

Chapter 1

Digitalisierung in der Audiovisuellen Translation

Sylvia Jaki^a

^aUniversität Hildesheim

Abstract

1 Einleitung

Der Reiz der verschiedenen Branchen im Bereich der Audiovisuellen Translation (AVT) besteht für viele Translator*innen insbesondere in den kreativen Anteilen ihrer Arbeit – wie lässt sich beispielsweise ein komplexes Wortspiel trotz der Platzbeschränkungen der Untertitelung in eine andere Sprache übertragen oder wie kann ich mithilfe von Audiodeskription einer sehgeschädigten Person die Emotionalität bestimmter visueller Eindrücke vermitteln? Nun hat sich die Arbeit in der AVT durch die vielfältigen Möglichkeiten der Digitalisierung jedoch in mancher Hinsicht tiefgreifend verändert. Die verschiedenen Softwares, die geeignet sind, um die Translate in den verschiedenen AVT-Branchen zu produzieren, sollen den Humanübersetzer*innen die Arbeit merklich erleichtern. Mit der Automatisierung mancher Arbeitsschritte, beispielsweise mit Hilfe von Maschineller Übersetzung (MÜ), können Produktionszeiten noch einmal deutlich verringert werden. Jedoch werden Übersetzenden zum Teil auch Aufgaben abgenommen, die sie als für ihren Beruf zentral einschätzen.

In dieser kurzen Einführung zur Digitalisierung in der AVT soll vor dem Hintergrund der verschiedenen Felder der AVT ein Überblick gegeben werden, welche Technologien die moderne AVT prägen, wenngleich aufgrund der zahlreichen verschiedenen Branchen eine gewisse Schwerpunktsetzung erfolgen muss, die hier vor allem im Bereich der Untertitelung liegt – ein Feld, das bereits sehr

stark durch Automatisierungstendenzen geprägt ist und auf das sich auch der Großteil der Forschung konzentriert. Darüber hinaus soll jedoch auch auf weitere Bereiche eingegangen werden. Es geht in diesem Beitrag folglich weniger darum, die umfangreiche Entwicklung der Digitalität von AVT historisch nachzuzeichnen (vgl. für einen historischen Überblick z.B. *Díaz-Cintas2013*), sondern die Entwicklungen, die die Branche(n) in den letzten beiden Jahrzehnten oder in manchen Fällen sogar erst in den letzten Jahren zunehmend prägen, zu beleuchten. Überdies wird das Augenmerk auf diejenigen hinter dem Translat gelenkt – die Translator*innen, da sich die Frage stellt, welche Auswirkungen die Digitalisierung der Arbeitsprozesse auf die Arbeitszufriedenheit der in der AVT Beschäftigten ausübt. Der Beitrag schließt mit einem Ausblick auf Forschungsperspektiven, die sich durch die erläuterten Tendenzen in Bezug auf Forschungsgegenstände und -methoden der AVT-Forschung ergeben.

2 Die Felder der Audiovisuellen Translation

Bei der Audiovisuellen Translation (AVT) handelt es sich um ein äußerst vielfältiges Betätigungsfeld, in dem es um die Translation von Medienprodukten geht, die von den Rezipierenden typischerweise sowohl über den visuellen als auch über den auditiven Kanal wahrgenommen werden. Das trifft insbesondere auf Filme zu, aber auch auf Videospiele oder Theaterstücke (vgl. z.B. *Jüngst2020* für einen ausführlichen Überblick über die verschiedenen Felder der AVT). Bei der audiovisuellen Übersetzung wird konkreter „das ursprünglich vorliegende Material durch die Übersetzung verändert und meist ergänzt“, wobei „Teile des ursprünglichen Materials erhalten“ bleiben (*Jüngst2020*).

