibt somit weiterhin bestehen.

Welche Parameter werden in der Literatur nun als besonders wichtig erachtet? Bühler (1986) schlägt eine erste Dolmetschleistungsbeurteilung unter Berücksichtigung zweier Kategorien, nämlich der sprachlichen und der außersprachlichen vor²¹. Der sprachliche Aspekt (*linguistic (semantic) criteria*) umfasst dabei: Akzent, angenehme Stimme, Flüssigkeit der Ausgabe, logische Kohäsion der Ausgabe, Sinnübermittlung des Originaltextes, Vollständigkeit, grammatikalisch korrekte Wiedergabe, korrekte Terminologie und adäquaten Sprachstil. Außersprachliche Bewertungsfaktoren (*extralinguistic (pragmatic) criteria*) wurden wie folgt zusammengefasst: Vorbereitung der Konferenzdokumente, Ausdauer, sicheres Auftreten, angenehmes Auftreten, Verlässlichkeit, Teamarbeit-Kompetenz, positives Feedback von Auftraggebern.

Kurz (1989, 1993, 1996) griff später, mit Ausnahme des Kriteriums „adäquater Sprachstil“, die sprachlichen Parameter Bühlers in einer eigenen Studie auf. Ihr Ziel war die Ermittlung dessen, ob die Erwartungen der ZuhörerInnen je nach Dolmetschsituation variieren. Diese These konnte in großem Maß bestätigt werden, zumal die verschiedenen Gruppen „einige Kriterien unterschiedliche gewichteten“ (Fünfer 2013:33).

Pöchhacker (1994b) bezog, sich eher allgemeiner haltend, in einer späteren Studie fünf Parameter mit ein: Gesamteindruck, Qualität des sprachlichen Ausdrucks, Beherrschung der Fremdsprache, Stimmqualität und Akzent sowie Rhythmus und Intonation. Eine spätere ausführlichere Studie Pöchhackers in Zusammenarbeit mit Zwischenberger (2010), unter AIIC-Mitgliedern²² beinhaltete 11 Kriterien zur Beurteilung, wobei neun dieser Parameter Bühlers Studie entnommen und zwei, nämlich *Synchronität* und *lebhaftes Intonation*, hinzugefügt wurden. Das Ergebnis bestätigte größtenteils Bühlers Resultat, zumal festgestellt

²¹ Bühlers Studie „Linguistic semantic and extra-linguistic (pragmatic) criteria for the evaluation of the conference interpretation and interpreters“ (1986) bezieht sich, wie der Titel bereits andeutet, auf die Leistungsbeurteilung von KonferenzdolmetscherInnen und wurde unter 47 Mitgliedern (in der Rolle von ZuhörerInnen) des AIIC-Berufsverbandes durchgeführt. (vgl. Fünfer 2013:31)

²² Dieser webbasierte Fragebogen wurde von insgesamt 704 DolmetscherInnen ausgefüllt.

wurde: „Die sinngemäße Übereinstimmung mit dem Original und [der] [logische] Zusammenhang“ (Fünfer 2013:38) sind als die wichtigsten Parameter zu sehen.

Um noch einen letzten Ansatz zum Vergleich darzulegen, sollen auch die 2009 präsentierten Beurteilungskriterien angeführt werden, die im Rahmen einer Studie zum Thema „Forschungsmethoden zur Qualitätsbewertung von Dolmetschleistungen“ (Research Methods in Interpreting Quality Assessment) angewendet wurde. Diese Evaluierung wurde durch das ECIS-Team der Universität in Granada (Evaluación de la calidad en Interpretación Simultánea, Universidad de Granada) in Zusammenarbeit mit der Universität Wien durchgeführt. Dabei verfolgte man die Ziele einer Verbesserung in Bezug auf eingesetzte Methodologien sowie der Ermöglichung von Zusammenarbeit zwischen Universitäten, zur Umsetzung von Forschungsprojekten. Auch im Rahmen dieser Arbeit lassen sich im Groben zu manchen der oben bereits genannten, ähnliche Kriterien erkennen: Flüssigkeit der Dolmetschung, Akzent, Intonation, erste (Gesamt)eindrücke und Stimme. (vgl. Universität Wien & Universidad de Granada 2007)

Diese sehr allgemein gehaltene Zusammenfassung durchgeführter Studien zum Thema ZuhörerInnenbefragungen stellt die facettenreiche wissenschaftliche Auseinandersetzung mit der Thematik in den Vordergrund. Selbiges geschieht auch im Rahmen von Textanalysen bzw. AT- und ZT-Vergleichen. Dabei kommen wiederum einige Nebenfaktoren, wie das vorauszusetzende Vertrauen in die DolmetscherInnen oder die in Transkriptionen nicht berücksichtigte Intonation mit ins Spiel. Gleichzeitig können, besonders auf grammatikalischer, semantischer und lexikalischer Ebene, Grundzüge auf angemessene Art und Weise analysiert werden. Manche Evaluierungen, wie jene nach Falbo (vgl. Fünfer 2013:44), stellen den Zieltext in Bezug auf Kohäsion und Kohärenz in den Vordergrund, andere (vgl. Barik 1994) (Altmann 1994) konzentrieren sich wiederum mehr auf die Inhaltsübertragung. Ein Beispiel für eine Analyse stellt u.a. Riccardis (2005:108ff.) auf Viezzis Vorschlag aufbauende Ausarbeitung von Kriterien dar. Diese umfassen eine Analyse auf Makroebene:

1. Erste Makroebene: Art der Wiedergabe, die sich aus Akzent, Prosodie, Ausgabe, Phonetik, Verzögerungsphänomenen und Pausen zusammenstellt.

2. Zweite Makroebene: Beurteilung von morphosyntaktischen, lexikalischen und terminologischen Kriterien wie Allgemeinwortschatz, Terminologie, Internationalismen, Interferenzen, Morphosyntax und Syntax.
3. Dritte Makroebene: Inhaltliche Entsprechung unter Berücksichtigung von Ersatzinformationen, Auslassungen, Zusatzinformationen, Register und logische Verbindungen.
4. Vierte Makroebene: Dolmetschstrategien wie Reformulierungen, Antizipation, *décalage*, Technik und Gesamtleistung.

Riccardi setzt voraus, dass die evaluierende Person, den Ausgangstext verstehen kann, zumal nur dann beispielsweise eine Inhaltsanalyse möglich wäre. Dies zeigt, dass nicht jede/r ZuhörerIn solch ein Analyse durchführen könnte, was sich wiederum auf die Objektivität der Ergebnisse auswirken könnte. Zusammenfassend sollte es sich mit der Textanalyse und dem Textvergleich wie folgt verhalten: „Idealerweise sollte eine Fehleranalyse durch verschiedene Personen und anhand mehrerer Textstellen durchgeführt werden, um zu einem möglichst objektiven Ergebnis zu kommen“ (Fünfer 2013:47).

Nachdem die soeben gelieferte Zusammenfassung auf die Qualitätsbestimmung von Dolmetschleistungen humaner DolmetscherInnen Bezug nimmt, muss diese, um vorliegender Arbeit gerecht zu werden, um den Aspekt des maschinellen Dolmetschens erweitert werden. Die erste große Frage, die sich dabei ergibt, lautet: Können die für die Beurteilung des Humandolmetschens geltenden Qualitätsparameter eins zu eins für die Bewertung maschinell produzierter Zieltexte übernommen werden? Um hier eine Antwort zu liefern, muss zuerst eine Unterscheidung vorgenommen und festgestellt werden: Handelt es sich um eine Qualitätsbestimmung der Dolmetschung als Produkt – und dies möglicherweise in Rahmen eines Vergleiches mit der Leistung von Humandolmetschungen? Oder soll einzig die Qualität bzw. Kapazität der Maschine ermittelt werden? Beide Situationen bedürfen zweifelsohne unterschiedlicher Herangehensweisen und Methodologien. Im Falle eines Vergleiches zwischen Mensch und Maschine werden wohl, mit Ausnahme vereinzelter, auf die Human- oder die maschinelle Dolmetschung bezogener Zusatzfragen, die gleichen Parameter eingesetzt (vgl. Kapitel 4). Ist hingegen von einer Leistungsevaluierung eines maschinellen Dolmetschsystems die Rede, so werden die Bewertungskriterien wiederum anders bestimmt. In vorliegender Arbeit wurde bereits darauf Bezug genommen, woran die Qualität einer

maschinellen Übersetzung und Dolmetschung gemessen wird bzw. können jegliche Qualitätsparameter insbesondere aus Kapitel 2 abgeleitet werden. Aus diesem Grund soll an dieser Stelle auf eine wiederholte Auseinandersetzung mit dieser Thematik verzichtet werden.

3.3. Methodologien der Qualitätsmessung

Die Frage der Qualitätsmessung kann wieder mit der Definition laut der DIN-Norm EN ISO 9000 eröffnet werden. Qualität als „Grad, in dem ein Satz inhärenter Merkmale Anforderungen erfüllt“ (zit. in Piechotta 2008:6) bedarf allen voran, wie zuvor bereits angedeutet, der Festlegung inhärenter Merkmale. In einem zweiten Schritt ist es notwendig, Anforderungen zu definieren, die sowohl an die Merkmale, als auch an die Dolmetschleistungen zu stellen sind. Schließlich kann auch das Maß, in dem diese Anforderungen erfüllt werden, bestimmt werden. Dabei handelt es sich prinzipiell um einen leicht nachvollziehbaren Prozess. Doch wurde bereits festgestellt, dass die Qualitätsmessung von Dolmetschleistungen in der Praxis eine große Herausforderung darstellt, zumal es sich um eine Handlung mit mehreren AkteurInnen handelt, die sich durch verschiedene Perspektiven und Erwartungen auszeichnen. Wahrscheinlich haben sich im Laufe der Zeit, eben um diese Hürde möglichst gut zu bewältigen, verschiedene Methodologien herauskristallisiert. Man unterscheidet zwischen folgenden Ansätzen:

1. Befragungen
2. Experimente
3. Korpusbasierte Analysen
4. Fallstudien

Den bis dato beliebtesten Ansatz, der zur Qualitätsermittlung eingesetzt wurde, stellen Befragungen dar (vgl. Pöchhacker 2002:99). Diese können sich unter anderem an DolmetscherInnen oder an LaiInnen bzw. RezipientInnen (vgl. Fünfer 2011) richten. Hier geht es beispielsweise um die konkrete Beantwortung der Fragen danach, was gute DolmetscherInnen und gute Dolmetschungen ausmacht (vgl. Pöchhacker 2002:99).

Im Rahmen des experimentellen Ansatzes wird größeres Augenmerk auf die Leistung der Dolmetschenden gelegt. Dabei steht im Vordergrund, welche Ergebnisse der/die

Dolmetschende unter bestimmten Voraussetzungen erzielt. Dafür werden zum Beispiel Parameter wie Genauigkeit und Vollständigkeit, Synchronität, Informationsgehalt sowie Adäquatheit des Ausgangstextes bewertet.

Die korpusbasierte Methode ist zweifelsohne eine Methodologie, die sich sehr gut für die Ermittlung maschineller Dolmetschleistungen einsetzen lässt²³. Hierbei werden in großem Ausmaß Fehler berücksichtigt und womöglich in deren Gewichtung bewertet. Zwar bringt diese Methodologie, die weiter unten angeführte Limitationen mit sich, doch entzieht diese Tatsache, Pöchhacker's Annahme nicht die Gültigkeit.

“On the assumption that quality is a multidimensional socio-psychological as well as textual phenomenon within a specific institutional and situational context of interaction, the observational study of quality is arguably best served by an approach which allows the researcher to collect a maximum of information on a single case.” (Pöchhacker 2002:105)

All die bis jetzt angeführten Ansätze weisen die gleichen Limitationen vor. Die Ergebnisse solcher Qualitätsermittlungen werden immer nur eine beschränkte Anzahl an Kriterien bewerten können und somit kein größeres Bild auf multi-dimensionaler Ebene zeichnen.

Als letzter Ansatz soll die Fallstudie erwähnt werden. Diese kann mit Einsatz verschiedener Methoden erfolgen. Dabei können diese miteinander kombiniert werden. Dabei kann es zur Anwendung korpusbasierter Analysen sowie von Befragungen bzw. Umfragen oder Forschungsbeobachtungen seitens an der Fallstudie Beteiligten kommen.

Bezieht man sich einzig auf maschinelles Dolmetschen, so ergeben sich in Bezug auf die Qualitätsmessung hier und da zusätzliche Fragestellungen. Beim maschinellen Dolmetschen wird im Vergleich zu Humanleistungen, meistens bis immer davon ausgegangen werden, dass die Korrektheit des Ausgangstextes in Bezug auf Grammatik, Terminologie, Intonation usw. sehr kritisch zu betrachten sei. Je nach Anwendungsbereich eines Systems können solche Fehler in den Hintergrund gesetzt werden, sofern die kontextbezogene Übermittlung sinngemäß übertragen wird. Mit anderen Worten kann eine Übertragung trotz sprachlichem Qualitätsmangel, Sinn ergeben. Der jetzige Reifegrad hat, wie in Kapitel 2 bereits beleuchtet noch bei weitem nicht den „idealen“ Qualitätsgrad erreicht, doch können Maschinen wie der Lecture Translator oder der Skype Translator schon eine große Unterstützung in der Beseitigung sprachlicher Barrieren darstellen. Je innovativer und fortgeschrittener Sprachtechnologien sein werden, desto mehr wird man sich wahrscheinlich in Bezug auf die

²³ Der korpusbasierte Ansatz wird verhältnismäßig selten für die Bewertung von Humandolmetschleistungen eingesetzt. (vgl. Pöchhacker 2002:104)

Qualitätsmessung maschineller Dolmetschleistungen mit folgender Frage auseinandersetzen: Wie soll, wenn überhaupt möglich, die Schere zwischen kontextbezogenen sinngemäßen Dolmetschungen und sprachlicher Qualität reduziert werden? Und inwieweit ist es weiters wichtig, wie hoch die Qualität der Maschinendolmetschung ist, wenn der Sinn richtig wiedergegeben wird?

