

Kapitel 12

Digitalisierung in der Audiovisuellen Translation

Sylvia Jaki

Universität Hildesheim

Der Beitrag nimmt das Feld der Audiovisuellen Translation in den Blick, das trotz seiner kreativen Komponente in den letzten Jahren mehr und mehr durch verschiedene Formen der Digitalisierung geprägt wird. Er bietet einen Einblick sowohl in die verschiedenen Bereiche der Audiovisuellen Translation als auch in die Bandbreite an Möglichkeiten der Digitalisierung mit Hilfe von Maschinellem Übersetzung, Spracherkennung u.a. Nicht zuletzt legt der Beitrag den Fokus auch auf die Translator*innen selbst und beleuchtet, welchen Einfluss die zunehmende Digitalisierung auf deren Berufsbild besitzt.

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 mit Hilfe 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 zahlreichen 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 Maschinellem Übersetzung (MÜ), können Produktionszeiten womöglich verringert



werden.¹ 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-Cintas 2013), sondern die Entwicklungen, die die Branche 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. Jüngst 2020 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 (Jüngst 2020: 1).

Wenn Díaz-Cintas (2013: 119) 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

¹Bei der Frage, inwieweit dies wirklich der Fall ist, herrscht in der AVT-Branche jedoch Uneinigkeit bzw. Skepsis (vgl. z. B. Agnetta 2025 oder Jaki u. a. 2024).

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 Georgakopoulou 2019: 521–522). 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 u. a. 2022: 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. Reinart 2018: 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.² 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 die Lautstärke des Originaltons gesenkt, 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 abzeichnen scheint (vgl. Flis & Szarkowska 2024). Ein Gebiet, das in jüngerer Zeit verstärkt Aufmerksamkeit erfährt, ist das der Videospiellokalisierung, für das es ebenfalls eine hohe Nachfrage im Bereich der Übersetzung zwischen verschiedenen Sprachen gibt. Hier finden verschiedenste Formen der AVT Anwendung (vgl. De Wille 2024).

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 & Wünsche 2018). Untertitel, die nicht im Vorhinein vorproduziert sind, werden als *Live-Untertitel* bezeichnet. Eine weitere intralinguale Translationstechnik ist die Audiodeskription für Sehgeschädigte (vgl. Maszerowska u. a. 2014).

²Weitere Synchronisationsländer in Europa sind beispielsweise Frankreich und Spanien.

3 Möglichkeiten der Digitalisierung in der AVT

Zunächst einmal sind zahlreiche Schritte in der AVT insofern ohnehin digital, als sie typischerweise mit Hilfe 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, Ooona, EZTitles oder AegiSubs erstellt.³ Von daher wäre es beinahe irreführend 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 Kurch (2024: 363) 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 überblickshaft eingegangen werden, wobei im Rahmen einer notwendigen Fokussierung 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 & Mejías-Climent 2023: 3). Sie wird in der AVT insbesondere für die Übersetzung von Untertiteln verwendet (siehe z. B. Jaki u. a. 2024), findet jedoch auch Einsatz in anderen Feldern, zum Beispiel bei der Übersetzung von Drehbüchern oder von Audiodeskriptionsskripten (vgl. Vercauteren u. a. 2021). 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. Etchegoyhen u. a. 2014). Hier kamen jedoch noch keine Neuronalen Netzwerke ins Spiel; stattdessen wurde auf Basis statistischer MÜ gearbeitet. Mit EMMA und TraMOOC

³Subtitle Edit stellt laut der European Language Industry Study aktuell sogar das meistgenutzte Untertitelungstool dar (ELIS Research 2024: 39).

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 & Mejías-Climent 2023: 4–5 und Georgakopoulou 2019: 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 mit Hilfe 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. Reinart 2018: 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 u. a. 2019, Matusov u. a. 2019). Auch die Übersetzung mittels LLMs, so insbesondere mit dem häufig genutzten ChatGPT, birgt durchaus die Option, mit Hilfe von Promptings Richtlinien zu implementieren (vgl. z. B. Lapshinova-Koltunski u. a. 2025). 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 Templates⁴ 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 u. a. 2022). In der Regel ist bei allen angewendeten Verfahren auf den sog.

⁴Ein Template kann definiert werden als „subtitle file with time codes and dialogue segmented into subtitle chunks“ (Oziembewska & Szarkowska 2022: 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. Oziembewska & Szarkowska 2022: 435).