Wenn *Díaz-Cintas2013* von einer „audiovisualisation of our communicative environment, where sounds and visuals coalesce in a winning combination over other formats, particularly among younger generations“ spricht, verwundert es wenig, dass der Bedarf an AVT enorm gestiegen ist. Allein wenn wir an Streaming-Dienste wie Netflix, Amazon Prime oder Disney Plus denken, wird deutlich, wie stark der Markt für *interlinguale* Übersetzungen in den letzten Jahren gewachsen ist (vgl. auch *Georgakopoulou2019*: 521f). Obgleich englische Produkte den Markt weiterhin dominieren, sind es gerade die Streaming-Dienste, die den Weg für ein breiteres internationales Angebot geebnet haben, allen voran für zahlreiche beliebte asiatische Filme und Serien, hauptsächlich aus Südkorea und Japan (vgl. *Kluge/Camacho/Lins2022*: 147/149). In diesem Zusammenhang ist auch der Bedarf an interlingualen Untertiteln stark gestiegen, da Rezipierende auf den Streaming-Plattformen Produkte vermehrt in der Originalsprache mit deutschen

Untertiteln konsumieren. In Theater und Oper beispielsweise werden zur besseren Sichtbarkeit alternativ Übertitel eingeblendet (vgl. Reinart2018: 137 für eine Übersicht zahlreicher weiterer Anwendungsfälle für Untertitel, die mittlerweile deutlich über Film und Fernsehen hinausgehen). Gleichzeitig dominiert an anderen Stellen, allen voran im Fernsehen, in einigen Ländern weiterhin die Synchronisation als Standardform der interlingualen Translation von Filmen. So auch in Deutschland, das nach wie vor als Synchronisationsland gilt.¹ Neben Untertitelung und Synchronisation stellt die Voice-Over-Übersetzung eine dritte Form der interlingualen AVT dar. In Deutschland kommt sie standardmäßig in Dokumentationen zum Einsatz, wenn Äußerungen fremdsprachiger Personen übersetzt werden sollen. Dabei wird der Originalton heruntergedreht, während die Übersetzung lauter darübergelegt wird. In einigen osteuropäischen Ländern wird Voice-Over (oft lapidar auch als *slawische Synchro* bezeichnet) auch für Spielfilme eingesetzt, wobei sich mittlerweile ein rückläufiger Trend abzuzeichnen scheint (vgl. Flis/Szarkowska2024). Ein Gebiet, das in jüngerer Zeit verstärkt Aufmerksamkeit erfährt, ist das der Videospielokalisierung, für das es ebenfalls eine hohe Nachfrage im Bereich der Übersetzung zwischen verschiedenen Sprachen gibt. Hier finden verschiedenste Formen der AVT Anwendung (vgl. De Wille2024).

Den zweiten großen Bereich in der AVT bilden *intralinguale* Formen der Translation. Der primäre Zweck intralingualer audiovisueller Translate besteht in der Reduktion von Barrieren für Menschen, die nicht in der Lage sind, audiovisuelle Medien über den visuellen oder den auditiven Kanal zu rezipieren. Es handelt sich folglich um eine Übersetzungen innerhalb einer Sprache. Das größte Betätigungsfeld stellt hierbei die Untertitelung für Hörgeschädigte dar (z.B. Mälzer/Wünsche2018). Untertitel, die nicht im Vorhinein vorproduziert sind, werden als *Live-Untertitel* bezeichnet. Eine weitere intralinguale Translationstechnik ist die Audiodeskription für Sehgeschädigte (vgl. Maszerowska/Matamala/Orero2014).

3 Möglichkeiten der Digitalisierung in der AVT

Zunächst einmal sind zahlreiche Schritte in der AVT insofern ohnehin digital, als sie typischerweise mithilfe spezieller Softwares durchgeführt werden. Dies wird am deutlichsten, wenn wir an die Untertitelung denken – interlinguale Untertitel werden in der Regel mit speziellen Untertitelungssoftwares wie Subtitle Edit, Oona, EZTitles oder Aegisubs erstellt.² Von daher wäre es beinahe irreführend

¹Weitere Synchronisationsländer in Europa sind beispielsweise Frankreich und Spanien.

²Subtitle Edit stellt laut der European Language Industry Study aktuell sogar das meistgenutzte Untertitelungstool dar (ELIS2024).

zu behaupten, dass die Digitalisierung in der AVT an sich etwas gänzlich Neues darstellt. Was hier vor allem gemeint ist, ist die Automatisierung von Arbeitsschritten, beispielsweise durch den Einsatz von Künstlicher Intelligenz (KI) zur Übersetzung eines Drehbuchs oder von Untertiteln.