4. MD vs. Humandolmetschen

Anhand Abbildung 17 wurde bereits auf den vergleichsweise hohen Marktanteil maschineller Translation im Vergleich zur Humantranslation aufmerksam gemacht. Es soll an dieser Stelle jedoch erinnert werden, dass es sich bei der maschinellen Arbeit größtenteils um Übertragungen handelt, die üblicherweise nicht von HumandolmetscherInnen übernommen werden. Darunter fielen beispielsweise Übersetzungen von *Facebook*-Postings oder Webseitenübersetzungen wie auf *ebay*.

Ein weiterer Aspekt, der nicht unter den Teppich gekehrt werden darf, ist der Kostenfaktor. Interessanterweise kann sich dieser je nach Situation für einen der beiden, Mensch oder Maschine, ungünstig auswirken. Einerseits kommt es zu Situationen, in denen der Einsatz von HumandolmetscherInnen auf Dauer zu teuer ausfallen würde und diese somit durch die Maschine ersetzt werden. Andererseits müssen auch die hohen mit der Entwicklung der Sprachtechnologien verbundenen Kosten berücksichtigt werden, durch die der technologische Fortschritt auf diesem Gebiet oft verlangsamt wird. Aus diesem Grund wird je nach Einsatzgebiet weiterhin nur auf Humandolmetscher zurückgegriffen. Zweifelsohne ist der menschliche Einsatz nicht einzig kostenbedingt, sondern wird dieser hauptsächlich von der Qualität bestimmt. Ein beliebtes Beispiel der zur Hervorhebung der Bedeutung humaner DolmetscherInnen herbeigezogen wird, ist die Investition der Institutionen der Europäischen Union in sprachliche Dienstleistungen. So beträgt das für Dolmetschdienstleistungen in den EU-Institutionen verfügbare Budget insgesamt um die 1,3 Mrd. Euro/Jahr. Davon gibt allein das Europäische Parlament 300 Mio. Euro/Jahr aus (vgl. Waibel 2012: 7:00) während sich die Ausgaben des Rates der Europäischen Union auf 79,8 Mio. Euro/Jahr belaufen. Davon werden vom Rat weiters 34,5 Mio. für allgemeine Dolmetschleistungen (inkl. unvorhergesehene Ausgaben) und 45,3 Mio. „für sprachenbezogene Mittelzuweisungen für

Dolmetschleistungen „auf Anforderung““ [Hervorhebung im Original] (Rat der EU 2017:8f.) ausgegeben.

Kosten, Einsatzbereich und Qualität zählen sicherlich zu den Hauptfaktoren, die über die finale Entscheidung darüber, ob es der Mensch oder doch die Maschine wird, bestimmen. Die Tatsache, dass die Möglichkeit solch einer Entscheidung überhaupt besteht, ist bereits ein Indikator für den aktuellen Reifegrad maschineller Dolmetschsysteme. Wie der Entwicklungsstand im Vergleich aussieht, soll anhand folgender Beispiele beleuchtet werden.

4.1. Dolmetschleistungen im Vergleich

4.1.1. Lecture Translator vs. Humandolmetscher

Zumal in vorherigen Kapiteln bereits auf die technischen Komponenten und die Verarbeitungsprozesse des Lecture Translator (LT) eingegangen wurde, soll an dieser Stelle auf eine detaillierte Auseinandersetzung mit dieser Thematik verzichtet werden. Umso mehr soll stattdessen auf den Vergleich von Dolmetschleistungen am Beispiel der Studie nach Fünfer (2013) eingegangen werden. Um einem besseren Verständnis und genaueren Überblick gerecht zu werden, ist dem Anhang der zu Fünfers Studie dazugehörige Evaluierungsbogen²⁴ hinzugefügt.

Obwohl Fünfer (2013) mit ihrer Studie²⁵ zweifelsohne konkrete Ergebnisse in Bezug auf den Leistungsvergleich zwischen Mensch und Maschine beim Dolmetschen geliefert hat, gibt es einige Kritikpunkte, die nicht zuletzt in Hinblick auf zukünftige Studien zu dieser Thematik berücksichtigt werden müssen. Allen voran stellt sich das Problem des technischen Ausfalls. Trotz gründlicher Pretests der LT-Anlage kam es am Tag der Konferenz schon zu dessen Beginn zu Ausfällen des Systems. In Bezug auf die Studie lässt sich hier argumentieren, dass dies zu keinen Veränderungen der Ergebnisse führte, zumal Fünfer das Ziel verfolgte, die Leistung des LT und nicht dessen technische Zuverlässigkeit zu prüfen. Somit sollte eine qualitative und keine quantitative Bewertung geliefert werden. Trotzdem muss bedacht

²⁴ dieser besteht aus insgesamt fünf Teilen (A, B, C, D und E) und aus Fragen betreffend die untersuchten Dolmetschleistungen

²⁵ Das Experiment wurde 2009 am Fachbereich für Translations-, Sprach- und Kulturwissenschaft der Universität Mainz in Germersheim (FTSK) durchgeführt; die Ausgangsrede wurde vom LT-Projektleiter, Prof. Alexander Waibel gehalten

werden, dass es bei Studien wichtig ist, diese unter möglichst realitätsnahen Bedingungen durchzuführen. Hätte keine/r der Zuhörenden der maschinellen Dolmetschung die englische Ausgangssprache beherrscht, so wäre ihnen durch den Ausfall ein langer Teil des Vortrages vorenthalten geblieben. Dies hätte sich sehr stark auf das Verständnis und folglich auf die Inhaltsfragen im Bewertungs-Fragebogen auswirken können. Da dies jedoch nicht der Fall war und es zusätzlich eine Powerpoint-Präsentation zum Mitverfolgen gab, führte das technische Problem zu keinen größeren Veränderungen der Studienergebnisse. Ein weiterer technischer Aspekt war die Positionierung des Mikrophons, über das die akustischen Signale an den LT weitergeleitet wurden. Dieses war als Nahbesprechungsmikrophon nur für den Redner bestimmt. Jegliche Fragen aus dem Publikum wurden vom LT folglich nicht wahrgenommen und nicht gedolmetscht. Dazu hätte sich für das System die zusätzliche Herausforderung ergeben, dass die Publikumsfragen in verschiedenen Sprachen gestellt wurden. So hätte man die Ausgangs- und Zielsprache des Systems ständig aktualisieren müssen. Um dies zu verhindern, wurde der LT bei Beginn der Diskussionsrunde abgedreht.

Insgesamt wurde versucht, eine möglichst realitätsnahe Ausgangssituation zu schaffen, obwohl alle am Vortrag Teilnehmenden über die Versuchsdurchführung vorab informiert wurden. Inwiefern sich die Dolmetschenden auf den Einsatz vorbereitet hatten, kann nur schwer nachvollzogen werden. Fest steht, dass ihnen diese Möglichkeit durch Übersenden von Materialien und Folien geboten wurde. Bei den Humandolmetschenden handelte es sich um fortgeschrittene Studierende. Somit ist es zwar wahrscheinlich, bleibt aber trotzdem noch offen, ob und inwieweit sich die Leistungen erfahrenerer, bereits berufstätiger DolmetscherInnen auf die Ergebnisse der vergleichenden Analyse zwischen Ausgangs- und Zieltext ausgewirkt hätten. Auch die RezipientInnengruppe bestand hauptsächlich, nämlich zu 88%, aus (ehemaligen) Studierenden der Translationswissenschaft. Von den insgesamt 37 VersuchsteilnehmerInnen aus dem Publikum waren nur zwei externe Zuhörende. Einerseits müsste man diesen Aspekt wieder neu aufgreifen und bei einer ähnlichen Studie anpassen, da davon auszugehen ist, dass sich die Erwartungen an Dolmetschleistungen je nach RezipientInnengruppe unterscheiden (vgl. Kurz 1993, 1996). Andererseits handelt es sich beim LT um eine Maschine, die für einen Anwendungsbereich generiert wurde, in dem davon ausgegangen werden kann, dass die meisten Zuhörenden vom Fach sind und sich mit dem präsentierten Thema auskennen. Sieht man aber davon ab und legt man das Augenmerk nur auf die Leistungsprüfung des LT, so könnte oder sollte man eine ZuhörerInnenschaft mit anderem Hintergrund in Erwägung ziehen. Dabei spielt gleichzeitig die Vertrautheit der Zuhörenden mit dem Thema eine große Rolle. Je weniger verstanden

wird, desto schneller und in umso größerem Ausmaße erfolgt die Schwächung der Konzentrationsleistung und somit die Ermüdung.

Insgesamt wurden 36 von den 37 Fragebögen gewertet. Ein Fragebogen wurde als ungültig gezählt, da der/die TeilnehmerIn sowohl bei der humanen als auch der maschinellen Dolmetschung zugehört hatte. Folglich standen 17 Fragebögen der HD-Versuchsgruppe und 19 der MD-Gruppe für die Bewertung zur Verfügung.

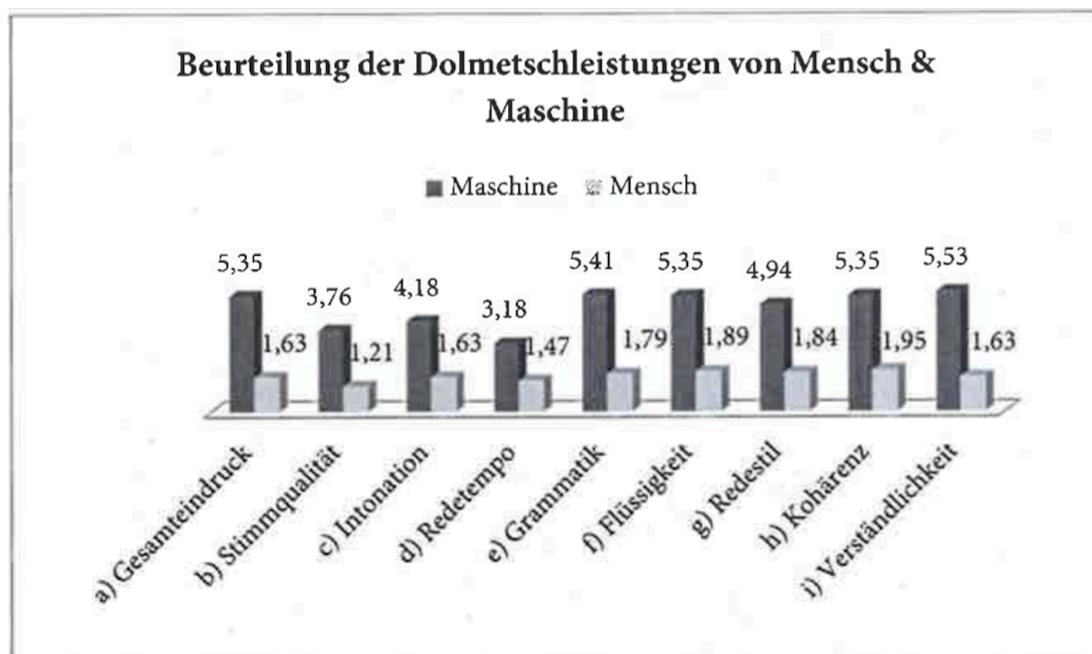


Abbildung 20: durchschnittliche Beurteilung der von Fünfer (2013) abgefragten Kriterien durch die ZuhörerInnen der MD und der HD auf einer Skala von 1 (positivster Wert) bis 6 (negativster Wert)

In Bezug auf den Inhalt der Fragebögen (vgl. Anhang) sind einige Anmerkungen hinzuzufügen – allen voran zu Teil A (allgemeine Fragen), in dem die VersuchsteilnehmerInnen gebeten wurden, ihre Sprachkombination anzugeben: Die meisten ZuhörerInnen waren StudentInnen der Translationswissenschaft an der Johannes Gutenberg Universität in Gernersheim (FTSK) und konnten hier eine klare Antwort liefern. Die oben bereits erwähnten externen TeilnehmerInnen mussten diese Frage auslassen. Es hätte sich hier wahrscheinlich eine offenere Fragestellung angeboten, wie beispielsweise: Beherrschen Sie die englische Sprache und wenn ja, dann auf welchem Niveau? Weiters ist allein durch die Tatsache selbst, dass, wie vermutet wird (Fünfer 2013:129), die meisten ZuhörerInnen Englisch verstanden, anzunehmen, dass dies Einfluss auf die Inhaltsfragen in Teil D des Fragebogens haben konnte – nicht zuletzt aufgrund des technischen Ausfalls, während dem

sich die MD-ZuhörerInnen einzig auf den englischen Originaltext des Vortragenden stützen konnten und sich außerdem umso mehr auf die Präsentationsfolien konzentrierten. Weiters ist betreffend Teil D hinzuzufügen, dass die Inhaltsfragen erst kurzfristig eingetragen und dem Redner eingereicht wurden. Aufgrund von Zeitmangel war es folglich nicht möglich, einen Pre-Test mit dem Vortragenden durchzuführen. Es bleibt folglich unklar, worauf falsche Antworten zurückzuführen sind. Genauso wie die Zuhörenden die DolmetscherInnen oder das MD-System nicht verstanden haben könnten, hätten sie Schwierigkeiten haben können, den Sinn, den der Redner mittels Vortrag und Folien übermitteln wollte, nachzuvollziehen.

