

# Von der klassischen Filmsynchronisation zur KI- Revolution:

Wie künstliche Intelligenz die Zukunft  
in der Tonpostproduktion prägt und  
vergangene Standards verändert

Fakultät Electronic Media  
Audiovisuelle Medien  
Hochschule der Medien

## Bachelorarbeit

vorgelegt von

**Leonie Herz**  
Matrikelnummer 40956

am 14.04.2025  
zur Erlangung des akademischen Grades eines  
Bachelor of Engineering

**Erstprüfer:** Prof. Oliver Curdt  
**Zweitprüfer:** Tilo Ehmann  
**Praxisbetreuer:** Jochen Fenzl

# Inhaltsverzeichnis

EHRENWÖRTLICHE ERKLÄRUNG	4
KURZFASSUNG	5
ABSTRACT	5
ABBILDUNGSVERZEICHNIS	6
ABKÜRZUNGSVERZEICHNIS	7
GENDERHINWEIS	8
1. EINLEITUNG	9
1.1 PROBLEMSTELLUNG UND RELEVANZ DES THEMAS	9
1.2 FORSCHUNGSFRAGE UND ZIELSETZUNG	10
1.3 METHODIK UND HERANGEHENSWEISE	11
1.4 AUFBAU DER ARBEIT	12
2. GRUNDLAGEN DER KLASSISCHEN FILMSYNCHRONISATION	13
2.1 GESCHICHTE UND ENTWICKLUNG DER FILMSYNCHRONISATION	14
2.2 TECHNISCHE UND KREATIVE ANFORDERUNGEN AN SYNCHRONSPRECHER UND REGISSEURE	15
2.3 ARBEITSABLAUF IN DER TRADITIONELLEN SYNCHRONISATION	17
3. DER AUFSTIEG DER KÜNSTLICHEN INTELLIGENZ IN DER MEDIENPRODUKTION	20
3.1 DEFINITION UND GRUNDLAGEN DER KÜNSTLICHEN INTELLIGENZ	20
3.2 ENTWICKLUNG DER KI-TECHNOLOGIEN IN DER FILM- UND MEDIENINDUSTRIE	22
4. TECHNOLOGISCHE TRANSFORMATIONEN DURCH KI IN DER TONPOSTPRODUKTION	24
4.1 AUTOMATISIERTE SPRACHSYNTHESE UND TEXT-TO-SPEECH (TTS)	24
4.2 VOICE CLONING UND STIMMKLON-TECHNOLOGIEN	28
4.3 KI-GESTÜTZTE LIPPENSYNCHRONISATION	34
4.4 AUTOMATISIERTE ÜBERSETZUNG UND KULTURADAPTION DURCH KI	40
5. AUSWIRKUNGEN AUF DIE SYNCHRONBRANCHE	44
5.1 VERÄNDERUNGEN IN ARBEITSABLÄUFEN UND BERUFSBILDERN	44

5.2	QUALITÄT UND KONSISTENZ DER KI-GESTÜTZTEN SYNCHRONISATION	46
5.3	ETHISCHE UND RECHTLICHE IMPLIKATIONEN	50
6.	<u>ZUKUNFTSPERSPEKTIVEN DER KI IN DER SYNCHRONISATION</u>	<u>54</u>
6.1	TECHNOLOGISCHE TRENDS UND MÖGLICHE WEITERENTWICKLUNGEN	54
6.2	CHANCEN UND RISIKEN FÜR DIE INDUSTRIE	56
6.3	MÖGLICHE VERÄNDERUNGEN DER KONSUMENTENWAHRNEHMUNG	58
7.	<u>FAZIT</u>	<u>60</u>
7.1	ZUSAMMENFASSUNG DER ERGEBNISSE	60
7.2	BEANTWORTUNG DER FORSCHUNGSFRAGE	61
7.3	AUSBLICK AUF ZUKÜNFTIGE ENTWICKLUNGEN	62
	<u>QUELLENVERZEICHNIS</u>	<u>63</u>
	BILDQUELLEN	71
	<u>ANHANG</u>	<u>72</u>
	GESPRÄCH MIT CORNELIUS FROMMANN AM 26.03.2025	72
	GESPRÄCH SUSANNE VON MEDVEY AM 28.03.2025	79

## Ehrenwörtliche Erklärung

Hiermit versichere ich, Leonie Herz, ehrenwörtlich, dass ich die vorliegende Bachelorarbeit mit dem Titel: „Von der klassischen Filmsynchronisation zur KI-Revolution: Wie künstliche Intelligenz die Zukunft in der Tonpostproduktion prägt und vergangene Standards verändert“ selbstständig und ohne fremde Hilfe verfasst und keine anderen als die angegebenen Hilfsmittel benutzt habe. Die Stellen der Arbeit, die dem Wortlaut oder dem Sinn nach anderen Werken entnommen wurden, sind in jedem Fall unter Angabe der Quelle kenntlich gemacht. Ebenso sind alle Stellen, die mit Hilfe eines KI-basierten Schreibwerkzeugs erstellt oder überarbeitet wurden, kenntlich gemacht. Die Arbeit ist noch nicht veröffentlicht oder in anderer Form als Prüfungsleistung vorgelegt worden.

Ich habe die Bedeutung der ehrenwörtlichen Versicherung und die prüfungsrechtlichen Folgen (§ 24 Abs. 2 Bachelor-SPO, § 23 Abs. 2 Master-SPO (Vollzeit)) einer unrichtigen oder unvollständigen ehrenwörtlichen Versicherung zur Kenntnis genommen.