Die in der AVT genutzten Automatisierungsmöglichkeiten sind mannigfaltig, wie bereits die in **Kurch2024** dargestellte (nicht-exhaustive) Liste von KI-gestützten Technologien illustriert: Audiosignalverarbeitung, automatische Spracherkennung, alternative Textgenerierung, Neuronale Maschinelle Übersetzung, Sprachsynthese, Direktübersetzung gesprochener Sprache, automatische Bilderkennung, Bildgenerierung und Bildmanipulation. Nachstehend soll auf einige der genannten Technologien überblickhaft eingegangen werden, wobei im Rahmen einer notwendigen Fokussetzung der Bereich der Maschinellen Übersetzung eingehender betrachtet wird und Spracherkennung bzw. Sprachsynthese jeweils punktuell behandelt werden

Seit die MÜ durch die Nutzung Neuronaler Netzwerke enorm an Qualität gewonnen hat, eignet sie sich auch zunehmend für mehr oder weniger kreative Übersetzungsleistungen (vgl. de los Reyes Lozano/**Mejímas-Climent2023**: 3). Sie wird in der AVT insbesondere für die Übersetzung von Untertiteln verwendet (vgl. z.B. Jaki/Bolz/Röther im Dr.), findet jedoch auch Einsatz in anderen Feldern, zum Beispiel bei der Übersetzung von Drehbüchern oder von Audiodeskriptionsskripten (vgl. Vercauteren/Reviers/**Steyaert2021**). Wir sprechen hier zunächst von einer MÜ von Text zu Text.

Ein Forschungszweig, der sich im Speziellen mit der Analyse und/oder Verbesserung der Qualität von mittels MÜ erstellten Untertiteln befasst, hat insbesondere im Zuge des europäischen Projekts SUMAT Aufwind erfahren. Im Projekt wurde schwerpunktmäßig die Qualitäts- und Produktivitätssteigerung von maschinell übersetzten Untertiteln in neun Sprachen angestrebt (vgl. **Etchegoyen2014**). Hier kamen jedoch noch keine Neuronalen Netzwerke ins Spiel; stattdessen wurde auf Basis statistischer MÜ gearbeitet. Mit EMMA und TraMOOC starteten zwei weitere internationale Projekte, die es zum Ziel hatten, die Untertitelung mit Hilfe von MÜ voranzubringen, beide am Beispiel von Lehrvideos (vgl. de los Reyes Lozano/**Mejímas-Climent2023**: 4f und **Georgakopoulou2019**: 526-528 für einen Überblick an Projekten). Seither bringen insbesondere die großen computerlinguistischen Konferenzen eine Vielzahl von Versuchen zur Steigerung der Übersetzungsqualität hervor.

Eine Schwierigkeit bei der Untertitelung mithilfe von MÜ-Technologien, die ausschließlich die sprachliche Komponente berücksichtigen, liegt darin, der Notwendigkeit der Kürzung in der Untertitelung nachzukommen. Zwar ist das Ausgangsprodukt häufig ebenfalls ein Untertitel, jedoch handelt es sich hierbei oft um Verbatim-

Untertitel, also um Texte ohne die erforderlichen Kürzungen, die bei der interlingualen Untertitelung zwangsläufig zum Tragen kommen müssen. Bei der Humanübersetzung ergeben sich viele Kürzungsmöglichkeiten infolge intersemiotischer Redundanzen, über die AVT-Produkte von Natur aus verfügen und die bewirken, dass meist nicht der vollständige Text übertragen werden muss (vgl. Reinart2018: 136). Allerdings beschäftigt sich die Forschung bereits seit Jahren mit der Frage, wie der Output von MÜ so gestaltet werden kann, dass Aspekte wie Textmenge, Standzeiten, Umbrüche usw. berücksichtigt werden können (vgl. z.B. Lakew/Gangi/Federico2019, Matusov/Wilken/Georgakopoulou2019). Auch die Übersetzung mittels LLMs, so insbesondere mit dem häufig genutzten ChatGPT, birgt durchaus die Option, mithilfe von Promptings Richtlinien zu implementieren. Insbesondere bei ChatGPT ist jedoch zu beachten, dass sich aufgrund der Urheberrechte oft rechtliche Probleme ergeben, die eine Nutzung des Tools für manche Zwecke von vornherein ausschließen.