Obwohl mehrere HD-Zieltexte für den Vergleich mit dem LT-Text zur Auswahl standen, wurde nur ein einziger Zieltext zu diesem Zwecke herangezogen. Ein Vergleich mit mehreren Humandolmetschungen hätte vermutlich detailliertere Ergebnisse liefern können. Für den segmentweisen Vergleich zwischen den Texten von Mensch einerseits und Maschine andererseits, mussten Fehlerkriterien bestimmt werden. Letztere waren jedoch nicht immer eindeutig abgrenzbar, da diese aufgrund der großen Fehleranzahl des LT nicht eins zu eins für die MD-Dolmetschung übernommen werden konnten. Umso mehr trifft dies auf inhaltliche Fehler zu. Hier kann die Dolmetschleistung der Maschine beispielsweise nicht auf Basis eingesetzter Dolmetschstrategien beurteilt werden, da der Einsatz von Strategien weit über die Kapazitäten eines MD-Systems hinausgeht.

Ungeachtet der Methodenkritik hat Fünfers (2013) Studie einen ersten Einblick in die Dolmetschleistung des Lecture Translator (Stand 2011) geliefert und einen Ausgangspunkt für zukünftige Studien und Versuche mit der LT-Anlage bereitgestellt. Für den direkten Vergleich und die Textanalyse(n) wurden außerdem Parameter entwickelt und eingesetzt, die zur Qualitätsbestimmung der Dolmetschleistungen im Vergleich verhalfen. Im Großen und Ganzen hat Fünfer einen aufschlussreichen Überblick über die Erfahrungen der VersuchsteilnehmerInnen sowie über die (vergleichenden) Analysemöglichkeiten und –methoden bei solch einer Studie geliefert.

Fünfer konzentriert sich in Ihrer Arbeit vor allem auf textanalytische und technische Aspekte, mit dem Ziel, einen Leistungsvergleich zwischen „Mensch und Maschine“ (Fünfer 2013) vorzunehmen. Sie liefert ein Bild darüber, was in zukünftigen, demselben Thema gewidmeten Versuchsdurchführungen besonders beachtet werden sollte. Auch eine Sensibilisierung auf unvorhersehbare Ereignisse wie technische Ausfälle muss stets vorhanden sein.

Aus textanalytischer Sicht lassen sich je nach Qualitätsparameterbestimmung zweifelsohne Unterschiede erkennen. So ist die MD-Maschine beispielsweise nicht in der Lage, Dolmetschstrategien anzuwenden und den Zieltext somit womöglich auch spannender zu gestalten. Ziel dieser Studie aber sollte die Leistung der Dolmetschenden und des MD-Systems sein – u.a. basierend auf dem inhaltlichen Verständnis seitens der ZuhörerInnenschaft. Im Großen und Ganzen kann, den technischen Ausfall nicht miteinbeziehend, festgestellt werden, dass der LT auf jeden Fall seinen Zweck, nämlich eine Verstehenshilfe für fremdsprachige Studierende zu liefern, erreicht, was auch Waibel (2012: 7:45) bestätigt: Die Maschine solle „nicht besser als Mensch, aber besser als nichts“ sein.

4.1.2. TC-Star-System vs. Humandolmetscher

Beim TC-Star-Projekt handelte es sich um ein von der Europäischen Kommission gefördertes Programm, das den Forschungsschwerpunkt auf die Hauptelemente (Spracherkennung, MÜ und Sprachsynthese) von maschinellen Translationssystemen setzte. Dabei lag der Fokus auf der multilingualen Speech-to-Speech-Komponente. Das Ziel des Projektes war, die Schere zwischen humaner und maschineller Dolmetschleistung stark zu reduzieren (vgl. FBK 2017) und ein möglichst domänenunabhängiges System für das Echtzeit-Simultandolmetschen zu entwickeln. Das Projekt wurde somit im Bereich von Spontansprache sowie beispielsweise bei Nachrichtensendungen oder Wettervorhersagen getestet. Die eingesetzten Sprachen waren Englisch, Spanisch und Chinesisch.

TC-Star kann als eine Art Vorstufe zum Lecture Translator gesehen werden, zumal letzterer auf aus dem TC-Star-Projekt gewonnenen Erkenntnissen aufbaut. So wurde das System, im Rahmen dieses EU-Bridge-Projektes zur Forschung im Bereich aller Hauptkomponenten eines MD-Systems (ASR, MÜ und T2S), unter anderem auch im EU-Parlament getestet. Die ambitionierte Forschungsfrage lautete dabei, ob die Maschine den Menschen beim Dolmetschen ersetzen könne. Die in diesem Setting und Einsatzbereich vorhersehbare Antwort fiel negativ aus. So können auch die DolmetscherInnen erleichtert sein und sich gleichzeitig aber auch umso mehr freuen, da das Experiment ergab, dass die Maschine im Stande sei, eine unterstützende Funktion für die SimultandolmetscherInnen einzunehmen. Gleichzeitig erkennt man einen erstaunlichen Fortschritt in Bezug auf die Inhaltsübermittlung (vgl. Abbildung 21). Weiters zeigt sich, dass die Schere im Bereich Inhaltsübermittlung nicht allzu weit auseinanderdriftet. Dies ist damit zu begründen, dass

HumandolmetscherInnen aufgrund der Schnelligkeit der Echtzeitsprache, nicht jede einzelne Information wiedergeben und sich nur auf die wichtigsten Inhalte konzentrieren können.

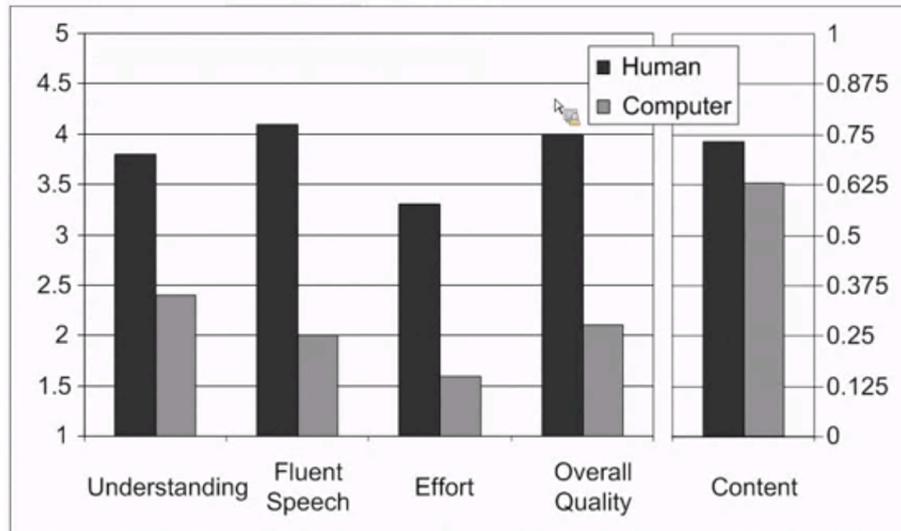


Abbildung 21: HD vs. MD im Rahmen des TC-Star-Projektes (Waibel 2011: 17:06)

Grundsätzlich erkannte man, dass die Maschine die Humandolmetscher des Europäischen Parlaments in dreierlei Hinsicht unterstützen könnte:

1. Terminologie
2. Namen, Benennungen
3. Abstimmungssitzungen

Dank der Terminologiekomponente bietet das System einerseits die Terminologiesuche, die auch als Terminologievorbereitung fungiert und andererseits eine Komponente zum Einsetzen der vorbereiteten Terminologie. Der vorbereitende Teil funktioniert dabei mittels automatischer webbasierter Suche.

Die zweite Komponente dient der Namens- oder auch Zahlenerkennung und –speicherung, was als eine Art Echtzeit-Translation-Memory zu sehen ist.

Besonders erfolgsversprechende Ergebnisse erreichte man beim Einsatz des Systems bei den oft sehr langen, stressigen und teilweise auch repetitiven Abstimmungssitzungen des

EU-Parlaments. Hier konnte das System insofern Hilfe leisten, als es Grußformeln, Namen, Zahlen und dergleichen problemlos übersetzen konnte, was für die Humandolmetschenden zu momentanen Erleichterungen in den Kabinen führt.

5. Berufsstand „DolmetscherIn“ in Österreich

Zwecks Ermittlung eines möglichst präzisen Überblickes des Berufsstandes Dolmetschender, sind einige Begriffserklärungen als unabdingbar zu betrachten. Um der Genauigkeit in möglichst hohem Maße zu Gute zu kommen, sind im Folgenden dargelegte Beobachtungen von sich entwickelten Berufsbildern, am Beispiel Österreichs angestellt worden.

Allen voran wird an dieser Stelle der Begriff „DolmetscherIn“ als Ausgangspunkt weiterer Beobachtungen betrachtet. Dabei soll nicht einzig auf translationswissenschaftliche Definitionen, sondern auch und vor allem auf Begriffsbeschreibungen aus anderen Quellen Bezug genommen werden. Dies soll eine tunlichst wahrheitsgetreue Wahrnehmungsdarstellung des Dolmetschbegriffes ermöglichen.

Eine der wohl meist zitierten Definitionen führt uns auf die in Kapitel 1 bereits dargelegte Ausführung Kades (1968:35), aus der resultiert, dass die Dolmetschenden jene seien, die einen „einmalig (in der Regel mündlich) dargebotenen [Text] der AS in einen nur bedingt kontrollierbaren und infolge Zeitmangels kaum korrigierbaren Text der ZS [übertragen].“ Dieser translationswissenschaftliche Ansatz bietet uns, abgesehen vom Bezug auf die beim Dolmetschen ausgeübte Tätigkeit auch Zusatzinformationen zum Zieltext („kaum korrigierbar“) sowie zu den Bedingungen, unter denen Dolmetschende arbeiten („Zeitmangel“). Eine weitere Definition, dessen Autorenschaft den (translations)wissenschaftlichen bzw. akademischen Kreisen zuzuschreiben ist, führt zum aus dem österreichischen Universitätsgesetz 2002 entsprungenen Curriculum für das Masterstudium *Dolmetschen*. Dolmetschende sind demnach als „Fachleute für die Kommunikation zumeist zwischen Angehörigen unterschiedlicher Kulturen und Sprachen“ zu sehen (Universität Wien 2007). Derselbe Text wird einem österreichischen Portal für Studium und Beruf (unikat 2005) in erweiterter bzw. detaillierterer Form übernommen. Dort heißt es:

„DolmetscherInnen (TranslatorInnen) sind Fachleute für die Kommunikation zumeist zwischen Angehörigen unterschiedlicher Kulturen und Sprachen, die Bewältigung fachübergreifender Problemstellungen, und sie übernehmen als ExpertInnen die Verantwortung für die von ihnen erbrachte Leistung.“

Im zuletzt Zitierten kommt folglich sowohl die Kulturkomponente, als auch ein pragmatikbezogener Aspekt und nicht zuletzt ein berufsethischer Bezug hinzu. Dies ist wahrscheinlich in hohem Maße auf die Zielgruppe des Textes zurückzuführen, die in Hinblick auf eine Studienrichtungs- oder Berufswahl, Informationen zu gewissen Berufsprofilen sucht. Bei folgender Definition aus einer Webseite der Wirtschaftskammer Österreich (2017) wird von linguistisch-pragmatischen oder berufsethischen Aspekten abgesehen. Dennoch wird die oft wenig wahrgenommene Abgrenzung des Dolmetschens vom Übersetzen spezifiziert, die sich in der mündlichen Translation äußert: „DolmetscherInnen übertragen gesprochene Texte mündlich und direkt in eine andere Sprache oder sie übersetzen schriftliche Texte direkt mündlich in eine andere Sprache.“ Die kontroverse Gegenüberstellung von mündlicher Übersetzung möge hier analog zum sich eingebürgerten Begriff der Stehgreifübersetzung (vgl. Sandrini 2011:1096) verstanden werden. Inwieweit man aus translationswissenschaftlicher Sicht von „mündlicher Übersetzung“ sprechen kann, ist im Allgemeinen in Frage zu stellen, zumal spätestens nach Kades (1968) Definitionsfestlegung eine strengere Differenzierung vorgenommen wurde und somit auch der Übersetzungsterminus in hohem Maße im Widerspruch mit der Mündlichkeit steht. Dieses Differenzierungsproblem kann noch in online-Wörterbüchern gefunden werden. So lautet beispielsweise der Eintrag zum Vokabel „Dolmetscher“ im Duden-Wörterbuch wie folgt: „jemand, der Äußerungen in einer fremden Sprache übersetzt (Berufsbezeichnung)“ (Duden 2017). Ähnlich verhält es sich mit dem Eintrag aus einer anderen online-Quelle: „Person, die gesprochenen Text mündlich übersetzt“. Hinzu kommt noch eine zweite Bedeutung: „elektronisches Gerät zur Übersetzung von Texten“ (Wortbedeutung 2017)

Bereits anhand dieser wenigen Beispiele lässt sich einiges in Bezug auf die Wahrnehmung des Dolmetschbegriffes ableiten. Allen voran kann, auf der Suche nach einem gemeinsamen Nenner oben genannter Definitionen und Kontexteinträge, festgestellt werden, dass das Bewusstsein, es handle sich beim Dolmetschen um die Übertragung eines Textes einer Sprache in eine andere, stets gegeben ist. Die bereits erwähnte Differenzierung zwischen dem dem Hyperonym der Translation untergeordneten Begriffen „Übersetzen“ und „Dolmetschen“, erscheint jedoch mangelhaft. Einen sehr wichtigen Aspekt stellt zweifelsohne die RezipientInnengruppe der jeweiligen Definition dar, zumal der Eintrag im Idealfall dem Skopos angepasst und somit je nachdem, umfassender und praxisbezogener oder eher oberflächlich sein wird. So ist beispielsweise der o.g. Eintrag aus dem Portal *unikat* besonders

genau, zumal davon ausgegangen wird, dass die Lesenden bestimmte Erwartungen in Bezug auf Informationen zum Beruf hegen.

