

TRANSKULTURALITÄT - TRANSLATION -
TRANSFER



Mensch oder Maschine?

Dolmetscher und maschinelles Dolmetschsystem
im Vergleich

Sarah Fünfer

F Frank & Timme

Verlag für wissenschaftliche Literatur

Sarah Fünfer
Mensch oder Maschine?

Transkulturalität – Translation – Transfer, Band 2

Herausgegeben von

Dörte Andres / Martina Behr / Larisa Schippel / Cornelia Zwischenberger

Sarah Fünfer

Mensch oder Maschine?

Dolmetscher und maschinelles Dolmetschsystem im Vergleich

FFrank & Timme
Verlag für wissenschaftliche Literatur

Umschlagabbildung: © Composer – Fotolia.com

ISBN 978-3-86596-548-6

ISSN 2196-2405

© Frank & Timme GmbH Verlag für wissenschaftliche Literatur
Berlin 2013. Alle Rechte vorbehalten.

Das Werk einschließlich aller Teile ist urheberrechtlich geschützt.
Jede Verwertung außerhalb der engen Grenzen des Urheberrechts-
gesetzes ist ohne Zustimmung des Verlags unzulässig und strafbar.
Das gilt insbesondere für Vervielfältigungen, Übersetzungen,
Mikroverfilmungen und die Einspeicherung und Verarbeitung in
elektronischen Systemen.

Herstellung durch das atelier eilenberger, Taucha bei Leipzig.

Printed in Germany.

Gedruckt auf säurefreiem, alterungsbeständigem Papier.

www.frank-timme.de

Inhaltsverzeichnis

1 Einleitung	9
2 Theoretische Grundlagen	13
2.1 Maschinelles Dolmetschen (MD).....	13
2.1.1 Begriffsklärung	13
2.1.2 Aufbau eines MD-Systems	14
2.1.3 Geschichtliche Entwicklung des maschinellen Dolmetschens..	18
2.1.4 Der Lecture Translator.....	21
2.2 Dolmetschqualität und Evaluierung von Leistungen	26
2.2.1 Der Qualitätsbegriff.....	26
2.2.2 Studien zur Qualität im Dolmetschen	31
2.2.3 Linguistische Evaluierung von maschinellen Übersetzungssystemen.....	48
2.2.4 Evaluierung der Leistungen von maschinellen Dolmetschsystemen.....	48
3 Fallstudie Mensch vs. Maschine – Versuchsaufbau und Durchführung ..	53
3.1 Fragestellung und Hypothesen	53
3.2 Versuchsaufbau.....	54
3.3 Methodik.....	57
3.3.1 Aufbau des Fragebogens	58
3.3.2 Kriterien zur Analyse des Originals und der verdolmetschten Texte	60
3.4 Stichprobenkonstruktion	73
3.5 Versuchsdurchführung.....	73

4 Ergebnisse zu den einzelnen Fragestellungen und Hypothesen	77
4.1 Rezipientenbefragung	77
4.1.1 A: Allgemeine Angaben.....	77
4.1.2 B: Beurteilung der Dolmetschleistungen	79
4.1.3 C: Nützlichkeit der Untertitel.....	81
4.1.4 D: Inhaltsfragen.....	82
4.1.5 E: Kommentare.....	83
4.1.6 Zusammenfassung	84
4.2 Beispielanalyse – Ausschnitt D1	85
4.2.1 Ausgangstext.....	85
4.2.2 Verdolmetschung des Humandolmetschers	86
4.2.3 Verdolmetschung des Lecture Translators.....	87
4.2.4 Zusammenfassung und Vergleich	89
4.3 Vergleich der Analysen von Ausgangstext und verdolmetschten Texten.....	90
4.3.1 Redetempo und Pausen.....	90
4.3.2 Gefüllte Pausen und Abbrüche/Neuplanungen.....	92
4.3.3 Unbekannte Wörter.....	93
4.3.4 Kohäsion und Kohärenz	94
4.3.5 Zusammenfassung	94
4.4 Segmentweiser Vergleich von Ausgangstext und Zieltexten.....	96
4.4.1 Überprüfung der Verständlichkeit	96
4.4.2 Überprüfung der inhaltlichen Vollständigkeit.....	100
4.4.3 Linguistische Fehleranalyse	108
4.4.4 Änderung der kommunikativen Funktion/andere Fehler	119
4.4.5 Zusammenfassung	122
5 Schlussbetrachtung und Ausblick	125
5.1 Zusammenfassung.....	125
5.2 Methodenkritik	129
5.3 Aktuelle Projekte und Ausblick	131