Herrenberg, 09.04.2025

---

Ort, Datum



---

Unterschrift

## Kurzfassung

Diese Bachelorarbeit untersucht den Wandel in der Tonpostproduktion von der klassischen Filmsynchronisation hin zu einer durch künstliche Intelligenz (KI) geprägten Zukunft. Im Fokus stehen die traditionellen Methoden der Synchronisation, ihre technischen und künstlerischen Grundlagen sowie die Herausforderungen, die mit ihrer Umsetzung verbunden sind. Anschließend wird beleuchtet, wie KI-basierte Technologien wie automatische Sprachsynthese, Stimmklonung und automatisiertes Lippen-Synchronisieren die etablierten Standards verändern. Durch eine Analyse von Praxisbeispielen und Interviews mit Branchenexperten wird deutlich, dass KI nicht nur die Effizienz steigert, sondern auch neue kreative Möglichkeiten eröffnet. Gleichzeitig werden ethische Fragen und die Auswirkungen auf die Arbeitsweise von Synchronsprechern und Toningenieuren diskutiert. Die Arbeit zeigt, wie KI die Tonpostproduktion transformiert und gleichzeitig traditionelle Ansätze beeinflusst.

## Abstract

This bachelor thesis explores the transformation of audio post-production from classical film synchronization to a future shaped by artificial intelligence (AI). The focus lies on traditional synchronization methods, their technical and artistic foundations, and the challenges associated with their implementation. Subsequently, the thesis examines how AI-based technologies such as automatic speech synthesis, voice cloning, and automated lip-syncing are altering established standards. Through the analysis of case studies and interviews with industry experts, it becomes clear that AI not only increases efficiency but also opens up new creative possibilities. At the same time, ethical questions and the impact on the workflows of voice actors and sound engineers are discussed. The study illustrates how AI is revolutionizing audio post-production while influencing traditional approaches.

## Abbildungsverzeichnis

ABBILDUNG 1 DUBBING TOOL ELEVENLABS (QUELLE: ELEVENLABS (2025))	27
ABBILDUNG 2 DUBBING TOOL DEEPDUB (QUELLE: DEEPDUB (2025))	27
ABBILDUNG 3 IVC TOOL ELEVENLABS (QUELLE: ELEVENLABS (2025))	33
ABBILDUNG 4 PVC TOOL ELEVENLABS (QUELLE: REDDIT (2025))	33
ABBILDUNG 5 VOICE CLONING TOOL RESEMBLE AI (QUELLE: RESEMBLE AI (2025))	34
ABBILDUNG 6 LIPPENBEWEGUNGEN FÜR BESTIMMTE LAUTE (QUELLE: VIDNOZ (2025))	36
ABBILDUNG 7 FLAWLESS TRUESYNC TOOL (QUELLE: FLAWLESS (2025))	40

## Abkürzungsverzeichnis

DSP	Digital Signal Processing
IVC	Instant Voice Cloning
KI	Künstliche Intelligenz
LLM	Large Language Model
ML	Machine Learning/Maschinelles Lernen
NLP	Natural Language Processing
PVC	Professional Voice Cloning
TTS	Text-To-Speech

## Genderhinweis

In dieser Arbeit wird aus Gründen der besseren Lesbarkeit überwiegend die männliche oder weibliche Sprachform verwendet. Sämtliche Personenbezeichnungen gelten jedoch gleichermaßen für alle Geschlechter. Diese Vereinfachung dient ausschließlich der sprachlichen Klarheit und stellt keine Wertung dar.

# 1. Einleitung

## 1.1 Problemstellung und Relevanz des Themas

Die Synchronisation von Filmen ist seit langem ein zentraler Bestandteil der audiovisuellen Postproduktion. Durch sie können audiovisuelle Werke über sprachliche und kulturelle Grenzen hinweg zugänglich gemacht werden, wodurch sie einen großen Beitrag zur Internationalisierung der Filmindustrie geleistet hat.

Traditionell ist die Filmsynchronisation ein aufwendiger, manueller Prozess, der sowohl technische Fähigkeiten als auch kreative Interpretationen von Dialogen erfordert. Neue Technologien, besonders die aufsteigende Künstliche Intelligenz (KI), sind dazu in der Lage, diesen Bereich grundlegend zu verändern.

KI-basierte Systeme versprechen, den Synchronisationsprozess schneller, kostengünstiger und präziser zu machen. Algorithmen für Sprachsynthese, maschinelles Lernen und neuronale Netzwerke können Stimmen naturgetreu nachbilden, Lippenbewegungen an Dialoge anzupassen und sogar emotionale Nuancen in synthetisierten Stimmen integrieren (Weitzman, 2023). Diese Entwicklungen haben nicht nur das Potenzial, in die gängigen Standards der Tonpostproduktion einzugreifen, sondern stellen auch grundlegende Fragen hinsichtlich der künstlerischen und ethischen Qualität dieser Produktionen.

Die Relevanz dieses Themas zeigt sich in der wachsenden Nachfrage nach globaler Medienanpassung. Laut einem Bericht der Motion Picture Association (2020) ist der internationale Marktanteil für audiovisuelle Inhalte in den letzten Jahren stark gestiegen. KI-gestützte Synchronisation könnte der hohen Nachfrage nach übersetzten Inhalten gerecht werden. Jedoch stellt sich die Frage, ob KI die menschliche Kreativität in der Synchronisation vollständig ersetzen kann oder ob der Einsatz solcher Technologien lediglich als unterstützendes Werkzeug gesehen werden kann. Auch der Arbeitsmarkt steht unter enormem Einfluss durch die Nutzung von KI in der Synchronbranche.