Nachdem nun auf den Dolmetschbegriff eingegangen wurde, konnte ein allgemeines Bild zum „Definitionsbestand“ des Berufes „DolmetscherIn“ geliefert werden. Gleichzeitig lässt sich hervorheben, dass gröbere Unterschiede in der Definitionsbildung vorhanden sind. Diese Divergenzen könnten nicht zuletzt durch äußere Einflüsse bedingt sein. Als äußere Einflüsse sind hier Darstellungs- und Wahrnehmungsmuster gemeint, die einen Berufsstand bestimmen. Oben genannter Überblick vermittelt folglich noch bei weitem kein Verständnis über das Berufsbild Dolmetschender. Um dieses ermitteln zu können, ist es allen voran notwendig, sich die Frage zu stellen, wie ein solches Berufsbild entsteht und inwieweit letzteres Auswirkungen auf die Wahrnehmung des Berufsstandes hat. Die Antwort auf diese Frage kann anhand einer Analyse von sich gebildeten Berufsbildern ermittelt werden. Mit Ahmanns (2012:64) Worten ausgedrückt soll hier die „intrakulturelle Wahrnehmung eines interkulturell agierenden Berufsstandes“²⁶ nähergebracht werden. Auf weitere Problematiken in Zusammenhang mit der Entstehung von solchen Bildern soll in weiterer Folge eingegangen werden. Vorerst ist es unabdingbar, festzustellen, welche Phänomene als Basis für eine weitere Auseinandersetzung mit dieser Thematik gelten. Ist von einem *Bild* die Rede, so spricht man auch von einem *Image*. Letzterer Begriff kann sich in unserem Fall allerdings hyperonymisch zum Begriff *Bild* verhalten, zumal ein *Image* als „Gesamtheit der Aussagen zu einem Berufsstand“ (Ahmann 2012:81) zu sehen ist und sich aus mehreren Einzelaspekten, das heißt Bildern, zusammensetzen kann. Diese Annahme bestätigt auch der Eintrag aus dem Duden-Wörterbuch, in dem ein *Image* als „Vorstellung, Bild, das ein Einzelner oder eine Gruppe von einer anderen Einzelperson, Gruppe oder Sache hat; [idealisiertes] (sic!) Bild von jemandem, etwas in der öffentlichen Meinung“ beschrieben wird. Wovon letztere Definition jedoch absieht, ist die Berücksichtigung einer potentiellen Entstehung eines Autoimages, also eines Selbstbildes – ein Aspekt, der in vorliegender Arbeit an Bedeutung gewinnt. Bezugnehmend auf den Analysegegenstand *Berufsstand DolmetscherIn* muss folgende Unterscheidung vorgenommen werden:

1. Heteroimage

²⁶ nach Ahmann (2012) auf den Beruf der ÜbersetzerInnen bezogen; diese Aussage kann jedoch auch analog für Dolmetschende gelten

2. Auto- und Heteroimage
3. Autoimage

Unter die erste Kategorie lassen sich in erster Linie Presstexte einordnen. Beim Begriff des Heteroimages handelt es sich also vielmehr um eine Medienwirklichkeit, die in hohem Ausmaß fiktional und somit im besten Fall ein „begrenzter Ausschnitt der Realität“ (Ahmann 2012:71) ist. Die RezipientInnen stehen folglich vor der Herausforderung, den Darstellungen aus den Medien Vertrauen zu schenken, in denen nicht das wirklich ist, „was wirklich ist, sondern das, was die Medien als wichtig darstellen“ (Merten 2007³:253). Schlussfolgernd lässt sich behaupten, dass das Heteroimage von Dolmetschenden auf mediale Ereignisse stützt und von diesen bestimmt wird, zumal seitens der RezipientInnen kein Drang dazu besteht, den Wahrheitsgehalt zu verifizieren. Somit wird der „Mangel an autoptischer Erfahrung gar nicht als solcher empfunden“ (Ahmann 2012:77).

Dem gegenübergestellt ist das Selbstbild bzw. Autoimage zu erwähnen, das auf das Bild zurückzuführen ist, das Dolmetschende als Kollektiv von sich selbst haben und (öffentlich) verbreiten. Dieses ist beispielsweise von universitären Informationsquellen (Broschüren, Webseiten usw.), der Selbstvermarktung durch Berufsverbände oder von Erkenntnissen der noch relativ jungen Translationswissenschaft abzuleiten. Ein Autoimage kann sich zweifelsohne auch mit einem Heteroimage vermischen, wodurch eine Hybridform des Berufsbildes entsteht. Dies könnte zum Beispiel bei StudieneinsteigerInnen des Faches *Translationswissenschaft* bzw. *Dolmetschen* der Fall sein, die zumeist noch ein Fremdbild des Berufes besitzen, aber mit einem oft in Kontrast dazu stehenden Autoimage in Berührung kommen (vgl. Ahmann 2012:79f.).

Analog zu diesen Überlegungen ließe sich das Konzept der diskursiv-semantischen Grundfigur nach Busse (1997:17ff.) anführen. Zwar liegt Busses Fokus dabei auf dem Phänomen der Diskursanalyse, doch können seine Überlegungen am Beispiel des vorliegenden Analysegegenstandes übernommen werden. So kann der Entstehungsprozess von Auto- und Heteroimages mit Busses Dichotomie des „Eigenen und des Fremden“ verglichen werden. In der Praxis ließe sich diese Grundfigur zum Beispiel an Personalpronomina wie „wir“ oder „sie“ erkennen, die Busse (1997:21) als „Funktionswörter“ bezeichnet. Die Grundfigur des Eigenen und des Fremdes geht allerdings Hand in Hand mit der Frage nach dem Entstehungsprozess dieses Konzeptes. Ist hier von einer „anthropologischen Konstante oder [von] einem kulturspezifischen Phänomen“ (Busse

1997:22) die Rede? Ginge man von ersterem aus, so würde sich jegliche Auswirkung o.g. Grundfigur als selbstverständlich, gar unvermeidlich erweisen.

„Wichtig in diskursanalytischem Zusammenhang ist dabei, daß schon die Tatsache einer dichotomischen Gegenüberstellung von „Eigenem“ und „Anderem“ (durchaus auch schon im Sinne des „Fremden“) offenbar schon anthropologisch vorangelegt ist. Dies würde darauf hindeuten, daß es diskursiv grundsätzlich überhaupt nicht möglich ist, das „Eigene“ (also Ich, die eigene Identität, konkrete Merkmale, die man „sich“ zuschreibt usw.) ohne den Gegenpol des „Anderen“/„Fremden“ und ohne die Abgrenzung von ihm zu denken bzw. auszusprechen und umgekehrt. Beide Konzepte wären sozusagen „eingebaute Konnotationen“ des jeweilig anderen Konzepts.“ (Busse 1997:23)

Folglich kann analog dazu, in Bezug auf die Entstehung eines Heteroimages, behauptet werden, dass die durch die Medien vermittelte oben bereits genannte „fiktive“ Wirklichkeit eine unentrinnbare Realität sei, da durch die anthropologisch vorbestimmte Tendenz vorab bestimmt. Gleichzeitig darf auch eine erkenntnisbasierte Wahrnehmung und Imagebildung nicht ausgeschlossen werden, zumal beachtet werden muss, dass der „Anderer“, aus soziopsychologischer Sicht, keine konkrete Figur darstellt, sondern als Konsequenz bestimmter „lebenspraktischer“ (Busse 1997:23) Erfahrungen zu sehen ist.

Schließlich stellt sich jeweils die Frage, inwieweit ein gegebener Text bzw. eine gegebene Darstellung – in diesem Fall, des Berufsbildes Dolmetschender – ein intentionell produzierter Text sei oder inwieweit selbiger Text aus Sicht eines/r (unbewusst) voreingenommenen Produzenten/In verfasst wurde. Um jedoch den Rahmen der vorliegenden Arbeit nicht zu sprengen, wird von einer näheren Auseinandersetzung mit dieser Problematik abgesehen.

Es steht zweifelsohne fest, dass ein „Ich“ bzw. ein „Wir“ entsteht, das u.a. durch die jeweilige Sprache sowie durch die Kultur bzw. Mentalität geprägt ist, wobei eine eventuelle Verschmelzung beider Identitäten keineswegs auszuschließen ist. Dies setzt voraus, dass das „Ich zum Wir wird und das Wir das zu determinieren beginnt, was zuvor nur das Ich war“ (Busse 1997:24). Konkret bedeutet dies, dass analog dazu ein individuelles Autoimage mit einem kollektiven Autoimage übereinstimmen kann. In gegebenem Diskurs wäre dies beispielsweise mit dem Autoimage eines Dolmetschenden als Individuum vergleichbar, das mit dem kollektiven Autoimage eines Berufsverbandes Dolmetschender gleichgesetzt werden könnte. Bildlich kann die Dichotomie in beiden Ebenen kreuzweise wie folgt veranschaulicht werden:

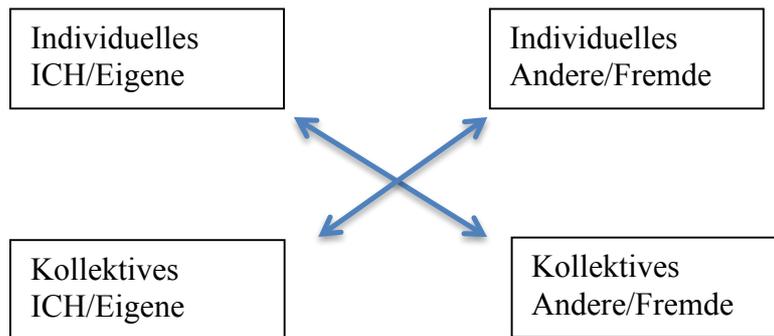


Abbildung 22: Beziehung 1 von Auto- und Heterobild

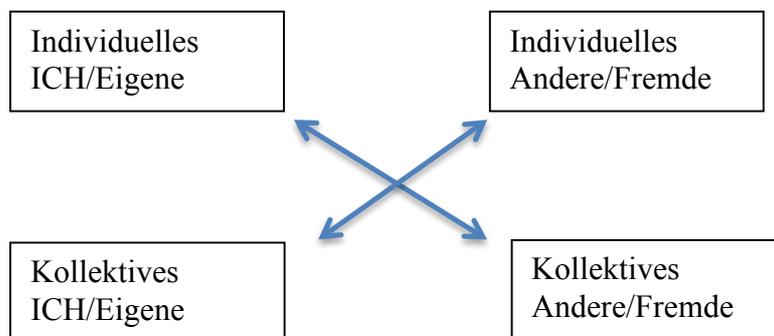


Abbildung 23: Beziehung 2 von Auto- und Heterobild

Weiters ist festzuhalten, dass das kollektive Autoimage wiederum das individuelle bzw. sowie das kollektive Heteroimage beeinflusst. Wichtig erscheint in diesem Zusammenhang auch, wie sich der Entstehungsprozess eines persönlichen Images gestaltet. Dabei ist das Zusammenspiel von Eigenem und Fremden beim Akt der Informationszufuhr nicht zu unterschätzen. So ist die Abhängigkeit eines Images nach Merten (2007:249) folgendermaßen zu veranschaulichen:

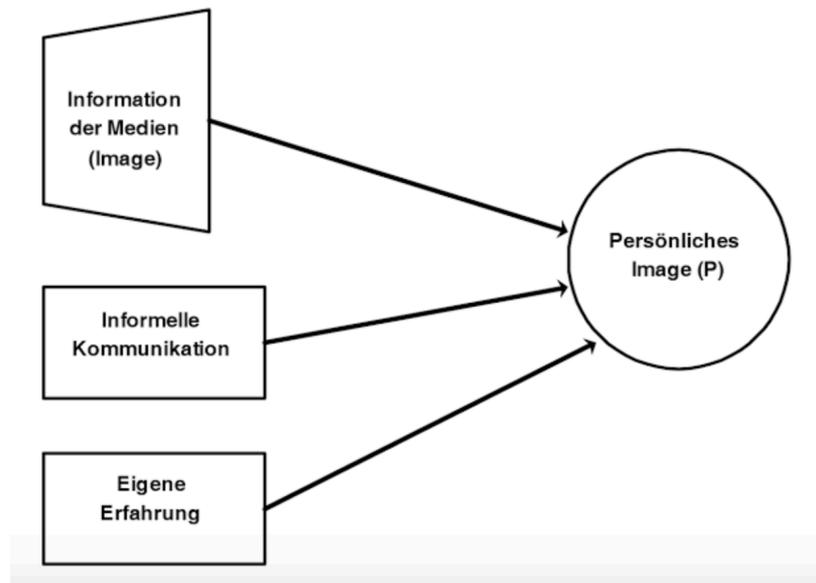


Abbildung 24: Abhängigkeit eines Image von eigener/fremder Information (Merten 2007:249)

Je nach Bestand und Art der eigenen Erfahrung wird das persönliche Image (P) in höherem oder geringerem Ausmaß von der Information der Medien geprägt. Gesetzt den Fall, dass RezipientInnen wenig bis gar keine Erfahrung zum Thema *Berufsstand Dolmetschende* besitzen, wird a) dieser Mangel an eigener Erfahrung ihrerseits nicht als solcher wahrgenommen und b) das mediale Image an Bedeutung und Einfluss gewinnen, zumal für die RezipientInnen angesichts der mangelnden eigenen Erfahrung kein Grund dazu besteht, den Wahrheitsgehalt zu verifizieren.