Bei der Qualität des automatisiert erzeugten Outputs gilt, dass die Qualität des Ergebnisses maßgeblich von den verfügbaren Daten abhängt. Dies bedeutet im Fall der Untertitelung zum Beispiel, dass eine Übersetzung aus einer relativ seltenen Sprache ins Deutsche, wofür naturgemäß weniger Sprachdaten vorliegen, sehr wahrscheinlich zu einem schlechteren Ergebnis führen wird. Zu bedenken ist allerdings, dass in vielen Fällen mit (vorgespotteten) englischsprachigen Templates³ gearbeitet wird. Ein beispielsweise koreanisches Filmprodukt wird also meist nicht als solches ins Deutsche übersetzt, sondern über die Mittlersprache Englisch, die dann als Vorlage für verschiedene Sprachversionen genutzt wird – folglich liegt eine indirekte oder auch Pivot- bzw. Relais-Übersetzung vor (vgl. Kluge/Camacho/Lins2022). In der Regel ist bei allen angewendeten Verfahren auf den sog. *Human in the Loop* mitnichten zu verzichten, sondern es wird auf verschiedene Schritte des Post-Editings (vgl. O'Brien2022) zurückgegriffen, um die Qualität der maschinell erstellten Erzeugnisse zu optimieren. Dennoch kommen Hagström und Pedersen2022 anhand einer Korpusstudie zu dem Schluss, dass sich die seit den 2020ern verstärkt eingesetzte MÜ von Untertiteln negativ auf deren Qualität auswirkt:

[T]he 2020s subtitles in our material were more fast-paced, more oral, less cohesive, less complete, and constructed with less meticulous punctuation and line-breaks than their predecessors. They were also of significantly lower quality in all areas investigated.

³Ein Template kann definiert werden als „subtitle [FB01?]le with time codes and dialogue segmented into subtitle chunks“ (Oziemblewska/Szarkowska2022: 433). Die enthaltenen Time Codes können entweder adaptierbar sein oder nicht – im ersten Fall spricht man von *unlocked*, im zweiten Fall von *locked* Templates (vgl. Oziemblewska/Szarkowska2022: 435).

Eine Steigerung von *Text-to-Text*-Systemen stellen Systeme dar, die verschiedene Modelle vereinen und somit nicht nur Teile des Translationsprozesses (also beispielsweise die interlinguale Übersetzung durch MÜ oder die Transkription gesprochener Sprache per Spracherkennung), sondern den gesamten Prozess automatisieren. Wurde dies zunächst vor allem über ein Kaskadenmodell (also bspw. einer Kombination von Spracherkennung und MÜ) erreicht, werden heute direkte Ansätze immer gängiger (vgl. BentivogliEtAl2021 zu einem Vergleich der Ansätze). So arbeiten beispielsweise PapiEtAl2023 für sieben Sprachenpaare mit „direct speech-to-text translation models to fully automatize the subtitling process, including translation, segmentation into subtitles, and timestamp estimation“ (PapiEtAl2023: 1368). Zu erwähnen sind auch moderne multimodale Translationsmodelle, die zur Optimierung von maschinell übersetzten Texten visuelle Informationen mit einfließen lassen (z.B. LiEtAl2023).

In anderen Bereichen der AVT ist die Verwendung von KI-gestützten Technologien noch weniger Standard als in der Untertitelung, jedoch haben die letzten Jahre hier ebenfalls rapide Fortschritte gezeigt (vgl. Kurch2024 für einen aktuellen Überblick über KI-gestützt erstellte Arten von AVT-Produktionen). Eine zunehmende Rolle spielt beispielsweise Sprachsynthese, die vor allem in Synchronisation, Audiodeskription und Voice-Over-Übersetzung verwendet wird (Georgakopoulou 529f für einen Überblick). Zur Erzeugung intralingualer Untertitel wird häufig mit Spracherkennung gearbeitet, besonders bei Live-Untertiteln für Hörgeschädigte (Romero-Fresco/Fresno2023). Auch abseits der Untertitelung finden wir vermehrt Versuche, den kompletten Translationsprozess mittels multimodal arbeitender Systeme zu automatisieren. Dies betrifft beispielsweise die vollautomatische Erstellung von Audiodeskriptionen (z.B. in Wang2021). Auch bei der Synchronisation werden volumnfähige Möglichkeiten genutzt, angefangen bei der MÜ von Drehbüchern über die Imitation der originalen Sprecherstimme bis hin zur Anpassung der ursprünglichen Lippenbewegungen auf den Zieltext (z.B. FedericoEtAl2020; vgl. Granell/Chaume2023: 28-30 oder Baños2023 für einen Überblick).