Berufsbilder wie Synchronsprecher (auch Synchronschauspieler) oder Toningenieure könnten in ihrer traditionellen Form an Bedeutung verlieren, während gleichzeitig neue Kompetenzen in der Arbeit mit KI-Systemen erforderlich werden (Klingbeil-

Döring, 2023). Dieser Wandel stellt die Branche vor die Herausforderung, sich an neue Standards und Arbeitsweisen anzupassen.

Insgesamt ist die Untersuchung dieses Themenbereichs relevant, da sie nicht nur die technologische Transformation der Tonpostproduktion beleuchtet, sondern auch die damit einhergehenden sozialen, wirtschaftlichen und ethischen Fragestellungen betrachtet. Eine tiefgehende Auseinandersetzung mit der Frage, wie KI die Zukunft der Filmsynchronisation prägt, ist daher nicht nur für die Filmindustrie von Bedeutung, sondern auch für die Gesellschaft als Ganzes.

## 1.2 Forschungsfrage und Zielsetzung

Die stetige Entwicklung künstlicher Intelligenz (KI) hat in der Film- und Medienindustrie, besonders in der Tonpostproduktion, einen bedeutenden Wandel ausgelöst. Mit diesem Hintergrund beschäftigt sich diese Arbeit mit der Frage: **Wie verändert künstliche Intelligenz die Standards und Prozesse der Synchronisation, und welche Entwicklungen sind weiter zu erwarten?** Dabei sollen nicht nur die aktuellen Auswirkungen der KI auf traditionelle Synchronisationsverfahren betrachtet werden, sondern auch zukünftige Perspektiven zeigen, wie diese Technologie die Branche weiter prägen könnte.

Das Ziel dieser Untersuchung ist es, ein Verständnis darüber zu erlangen, wie KI-gestützte Tools und Verfahren die Synchronisation grundlegend transformieren.

Klassische Methoden, die sich bisher durch die enge Zusammenarbeit von Synchronsprechern, Regisseuren, Redakteuren, Dialogbuch-Autoren, Editoren und Tontechnikern auszeichneten, werden zunehmend durch KI-gestützte Technologien ergänzt oder sogar ersetzt. Systeme wie neuronale Sprachsynthesen, automatische Lippenanpassung und maschinelles Lernen ermöglichen es, Dialoge schneller und mit bemerkenswerter Präzision an neue Sprachen und Zielmärkte anzupassen.

Dabei stellt sich die Frage, wie diese Technologien bestehende Qualitätsstandards verändern und ob sie mit den kreativen Ansprüchen menschlicher Synchronisation mithalten können. Ein weiteres Ziel der Arbeit ist es, die ethischen und sozialen Bedeutungen dieser Entwicklung zu analysieren. Während KI eine erhebliche

Effizienzsteigerung verspricht, wirft ihr Einsatz auch Herausforderungen auf, etwa in Bezug auf die Wahrung künstlerischer Authentizität, mögliche Jobverluste oder die Gefahr der Monopolisierung durch große Technologieunternehmen (Arndt et al., 2024). Darüber hinaus wird untersucht, welche künftigen Innovationen in diesem Bereich zu erwarten sind und welche Rolle die Filmindustrie dabei spielen könnte, technologische und kreative Ansprüche in Einklang zu bringen.

Die Ergebnisse dieser Arbeit sollen nicht nur einen Beitrag zur Diskussion über die Rolle von KI in der Medienproduktion leisten, sondern auch Einblicke für Branchenakteure bieten. Durch die Untersuchung der Chancen und Herausforderungen, sollen mögliche Strategien zur Integration von KI in der Synchronisation gezeigt werden, die sowohl Effizienz als auch Qualität und Kreativität berücksichtigen.

### 1.3 Methodik und Herangehensweise

Um die eben genannte Forschungsfrage zu beantworten, stützt sich diese Arbeit auf eine Kombination aus Literaturrecherche und Interviews. Dadurch können Entwicklungen in der Filmsynchronisation sowohl aus technischer als auch aus kreativer Perspektive betrachtet werden.

Die Literaturrecherche bildet die Grundlage der Analyse. Hierbei werden wissenschaftliche Artikel, Fachbücher und Berichte erfasst, um die theoretischen Grundlagen zu den Themen Filmsynchronisation, künstliche Intelligenz und Medienproduktion zu erarbeiten. Dabei ist besonders die Untersuchung aktueller Technologien wie neuronaler Sprachsynthese, maschinellem Lernen und automatischer Lippenanpassung wichtig, deren Einsatz in der Postproduktion zunehmend relevant wird. Die Literaturrecherche dient zudem dazu, bestehende Standards und Prozesse der klassischen Synchronisation zu beschreiben, um Veränderungen durch KI sichtbar zu machen.

Um die Ergebnisse zu ergänzen und aktuelle Entwicklungen sowie Zukunftsperspektiven einzuschätzen, werden Interviews mit Fachleuten aus der Film- und Postproduktionsbranche durchgeführt. Dabei sollen die Chancen und

Herausforderungen durch die Meinungen und Einschätzungen von Synchronsprechern und Tontechnikern detailliert eingeschätzt werden.

Durch diese Herangehensweise wird sichergestellt, dass die Forschungsfrage aus unterschiedlichen Perspektiven betrachtet wird. Die Kombination aus theoretischer Analyse und Meinungen aus der Branche schafft eine stabile Basis um sowohl technologische als auch gesellschaftliche Aspekte des Themas zu berücksichtigen.