5.1. Exkurs: Darstellung in österreichischen (Online-)Presseartikeln

Um einen konkreten Bezug zu oben genannter Unterscheidung zwischen Hetero- und Autoimage sowie zu einer eventuellen Verschmelzung dieser beiden Aspekte herzustellen, wurde versucht, anhand von Online-Pressetexten zu ermitteln, wie sich die aktuelle Darstellung und Wahrnehmung des Berufsstandes *DolmetscherIn* aktuell gestaltet. Als Quellen dafür dienten insgesamt drei Online-Medien:

1. derstandard.at
2. diepresse.com

3. futurezone.at

In zwei Fällen, bezogen auf die Webseiten derstandard.at und diepresse.com, handelt es sich um Online-Medien mit einer gemäß ÖWA (Österreichische Webanalyse) (ÖWA 2017) durchschnittlich hohen LeserInnenschaft beziehungsweise monatlichen Besucheranzahl. Im Kontrast dazu steht die Online-Quelle futurezone.at, die zwar einen durchschnittlich sehr niedrigen Wert aufweist, jedoch über die Webseite kurier.at, die Eigentum der Telekurier Online Medien GmbH & Co KG ist, aufrufbar ist, die wiederum von einer sehr hohen Anzahl an LeserInnen besucht wird. Im Folgenden soll ein tabellarischer Auszug aus einer durch die ÖWA durchgeführten Medienanalyse, die jeweilige Webseiten-Besucheranzahl veranschaulichen. Zwar umfasst der gewählte Analysezeitraum, ab 1.Jänner 2015 bis zum Stichtag 16.Juli 2017 insgesamt neun Quartale, doch wurden Daten zu den Quartalen I. und III. des Jahres 2015 seitens ÖWA nicht angeboten. Folgende Durchschnittswerte beziehen sich folglich auf insgesamt 7 Quartale:

Quelle	Unique User Monat (Ø Jän. 2015-Juli 2017)
derstandard.at	2114
diepresse.com	1361
futurezone.com	325

Tabelle 3: Unique User/Monat (vgl. ÖWA 2017)

Ausschlaggebende Kriterien für die Quellenauswahl sind neben der vergleichsweise hohen LeserInnenschaft und dem oben genannten Zeitraum, die Zugehörigkeit der Medien zur österreichischen Presselandschaft. Als weiteres Kriterium wurde auch die RezipientInnengruppe berücksichtigt, die, wie auch aus einer Media Analyse-Studie für das Jahr 2015 (vgl. Der Standard 2015) resultierend, bei oben genannten Quellen, allen voran von AkademikerInnen gebildet wird. Dabei liegt die Reichweite unter AkademikerInnen bzw. Personen mit (Fach-)Hochschul- oder Universitätsabschluss bei 15,3% für die Tageszeitung *Der Standard* (Platz 1), bei 12,1% für den *Kurier* (Platz 3) und bei 10,9% bei der Zeitung *Die Presse* (Platz 4) (vgl. Der Standard 2015). Eine weitere Studie der Media Analyse aus dem Jahr 2016 liefert einen Überblick über die prozentuelle Gesamtreichweite österreichischer

Medien, darunter auch der drei für vorliegende Arbeit relevanten. So lesen gemäß o.g. Studie österreichweit 7,6% die Tageszeitung *Kurier*, 5,3% die Zeitung *Der Standard* und 4% *Die Presse*. Die höchste LeserInnenzahl betrifft jeweils folgende Altersgruppen: *Kurier* – ab 50 Jahren, *Der Standard* – zwischen 20 und 39 sowie zwischen 50 und 59 Jahren, *Die Presse* – ab 60 sowie zwischen 14 und 19 Jahren. Bei den Wochenendausgaben lassen sich jeweils höhere LeserInnenzahlen ermitteln.

Als Suchbegriff für die Presseartikelrecherche wurde auf den jeweiligen Webseiten der Terminus „Dolmetscher“ gewählt, der den in für gegebenen Analysegegenstand relevanten Preetexten oft vorkommenden Begriff „Dolmetsch“ miteinbezieht und in den weiblichen Formen „Dolmetscherin“ und „Dolmetscherinnen“ jeweils enthalten ist. Die Gesamtheit der Ergebnisse aus dem Online-Archiv wurde manuell unter Berücksichtigung von vorab festgesetztem Analysezeitraum (01.01.2015-16.07.2017) sowie in Bezug auf die jeweilige Themenrelevanz gefiltert. Letztere kann in folgende zwei Kategorien unterteilt werden: Zentralthema (ZT) und Randthema (RT). Zur Kategorie ZT zählen Artikel mit ausschließliehem und unmittelbarem Bezug auf das Berufsbild und den damit zusammenhängenden Aspekten Dolmetschender. Texte hingegen, die im Rahmen eines anderen Themas auf den Beruf DolmetscherIn Bezug nehmen, wurden als Randthemen behandelt. Eine Verschmelzung dieser beiden Themenarten ist nicht auszuschließen. So wurden einige Artikel als Zentralthema und Randthema gleichzeitig kategorisiert. Dies betrifft jene Texte, die sich im Rahmen einer bestimmten Thematik mit dem Gebiet des Berufes DolmetscherIn befassten. Als Beispiel lässt sich die Problematik in Bezug auf den Einsatz von LaiendolmetscherInnen im Rahmen der Flüchtlingskrise anführen. Weiters wurden die Artikel auch als Darstellung eines Hetero- oder eines Autoimages bzw. einer Verschmelzung dieser beiden eingestuft. Insgesamt ergaben die Suchen ein Gesamtergebnis²⁷ von 730 Artikeln, von denen 27 Artikel auf das Berufsbild Dolmetschender Bezug nahmen und für vorliegende Arbeit somit relevant sind. Dies macht einen Anteil von 3,7% aus. Ob dieser Prozentsatz ein vergleichsweise hoher oder niedriger in Bezug auf die Thematik von Berufsbilddarstellungen in der Presse ist, stellt eine Fragestellung dar, die den Rahmen dieser Arbeit sprengen würde. Folgende Tabelle soll oben erwähnte Themen- und Imagetypen im Vergleich veranschaulichen:

²⁷ bezieht sich auf alle drei o.g. Quellen

Titel (Jahr)	Quelle	Thema	Image
1 Dolmetscher ohne Töne (2017)	diepresse.com	ZT	Auto/Hetero
2 Die (Ohn)Macht der Trump-Übersetzer (2017)	diepresse.com	ZT	Auto/Hetero
3 Ostasien: Eintauchen in das Reich der Mitte (2016)	diepresse.com	RT	Auto/Hetero
4 Spielfeld: Nur Laiendolmetscher im Einsatz (2016)	diepresse.com	RT/ZT	Hetero/Auto
5 Londoner Übersetzungsbüro sucht Emoji Dolmetscher (2016)	futurezone.at	RT/ZT	Hetero/Auto
6 Vermitteln, was Donald Trump den Leuten sagen will (2017)	diepresse.com	ZT	Auto
7 Asyl: Mangel an Gerichtsdolmetschern (2016)	diepresse.com	ZT	Auto
8 Neues Masterprogramm für Übersetzer (2016)	diepresse.com	ZT	Auto
9 Alpbachs Dolmetscher: (Fast) 100 Prozent Frauen (2015)	diepresse.com	ZT	Auto
10 Videodolmetschen, wenn der Arzt kommt (2016)	diepresse.com	ZT	Hetero
11 Arbeitsmarkt: Gute Chancen vor allem im Ausland (2016)	diepresse.com	ZT	Hetero
12 Gebärdensprache: Ihre Avatare zeigen es allen (2015)	diepresse.com	ZT	Hetero
13 Wann Ärzte Dolmetscher brauchen (2015)	diepresse.com	ZT	Hetero
14 Aus stillem Elternhaus (2017)	diepresse.com	RT	Hetero
15 Diese 32 gemeinnützige Tätigkeiten dürfen Asylwerber ausüben (2016)	diepresse.com	RT	Hetero
16 Absolventenstudie: Am besten verdienen Pharmazie-Absolventen (2016)	diepresse.com	RT	Hetero
17 MINT-Studie als Karrieresprungbrett (2016)	diepresse.com	RT	Hetero
18 Tiroler SPÖ: Mehr Migranten im Staatsdienst "bedenklich" (2016)	diepresse.com	RT	Hetero
19 Wiener Polizeidolmetscherin unter Betrugsverdacht (2016)	diepresse.com	RT/ZT	Hetero
20 Wenn eine Elfjährige bei Gericht übersetzt (2016)	diepresse.com	RT/ZT	Hetero
21 Spielfeld: Volksanwaltschaft ortet "Fähigkeitsdefekt" (2016)	diepresse.com	RT/ZT	Hetero
22 Digitaler, impulsiver, individueller (2016)	diepresse.com	RT/ZT	Hetero
23 "Was hab´ich": Online-Portal übersetzt Arzt-Befunde (2015)	diepresse.com	RT/ZT	Hetero
24 Sprachbarriere erschwert medizinische Versorgung von Flüchtlingen (2016)	derstandard.at	ZT/RT	Hetero
25 Ear-Translator: Start Ups versprechen Echtzeitdolmetscher für unterwegs (2017)	futurezone.at	RT/ZT	Hetero
26 Skype startet Live Übersetzung in seiner Windows-App (2016)	futurezone.at	RT/ZT	Hetero
27 Salzburger Sozialamt setzt Videodolmetscher ein (2015)	futurezone.at	RT/ZT	Hetero

Tabelle 4: Gliederung nach Themen- und Imageart

Anhand oben angeführter Tabelle wird der Anteil an Zentral- und Randthemen ersichtlich, deren Zugehörigkeit nach oben bereits erwähnten Kategorien bewertet wurde. Am häufigsten lässt sich eine Verschmelzung beider Themenarten (RT/ZT) erkennen, die rund 41% ausmacht. Fast genauso dominant sind Zentralthemen (37%). Die anteilmäßig letzte Position wird schließlich von Randthemen (22%) eingenommen.

Nähere Aufmerksamkeit verdient an dieser Stelle die Auseinandersetzung mit den Imagearten sowie mit den dominierenden Thematiken.

5.1.1. Imagearten

Die Erkenntnisse in Bezug auf das aus oben geliefertem Überblick resultierende Image Dolmetschender, beantworten zum Teil die Frage, wovon das entstehende Berufsbild abhängig ist. Gleichzeitig bestätigen sie die Annahme einer Wirklichkeitsdarstellung durch die Medien und deren Einfluss auf die Bildung von Heteroimages, die nicht zwingendermaßen als realitätsgetreue Darstellungen zu sehen sind. Demnach ist die Medienwirklichkeit mit Ahmanns (2012:71) Worten ausgedrückt „im besten Falle ein begrenzter Ausschnitt der Realität“. Es darf an dieser Stelle auch nicht aus den Augen

gelassen werden, dass der Bezug zur Realität meist auch gesellschafts- und kulturbedingt ist. Foucault, der Vater dieses Ansatzes deutete bereits darauf hin, dass jede Gesellschaft einen jeweils anderen Zugang zur Wirklichkeit und zur damit verbundenen Glaubwürdigkeit beziehungsweise „Wahrheit“ habe. Somit gelte:

„Die Wahrheit ist von dieser Welt; in dieser wird sie aufgrund vielfältiger Zwänge produziert, verfügt sie über geregelte Machtwirkungen. Jede Gesellschaft hat ihre eigene Ordnung der Wahrheit, ihre ‚allgemeine Politik‘ der Wahrheit: d.h. sie akzeptiert bestimmte Diskurse, die sie als wahre Diskurse funktionieren lässt.“ (zit. in Keller 2011:139)

Foucaults Ansatz impliziert – auf vorliegende Arbeit bezogen – , dass die Wahrnehmung eines Berufsbildes, sowohl ein variabler Analysegegenstand innerhalb eines kulturellen Umfelds ist, als auch je nach gesellschaftlichem und kulturellem Umfeld, eine jeweils andere Form annehmen kann. Umso wichtiger erscheint hier der Bezug des Berufsbildes am Beispiel Österreichs. Die Ergebnisse der Berufsbildanalyse anhand der österreichischen Pressenlandschaft bestätigen die medien- und kommunikationswissenschaftliche Annahme, dass Medien ein Berufsbild in großem Ausmaß formen können. So bekräftigt auch Ahmann (2012:65), dass „die Darstellung des Berufs in der Presse, im Fernsehen und im Radio einen wesentlichen Beitrag zum Berufsbild leisten.“ Folglich überrascht nicht, dass die Darstellung des Berufsbildes als Heteroimage die Oberhand gewinnt (vgl. Tabelle 3) und rund 67% der Gesamtheit ausmacht. Ob dieses Heteroimage in den Presseartikeln bewusst oder unbewusst bzw. intentionelle oder nicht intentionell geformt wird, könnte separat ermittelt werden. Für vorliegende Arbeit scheint dieser Aspekt nicht so relevant zu sein, wie die Frage nach dem Wie. Welche wiederkehrende, zur Bildung eines Heteroimages beitragende Muster, lassen sich am Analysegegenstand erkennen? Folgende Eigenschaften der untersuchten Presseartikel sollen als Beispiele für solche Muster angeführt werden:

- Bezug zum Beruf Dolmetschender im Rahmen der Behandlung einer anderen Thematik; Dolmetschende treten in einer Nebenrolle (zB: als Zeugen) auf bzw. werden sie nur nebenbei erwähnt.
- Keine nähere Erläuterung der Bedeutung und Funktion Dolmetschender im Rahmen einer dargestellten Handlung oder Problematik bzw. Thematik (zB: Rolle Dolmetschender im Rahmen von Gerichtsverhandlungen)
- Mangel an Zitatführung Dolmetschender

- Auf Mikroebene: synonyme Verwendung der Begriffe „Übersetzen“ und „Dolmetschen“

Eine etwas andere Wirklichkeitsdarstellung ist zu beobachten, sobald das Heteroimage um eine oder mehrere Autoimage-Komponente/n ergänzt wird. Dies geschieht durch Bekräftigung eines behandelten Themas mittels Zitaten beziehungsweise Aussagen berufstätiger DolmetscherInnen. Zu solch einer Verschmelzung kommt es im Rahmen vorliegenden Analysegegenstandes in 18,5% der Fälle.