Human in the loop mitnichten zu verzichten, sondern es wird auf verschiedene Schritte des Post-Editings (vgl. O'Brien 2022) zurückgegriffen, um die Qualität der maschinell erstellten Erzeugnisse zu optimieren. Dennoch kommen Hagström & Pedersen (2022: 223) 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.

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. Bentivogli u. a. 2021 zu einem Vergleich der Ansätze). So arbeiten beispielsweise Papi u. a. (2023) für sieben Sprachenpaare mit „direct speech-to-text translation models to fully automatize the subtitling process, including translation, segmentation into subtitles, and timestamp estimation“ (Papi u. a. 2023: 1368). Zu erwähnen sind auch moderne multimodale Translationsmodelle, die zur Optimierung von maschinell übersetzten Texten visuelle Informationen mit einfließen lassen (z. B. Li u. a. 2023).

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. Kurch 2024 für einen aktuellen Überblick über KI-gestützt erstellte Arten von AVT-Produktionen). Eine zunehmende Rolle spielt beispielsweise die Sprachsynthese, die vor allem in Synchronisation, Audiodeskription und Voice-Over-Übersetzung verwendet wird (Georgakopoulou 2019: 529f. für einen Überblick). Zur Erzeugung intralingualer Untertitel wird häufig mit Spracherkennung gearbeitet, besonders bei Live-Untertiteln für Hörgeschädigte (Romero-Fresco & Fresno 2023). 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 Wang u. a. 2021). Auch bei der Synchronisation werden vollumfängliche 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. Federico u. a. 2020; vgl. Granell & Chaume 2023: 28-30 oder Baños 2023 für einen Überblick).

4 Digitalisierung und Berufsbild

Ziel der zunehmenden Digitalisierung verschiedener Arbeitsschritte in der AVT ist in erster Linie das Erzielen kürzerer Produktionszeiten. Bei tiefgreifenden Veränderungen wie den in Abschnitt 3 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. Reinart 2018: 146 zur Untertitelung). Neben den oben genannten Entwicklungen gehört hierzu auch das immer häufiger genutzte cloudbasierte Arbeiten (vgl. Bolaños-García-Escribano & Díaz-Cintas 2020), das u. a. den Vorteil birgt, relativ einfach Tools wie Terminologiemanagement oder MÜ integrieren zu können (vgl. Oziemblewska & Szarkowska 2022: 435). Cloudbasiertes Arbeiten im Vergleich zur Nutzung von Desktop-Lösungen ist besonders beliebt bei internationalen Unternehmen (vgl. Oziemblewska & Szarkowska 2022: 444). Wie Tardel (2023) 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 (Risku 2024: 67) und auf eine *humanization* der Übersetzungswissenschaft abgezielt wird (Risku 2024: 72). Wie Künzli (2024: 104) 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), persönliche 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 & Mäkisalo (2018), 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 (vgl. Ruokonen & Mäkisalo 2018: 11–12). 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 & Mäkisalo 2018: 14), was wiederum eine vergleichsweise niedrige Arbeitszufriedenheit vermuten lässt. Weitere nationale Studien sind beispielsweise Jankowska (2012) zur AVT in Polen, Nikolić (2010) zur Untertitelung in Kroatien und Silvester (2022) zu einer Gruppe von Untertitler*innen in Frankreich. Wie Künzli (2024: 106) 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.⁵ Für das Jahr 2024 nahmen 17 Unternehmen und 57 Einzelpersonen aus Deutschland teil (vgl. ELIS Research 2024: 6). Die Ergebnisse der Umfragerunde des Jahres 2024 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. ELIS Research 2024: 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. ELIS Research 2024: 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 3).

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 u. a. (2024). 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

⁵Vgl. die verschiedenen Durchgänge auf <https://elis-survey.org/>.

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 Translaten fragt, die KI-gestützt (besonders durch MÜ) erzeugt werden, erscheint das Bild auffallend negativ.⁶ Ähnlich wie in ELIS Research (2024) wird in Jaki u. a. (2024) die Sorge vor Preisdumping aufgrund von vermehrten, und schlecht bezahlten, Post-Editing-Aufgaben verbalisiert.⁷ Auch wird vereinzelt der Verlust kreativen Arbeitens beklagt. Teilweise besteht sogar die Befürchtung, dass Humanübersetzer*innen zukünftig überhaupt nicht mehr benötigt werden.