#### 1.4 Aufbau der Arbeit

Die vorliegende Arbeit gliedert sich in sieben Kapitel, die eine schrittweise Annäherung an die Forschungsfrage und Zielsetzung ermöglichen. Der Aufbau verfolgt dabei einen logischen und thematisch kohärenten Verlauf, um das Thema umfassend zu bearbeiten.

Das erste Kapitel bildet die **Einleitung**, in der die Problemstellung und Relevanz des Themas erläutert werden. Es wird verdeutlicht, warum die technologischen Fortschritte im Bereich der künstlichen Intelligenz (KI) insbesondere für die Synchronbranche von großer Bedeutung sind. Darauf aufbauend werden die Forschungsfrage und die Zielsetzung der Arbeit formuliert. Abschließend wird in diesem Kapitel der Aufbau der Arbeit beschrieben, um einen Überblick über die Herangehensweise zu geben.

Im zweiten Kapitel werden die **Grundlagen der klassischen Filmsynchronisation** aufgezeigt. Zunächst wird ein Überblick über die Geschichte und Entwicklung der Synchronisation gegeben, bevor auf die technischen und kreativen Anforderungen an Synchronsprecher und Regisseure eingegangen wird. Abschließend wird der traditionelle Arbeitsablauf in der Synchronisation beschrieben, um ein solides Verständnis der bisherigen Standards und Prozesse zu schaffen. Kapitel drei beschäftigt sich mit dem **Aufstieg der künstlichen Intelligenz in der Medienproduktion**. Nach einer allgemeinen Definition und Einführung in die Grundlagen der KI, werden die Entwicklungen dieser Technologien speziell in der Film- und Medienindustrie untersucht. Ein besonderer Fokus liegt auf den KI-

Technologien, die für die Synchronisation relevant sind. Dazu zählen beispielsweise die Sprachsynthese und das maschinelle Lernen.

Das vierte Kapitel beleuchtet die **technologischen Transformationen durch KI in der Tonpostproduktion**. Es werden zentrale Technologien wie automatisierte Sprachsynthese, Voice Cloning, KI-gestützte Lippensynchronisation sowie

automatische Übersetzungs- und Kulturadaptionssysteme betrachtet. Dabei wird auch auf die Funktionsweise und das Training dieser Systeme eingegangen.

In Kapitel fünf werden die **Auswirkungen auf die Synchronbranche** untersucht. Es wird analysiert, wie sich Arbeitsabläufe und Berufsbilder durch den Einsatz von KI verändern. Außerdem werden die Auswirkungen auf die Qualität und Konsistenz der Synchronisation bestimmt und die entstehenden ethischen und rechtlichen Fragen in Hinblick auf KI-Synchronisationen aufgezeigt.

Kapitel sechs stellt die **Zukunftsperspektiven der KI in der Synchronisation** und Tonpostproduktion in den Vordergrund. Es werden die technologischen Trends und die potenziellen Weiterentwicklungen sowie die Chancen und Risiken für die Branche beleuchtet. Zudem wird untersucht, wie sich diese Entwicklungen auf die Wahrnehmung der Konsumenten auswirken könnte.

Das abschließende siebte Kapitel bildet das **Fazit**. Hier werden die wichtigsten Erkenntnisse zusammengefasst und die Forschungsfrage beantwortet. Darüber hinaus wird ein Ausblick auf zukünftige Entwicklungen gegeben.

Der Aufbau bietet einen klaren Überblick, der von den traditionellen Methoden der Filmsynchronisation bis zu den technologischen und gesellschaftlichen Transformationen durch Künstliche Intelligenz reicht.

## 2. Grundlagen der klassischen Filmsynchronisation

Die klassische Filmsynchronisation ist ein zentraler Bestandteil der Filmproduktion. Hierbei werden audiovisuelle Inhalte sprachlich und kulturell an verschiedene Zielgruppen angepasst. Dieser Prozess beinhaltet die Übersetzung und Anpassung von Dialogen, welche in die deutsche Sprache adaptiert bzw. nachgesprochen werden. Diese können dann in die filmischen Gegebenheiten integriert werden, wobei ein hohes Maß an Präzision erforderlich ist. Dabei spielen neben technischer

Expertise auch kreative und dramaturgische Entscheidungen eine wesentliche Rolle, um die Intention des Originals möglichst authentisch zu halten. Filmsynchronisation ist somit nicht nur eine technische Herausforderung, sondern erfordert auch künstlerische Fähigkeiten, die wesentlich zur internationalen Verarbeitung von Filmen beiträgt.

## 2.1 Geschichte und Entwicklung der Filmsynchronisation

Die Geschichte der Filmsynchronisation ist eng mit der Entwicklung des Tonfilms verbunden, der Ende der 1920er Jahre seinen Durchbruch erlebte. Mit dem ersten vollständig vertonten Film *The Jazz Singer* (1927) begann eine neue Ära des Kinos, die es erforderlich machte, Filme für ein internationales Publikum verständlich zu machen. Während sich Untertitelte internationale Produktionen immer weiter verbreiteten, entstanden die ersten Ansätze zur Synchronisation von Filmen bereits 1929, als Alfred Hitchcocks Film *Erpressung* (*Blackmail*) mit einer zusätzlichen deutschen Tonspur versehen wurde (Wikipedia, 2023).