Wenig überraschend ist der mit rund 15% verhältnismäßig geringe Anteil an Artikeln mit Autoimage-Darstellungen. Besonders dominierend sind dabei direkte Einblicke in die Dolmetsch- oder Übersetzungstätigkeit sowie die Bewertung (in Zitaten) der Stellung dieser Berufe am österreichischen Arbeitsmarkt durch DolmetscherInnen. Dies lässt sich beispielsweise an Wörtern wie „unzumutbar“ (Die Presse 26/10/2016) in Bezug auf den Mangel an Gerichts- und Polizeidolmetschern erkennen. Ein weiteres Muster stellen die oben bereits erwähnten Grundfiguren nach Busse dar, die in zweierlei Form besonders häufig auftreten:

- Als Funktionswörter, an denen sich die Abgrenzung zwischen individuellen oder kollektivem Ich/Eigenem und kollektivem Anderen/Fremden erkennen lässt. Dazu gehören meist Pronomina wie „wir“, „uns“ usw.
- Als ein Exklusions- beziehungsweise Inklusionsmuster: Dabei entsteht ein Selbstbild i.S. einer Identifizierung bzw. Zugehörigkeitsdarstellung, die sich auf Mikroebene anhand einzelner Wörter oder Aussagen erkennen lässt (z.B: „You have no idea how impossible it is to translate Trump into something articulate“ (Die Presse 01/02/2017), „Kollegen“ usw.)

5.1.2. Dominierende Thematiken

Das Berufsbild Dolmetschender tritt im Rahmen vorliegender Analyse in Zusammenhang mit vier Thematiken auf, wobei Verschmelzungen der Themen beziehungsweise deren gleichzeitiges Auftreten in einem Artikel nicht auszuschließen sind. Angesichts des untersuchten Zeitfensters (2015-2017), überrascht nicht, dass die Thematik um den Beruf

Dolmetschender sehr häufig in Zusammenhang mit der Flüchtlingskrise und deren Folgen behandelt wird. Insgesamt macht dieses Thema einen prozentuellen Anteil von rund 30% aus und erscheint zumeist gemeinsam mit der in Verbindung mit dem Einsatz von LaiendolmetscherInnen stehenden Problematik der Unterbesetzung durch qualifizierte Dolmetscher, die als genauso häufig auftretendes Thema anzuführen ist. Ebenfalls in rund 30% aller untersuchten Artikel wird auf Dolmetschmodi Bezug genommen. Dies ist besonders interessant, da die Behandlung dieses Themas darauf hindeutet, dass eine Sensibilisierung der RezipientInnen mit Details zu berufsrelevanten Aspekten (auch zukünftig) nicht auszuschließen ist. In rund 26% der untersuchten Artikel bzw. an zweiter Stelle der Liste von meistbehandelten Themen in Zusammenhang mit dem Berufsbild, steht der Bezug zur Bezahlung von DolmetscherInnen in Österreich. Dabei wird die Gebührenfrage bewertend (zB: mit Bezug zum Einsatz von LaiendolmetscherInnen statt qualifizierter DolmetscherInnen aufgrund von dadurch niedrigerer Kosten) oder einzig als zusätzliche Informationskomponente dargestellt.

5.2. Bedeutung des MD für den Berufsstand „DolmetscherIn“

Man kann nicht leugnen, dass der technologische Fortschritt im Bereich der Computerlinguistik und Sprachtechnologien immer schneller vorangetrieben wird. Es lässt sich eine immer größer werdende Marktpräsenz von MT- und MD-Systemen sowie von Startups aus dieser Branche beobachten. So werden auch alle paar Jahre neue, innovative T2S- oder S2S-Systeme vorgestellt und von der Öffentlichkeit zumeist mit Freude willkommen geheißen. Als sehr aktuelles Beispiel gilt sicherlich der 2014 erschienene Skype Translator. Es besteht also kein Zweifel darin, dass das Angebot im Bereich maschineller Translation facettenreich und breit gefächert ist. Es umfasst nicht nur verschiedenste Einsatzbereiche, sondern auch die zwei grundlegenden Dolmetschmodi der Simultan- und Konsekutivdolmetschung. Weiters sind manche Systeme wie zum Beispiel JIBBIGO sowohl als online- als auch als offline-Lösung erhältlich.

Was bedeuten diese Entwicklungen nun aber für humane Dolmetschende? Inwieweit übt die Maschine einen Einfluss auf den Berufsstand „DolmetscherIn“ aus? Die Beantwortung dieser Fragen erfordert vorerst die Berücksichtigung der bisherigen Situation Dolmetschender am (Arbeits)Markt. Vor allem die Autoimage-Darstellungen seitens TranslatorInnen selbst, haben schon immer eine Unterbewertung oder ein in der Öffentlichkeit mangelndes Bewusstsein für diesen Beruf impliziert. DolmetscherInnen, vor

allem aber ÜbersetzerInnen wurden nicht selten als „unsichtbare“ MittlerInnen bezeichnet, wobei diese Beschreibung im Grunde nicht als Unterbewertung, sondern viel mehr als Anforderung und Voraussetzung für ein gutes Produkt gesehen werden sollte, zumal wie Venuti (1995) feststellt, eine (gute) Übersetzung, wie ein Original wahrgenommen werden sollte:

„A translated text, whether prose or poetry, fiction or nonfiction, is judged acceptable by most publisher, reviewers and reader when it reads fluently, when the absence of any linguistic or stylistic peculiarities makes it seem transparent, giving the appearance that it reflects the foreign writer's personality or intention or the essential meaning of the foreign text – the appearance, in other words, that the translation is not in fact a translation, but the “original.” (Venuti 1995:1)

Ähnliches trifft in der Praxis auch auf (Konferenz)DolmetscherInnen zu, zumal diese oft von der Kabine aus arbeiten und von den Zuhörenden nur gehört, nicht aber gesehen werden. Dieser Faktor der „Unsichtbarkeit“ ist sicherlich ein Aspekt, der sich auf das Berufsbild Dolmetschender ausgewirkt hat und in weiterer Folge Einfluss auf den Berufsstand hatte. So schien sich auch das Bewusstsein in Bezug auf die Bedeutung dieses Berufes lange Zeit in Grenzen zu halten, nach dem Motto: „Beherrscht man zwei Sprachen, so kann man schon dolmetschen und übersetzen.“ Diese Einstellung führt in der Praxis hier und da dazu, dass auf LaiendolmetscherInnen zurückgegriffen wird, was sich des Öfteren negativ auf die Qualität der Dolmetschung auswirkt. Es stellt sich folglich die Frage, ob dieses verschobene Verständnis des Berufes zu einem Ersatz des Menschen, genauso wie im Falle der LaiendolmetscherInnen, auch durch maschinelle Dolmetschsysteme stattfindet. Allen voran kann festgestellt werden, dass sich das auf den Dolmetschberuf geworfene Licht nicht zuletzt durch immer größer werdendes Engagement seitens verschiedener Berufsverbände sowie durch die vermehrte Medienpräsenz stark verändert hat und die Öffentlichkeit verhältnismäßig mehr Zugriffsmöglichkeiten zu Informationen zum Thema „Berufsstand Dolmetschender“ erhalten kann.

Das maschinelle Dolmetschen scheint dabei keine direkten ungünstigen Einflüsse auf den Berufsstand zu haben. Vielmehr herrscht die Tendenz seitens der Sprachtechnologie, Lücken zu füllen. So werden T2S- und S2S-Maschinen hauptsächlich dort eingesetzt, wo der Einsatz von HumandolmetscherInnen unrealistisch wäre. So zum Beispiel im Falle von Social-Media-Übersetzungen oder Echtzeitdolmetschungen zu privaten Zwecken wie beim Skype Translator. Natürlich können solche Systeme stets weiterentwickelt werden, doch ob sie dann auch tatsächlich die Humandolmetschungen ersetzen, ist vom jetzigen Standpunkt heraus noch zu bezweifeln.

Ein weiterer Aspekt, dessen Ermittlung in diesem Zusammenhang interessant wäre, ist die Bedeutung der persönlichen Präsenz: Inwieweit spielt für die Auftraggeber die persönliche Präsenz humaner DolmetscherInnen eine Rolle? Zwar soll diese Fragestellung in vorliegender Arbeit nicht weiter ausgeführt werden, doch ist davon auszugehen, dass viele (potentielle) Auftraggeber diesen Aspekt keineswegs unterschätzen.

Der aktuelle technische Stand beweist in vielerlei Hinsicht, dass MD-Systeme derzeit vermehrt den Trend verfolgen, neue Marktlücken zu füllen und Translationsbereiche zu übernehmen, für die ein beschränkter Reifegrad ausreichend ist und in denen kein Ersatz humaner DolmetscherInnen stattfindet.

5.2.1. Voraussichtliche Erwartungen und Entwicklung

Der in Kapitel 5.2 angeführte Entwicklungsstand scheint vom jetzigen Standpunkt heraus auch für die Zukunft zu gelten. Am Beispiel der innovativsten Systeme kann gemäß Waibel bestätigt werden, dass die MD-Maschinen nicht perfekt sind, aber dort Hoffnung bringen, wo es keine andere Möglichkeit zu geben scheint, Sprachbarrieren zu durchbrechen (Waibel 2012: 31:30). Waibel hebt auch hervor, dass die Systeme keineswegs für den Ersatz humaner DolmetscherInnen gedacht sind (Waibel 2012: 7:40). Vielmehr sollen MD-Systeme eine unterstützende Funktion einnehmen sowie Bereiche abdecken, für die keine HumandolmetscherInnen eingesetzt werden (können). Selbige Tendenz bestätigt auch Vogel (2016: 14:02) „MT is not so much replacing human translation, but covering before uncovered territory.“ Die im Folgenden aufgelisteten Faktoren, sollen zusammenfassend hervorheben, wodurch diese Unmöglichkeit des (vollkommenen) Ersatzes des Menschen durch die MD-Maschine unrealistisch und voraussichtliche nicht gegeben ist:

- **Qualität:** MD-Systeme besitzen noch bei weitem nicht den notwendigen Reifegrad. Auf diesem Bereich bestehen noch zu viele Herausforderungen und Limitationen.
- **Einsatzbereich:** MD-Systeme werden vermehrt in Bereichen eingesetzt, die Humandolmetschende nicht abdecken. Ein aktuelles Beispiel wäre hier das BULB-Projekt, im Rahmen dessen eine Technologie für einzig in Wort lebende Sprachen entwickelt werden soll (BULB-Partners 2017).

- **Physische Präsenz von HumandolmetscherInnen:** inwieweit ist dies in einem gegebenen Setting und für die jeweiligen AuftraggeberInnen relevant?
- **Kostenfaktor:** Die Entwicklung eines MD-Systems ist mit sehr hohen Kosten verbunden (vgl. Kapitel 2.3.2).

Voraussichtlich werden MD-Systeme also als CAT-Tools für Dolmetschende oder Verstehenshilfen²⁸ für RezipientInnen eingesetzt werden. Zusammenfassend kann mit Waibels (2015a: 7:50) Worten auch auf die Zukunft bezogen vorerst vorausgesehen werden: MD-Systeme sind „sicher nicht besser als der Mensch, aber besser als nichts.“

6. Zusammenfassung

In vorliegender Arbeit wurde versucht zu ermitteln, welche zukünftigen Entwicklungen im Bereich des MD und der MÜ voraussichtlich zu erwarten sind. Weiters stand auch die Frage im Mittelpunkt, ob sich der technologische Fortschritt auf den Beruf des/der DolmetscherIn auswirken wird und wenn ja, dann inwieweit.

Es kann zweifelsohne gesagt werden, dass der Fortschritt in der Sprachtechnologie und Computerlinguistik immer rapider umgesetzt wird. Gleichzeitig lassen sich seit Beginn der MD/MÜ-Entwicklung bis heute auch kontinuierliche, bis dato unüberwindbare Hürden beobachten, die es den MD-Systemen nicht erlauben, einen „humanen“ Reifegrad zu erreichen. Trotzdem haben viele Systeme bereits bewiesen, dass sie (im Rahmen der jeweils vorab festgelegten Ziele) erfolgreich umgesetzt werden können. Aktuellste Beispiele dafür sind der Lecture Translator und der Skype Translator. Beide Systeme erfüllen bereits große Anforderungen wie das Simultandolmetschen in Echtzeit. Doch nicht zuletzt deren noch zu hohe Fehlerraten lassen darauf schließen, dass MD-Systeme vorerst keine Bedrohung für HumandolmetscherInnen darstellen werden. Es gibt sehr wohl aber einen erfolgsversprechenden Bereich, in dem die MD-Maschine dem Menschen sogar um einen Schritt voraus ist: Die quantitative Kapazität. MD-Systeme können innerhalb kürzester Zeit viel mehr Daten als ein Mensch speichern. Dies stellt vor allem beim simultanen Modus des

²⁸ Dabei ist die Verstehenshilfe auch tatsächlich nur als solche zu sehen. Das Durchbrechen von Sprachbarrieren (wie z.B. beim LT) funktioniert dabei in Form einer Hilfe, die den Zuhörenden die Möglichkeit einer Mitverfolgung des kontextuellen Gerüsts, nicht aber ein einwandfreies, vollständiges Verständnis des Ausgangstextes ermöglicht.

Dolmetschens einen Vorteil dar, zumal man hier vor der Herausforderung steht, möglichst alle Informationen des Ausgangstextes in Echtzeit wiederzugeben. Experimente wie jene im Rahmen des TC-Projektes haben gezeigt, dass die Schere zwischen Mensch und Maschine in Bezug auf die Vermittlung auf inhaltlicher Ebene verhältnismäßig gering ist.

Es kann aber nicht geleugnet werden, dass MD-Entwickler noch vor einer langen Liste an technischen Herausforderungen stehen. Vor allem auf morphologisch-syntaktischer Ebene trifft man immer wieder auf dieselben Probleme.