Literaturverzeichnis	135
ANLAGE: FRAGEBOGEN	141
Abkürzungsverzeichnis	147

1 Einleitung

Seit einigen Jahren arbeitet eine Forschungsgruppe des Karlsruher Instituts für Technologie (KIT) um Prof. Dr. Alex Waibel zusammen mit Wissenschaftlern¹ der Carnegie Mellon University, Pittsburgh, Pennsylvania, USA, am *Lecture Translator*, einem System, das gesprochene Sprache erkennt, maschinell übersetzt und dann als Sprache und/oder Text wiedergibt. „Unsere Neuentwicklung kann nun unbeschränkt Themen und Gesprächssituationen wie Vorträge oder Ansprachen simultan übersetzen – eine absolute Neuheit.“² Diese Worte sind in einer Pressemitteilung aus dem Jahr 2005 zu lesen, in der das System vorgestellt wird. Der Beschreibung zufolge ist die Arbeit und Leistung dieser Maschine vergleichbar mit der eines menschlichen Dolmetschers. Wird also in naher Zukunft eine Maschine die Aufgaben des Konferenzdolmetschers übernehmen und zwischen beliebig vielen Sprachen und zu beliebigen Themen simultan dolmetschen?

In der Tat werden in der Dolmetschgemeinschaft automatisierte Systeme oft als Bedrohung wahrgenommen, da befürchtet wird, sie könnten die Person des Dolmetschers überflüssig machen. Im Folgenden soll deshalb anhand einer umfangreichen, dolmetschwissenschaftlichen Evaluierung die Frage beantwortet werden, was ein solches System im Vergleich zu menschlichen Dolmetschern heute tatsächlich leisten kann.

Hierzu werden in Kapitel 2 die theoretischen Grundlagen für die nachfolgende Untersuchung gelegt. Diese betreffen zunächst die technischen Hintergründe wie die Geschichte und den aktuellen Entwicklungsstand des maschinellen Dolmetschens sowie eine kurze Beschreibung des Aufbaus des zu untersuchenden Systems. Anschließend wird näher auf dolmetschwissenschaftliche Erkenntnisse und Fragestellungen eingegangen. Wie wird Qualität in der Dolmetschwissenschaft definiert? Und welche Möglichkeiten existieren, um eine Dolmetschleistung möglichst objektiv zu beurteilen? In der Literatur finden sich zwei Ansätze zur Evaluierung von Dolmetschleistungen, die sich durch ihre Perspektive und Herangehensweise unterscheiden. Zum einen

.....
1 Die männliche Form ist hier und im Folgenden inkludierend zu verstehen.

2 <http://www.uebersetzerportal.de/nachrichten/n-archiv/2005/2005-10/2005-10-29.html>

besteht die Möglichkeit der Befragung von Zuhörern. Sie gibt Aufschluss darüber, welche Anforderungen Rezipienten an eine qualitativ hochwertige Verdolmetschung stellen, oder wie sie eine bestimmte Leistung beurteilen. Dementsprechend werden die wichtigsten Studien aus dem Bereich der Rezipientenbefragung vorgestellt, wobei der Schwerpunkt auf den dabei verwendeten Kriterien liegt. Sollen jedoch Aussagen bezüglich der inhaltlichen Vollständigkeit und Korrektheit getroffen werden, sind Rezipientenbefragungen weniger geeignet, da die Zuhörer in der Regel die Ausgangssprache nicht oder nicht ausreichend verstehen. Eine weitere Möglichkeit besteht deshalb darin, den Ausgangstext mit dem Zieltext zu vergleichen. Auch hierzu gibt es bereits verschiedene Forschungsarbeiten mit unterschiedlichen Herangehensweisen. Ramlow (2008) nimmt eine systematische linguistische Evaluierung verschiedener maschineller Übersetzungssysteme vor und entwickelt hierfür eine spezielle Fehlerklassifizierung. Er spannt so eine Brücke zwischen Translationswissenschaft und maschineller Translation, die in Kapitel 2 ebenso vorgestellt wird wie eine Evaluierung des Karlsruher Systems aus dem Jahr 2007.