In den 1930er Jahren entwickelte sich die Filmsynchronisation zu einer standardisierten Praxis. Während in einigen Ländern Untertitel bevorzugt wurden, etablierte sich in Deutschland, Frankreich und Italien die Synchronisation als dominierende Methode, um Filme für das jeweilige Publikum kulturell anzupassen. Diese Vorliebe lässt sich historisch durch den Wunsch nach Sprachbarrierefreiheit und die starke Betonung nationaler Identität erklären (Naumann, 2015).

Besonders in Deutschland führte die nationalsozialistische Kulturpolitik dazu, dass ausländische Filme systematisch synchronisiert wurden, um die Fremdsprachigkeit zu eliminieren und ideologische Inhalte anzupassen (Goethe-Institut, 2023).

Technologisch und kreativ entwickelte sich die Synchronisation in den folgenden Jahrzehnten stetig weiter. In den 1950er Jahren war die Technik so weit fortgeschritten, dass Dialoge präzise an die Lippenbewegungen der Schauspieler angepasst werden konnten. Diese Weiterentwicklung der sogenannten Lippensynchronität setzte neue Standards, die bis heute eine zentrale Anforderung an die Filmsynchronisation darstellen. Neben der technischen Präzision rückte auch

die kreative Anpassung an kulturelle und sprachliche Eigenheiten in den Fokus, um den Charakter und die Authentizität der Originalwerke zu wahren (Naumann, 2015).

Seit den 1990er Jahren hat die Globalisierung der Film- und Medienindustrie die Bedeutung der Filmsynchronisation weiter verstärkt. Mit dem Aufkommen internationaler Streaming-Dienste wie Netflix und Amazon Prime hat die Nachfrage nach synchronisierten Inhalten einen neuen Höhepunkt erreicht. Dies führte nicht nur zu einer Professionalisierung der Branche, sondern auch zu verstärkten Bemühungen, die Qualität und kulturelle Sensibilität von Synchronisationen zu verbessern (Goethe-Institut, 2023).

Zusammenfassend lässt sich sagen, dass die Filmsynchronisation eine dynamische Entwicklung durchlaufen hat, die sowohl durch technologische Innovationen als auch durch kulturelle und wirtschaftliche Anforderungen geprägt wurde. Von den ersten Tonfilmen bis hin zur heutigen Ära der Streaming-Plattformen bleibt die Synchronisation ein unverzichtbarer Bestandteil der internationalen Film- und Medienindustrie.

## 2.2 Technische und kreative Anforderungen an Synchronsprecher und Regisseure

Die Arbeit in der Filmsynchronisation stellt sowohl für Synchronsprecher als auch für Regisseure hohe technische und kreative Anforderungen, die entscheidend für die Qualität der Synchronisation sind. Synchronsprecher benötigen nicht nur eine präzise stimmliche Kontrolle und Ausdrucksfähigkeit, sondern auch die Fähigkeit, sich in die emotionalen und körperlichen Nuancen der dargestellten Charaktere einzufühlen. Diese Anforderungen erfordern intensives Training in Schauspielkunst sowie eine stimmliche Flexibilität, um unterschiedliche Rollen glaubhaft darstellen zu können (m-Voices CAMPUS, 2014).

Ein zentraler technischer Aspekt der Arbeit von Synchronsprechern ist die sogenannte **Lippensynchronität**. Dabei müssen gesprochene Dialoge exakt auf die Mundbewegungen der Originaldarsteller abgestimmt werden, um eine glaubwürdige

Illusion zu schaffen. Dies erfordert ein präzises Timing und die Fähigkeit, sich schnell auf den Rhythmus und die Betonung der Originalsprache einzustellen (Naumann, 2015).

Synchronregisseure wiederum tragen die Verantwortung für die kreative und künstlerische Leitung des Synchronisationsprozesses. Sie müssen sicherstellen, dass die Übersetzung der Dialoge sowohl inhaltlich als auch kulturell authentisch ist und gleichzeitig die emotionalen und dramaturgischen Feinheiten des Originalwerks aufrechterhalten. Dies erfordert ein tiefes Verständnis für Filmkunst sowie ein gutes Gespür für sprachliche und kulturelle Nuancen (Bundesverband Regie, o.D.). Darüber hinaus ist die Zusammenarbeit mit Übersetzern und Dialogbuchautoren von großer Bedeutung, um den Dialog nicht nur wortgetreu, sondern auch stilistisch passend zu adaptieren.

Neben der kreativen Arbeit steht der Regisseur auch vor organisatorischen Herausforderungen. Dazu gehört unter anderem die Sicherstellung, dass alle Beteiligten eng mit der technischen Abteilung zusammenarbeiten, um ein hochwertiges Ergebnis zu erzielen (Bundesverband Regie, o.D.).

Auch die Globalisierung der Medienbranche prägt die Arbeit von Synchronsprechern und Regisseuren zunehmend. Internationale Produktionen und Streaming-Dienste erfordern eine größere kulturelle Sensibilität und Sprachvielfalt. Synchronisationen müssen nicht nur sprachlich korrekt sein, sondern auch kulturell angepasst werden, um für ein breites Publikum akzeptabel zu sein (Amberscript, 2023).

Zusammenfassend lässt sich sagen, dass Synchronsprecher und Regisseure eine anspruchsvolle Kombination aus technischer Präzision, künstlerischer Kreativität und kulturellem Feingefühl mitbringen müssen, um den hohen Ansprüchen der Filmsynchronisation gerecht zu werden. Diese Fähigkeiten sind entscheidend, um das Publikum in die erzählte Geschichte einzubinden und die Authentizität des Originals zu bewahren.