Fazit ist, dass MD-Maschinen immer öfter eine unterstützende Funktion für den Menschen einnehmen werden. Sowohl für den/die Auftragnehmerin, als auch für den/die Humandolmetscherin (z.B.: CAT-Tools). Dadurch können in vielen Bereichen, wie zum Beispiel auf Universitäten (Lecture Translator) Sprachbarrieren beseitigt und zumindest kontextuelle Verstehenshilfen geschaffen werden. Weiters zeigen die bisherigen Entwicklungen, dass MT-Systeme hauptsächlich Nischenbereiche des Sprachdienstleistungsmarktes einnehmen (z.B. Übersetzungen von Facebook-Postings, auf ebay usw.). Eine direkte Bedrohung des Berufsstandes Dolmetschender ist derzeit somit nicht in Sicht. Hinzu kommt noch der „menschliche“ Aspekt, der vor allem beim Konsektivdolmetschen durch die physische Präsenz des/der Humandolmetschers/In zu tragen kommt. Inwieweit wären Auftraggeber – gesetzt den Fall, es gäbe qualitativ ausreichend entwickelte MD-Systeme – bereit, den humanen Aspekt mit künstlicher Intelligenz zu ersetzen?

7. Bibliographie

Ahmann, Heiko (2012) *Das Trügerische am Berufsbild des Übersetzers*. Berlin: Logos Verlag Berlin GmbH

AIIC. 2012. Practical guide for professional interpreters. In: <https://aiic.net/page/628/practical-guide-for-professional-conference-interpreters/lang/1>, Stand: 15.07.2017

Altman, Janet. 1994. Error Analysis in the Teaching of Simultaneous Interpreting: A Pilot Study. In: Lambert, Sylvie & Moser-Mercer, Barbara (Hg.), 25-38

Barik, Henri C. 1994. A Descriptive of Various Types of Omissions, Additions and Errors of Translation Encountered in Simultaneous Interpretation. In: Lambert, Sylvie & Moser-Mercer, 121-137

Becher, Johann Joachim. 1661. *Character, pro notitia linguarum universali*. Frankfurt am Main: Johannes Wilh. Ammon & Wilhelmi Serlin

Böke, Karin (Hg.) *Die Sprache des Migrationsdiskurses. Das Reden über „Ausländer“ in Medien, Politik und Alltag*. Opladen: Westdeutscher Verlag, 17-35

Bub, Thomas & Wahlster, Wolfgang & Waibel, Alex. 1997. Verbmobil: The combination of deep and shallow processing for spontaneous speech translation. In: http://isl.anthropomatik.kit.edu/cmu-kit/downloads/Verbmobil_the_combination_of_deep_and_shallow_processing_for_spontaneous_speech_translation.pdf, Stand: 12.01.2017

BULB-Partners. 2017. Breaking the Unwritten Language Barrier (BULB). In: <http://www.bulb-project.org/>, Stand: 14:08.2017

BusinessWire. 2012. Jibbigo führt kostenlos Übersetzer-App mit mehr als 20 Sprachen und Offline-Funktionalität ein. In: <http://www.businesswire.com/news/home/20121018005975/de/>, Stand: 17.07.2017

Busse, Dietrich. 1997. Das Eigene und das Fremde – Annotationen zu Funktion und Wirkung einer diskurssemantischen Grundfigur. In: Jung, Matthias & Wengeler, Martin & Böke, Karin (Hg.). *Die Sprache des Migrationsdiskurses. Das Reden über „Ausländer“ in Medien, Politik und Alltag*. Opladen: Westdeutscher Verlag, 17-35

Bühler, Hildegard. 1986. Linguistic (Semantic) and Extralinguistic (Pragmatic) criteria for the Evaluation of Conference Interpretation and Interpreters. *Multilingua – Journal of Cross-Cultural Interlanguage Communication* 5:4, 231-235

Carbonell, Jamie. 2016. Massively Multilingual Language Technologies. [20:34]. In: http://videlectures.net/interACT2016_carbonell_multilingual_technologies/, Stand 14.05.2017

Carstensen, Kai-Uwe & Ebert, Christian & Endriss, Cornelia & Jekat, Susanne & Klabunde Ralf & Langer, Hagen (Hg.). 2001. *Computerlinguistik und Sprachtechnologie – Eine Einführung*. Berlin/Heidelberg: Spektrum Akademischer Verlag GmbH

Chong, Theresa. 2014. How social media teaches skype to speak. In: <http://ieeexplore.ieee.org/document/6882973/#full-text-section>, Stand: 19.07.2017

Cohen, David. 2013. Not Lost in Translation: Facebook Acquires Jibbig Developer Mobile Technologies. In: <http://www.adweek.com/digital/jibbig-mobile-technologies/>, Stand: 29.07.2017

Collados Aís, Ángela & García Becerra, Olalla. 2015. Quality. In: Mikkelson, Holly & Jourdenais, Renée (Hg.) *The Routledge Handbook of Interpreting*. London/New York: Routledge Taylor & Francis Group, 368-384

Der Standard. 16/10/2015. STANDARD bei Akademikern meistgelesene Zeitung. VerfasserIn: (prie). In: <http://derstandard.at/2000023906505/STANDARD-bei-Akademikern-meistgelesene-Zeitung>, Stand: 23.07.17

Die Presse. 2016. 26/10/2016. Asyl: Mangel an Gerichtsdolmetschern. Verfasser: Martin Stuhlpfarrer. In: http://diepresse.com/home/panorama/oesterreich/5108355/Asyl_Mangel-an-Gerichtsdolmetschern?from=suche.intern.portal, Stand: 24.06.2017

Die Presse. 01/02/2017. Die (Ohn)Macht der Trump-Übersetzer. Verfasserin: Anne-Chatherine Simon. In: http://diepresse.com/home/kultur/literatur/5163316/Feuilleton_Die-OhnMacht-der-TrumpUebersetzer?from=suche.intern.portal , Stand: 24.06.2017

Diño, Gino. 2015. Disruption In Progress: 10 Years and 190 Translation Startups. In: <https://slator.com/features/disruption-in-progress-10-years-and-190-translation-startups/>,

Duden Online-Wörterbuch. 2017. Bibliographisches Institut GmbH, Dudenverlag (Hg.) In: <http://www.duden.de/rechtschreibung/Dolmetscher>, Stand: 20.06.2017

EU-Bridge. 2015. EU-BRIDGE. In: <http://project.eu-bridge.eu/>, Stand: 20.07.2017

FBK Fondazione Bruno Kessler. 2017. TC-Star (Technology and Corpora for Speech to Speech Translation). In: <https://hlt-mt.fbk.eu/projects/tc-star-technology-and-corpora-speech-speech-translation> , Stand: 13:07.2017

Fügen, Christian. 2008. *A System for Simultaneous Translation of Lectures and Speeches*. Universität Fridericiana zu Karlsruhe (TH): Dissertation

Fünfer, Sarah. 2013. *Mensch oder Maschine? Dolmetscher und maschinelles Dolmetschsystem im Vergleich*. Berlin: Frank & Timme

Gardt, Andreas. 1999. *Geschichte der Sprachwissenschaft in Deutschland – Vom Mittelalter bis ins 20. Jahrhundert*. Berlin/New York: Walter de Gruyter GmbH & Co. KG

Gile, Daniel. 2003. Quality Assessment in Conference Interpreting. In: Collados Aís, Angela (Hg.) *La Evaluación de la Calidad en Interpretación*. Granada: Comares, 109-123

Grbič, Nadja. 2008. Constructing Interpreting Quality. *Interpreting* 10:2, 232-257

Gundacker, Roman. 2016. E-Mail Korrespondenz am 11.10.2016 zum Thema altägyptischer Darstellungen Dolmetschender

Hutchins, John. 2004. Two precursors of machine translation: Artrouni and Trojanskij. In: <http://www.hutchinsweb.me.uk/IJT-2004.pdf>, 07.02.2017

Hutchins. 2015. Machine translation: history and research of application. In: Sin Wai, Chan. *The Routledge Encyclopedia of Translation Technology*. New York: Routledge, (120 ff.)

Kade, Otto. 1968. *Zufall und Gesetzmäßigkeiten in der Übersetzung*. Leipzig: Verl. Enzyklopädie

Kalina, Sylvia. 2005. Quality Assurance for Interpreting Processes. *Meta: Translators' Journal* 50:2, 768-784

Kannenberg, Axel. 2016. Übersetzer: Skype Translator nun in Desktop-App integriert. In: <https://www.heise.de/newsticker/meldung/uebersetzer-skype-translator-nun-in-desktop-app-integriert-3070629.html>, Stand: 20.07.2017

Keller, Reiner. 2011³. *Wissenssoziologische Diskursanalyse: Grundlegung eines Forschungsprogramms*. Wiesbaden: VS Verlag für Sozialwissenschaften

KIT. 2015. Lecture Translator. In: <http://isl.anthropomatik.kit.edu/english/1520.php>, Stand: 20.07.2017

Kuhn, Thomas. 1976². *Die Struktur wissenschaftlicher Revolutionen*. Frankfurt: Suhrkamp

Kurz, Ingrid. 1989. Conference interpreting: User Expectations. In: Hammond, Deanna (Hg.) *Coming Of Age. Proceedings of the 30th Annual Conference of the ATA*. Medford, NJ: Learned Information Inc., 143-148

Kurz, Ingrid. 1993. Conference Interpretation: Expectations of Different User Groups. *The Interpreters' Newsletter* 5, 13-21

Kurz, Ingrid. 1996. *Simultandolmetschen als Gegenstand der interdisziplinären Forschung*. Wien: WUV-Universitätsverlag

Kurz, Ingrid. 2001. Conference interpreting: quality in the ears of the user. *Meta: Translators' Journal* 46:2, 349-409

Küffer, Alexandra. 2017. Zeittafel des Alten Ägypten. In: <http://www.e-coffins.ch/index.php/zeittafel-des-alten-aegypten>, Stand 13.03.2017

Lambert, Sylvie & Moser-Mercer, Barbara. 1994. *Bridging the gap – Empirical research in simultaneous interpretation*. Amsterdam/Philadelphia: John Benjamins Publishing Co.

Mauthner, Fritz. 1923. *Wörterbuch der Philosophie – Universalsprache*. In: <http://www.zeno.org/Mauthner-1923/A/Universalsprache>, 07.02.2017

Merten, Klaus. 2007³. *Einführung in die Kommunikationswissenschaft*. Berlin: LIT Verlag

Moser-Mercer, Barbara. 1996. Quality in Interpreting: Some Methodological Issues. *The Interpreter's Newsletter* 7, 43-55

Müller, Markus & Son Nguyen, Thai & Niehues, Jan & Cho, Eunah & Krüger, Bastian & Ha Than-Le & Kilgour, Kevin & Sperber, Matthias & Mediani, Mohammed & Stüker, Sebastian & Waibel, Alex. 2016. Lecture Translator – Speech translation framework for simultaneous lecture translation. In: <http://www.aclweb.org/anthology/N16-3017>, Stand: 10.02.2017

ÖWA (Österreichische Webanalyse). 2017. Medienanalyse. In: <http://www.oewa.at/plus/medienanalyse>, Stand: 23.06.2017

Paul, Ian. 2015. Skype's amazing real-time Translator Preview now available to all. In: <http://www.pcworld.com/article/2921129/skypes-amazing-real-time-translator-preview-now-available-to-all.html>, Stand: 17.07.2017

Pöchhacker, Franz. 1994a. Quality Assurance in Simultaneous Interpreting. In: Dollerup Cay & Lindegaard Annette (Hg.) *Teaching Translation and Interpreting 2*. Amsterdam/Philadelphia: John Benjamins Publishing Company, 233-242

Pöchhacker, Franz. 1994b. *Simultandolmetschen als komplexes Handeln*. Tübingen: Günter Narr

Pöchhacker, Franz. 2002. Researching Interpreting quality. In: Garzone, Giuliana & Viezzi, Mairizio (Hg.) *Interpreting in the 21st Century: Challenges and opportunities*. Amsterdam/Philadelphia: John Benjamin Publishing Company, 95-107

Pöchhacker, Franz. 2016². *Introducing Interpreting Studies*. London/New York: Routledge Taylor & Francis Group

Pöchhacker, Franz & Cornelia, Zwischenberger. 2010. Survey on Quality and Role: Conference Interpreters' Expectations and Self-Perceptions. In: <https://aiic.net/page/3405/survey-on-quality-and-role-conference-interpreters-expectations-and-self-perceptions/lang/1>, Stand: 30:07.2017

Piechotta, Beatrie. 2008. *Qualitätsmanagement für psychotherapeutische Praxen*. Berlin/Heidelberg: Springer-Verlag

Rat der Europäischen Union. Voranschlag der Ausgaben und Einnahmen für das Haushaltsjahr 2018 für den Europäischen Rat (Einzelplan II des EU-Haushaltsplans)-Begründung. In: https://www.parlament.gv.at/PAKT/EU/XXV/EU/13/64/EU_136491/imfname_10704033.pdf, Stand: 30.07.2017

Riccardi, Alessandra. 2005. Qualitätskriterien beim Simultandolmetschen. In: Zybatow, Lew N. (Hg.) *Translatologie*. Frankfurt am Main: Lang, 99-122

Sandrini, Peter. 2011. Lexikoneintrag „Translationswissenschaft“. In: Reinalter, H.; Brenner, P. (Hg.) *Lexikon der Geisteswissenschaften. Sachbegriffe – Disziplinen – Personen*. Wien: Böhlau, 1195-1100.

Schmidt-Biggemann, Wilhelm. 1983. *Topica universalis – Eine Modellgeschichte humanistischer und barocker Wissenschaft*. Hamburg: Felix Meiner Verlag GmbH

Schwanke, Martina. 1991. *Maschinelle Übersetzung – Ein Überblick über Theorie und Praxis*. Berlin/Heidelberg/New York/London/Paris/Hong Kong/Tokyo/Barcelona/Budapest: Springer-Verlag

Skype Team. 2016. Skype Translator empowers more people. In: <https://blogs.skype.com/news/2016/01/13/skype-translator-empowers-more-people/>, Stand: 20.07.2017

Tessiore, Lorenzo & Hahn, Walther v. 2000. Functional Validation of a Machine Interpretation System: Verbmobil. In: Wahlster, Wolfgang. *Verbmobil: Foundations of Speech-to-Speech Translation*. Berlin/Heidelberg/New York/Barcelona/Hong Kong/London/Mailand/Paris/Singapur/Tokyo: Springer, 611-635.