4 Digitalisierung und Berufsbild

Ziel der zunehmenden Digitalisierung verschiedener Arbeitsschritte in der AVT ist in erster Linie eine Arbeitserleichterung, die mit schnelleren Produktionszeiten einhergeht. Bei tiefgreifenden Veränderungen wie den in Abschnitt 3 beschriebenen stellt sich jedoch auch immer die Frage, wie sich diese Entwicklungen auf die Berufstätigen auswirken. Dies ist umso relevanter, als die zahlreichen Möglichkeiten der Digitalisierung auch maßgeblich auf die Workflows in der Branche Einfluss

nehmen (vgl. z.B. Reinart2018: 146 zur Untertitelung). Neben den oben genannten Entwicklungen gehört hierzu auch das immer häufiger genutzte cloudbasierte Arbeiten (vgl. Bolaños-García-Escribano/Díaz-Cintas2020), das u.a. den Vorteil birgt, relativ einfach Tools wie Terminologiemanagement oder MÜ integrieren zu können (vgl. Oziemblewska/Szarkowska2022: 435). Cloudbasiertes Arbeiten im Vergleich zur Nutzung von Desktop-Lösungen ist besonders beliebt bei internationalen Unternehmen (vgl. Oziemblewska/Szarkowska2022: 444). Wie Tardel2023 für die inter- und intralinguale Untertitelung eindrücklich demonstriert, handelt es sich generell um zunehmend ausdifferenzierte Arbeitsprozesse, die essentiell davon abhängen, welche Grade und Möglichkeiten der Automatisierung an welchen Stellen eingesetzt werden.

Der Fokus auf die Perspektive der Translator*innen anstelle einer Produktanalyse wird auch als *translator studies* bezeichnet, in denen Translator*innen als Individuen mit ihren persönlichen Einstellungen in den Vordergrund gerückt werden (Risku2024) und auf eine *humanization* der Übersetzungswissenschaft abgezielt wird (Risku2024). Wie Künzli2024 betont, „ist das Feld bezüglich Audiovisual Translator Studies und insbesondere Subtitler Studies noch weitgehend unbestellt“. Aus diesem Grund ist es sinnvoll, dass sich die Forschung zur Digitalisierung in der AVT auch mit der Perspektive der Translator*innen befasst. Hierbei findet eine Reihe wissenschaftlicher Methoden Anwendung, die Aufschlüsse zur Arbeitszufriedenheit und zum wahrgenommenen Berufsbild liefern können, so zum Beispiel Umfragen (die sicherlich häufigste Methode zur Erhebung der Arbeitszufriedenheit), Interviews und ferner E-Mail-Interviews.

Da Arbeitsbedingungen immer kontextabhängig sind, analysieren einige Untersuchungen die Arbeitszufriedenheit von audiovisuellen Translator*innen vor einem konkreten nationalen Hintergrund. Eine umfangreiche nationale Fragebogenstudie mit 459 Literatur-, Fach- und Filmübersetzer*innen, durchgeführt von Ruokonen und Mäkisalo2018, beleuchtet beispielsweise den finnischen Kontext. Bei den 57 befragten audiovisuellen Translator*innen zeigte sich eine Korrelation zwischen Arbeitszufriedenheit und der Höhe des Einkommens sowie, allerdings bei allen drei Gruppen, eine proportional zur Länge der Arbeitserfahrung empfundene Wertigkeit der eigenen Arbeit – also je länger die Arbeitserfahrung, desto höher die empfundene Wertigkeit (vgl. Ruokonen und Mäkisalo2018: 11f). Ein interessanter Befund betrifft die Tatsache, dass von den drei Gruppen die audiovisuellen Translator*innen den am niedrigsten empfundenen Status der drei Übersetzer*innengruppen besaßen (vgl. Ruokonen und Mäkisalo2018: 14), was wiederum eine vergleichsweise niedrige Arbeitszufriedenheit vermuten lässt. Weitere nationale Studien sind beispielsweise Jankowska2012 zur AVT in Polen, Nikolić2010

zur Untertitelung in Kroatien und Silvester2022 zu einer Gruppe von Untertitler*innen in Frankreich. Wie Künzli2024 zu seiner Aufarbeitung des Forschungsstandes resümiert, zeigen die bisherigen Forschungsergebnisse, dass die Wahrnehmung der Arbeitsbedingungen vielfältig ist.