Noch spezialisierter ist die Untersuchung von Künzli (2024) mit 19 Untertitler*innen (ebenfalls aus dem deutschsprachigen Markt), in der die Daten mit Hilfe von E-Mail-Interviews erhoben werden. Im Gegensatz zu zahlreichen anderen Untersuchungen liegt der Fokus in Künzli (2024) 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 Auftraggebenden. 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 & Kaindl 2024: 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.

⁶Läubli & Orrego-Carmona (2017) erzielen mit Hilfe der Analyse von Social-Media-Daten ähnliche Befunde.

⁷Vgl. auch Szarkowska u. a. (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 Untertitelqualität verantwortlich machen.

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 (siehe z. B. Jaki u. a. 2024), 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 mittels KI-Technologien erzielt in der AVT bislang jedoch noch keine zufriedenstellenden Ergebnisse (vgl. Kurch 2024: 366). Daher wird die Forschung auch die Frage, wie die Performanz der Systeme verbessert werden kann, beispielsweise durch Pre-Editing (vgl. z. B. Hiraoka & Yamada 2019), 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)* 2024.⁸ 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. Burchardt u. a. 2016).

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 & Szarkowska 2020). Wie die Umfrageergebnisse in Jaki u. a. (2024) nahelegen, ist die Vorstellung der Translator*innen von den Qualitätsansprüchen der Rezipierenden eher negativ. Alleine schon deshalb muss künftige Forschung noch verstärkt zeigen,

⁸Einen Überblick über die verschiedenen *Shared Tasks* (z. B. zur automatisierten Synchronisation oder Untertitelung) erhält man auf <https://iwslt.org/2024/>.

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 & Pym (2024) 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 (Schierl 2023: 50). Eine Studie von Calvo-Ferrer (2023) mit spanischen Studierenden der Übersetzungs- und Dolmetschwissenschaft kommt sogar zu dem Ergebnis, dass nur die fortgeschrittenen Studierenden den Unterschied zwischen menschlich 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 u. a. (2024) fast durchweg eine sehr negative Einstellung zur Qualität maschinell erstellter Übersetzungen äußern. Dies legt zwei Dinge nahe: Erstens 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. Tardel 2023 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-Schirra 2024: 294-296) zu untersuchen, wie die Translator*innen mit MÜ, LLMs, automatisierter Spracherkennung und anderen Technologien interagieren, um zu erkennen, welche Herausforderungen und Erleichterungen sie bringen, aber zuletzt auch, um hieraus Lehren für die Ausbildung von Transla-

tor*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 existenzielle 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. Nagel 2024) in einem Feld, das maßgeblich von individuell arbeitenden Freelancer*innen geprägt ist, auf die Einstellungen und das Berufsbild womöglich sogar positiven Einfluss nehmen kann.

Literatur

- Agnetta, Marco. 2025. Zu Einschätzung und Akzeptanz von KI-Technologien in der AVT-Branche: Die Perspektive der Verbände. *Lebende Sprachen* 70(1). 149–176. DOI: 10.1515/les-2025-0018.
- Baños, Rocío. 2023. Key challenges in using automatic dubbing to translate educational YouTube videos. *Linguistica Antverpiensia, New Series: Themes in Translation Studies*, 22. 61–79.
- Bentivogli, Luisa, Mauro Cettolo, Marco Gaido, Alina Karakanta, Alberto Martinnelli, Matteo Negri & Marco Turchi. 2021. *Cascade versus direct speech translation: Do the differences still make a difference?* <https://arxiv.org/abs/2106.01045> (13 Oktober, 2024).
- Bolaños-García-Escribano, Alejandro & Jorge Díaz-Cintas. 2020. The cloud turn in audiovisual translation. In Łukasz Bogucki & Mikołaj Deckert (Hrsg.), *The Palgrave handbook of audiovisual translation and media accessibility*, 519–544. Cham: Springer International Publishing. https://link.springer.com/10.1007/978-3-030-42105-2_26 (10 Oktober, 2024).
- Burchardt, Aljoscha, Arle Lommel, Lindsay Bywood, Kim Harris & Maja Popović. 2016. Machine translation quality in an audiovisual context. *Target. International Journal of Translation Studies* 28(2). 206–221. DOI: 10.1075/target.28.2.03bur.
- Calvo-Ferrer, José Ramón. 2023. Can you tell the difference? A study of human vs machine-translated subtitles. *Studies in Translation Theory and Practice* 6(32). 1–18. DOI: 10.1080/0907676X.2023.2268149.