## 2.3 Arbeitsablauf in der traditionellen Synchronisation

Für die Synchronisation fremdsprachiger Produktionen werden verschiedene Stationen durchlaufen, welche in enger Zusammenarbeit stehen. Von der Auftragsstellung bis zum fertigen Endprodukt erfolgen sorgfältig aufeinander abgestimmte Schritte, um eine qualitativ hochwertige Anpassung fremdsprachiger Filme und Serien in die gewünschte Zielsprache zu erreichen.

### Skriptübersetzung

Zunächst wird das Originalskript beziehungsweise das Drehbuch benötigt, welches dann in die gewünschte Zielsprache übersetzt wird. Wichtig hierbei ist, dass alle gesprochenen Texte vorhanden und identisch zum abgedrehten Film sind. Dies ist für alle weiteren Schritte im Synchronisationsprozess essentiell. Nach der genauen Festlegung der Zielsprache (da beispielsweise europäisches Spanisch nicht gleich argentinisches Spanisch) folgt die Übersetzung des Skripts (Morrison, 2020). Hierbei muss nicht nur der gesprochene Text, sondern auch kulturelle Aspekte und Wortspiele berücksichtigt werden um der Authentizität des Inhalts gerecht zu werden.

### Dialogbucherstellung

Mit dem übersetzten Skript kann nun ein lippensynchrones Dialogbuch erstellt werden. Ein Dialogbuchautor modifiziert die Dialoge so, dass sie auf die Lippenbewegungen der Originaldarsteller angepasst sind, dabei aber nicht übermäßig vom Inhalt abweichen. Dies erzielt eine natürliche und glaubwürdige Darstellung. Anschließend werden die Dialoge in sogenannte Takes aufgeteilt. Dabei handelt es sich um kleinere Abschnitte, welche aus einzelnen Sätzen oder auch nur Wortlauten bestehen. Diese ermöglichen es dem Sprecher sich aufgrund ihrer Kompaktheit besser auf die Lippenbewegungen konzentrieren zu können. Außerdem lässt sich anhand der Takes der grobe Umfang des Projekts einschätzen.

## Sprecherbesetzung

Die Auswahl der Sprecher ist ebenfalls ein weiterer sehr wichtiger Schritt. Sie müssen nicht nur stimmlich zu den Originalstimmen passen, sondern auch deren Emotionen und Ausdrucksweisen authentisch nachbilden. Mit Hilfe von Stimmproben werden die Sprecher dann anhand verschiedener Kriterien wie ihres Geschlechts, Alters und Fähigkeiten ausgewählt.

## Aufnahmeprozess

Nach den bisher genannten Schritten wird ein Aufnahmeplan erstellt, bei dem Sprecher, Regisseure und Toningenieure berücksichtigt werden müssen. Die verschiedenen Rollen werden dann planmäßig der Reihe nach takeweise im Tonstudio aufgenommen. Dabei spricht der Sprecher in einem Aufnahmerraum, während Regisseur und Toningenieur in der Regie Anweisungen geben. In manchen Fällen wird auf einen separaten Aufnahmerraum verzichtet, um in einem direkteren Kontakt zum Sprecher zu stehen. Dadurch kann eine bessere Kommunikation und damit eine bessere Regiearbeit erzielt werden. Falls notwendig werden Tonspuren schon bei den Aufnahmen zeitlich angepasst und zurechtgeschnitten. Dies erfolgt jedoch üblicherweise aus zeitlichen Gründen erst nach den Aufnahmen.

## Abmischung

Sind die Aufnahmen abgeschlossen, folgt die Abmischung der Tonspuren, bei der die neuen Dialogspuren die alten ersetzen. Dabei werden die neuen Tonspuren in den Originalfilm eingebettet, sodass diese gut mit den anderen Elementen (Musik und Soundeffekte/Geräusche) harmonieren. Es werden unter anderem Lautstärken, Frequenzen, Raumeffekte und stilistische Eigenschaften mit Hilfe von geeigneten Tools bearbeitet.

## Herausforderungen

Einer der größten Herausforderungen im Synchronisationsprozess ist wohl die **Lippensynchronität**. Um einen natürlichen Dialog zu kreieren ist höchste Präzision bei Timing und Rhythmus gefragt. Dies ist nicht immer einfach und kann nur schwer in der Nachbearbeitung korrigiert werden.

Auch **kulturelle Aspekte** können zu Herausforderungen werden. Wortspiele, Redewendungen oder kulturelle Anspielungen lassen sich oft nicht direkt übersetzen. Hier ist Kreativität gefragt, um sinngemäße Alternativen zu finden, die dieselbe Wirkung erzielen, ohne den kulturellen Kontext zu verfremden (Pedersen, 2017).

Eine weitere große Herausforderung ist die **emotionale Authentizität**. Sprecher müssen nicht nur stimmlich zu den jeweiligen Charakteren passen, sondern auch deren Emotionen überzeugend transportieren.

**Mehrsprachigkeit und Dialekte** können in manchen Fällen ebenfalls zur Schwierigkeit werden. In manchen Filmen wird mehr als eine Sprache gesprochen oder es kommen Dialekte vor. Diese Aspekte müssen ebenfalls berücksichtigt werden, um die Authentizität zu wahren. Ein Charakter mit einem markanten Akzent im Original sollte im synchronisierten Film ebenfalls durch einen regionalen Akzent oder eine entsprechende Variation dargestellt werden (Díaz Cintas & Remael, 2021).

**Zeit- und Budgeteinschränkungen** können für qualitative Einbußen sorgen, indem ein hoher Zeitdruck sowie ein zu geringes Budget weniger Raum für kreative Lösungen und die Auswahl hochqualifizierter Sprecher lassen.