Der deutsche Kontext wird unter anderem in der bekanntesten und größten Studie abgedeckt: Die European Language Industry Study (ELIS) ist eine umfangreiche Fragebogenstudie, die jährlich weltweit durchgeführt wird.⁴ Für das Jahr2024 nahmen 17 Unternehmen und 57 Einzelpersonen aus Deutschland teil (vgl. ELIS2024: 6). Die Ergebnisse der Umfragerunde des Jahres2024 legen nahe, dass die Übersetzenden (und darunter insbesondere die älteren Translator*innen) durch den verstärkten Einsatz von KI bezüglich der Rentabilität des Freelance-Übersetzens pessimistischer geworden sind (vgl. ELIS2024: 22). Diese Sorge wird von den Befragten häufig explizit mit dem vermehrten Einsatz von MÜ bzw. anderen Technologien sowie mit einer Verschiebung des Aufgabenbereichs in Richtung weniger gut bezahlter Tätigkeiten im Post-Editing assoziiert (vgl. ELIS2024: 26). Allerdings deckt ELIS verschiedenste Translationsbranchen ab, so dass die Ergebnisse nicht exklusiv die AVT betreffen. Insgesamt enthält die Studie nur wenige Fragen, die gezielt in Bezug auf die AVT erhoben werden (so z.B. zu den verwendeten Tools, vgl. Fußnote 2).

Eine weitere, kleinere Fragebogenstudie (59 Teilnehmende), die auf die Nutzung von KI-gestützten Systemen und die damit verbundenen Einstellungen von Translator*innen speziell in der audiovisuellen Translation abzielt, ist Jaki/Bolz/Röther (im Dr.). Im Unterschied zu ELIS werden in dieser Untersuchung ausschließlich auf AVT spezialisierte Übersetzer*innen aus dem deutschsprachigen Raum befragt. Was die Einstellungen der teilnehmenden Translator*innen zu KI-gestützten Technologien sowie das damit verbundene Bild der AVT in Zeiten der Digitalisierung betrifft, sind die Befindlichkeiten auf den ersten Blick heterogen: Laut den Befragten ist das Feld positiver und negativer Einstellungen zu KI-gestützten Tools in der AVT recht ausgewogen. Wenn man jedoch nach der wahrgenommenen Qualität von Translates fragt, die KI-gestützt (besonders durch MÜ) erzeugt werden, erscheint das Bild auffallend negativ.⁵ Ähnlich wie in ELIS2024 wird in Jaki/Bolz/Röther (im Dr.) die Sorge vor Preisdumping aufgrund von vermehrten, und schlecht bezahlten, Post-Editing-Aufgaben verbalisiert.⁶ Auch wird vere-

⁴Vgl. die verschiedenen Durchgänge auf <https://elis-survey.org/>.

⁵Läubli/Orrego-Carmora2017 erzielen mit Hilfe der Analyse von Social-Media-Daten ähnliche Befunde.

⁶Vgl. auch Szarkowska/Díaz-Cintas/Gerber-Morón (2021: 666/669), die im Rahmen einer internationalen Befragung zum Qualitätsempfinden von professionellen Untertitler*innen aus 27 Ländern beobachten, dass einige der Befragten die fallenden Preise für den Rückgang der Un-

inzelt der Verlust kreativen Arbeitens beklagt. Teilweise besteht sogar die Befürchtung, dass Humanübersetzer*innen zukünftig überhaupt nicht mehr benötigt werden.

Noch spezialisierter ist Künzlis²⁰²⁴ Untersuchung mit 19 Untertitler*innen (ebenfalls aus dem deutschsprachigen Markt), in der die Daten mithilfe von E-Mail-Interviews erhoben werden. Im Gegensatz zu zahlreichen anderen Untersuchungen liegt der Fokus in Künzli²⁰²⁴ auf Faktoren, die sich positiv auf die Arbeitszufriedenheit der befragten Personen auswirken. Zu diesen Faktoren gehören beispielsweise die Aspekte Arbeitssicherheit durch regelmäßige Auftragserteilungen, die Möglichkeit zum direkten Austausch im Projektteam und zur Teilhabe an Entscheidungen sowie das Feedback zu den Untertiteln durch den Auftraggebernden. Ein weiterer Punkt, der positiv mit der Arbeitszufriedenheit der Befragten korreliert und sich auf die Fragmentierung der Arbeitsprozesse durch eine gesteigerte Digitalisierung bezieht, ist die Möglichkeit, am gesamten Untertitelungsprozess mitwirken zu können (also zum Beispiel keinen vorgespotteten Text zu erhalten).