Tiselius, Elisabeth. 2012. Research on quality in interpreting. In: <https://aiic.net/page/6318/research-on-quality-in-interpreting/lang/1>, Stand: 25.07.2017

unikat Hochschulinformationssysteme GmbH. 2005. Berufsbilder und Aussichten <http://www.wegweiser.ac.at/univie/studieren/zentrans/A063.html?klapp=2>, Stand: 23.05.2017

Universität Wien. 2007. Masterstudium an der Universität Wien – Studium Dolmetschen. https://transvienna.univie.ac.at/fileadmin/user_upload/z_translationswiss/Studium/Curricula/Curriculum_Masterstudium_Dolmetschen_2007.pdf, Stand: 23.05.2017

Universität Wien & Universidad de Granada. 2009. Research Methods in Interpreting Quality Assessment. In: http://quasi.univie.ac.at/fileadmin/user_upload/proj_quasi/Granada_WS.pdf, Stand: 24.07.2017

Universitat de Barcelona. o.J. Wer war Ramon Llull – Eine Einführung. In: http://quisestlullus.narpan.net/de/1_intro_de.html, Stand: 29.01.2017

Vashee, Kirti. 2016. How many words are translated every day by Google translate? In: <https://www.quora.com/How-many-words-are-translated-every-day-by-Google-translate>, Stand 11.07.2017

Venuti, Lawrence. 1995. *The Translator's Invisibility. A History of Translation*. New York: Routledge Taylor & Francis Group

Vogel, Stephan. 2016. Machine Translation – Winds of Change. [19:25]. In: http://videolectures.net/interACT2016_vogel_machine_translation/, Stand 11.07.2017

Wahlster, Wolfgang. 2000. VERBMOBIL – Erkennung, Analyse, Transfer, Generierung und Synthese von Spontansprache. In: <http://verbmobil.dfki.de/Vm.Info.Phase2.html>, Stand: 12.07.2017

Waibel, Alexander. 2011. A New Assault on the Language Divide. [57:07]. In: <https://www.youtube.com/watch?v=-859pN5dDIw>, Stand: 12.07.2017

Waibel, Alexander. 2012. Simultandolmetschen durch Maschinen. [34:56]. In: <https://www.youtube.com/watch?v=x5IL0wpr-88&t=15s> , Stand: 13.07.2017

Waibel, Alexander. 2015a. Riga Summit 2015. Plenary. Alexander Waibel. [20:27]. In: <https://www.youtube.com/watch?v=JJcxdj-eYOc>, Stand 12.03.2017

Waibel, Alexander. 2015b. Sprachbarrieren durchbrechen – Traum oder Wirklichkeit?. *Nova Acta Leopoldina* 122:410, 101-123.

Waibel, Alexander. 2016. Technologies for Multilingual Human Communication [16:26]. In: http://videolectures.net/interACT2016_waibel_language_barrier/, Stand: 13.07.2017

Watzlawick, Paul. 2011. *Menschliche Kommunikation – Formen, Störungen, Paradoxien*. Bern: Hans Huber, Hogrefe AG

Weinstein, C.J. 2002. Speech-to-Speech Translation: Technology and Applications Study.
In: https://www.ll.mit.edu/mission/cybersec/publications/publication-files/full_papers/020510_Weinstein.pdf, Stand: 12.01.2017

Werthmann, Antonina & Witt, Andreas. 2014. Maschinelle Übersetzung – Gegenwart und Perspektiven. In: *Translation and Interpretation in Europe. Contributions to the Annual Conference 2013 of EFNIL in Vilnius*. Frankfurt am Main/Berlin/Bern/Brüssel/New York/Oxford/Wien: Lang, 79-103 (Duisburger Arbeiten zur Sprach- und Kulturwissenschaft 102)

Wilkins, John. 1668. *An Essay towards a Real Character and a Philosophical Language*. London: S. Gellibrand

Wirtschaftskammer Österreich, 2017. DolmetscherIn – Berufsbeschreibung. In: <http://www.bic.at/berufsinformation.php?brfid=563>, Stand: 20.05.2017

Wortbedeutung Online-Wörterbuch. 2017. o.A. In: <http://www.wortbedeutung.info/Dolmetscher/>, Stand: 20:06.2017

Sophie & Garnier-Rizet Martine & Devillers Laurence (Hg.) *Natural Interaction with Robots, Knowbots and Smartphones*. New York: Springer, 51-57.

Xu Feiyu & Schmeier Sven & Ai Renlong & Uszkoreit Hans. 2014. Yochina: Mobile Multimedia and Multimodal Crosslingual Dialogue System. In: Mariani, Joseph & Rosset, Sophie & Garnier-Rizet, Martine & Devillers, Laurence. *Natural Interaction with Robots, Knowbots and Smartphones*. New York: Springer, 51-57

Yang, Jie & Gao, Jiang & Zhang, Ying & Chen, Xilin & Waibel, Alex. 2001. An automatic sign recognition and translation system. In: Workshop on Perceptive User Interfaces, ACM Digital Library (2001)

Ziliani, Alba Rosa. 2005. Aufforderungen zur Klärung. In: <http://www.cis.uni-muenchen.de/people/kristof/DIALOG05/sitzung11a.pdf>, Stand: 29.01.2017

Zingel, Hermann-Josef. 1978. Experiences with TITUS II. In: <http://mt-archive.info/IntClass-1978-Zingel.pdf>, 13.01.2017

Anhang

Verbmobil – Technische Ziele

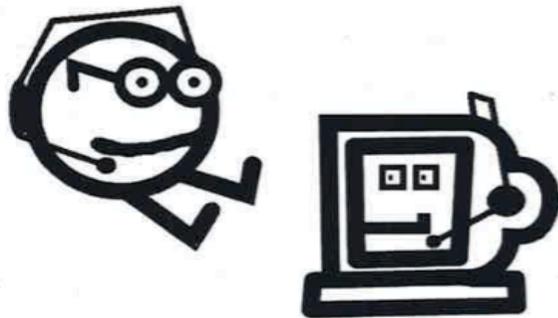
Die technischen Ziele der ersten VERBMOBIL-Phase

1. Erkennung fließend gesprochener Spontansprache für Deutsch, Japanisch und Englisch über Nahbesprechungsmikrophon
2. Wortschatz von ca. 2500 Wörtern für die Übersetzungsrichtung Deutsch nach Englisch
3. Sprecheradaptives System mit sprecherunabhängigem Kern
4. Linguistisch fundierte deutsche Basisgrammatik für Spontansprache mit tiefer und flacher semantischer Analyse
5. Gesprochene Klärungsdialoge zwischen dem Benutzer und dem Verbmobil- System bei Spracherkennungs- und Verstehensproblemen
6. Semantischer Transfer für Deutsch -Englisch und Japanisch - Englisch
7. Sprachgenerierung für Englisch und für deutsche Paraphrasen
8. mehr als 70% approximativ korrekte Übersetzungen bei der End-to-End Evaluation in der Domäne Terminverhandlung
9. Reine Softwarelösung für alle Module auf Standardhardware
10. Netto-Verarbeitungszeit < sechsfache Echtzeit bezogen auf die Länge des Eingabe-Sprachsignals

Die technischen Ziele der zweiten VERBMOBIL-Phase

1. Multifunktionalität: Verbmobil soll rasch auf neue Gesprächsdomänen einstellbar sein.
2. Multilingualität: Verbmobil soll spontane Dialoge in mehrere Sprachen übersetzen können
3. Multimedialität: Verbmobil soll in internationalen Multimedia-Anwendungen
4. Mobilität: Durch einen Sprachserver soll Verbmobil auch über Handy nutzbar sein
5. Multiparty-Funktionalität: Verbmobil soll nicht nur in Dialogsituationen, sondern auch in Telekooperationsanwendungen mit vielen Gesprächspartnern Übersetzungshilfe leisten

Fragebogen Fünfer Studie



MENSCH VS. MASCHINE

Fragebogen zur Bewertung der Qualität
der Verdolmetschungen im Rahmen der Freitagskonferenz
"Speech-to-Speech - Translation" am 1. Juli 2011

Masterarbeit von Sarah Fünfer an der



in Zusammenarbeit mit



HINWEISE ZUM AUSFÜLLEN DES FRAGEBOGENS

Bei den meisten Fragen brauchen Sie lediglich eines der vorgegebenen Kästchen ankreuzen:

<input checked="" type="checkbox"/>	<input type="checkbox"/>				
-------------------------------------	--------------------------	--------------------------	--------------------------	--------------------------	--------------------------

Möchten Sie etwas korrigieren, kennzeichnen Sie dies bitte folgendermaßen:

<input checked="" type="checkbox"/>					
-------------------------------------	-------------------------------------	-------------------------------------	-------------------------------------	-------------------------------------	-------------------------------------

Bei einigen Fragen haben Sie die Möglichkeit, eine eigene Antwort zu formulieren. Bitte verwenden Sie nach Möglichkeit Blockschrift.

A: Deutsch

Wenn Sie eine oder mehrere Fragen kommentieren oder ausführlichere Antworten geben möchten, benutzen Sie hierfür bitte die letzte Seite oder verwenden Sie ein zusätzliches Blatt Papier.

Noch einmal herzlichen Dank und viel Spaß beim Ausfüllen des Fragebogens!

A Allgemeine Fragen

1. Wem haben Sie zugehört?
 - Mensch (Dolmetscher)
 - Maschine (Computer)
2. In welchem Studiengang studieren Sie?
 - B.A. Sprache, Kultur, Translation
 - M.A. Sprache, Kultur, Translation
 - M.A. Konferenzdolmetschen
 - Andere _____
3. Welche Sprachen studieren Sie (Sprachkombination)?
 - A: _____
 - B: _____
 - C: _____
 - D: _____
4. In welchem Semester studieren Sie (Studiensemester insgesamt)?
 - _____

B Beurteilung der Dolmetschleistung

5. Auf einer Skala von 1 (positivster Wert) bis 6 (negativster Wert), wie beurteilen Sie ...

- | | 1 | 2 | 3 | 4 | 5 | 6 |
|--|--------------------------|--------------------------|--------------------------|--------------------------|--------------------------|--------------------------|
| a) ... die Dolmetschleistung insgesamt? | <input type="checkbox"/> |
| b) ... die Stimmqualität (Wie angenehm war die Stimme?)? | <input type="checkbox"/> |
| c) ... die Intonation? | <input type="checkbox"/> |
| d) ... das Redetempo der Verdolmetschung? | <input type="checkbox"/> |
| e) ... die grammatikalische Korrektheit? | <input type="checkbox"/> |
| f) ... die korrekte Benutzung von Fachtermini? | <input type="checkbox"/> |
| g) ... die Flüssigkeit der Verdolmetschung? | <input type="checkbox"/> |
| h) ... den Redestil? | <input type="checkbox"/> |
| i) ... den logischen Zusammenhang? | <input type="checkbox"/> |
| j) ... die Verständlichkeit der Verdolmetschung? | <input type="checkbox"/> |

C Fragen zu den Untertiteln

ACHTUNG: DIE FRAGEN 6 UND 7 RICHTEN SICH NUR AN DIEJENIGEN, DIE DER COMPUTER-VERDOLMETSCHUNG ZUGEHÖRT HABEN! FALLS DIES FÜR SIE NICHT ZUTRIFFT, BITTE DIREKT MIT FRAGE 8 FORTFAHREN.

6. Auf einer Skala von 1 (immer) bis 6 (nie), wie häufig haben Sie die Untertitel genutzt?

1	2	3	4	5	6
<input type="checkbox"/>					

7. Auf einer Skala von 1 (positiv) bis 6 (negativ), wie haben die Untertitel Ihr Verständnis beeinflusst?

1	2	3	4	5	6
<input type="checkbox"/>					

D Fragen zum Inhalt

8. Wofür will InterAct automatische Sprach-zu-Sprach-Übersetzungssysteme einsetzen?

9. Auch die Leistung menschlicher Dolmetscher ist begrenzt. Welche Gründe wurden hierfür genannt?

10. Wird es, nach Meinung des Redners, in der Zukunft übermenschliche, also maschinelle Simultanverdolmetschungen geben?

11. Warum ist Spracherkennung für eine Maschine so schwierig?

12. Welche Nachteile hat ein mobiles Dolmetschsystem?

13. Was sind die Herausforderungen bei der automatischen Übersetzung von Vorträgen?

E Platz für Kommentare / Meinungen

Abstract

In vorliegender Arbeit wird das Thema des maschinellen Dolmetschens (MD) und dessen Bezug zum Berufsstand Dolmetschender behandelt. Durch die rapiden Fortschritte im Bereich der Sprachtechnologie und Computerlinguistik, steht man immer öfter vor der großen Frage: Werden HumandolmetscherInnen durch die Maschine ersetzt werden? Eine eindeutige Beantwortung der Frage scheint aufgrund der kontinuierlichen, erfolgsversprechenden Entwicklungen auf diesem Forschungsgebiet nur bedingt möglich zu sein. Heute haben Sprachtechnologien zwar bereits einen hohen Reifegrad erreicht, doch stellen sie die Entwickler weiterhin noch vor große Herausforderungen, die sich seit Anfang der MD-Geschichte beobachten ließen. So ist es beispielsweise schwer, auf morphosyntaktischer Ebene, hohe Qualität zu erreichen. Letzterer Punkt führt schließlich zur Frage nach der Qualität, deren Definitionsbestimmung in der Translationswissenschaft bis heute eine Hürde darstellt.

Unter Berücksichtigung des oben Genannten sowie nach näherer Auseinandersetzung mit dem Berufsbild Dolmetschender in Österreich, wurde in vorliegender Arbeit versucht zu ermitteln, in welcher Beziehung (MD)-Sprachtechnologien zu HumandolmetscherInnen stehen und welche voraussichtlichen Entwicklungen zu erwarten sind.