- De los Reyes Lozano, Julio & Laura Mejías-Climent. 2023. Beyond the black mirror effect: the impact of machine translation in the audiovisual translation environment. *Linguistica Antverpiensia, New Series: Themes in Translation Studies* 22. 1–19.
- De Wille, Tabea. 2024. Videospielübersetzung. In Alexander Künzli & Klaus Kaindl (Hrsg.), *Handbuch Audiovisuelle Translation: Arbeitsmittel für Wissenschaft, Studium, Praxis*, 161–172. Berlin: Frank & Timme.
- Díaz-Cintas, Jorge. 2013. The technology turn in subtitling. In Marcel Thelen & Barbara Lewandowska-Tomaszczyk (Hrsg.), *Translation and Meaning – Part 9*, 119–132. Maastricht: Maastricht School of Translation & Interpreting, Zuyd University of Applied Sciences Maastricht.
- Díaz-Cintas, Jorge & Agnieszka Szarkowska. 2020. Introduction: Experimental research in audiovisual translation – cognition, reception, production. *The Journal of Specialised Translation* 33. 3–16.
- ELIS Research. 2024. *European Language Industry Survey 2024. Trends, expectations and concerns of the European language industry*. Techn. Ber. <https://elis-survey.org/wp-content/uploads/2024/03/ELIS-2024-Report.pdf> (3 August, 2024).
- Etchegoyhen, Thierry, Lindsay Bywood, Lindsay Fishel, Panayota Georgakopoulou, Jie Jiang, Gerard van Loenhout, Arantza del Pozo, Mirjam Sepesy Maučec, Anja Turner & Martin Volk. 2014. Machine translation for subtitling: A large-scale evaluation. In *LREC Conference Proceedings 2014*, 46–53. http://www.lrec-conf.org/proceedings/lrec2014/pdf/463_Paper.pdf.
- Federico, Marcello, Robert Enyedi, Roberto Barra-Chicote, Ritwik Giri, Umut Isik, Arvinth Krishnaswamy & Hassan Sawaf. 2020. From speech-to-speech translation to automatic dubbing. In *Proceedings of the 17th International Conference on Spoken Language Translation*, 257–264. Online: Association for Computational Linguistics. DOI: 10.18653/v1/2020.iwslt-1.31.
- Flis, Gabriela & Agnieszka Szarkowska. 2024. Voice-over country? Okay, Boomer. How young viewers are disrupting the AVT landscape in Poland. *The Journal of Specialised Translation* 42. 193–216. DOI: 10.26034/cm.jostrans.2024.5989.
- Georgakopoulou, Panayota. 2019. Technologization of AVT. In Luis Pérez-González (Hrsg.), *The Routledge Handbook of Audiovisual Translation*, 526–539. London/New York: Routledge.
- Granell, Ximo & Frederic Chaume. 2023. Audiovisual translation, translators, and technology: From automation pipe-dream to human-machine convergence. *Linguistica Antverpiensia, New Series: Themes in Translation Studies* 22. 20–40.