In den daran anschließenden Kapiteln wird näher auf den Versuch eingegangen, der den Mittelpunkt der Arbeit bildet und in dem die Leistung des in Karlsruhe entwickelten *Lecture Translators* erstmals direkt und unter möglichst realitätsnahen Bedingungen mit menschlichen Dolmetschleistungen verglichen wurde.

Kapitel 3 beschreibt den Versuchsaufbau, die Zusammensetzung des Fragebogens und die Versuchsdurchführung. Auch die in der Auswertung angewandte Methodik und Fehlertypologie werden erläutert. Der *Lecture Translator* wurde im Rahmen der Lehrveranstaltung „Freitagskonferenz“ des Fachbereichs für Translations-, Sprach- und Kulturwissenschaft der Universität Mainz getestet, da er, wie der Name schon andeutet, u.a. im universitären Umfeld eingesetzt werden soll. Ein Vortrag zum Thema „Speech-to-Speech Translation“ wurde dort sowohl von der Maschine als auch von Studierenden im Examensemester gedolmetscht, wobei Verdolmetschungen und der Originalvortrag zur späteren Analyse aufgezeichnet wurden. Außerdem wurden die Zuhörer gebeten, der Verdolmetschung von Mensch bzw. Maschine zu folgen und die Leistungen mittels eines Fragebogens zu bewerten, um eine subjektive Einschätzung des Systems und der Humandolmetscher durch Studierende, also durch die Zielgruppe, zu erreichen.

In Kapitel 4 erfolgt die Analyse der gesammelten Daten aus dolmetsch- bzw. translationswissenschaftlicher Perspektive. Zunächst werden die Fragebö-

gen ausgewertet. Die Ergebnisse geben darüber Aufschluss, wie die Zuhörer die allgemeine Qualität und bestimmte Eigenschaften einschätzen und inwieweit sie das Gehörte tatsächlich verstehen und behalten. Da davon ausgegangen wird, dass die Zuhörer die inhaltliche Vollständigkeit und Richtigkeit nicht einschätzen können und ihre Beurteilung von persönlichen Erfahrungen geprägt ist, werden anschließend ausgewählte Textabschnitte dolmetschwissenschaftlich evaluiert. Die Abschnitte werden hierzu in einem ersten Durchlauf als eigenständige Texte betrachtet und auf Kohäsion und Kohärenz sowie auf bestimmte Auffälligkeiten der Textoberfläche wie Häsitationslaute, lange Pausen etc. untersucht. Anschließend erfolgt ein segmentweiser, systematischer Vergleich der Textabschnitte des Ausgangstextes mit den entsprechenden Zieltexten. Grundlage hierfür ist die in Kapitel 3 erarbeitete Fehlertypologie. Überprüft werden unter anderem Verständlichkeit, inhaltliche Vollständigkeit, sprachliche Fehler sowie die Änderung bzw. Beibehaltung der kommunikativen Funktion. Die Analyse gibt Aufschluss über die Qualität sowie die Stärken und Schwächen der jeweiligen Dolmetschleistung.

In Kapitel 5 werden die Ergebnisse zusammengefasst. Da es sich um einen Versuch handelt, der in dieser Art erstmalig durchgeführt wurde, erfolgt eine kritische Auseinandersetzung mit den angewandten Methoden. Am Schluss steht ein kurzer Ausblick, in dem einige humanitäre Projekte vorgestellt werden, in denen maschinelle Dolmetschsysteme bereits angewendet werden.