Aber nicht nur auf Seiten der Auftraggeber, sondern auch vom **Publikum** kann großer Druck ausgeübt werden. Fehler in der Synchronisation, unpassende Stimmen oder falsche Übersetzungen können für großen Tumult sorgen und das Ansehen des Films/Serie schädigen.

### 3. Der Aufstieg der künstlichen Intelligenz in der Medienproduktion

Die Künstliche Intelligenz hat die Medienbranche in den letzten Jahren grundlegend verändert. Durch den immer häufigeren Einsatz von KI-Technologien wie dem maschinellen Lernen oder Sprachverarbeitungstools, können Prozesse immer schneller und effizienter erfolgen. KI ermöglicht es, riesige Datenmengen in Echtzeit zu durchforsten und dabei neue Muster und Trends zu erkennen. Von automatisierter Textgenerierung bis hin zur Bild- und Videoanalyse wird die Künstliche Intelligenz als vorantreibender Revolutionär der Medienbranche gesehen. Gleichzeitig stellt der Einsatz von KI in der Medienproduktion auch neue Herausforderungen in Bezug auf Urheberrecht, Transparenz und ethische Fragestellungen dar (Bippes, o.D.).

#### 3.1 Definition und Grundlagen der künstlichen Intelligenz

Künstliche Intelligenz (kurz KI) kann als eine Art technische Intelligenz beschrieben werden, also eine Intelligenz die auf technischen Maschinen basiert. Insbesondere war für diese künstliche Schaffung die kognitive Intelligenz des Menschen inspirierend. Diese ist momentan auch die Einzige, die technisch nachgebildet werden kann. Hierbei geht es darum diese menschliche kognitive Intelligenz auf technischen Maschinen zu simulieren, um diese für den gewinnerbringenden und nützlichen Einsatz für den Menschen einzusetzen (Otte, 2021). Diese künstlich intelligenten Systeme sollen Aufgaben übernehmen, bei denen normalerweise menschliche Intelligenz benötigt wird wie beispielsweise Lernen, Problemlösungen oder Sprachverstehen. Um diese Aufgaben erfüllen zu können müssen künstliche Intelligenzen mehrere Eigenschaften besitzen. Dazu gehören unter anderem das selbstständige Denken („**Deep thinking**“) und das Lernen (auch „**Deep learning**“). Die KI soll also ein Modell selbstständig erlernen und bei auftretenden Fehlern in der Lage sein, dieses eigenständig so anzupassen, dass diese bei der nächsten Aktion nicht mehr auftreten.

Im Bereich der künstlichen Intelligenz unterscheidet man zwischen schwacher und starker Intelligenz. **Schwache Intelligenz** zielt darauf ab die menschliche Intelligenz auf technischen Maschinen zu simulieren. Bei einer **starken Intelligenz** soll hingegen ein künstliches Bewusstsein, Willen oder sogar Emotionen erzeugt werden. Dies ist aber zur heutigen Zeit völlig unrealistisch und wird auch für die Zukunft als äußerst unwahrscheinlich gehalten (Otte, 2021).

Weitere grundlegende Aspekte der künstlichen Intelligenz basieren auf dem Konzept des Lernens. Dabei lassen sich ein Großteil der heutigen Erfolge auf das maschinelle Lernen (auch „Machine Learning“) zurückführen. „**Machine Learning**“ basiert auf dem Prinzip, Algorithmen zu schaffen, die aus Daten selbstständig und vollautonom lernen können. Man unterscheidet hierbei zwischen symbolischen Lernverfahren (z.B. Entscheidungsbäume oder Assoziationsregeln) und sub-symbolischen Lernverfahren. Ein großer und essentieller Bestandteil der sub-symbolischen Verfahren sind die künstlichen neuronalen Netze. **Neuronale Netze** sind von der Struktur des menschlichen Gehirns inspiriert und somit in der Lage komplexe Muster in großen Datensätzen zu erkennen (LeCun, Bengio & Hinton, 2015). Durch sehr einfache Rechenschritte kann das neuronale Netz Informationen sammeln und Ergebnisse berechnen. Parallele Datenverarbeitungen von mehreren Millionen Einzelementen ermöglichen es, eine enorme Leistungsfähigkeit im Gesamtsystem zu erzielen, indem die unzähligen Neuronen eine hohe Vernetzung untereinander aufweisen. Dabei kann die KI jedoch nur statistische Zusammenhänge finden und keine ursächlichen/begründeten. Die ersten Bausteine der KI wurden bereits 1950 gelegt, als man Überlegungen anstellte, ob Maschinen das menschliche Denken nachahmen können (Turing, 1950).

## 3.2 Entwicklung der KI-Technologien in der Film- und Medienindustrie

Wie bereits erwähnt, begann die Reise der KI maßgeblich in den 1950er Jahren mit der Entwicklung des Turing-Tests. Dieser war in der Lage maschinengenerierte und von Menschen stammende Ergebnisse zu unterscheiden. In den darauffolgenden Jahrzehnten gab es ein enormes Wachstum der Größe dieser KI-Modelle. Mit der Erfindung neuronaler Netze, maschinellem Lernen und der Verarbeitung natürlicher Sprache, konnten große Durchbrüche erzielt und somit größere (bessere) KI-Technologien entwickelt werden. Beschleunigt wurde dieser Fortschritt durch die Zusammenarbeit riesiger Technologieentwickler und Investitionen in Open-Source-Material (Glänzel, o.D.).