- Hagström, Hanna & Jan Pedersen. 2022. Subtitles in the 2020s: The influence of machine translation. *Journal of Audiovisual Translation* 5(1). 207–225. DOI: 10.47476/jat.v5i1.2022.195.
- Hiraoka, Yusuke & Masaru Yamada. 2019. Pre-editing plus neural machine translation for subtitling: Effective pre-editing rules for subtitling of TED talks. In *Proceedings of Machine Translation Summit XVII: Translator, Project and User Tracks*, 64–72. Dublin: European Association for Machine Translation. <https://aclanthology.org/W19-6710.pdf>.
- Jaki, Sylvia, Maren Bolz & Sophie Röther. 2024. KI-Technologien in der Audiovisuellen Translation. Umfrageergebnisse aus der deutschen Translationsindustrie. *trans-kom* 2(17). 320–342.
- Jankowska, Anna. 2012. “I do what I like, and I don’t have to go to work every day”: The status quo of audiovisual translators in Poland. In Silvia Bruti & Elena Di Giovanni (Hrsg.), *Audiovisual translation across Europe: An ever-changing landscape*, 35–58. Oxford u.a.: Peter Lang.
- Jüngst, Heike Elisabeth. 2020. *Audiovisuelles Übersetzen: ein Lehr- und Arbeitsbuch*. 2., überarbeitete und erweiterte Auflage (Narr Studienbücher). Tübingen: Narr Francke Attempto.
- Kluge, Bettina, Liliana Camacho & Antonia Lins. 2022. Indirekte Übersetzung, oder: warum auch Linguist*innen wissen sollten, wie Netflix seine Filme untertitelt. *Romanistik in Geschichte und Gegenwart* 28(2). 147–158. DOI: 10.46771/9783967694185_6.
- Künzli, Alexander. 2024. ”Aber ein toller Film reicht, und man ist wieder Feuer und Flamme für den Beruf”. Zur Arbeitszufriedenheit von Untertitel-ExpertInnen. In Mira Kadrić, Waltraud Kolb & Sonja Pöllabauer (Hrsg.), *Translation als Gestaltung: Beiträge für Klaus Kaindl zur translatorischen Theorie und Praxis*, 103–115. Tübingen: Narr Francke Attempto.
- Künzli, Alexander & Klaus Kaindl. 2024. Einleitung. In Alexander Künzli & Klaus Kaindl (Hrsg.), *Handbuch Audiovisuelle Translation: Arbeitsmittel für Wissenschaft, Studium, Praxis* (Audiovisual translation studies), 13–19. Berlin: Frank & Timme.
- Kurch, Alexander. 2024. Technologische Entwicklungen in der AVT. In Alexander Künzli & Klaus Kaindl (Hrsg.), *Handbuch Audiovisuelle Translation: Arbeitsmittel für Wissenschaft, Studium, Praxis* (Audiovisual Translation Studies 2), 361–375. Berlin: Frank & Timme.
- Lakew, Surafel Melaku, Mattia Di Gangi & Marcello Federico. 2019. *Controlling the output length of neural machine translation*. DOI: <https://arxiv.org/abs/1910.10408>.

- Lapshinova-Koltunski, Ekaterina, Sylvia Jaki, Maren Bolz & Merle Sauter. 2025. Human- or machine-translated subtitles: Who can tell them apart? In Catarina Farinha & Marco Gaido (Hrsg.), *Proceedings of Machine Translation Summit 2025, Vol. 1: Research Track*, 496–505. Geneva: ACL. <https://aclanthology.org/2025.mtsummit-1.39/>.
- Läubli, Samuel & David Orrego-Carmona. 2017. When Google Translate is better than some human colleagues, those people are no longer colleagues. In *Proceedings of the 39th Conference Translating and the Computer*, 59–69. London. DOI: 10.5167/UZH-147260.
- Li, Yihang, Shuichiro Shimizu, Chenhui Chu, Sadao Kurohashi & Wei Li. 2023. Video-helpful multimodal machine translation. In *Proceedings of the 2023 Conference on Empirical Methods in Natural Language Processing*, 4281–4299. Singapore: Association for Computational Linguistics. DOI: 10.18653/v1/2023.emnlp-main.260.
- Mälzer, Nathalie & Maria Wünsche. 2018. Untertitel für Hörgeschädigte (SDH). In Christiane Maaß & Isabel Rink (Hrsg.), *Handbuch Barrierefreie Kommunikation*, 327–344. Berlin: Frank & Timme.
- Maszerowska, Anna, Anna Matamala & Pilar Orero (Hrsg.). 2014. *Audio description: New perspectives illustrated* (Benjamins Translation Library). Amsterdam Philadelphia: John Benjamins Publishing Company. <https://doi.org/10.1075/btl.112>.
- Matusov, Evgeny, Patrick Wilken & Panayota Georgakopoulou. 2019. Customizing neural machine translation for subtitling. In *Proceedings of the Fourth Conference on Machine Translation (WMT), Volume 1: Research Papers*, 82–93. Florenz, Italien: Association for Computational Linguistics. <https://aclanthology.org/W19-5209.pdf>.
- Nagel, Silke. 2024. Berufsverbände in der audiovisuellen Übersetzungsbranche. In Alexander Künzli & Klaus Kaendl (Hrsg.), *Handbuch Audiovisuelle Translation: Arbeitsmittel für Wissenschaft, Studium, Praxis*, 265–272. Berlin: Frank & Timme.
- Nikolić, Kristijan. 2010. The subtitling profession in Croatia. In Jorge Díaz-Cintas, Anna Matamala & Josélia Neves (Hrsg.), *New insights into audiovisual translation and media accessibility*, 99–108. Amsterdam: Rodopi. DOI: 10.1163/9789042031814_009.
- O'Brien, Sharon. 2022. How to deal with errors in machine translation: Post-editing. In Dorothy Kenny (Hrsg.), *Machine translation for everyone. Empowering users in the age of artificial intelligence* (Translation and Multilingual Natural Language Processing 18), 105–120. Berlin: Language Science Press. <https://doi.org/10.5281/zenodo.6759982>.