2 Theoretische Grundlagen

2.1 Maschinelles Dolmetschen (MD)

2.1.1 Begriffsklärung

Gemäß den in der Translationswissenschaft geltenden Definitionen beschreibt „Translation“ als Oberbegriff die Übertragung eines Textes in einer Ausgangssprache (AS) in einen Text in einer Zielsprache (ZS). Während bei der Übertragung eines schriftlichen Ausgangstextes (AT) in einen schriftlichen Zieltext (ZT) von Übersetzung gesprochen wird, bezeichnet Dolmetschen die Übertragung „eines einmalig (in der Regel mündlich) dargebotenen Textes der Ausgangssprache in einen nur bedingt kontrollierbaren und infolge Zeitmangels kaum korrigierbaren Text der Zielsprache“ (KADE 1968: 35). Je nach Art und Weise, wie die Verdolmetschung erfolgt, lassen sich verschiedene Dolmetschmodi unterscheiden, zum Beispiel Flüsterdolmetschen, Simultandolmetschen, Konsektivdolmetschen oder Dialogdolmetschen.

Im Zusammenhang mit der vorliegenden Thematik muss des Weiteren eine Unterscheidung zwischen der vom Menschen verrichteten Humantranslation, die sich in Humandolmetschen (HD) und Humanübersetzung (HÜ) unterteilen lässt, und der maschinellen Translation, die sich wiederum in maschinelles Dolmetschen (MD) und maschinelles Übersetzen (MÜ) gliedert, vorgenommen werden.

In der computerwissenschaftlichen Literatur zum MD, die hauptsächlich auf Englisch vorliegt, wird das MD meist nicht als *Machine* oder *Automatic Interpretation*, sondern als *Speech-to-Speech-Translation* (SST) bezeichnet (vgl. KITANO 1994; WAHLSTER 2000; STÜKER et al. 2006). Auch die Begriffe *Automatic Simultaneous Translation* (vgl. HAMON et al. 2009) oder *Spoken Language Translation* (vgl. WAIBEL & FÜGEN 2008) beziehen sich auf das MD. Fügen et al. (2008) erklären dies folgendermaßen:

The reason [...] is that in the past the main focus of MT [Machine Translation] was the translation of text and only recently is SST attracting a wider interest. Therefore, [...] we use the term translation as in ‘SST’,

'simultaneous translation' or 'simultaneous speech translation' to mean the automatic interpretation of spoken language (FÜGEN et al. 2008: 212).

Im Folgenden wird dennoch bewusst der deutsche Ausdruck *maschinelles Dolmetschen* verwendet, da wie beim Humandolmetschen auch im Falle der Maschine ein gesprochener AT u.a. in gesprochene Sprache übertragen wird und Konventionen der Computerwissenschaft zurückstehen, die für die vorliegende Perspektive nur von untergeordneter Bedeutung sind. Auch Hauen-schild und Schmitz (1998) benutzen in ihrem Artikel den Ausdruck *maschinelles Dolmetschen*.

2.1.2 Aufbau eines MD-Systems

Ein MD-System besteht aus zwei Hauptkomponenten: einem System zur automatischen Spracherkennung (*automatic speech recognition, ASR*) und einem System zur maschinellen Übersetzung (*machine translation, MÜ*). Soll der ZT als gesprochene Sprache ausgegeben werden, schließt sich ein Modul zur Sprachsynthese an. Auch andere Formen der ZT-Präsentation wie Untertitel sind möglich (vgl. Kapitel 2.1.4). Hinzu kommen weitere Komponenten, die es ermöglichen, die verschiedenen Informationen so aufzubereiten, dass sie bestmöglich weiterverarbeitet werden können. So hängt die Übersetzungsleistung des MD stark von dem Erfolg der Spracherkennung ab (vgl. FÜGEN et al. 2008: 216).

Die Qualität des MD-Systems entspricht dabei nicht zwingend der Summe seiner Teile. Zwar kann die Leistung des Systems durch die Verbesserung einzelner Elemente gesteigert werden, doch ist es auch möglich, das System komponentenübergreifend auf Grundlage der endgültigen Ausgabe zu optimieren:

Since the goal is to produce output in a target language, the correctness of the components' output is of secondary concern. Uncertainty at the component level can be addressed by being noncommittal at their interface, linking components via near-miss hypothesis lattices. Such a view also offers the possibility of jointly optimizing components based on the overall output rather than each component independently (WAIBEL & FÜGEN 2008: 71).

Die verschiedenen Komponenten, ihre Aufgaben und die Herausforderungen, denen sie gegenüberstehen, werden im Folgenden kurz vorgestellt.