5 Ausblick

Die Digitalisierung in der AVT, „eine[m] der am schnellsten wachsenden Forschungsfelder in der Übersetzungswissenschaft“ (Künzli/Kaindl²⁰²⁴: 13), bringt neue Herausforderungen für alle der drei Grundperspektiven der Translationsforschung: die Produktforschung, die Produzenten- bzw. Prozessforschung und die Rezeptionsforschung. Die *Produktforschung* macht, und dies nicht nur in der AVT-Forschung, sicherlich den größten Teil der Untersuchungen aus. Waren lange Zeit Themen wie beispielsweise die Synchronisation von Humor oder die Untertitelung von Kulturspezifika „klassische“ Themen für studentische Abschlussarbeiten wie auch wissenschaftliche Artikel, tritt nun verstärkt die Performanz von MÜ bzw. die Übersetzung durch LLMs in den Fokus derartiger wissenschaftlicher Erzeugnisse. Mindestens bis der rapide Anstieg in der Qualität des Outputs solcher Technologien gesättigt ist, wird sich die Forschung ausführlich mit der Translationsqualität in Bezug auf verschiedenste sprachliche bzw. textuelle Merkmale, in Bezug auf verschiedenste Textsorten und Sprachenpaare befassen. Da gerade im Bereich der Untertitelung eine starke Tendenz zur MÜ besteht (vgl. z.B. Jaki/Bolz/Röther im Dr.), ist zu erwarten, dass dieser Bereich, der bereits im Vergleich zu den anderen Feldern am häufigsten erforscht wird, auch weiterhin stark im Zentrum des Interesses von Forschenden stehen wird. Eine vollautomatisierte Produktion

tertitelqualität verantwortlich machen.

mittels KI-Technologien erzielt in der AVT bislang jedoch noch keine zufriedenstellenden Ergebnisse (vgl. Kurch2024: 366). Daher wird die Forschung auch die Frage, wie die Performanz der Systeme verbessert werden kann, beispielsweise durch Pre-Editing (vgl. z.B. Hiraoka/Yamada2019), begleiten, und zwar insbesondere bei den intersemiotisch arbeitenden Technologien (z.B. *Image-to-Text*), bei denen im Vergleich zur *Text-to-Text*-Übersetzung noch größerer Verbesserungsbedarf besteht. Interessant sind in diesem Zusammenhang auch sogenannte *Shared Tasks*, die die Möglichkeit bieten, Systeme zur automatisierten Erfüllung einer Aufgabe im Bereich der AVT wettbewerbsmäßig zu erproben, wie zum Beispiel die verschiedenen Tracks der *International Conference on Spoken Language Translation (IWSLT) 2024*⁷. Vermutlich wird in den nächsten Jahren die Entwicklung und Evaluation von vollautomatischen oder gar multimodalen Systemen noch deutlich zunehmen. Wie sich bisher gezeigt hat, besteht auch eine Schwierigkeit in der Evaluation solcher, aber auch einfacher *Text-to-text*-Systeme. Die Evaluation von Translationsqualität für MÜ in der AVT stellt schließlich im Vergleich zu anderen Translationsarten sehr spezifische Anforderungen (vgl. z.B. BurchardtEtAl2016).