- Oziemblewska, Magdalena & Agnieszka Szarkowska. 2022. The quality of templates in subtitling. A survey on current market practices and changing subtitler competences. *Studies in Translation Theory and Practice* 30(3). 432–453. DOI: 10.1080/0907676X.2020.1791919.
- Papi, Sara, Marco Gaido, Alina Karakanta, Mauro Cettolo, Matteo Negri & Marco Turchi. 2023. Direct speech translation for automatic subtitling. *Transactions of the Association for Computational Linguistics* 11. 1355–1376. DOI: 10.1162/tacl_a_00607.
- Qiu, Juerong & Anthony Pym. 2024. Fatal flaws? Investigating the effects of machine translation errors on audience reception in the audiovisual context. *Perspectives: Studies in Translation Theory and Practice*. 1–17. DOI: 10.1080/0907676X.2024.2328757.
- Reinart, Sylvia. 2018. Untertitelung – über die etablierten Normen hinaus. In Marco Agnetta (Hrsg.), *Über die Sprache hinaus. Translatorisches Handeln in semiotischen Grenzräumen*, 131–155. Hildesheim u.a.: Olms.
- Risku, Hanna. 2024. Reflections on individualized and extended translator studies. In Mira Kadrić, Waltraud Kolb & Sonja Pöllabauer (Hrsg.), *Translation als Gestaltung: Beiträge für Klaus Kaindl zur translatorischen Theorie und Praxis*, 65–74. Tübingen: Narr Francke Attempto.
- Romero-Fresco, Pablo & Nazaret Fresno. 2023. Accuracy of automatic and human live captions in English. *Linguistica Antverpiensia, New Series: Themes in Translation Studies* 22. 114–133.
- Ruokonen, Minna & Jukka Mäkisalo. 2018. Middling-status profession, high-status work: Finnish translators' status perceptions in the light of their backgrounds, working conditions and job satisfaction. *The International Journal of Translation and Interpreting Research* 10(1). 1–17. DOI: 10.12807/ti.110201.2018.a01.
- Schierl, Federike. 2023. Reception of machine-translated and human translated subtitles: A case study. In *Proceedings of Machine Translation Summit XIX, Vol. 2: Users Track*, 42–53. Macau SAR, China. <https://aclanthology.org/2023.mtsummit-users.4.pdf>.
- Silvester, Hannah. 2022. Working conditions and collaborative practices in the translation of French film: Subtitling banlieue cinema. *Studies in Translation Theory and Practice* 30(3). 399–414. DOI: 10.1080/0907676X.2021.1903517.
- Szarkowska, Agnieszka, Jorge Díaz Cintas & Olivia Gerber-Morón. 2021. Quality is in the eye of the stakeholders: What do professional subtitlers and viewers think about subtitling? *Universal Access in the Information Society* 20(4). 661–675. DOI: 10.1007/s10209-020-00739-2.

- Tardel, Anke. 2023. A proposed workflow model for researching production processes in subtitling. *trans-kom* 16(1). 140–173.
- Tardel, Anke & Silvia Hansen-Schirra. 2024. Prozessorientierte Methoden. In Alexander Künzli & Klaus Kaindl (Hrsg.), *Handbuch Audiovisuelle Translation. Arbeitsmittel für Wissenschaft, Studium, Praxis*, 287–303. Berlin: Frank & Timme, Verlag für Wissenschaftliche Literatur.
- Vercauteren, Gert, Nina Reviers & Kim Steyaert. 2021. Evaluating the effectiveness of machine translation of audio description: The results of two pilot studies in the English-Dutch language pair. *Tradumàtica: tecnologies de la traducció* (19). 226–252. DOI: 10.5565/rev/tradumatica.288. (14 Oktober, 2024).
- Wang, Yujia, Wei Liang, Haikun Huang, Yongqi Zhang, Dingzeyu Li & Lap-Fai Yu. 2021. Toward automatic audio description generation for accessible videos. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems (CHI '21)*, 1–12. DOI: <https://doi.org/10.1145/3411764.344534>.