Vergleichsweise weniger Aufmerksamkeit als die Produktperspektive erhalten die *Produktions-* und *Rezeptionsperspektive*, wenngleich sie in den letzten Jahren einen deutlichen Aufschwung erfahren haben (vgl. Díaz-Cintas/Szarkowska2020). Wie die Umfrageergebnisse in Jaki/Bolz/Röther (im Dr.) nahelegen, ist die Vorstellung der Translator*innen von den Qualitätsansprüchen der Rezipierenden eher negativ. Alleine schon deshalb muss künftige Forschung noch verstärkt zeigen, wie automatisiert erstellte audiovisuelle Translate von der Zielgruppe wahrgenommen werden. Dazu gehört die Frage, unter welchen Umständen welchen Arten von Fehlern zu Missverständnissen führen oder welche Fehler vielleicht nicht einmal wahrgenommen werden – die Untersuchung von Qiu und Pym2024 legt beispielsweise nahe, dass Zuschauer*innen eher geneigt sind, Fehlern gegenüber „gnädig“ zu sein, wenn sie den Inhalt des Produkts als sehr unterhaltsam empfinden. Eine weitere Fallstudie mit finnischen und deutschen Untertiteln kommt zu dem Schluss, dass die human erzeugten Untertitel von den Rezipient*innen zwar als besser eingestuft wurden, aber es keinen Grund zur Annahme gibt, dass die Rezipient*innen längere Zeit für das Lesen maschineller erstellter Untertitel aufwenden (Schierl2023). Eine Studie von Calvo-Ferrer2023 mit spanischen Studierenden der Übersetzungs- und Dolmetschwissenschaft kommt sogar zu dem Ergebnis, dass nur die fortgeschrittenen Studierenden den Unterschied zwischen men-

⁷Einen Überblick über die verschiedenen Shared Tasks (z.B. zur automatisierten Synchronisation oder Untertitelung) erhält man auf <https://iwslt.org/2024/>.

schlich erstellten und von ChatGPT generierten Untertiteln erkennen. Diese Befunde scheinen in einem scharfen Kontrast zu den Empfindungen von professionellen Translator*innen zu stehen, die in Jaki/Bolz/Röther (im Dr.) fast durchweg eine sehr negative Einstellung zur Qualität maschinell erstellter Übersetzungen äußern. Dies legt zwei Dinge nahe: Ersten muss die Robustheit der oben genannten Befunde über einzelne Fallstudien hinaus, auch vor dem Hintergrund der Weiterentwicklung der Technologien, weiter getestet werden. Obwohl in diesem Bereich die Forschung zur Untertitelung noch dominiert, sind auch für die anderen Bereiche der AVT, in denen sich für die Automatisierung der Arbeitsprozesse eine immer höhere technologische Performanz beobachten lässt, künftig interessante Ergebnisse zu erwarten. Zweitens, und hiermit soll dieser Beitrag schließen, lenkt der beschriebene Kontrast das Augenmerk auch auf den Bereich der Produktionsforschung.

Hier eröffnet sich ein weites Feld an Perspektiven: Zunächst gilt es die neuen, diversifizierten und komplexen Workflows zu erforschen, die sich durch die unterschiedlichen Grade und Formen der Automatisierung in den verschiedenen Bereichen der AVT ergeben (vgl. Tardel2023 für die Untertitelung) – hier besteht gerade auch Bedarf außerhalb der Untertitelung, wo die Automatisierung zum Teil noch nicht in den Alltag der Translator*innen integriert ist (so werden Audiodeskriptionen teilweise noch in Word-Dokumenten geliefert). Überdies ist es sinnvoll, durch Prozessstudien mit Hilfe von empirischen Methoden wie Eye-Tracking, Keylogging oder Bildschirmaufzeichnungen (vgl. Tardel/Hansen-Schirra2024: 294-296) zu untersuchen, wie die Translator*innen mit MÜ, LLMs, automatisierter Spracherkennung und andere Technologien interagieren, um zu erkennen, welche Herausforderungen und Erleichterungen sie bringen, aber zuletzt auch, um hieraus Lehren für die Ausbildung von Translator*innen zu ziehen. Schließlich bleibt abzuwarten, wie sich die Einstellungen der Translator*innen zu KI-gestützten Technologien vor dem Hintergrund der Entwicklung dieser Technologien und der Routinisierung in der Arbeit mit ihnen im AVT-Sektor in den nächsten Jahren entwickeln werden. Ob die Technologien in Zukunft als echte Hilfe oder weiterhin als existentielle Bedrohung angesehen werden, ist aber natürlich auch mit abhängig von der Preisentwicklung in den Branchen. In jedem Fall ist es jedoch auch sinnvoll zu beleuchten, inwiefern die Repräsentation durch und der Austausch in Berufsverbänden wie dem AVÜ (vgl. Nagel2024) in einem Feld, das maßgeblich von individuell arbeitenden Freelancer*innen geprägt ist, auf die Einstellungen und das Berufsbild womöglich sogar positiven Einfluss nehmen kann.