Auch die Film- und Medienbranche profitierte von diesen revolutionierenden Entwicklungen, worauf eine immer steigendere Nutzung künstlich-intelligenter Systeme folgte. Dabei beeinflussen diese verschiedene Phasen der Produktionen. Von der Vorproduktion über die Produktion bis hin zur Postproduktion werden KI-Technologien immer stärker integriert.

### Vorproduktion und Drehbucherstellung

KI-Technologien werden zunehmend zur Analyse und Erstellung von Drehbüchern eingesetzt. Algorithmen können bestehende Drehbücher analysieren, um erfolgreiche Handlungsstränge zu identifizieren und Vorhersagen über den potenziellen Erfolg eines Films zu treffen. Beispielsweise nutzen Studios KI, um den Wert von Schauspielern für bestimmte Produktionen einzuschätzen und optimale Besetzungen zu ermitteln (Marketingscout, 2024).

### Produktion und Spezialeffekte

Während der Produktion kommen KI-gestützte Tools zum Einsatz, um Prozesse zu optimieren. Ein bemerkenswertes Beispiel ist die Verjüngung von Schauspielern mittels KI. So wurde der 87-jährige Sänger Roberto Blanco in einem geplanten Film über sein Leben durch KI-Technologie digital verjüngt, um ihn wie einen 30-

Jährigen erscheinen zu lassen (BILD, 2024). Ähnlich wurde Harrison Ford in "Indiana Jones and the Wheel of Destiny" mithilfe von KI verjüngt, indem die Software sein jüngeres Gesicht auf sein aktuelles Schauspiel projizierte (Marketingscout, 2024).

## Postproduktion und Bearbeitung

In der Postproduktion erleichtert KI Aufgaben wie die Farbkorrektur, das Editing und die Erstellung visueller Effekte. KI kann genutzt werden um Fehler zu erkennen und automatisch zu korrigieren, aber auch Farbanpassungen vorzunehmen und in manchen Fällen sogar ganze Szenen zu generieren (Chiappani, 2024). Zudem ermöglicht die Technologie die Erstellung realistischer visueller Effekte, die zuvor nur mit großem Aufwand realisierbar waren.

## Marketing und Vertrieb

KI findet auch im Marketing und Vertrieb Anwendung, indem sie Zuschauerpräferenzen analysiert und personalisierte Empfehlungen ausspricht. Dies ermöglicht es Studios, gezielte Werbekampagnen zu entwickeln und die Erfolgchancen von Filmen besser einzuschätzen (KI-Praxisbeispiele, o.D.).

KI-Technologien haben die Film- und Medienindustrie also tiefgreifend verändert. Von der Effizienzsteigerung in Produktionsprozessen bis hin zur Schaffung neuer kreativer Möglichkeiten bietet KI ein breites Spektrum an Anwendungen. Dennoch ist es essenziell, die ethischen Implikationen zu berücksichtigen und sicherzustellen, dass die Technologie verantwortungsbewusst eingesetzt wird, um sowohl die Kreativität als auch die Integrität der Filmkunst zu bewahren.

## 4. Technologische Transformationen durch KI in der Tonpostproduktion

Die Synchronisation von Filmen ist ein entscheidender Aspekt der internationalen Medienproduktion und wird durch den Einsatz künstlicher Intelligenz grundlegend verändert. Fortschrittliche KI-Technologien ermöglichen präzisere Lippenbewegungsanpassungen, realistischere Stimmklänge und effizientere Übersetzungsprozesse. Systeme wie die von Flawless AI oder DeepDub revolutionieren den Synchronisationsprozess, indem sie eine nahtlose Integration von Dialogen in verschiedenen Sprachen gewährleisten und gleichzeitig die emotionale Authentizität der Originalaufführung bewahren (Flawless AI, 2023; DeepDub, 2022). Diese Innovationen sind in der Lage, Filme einem globalen Publikum schneller und in höherer Qualität zugänglich zu machen. Dabei können Herausforderungen der traditionellen Synchronisation wie Zeitaufwand und Kosten, reduziert werden.

### 4.1 Automatisierte Sprachsynthese und Text-to-Speech (TTS)

Automatische Sprachsynthesen und Text-to-Speech-Technologien sind ein wichtiger Bestandteil der KI-Technologien zur Synchronisation fremdsprachiger Produktionen. Diese haben sich in den letzten Jahren weiterentwickelt und spielen eine immer größere Rolle in der Medienbranche. Technologien dieser Art nutzen maschinelles Lernen und neuronale Netzwerke, um menschenähnliche Stimmen zu erzeugen, die aus schriftlichen Eingaben gesprochen werden können. Ursprünglich wurde die computergenerierte Sprachsynthese dafür entwickelt, um Seh- und Sprachbehinderten Menschen die Kommunikation mit ihrer Umgebung zu vereinfachen (Stadler, 2024).

#### Funktionsweise

Ein Text-to-Speech-System setzt sich im Wesentlichen aus zwei Hauptkomponenten zusammen: der NLP-Komponente (Natural Language Processing) und der DSP-Komponente (Digital Signal Processing). Die NLP-

Komponente ist dafür verantwortlich, den eingegebenen Text zu analysieren und phonetisch zu interpretieren, während die DSP-Komponente das akustische Sprachsignal auf Basis dieser Analyse generiert.

### **NLP-Komponente**

Die NLP-Komponente (Natural Language Processing) wandelt den Text in eine phonetische<sup>1</sup> Repräsentation um. Dazu wird die gesamte Zeichenkette in einzelne Tokens oder Laute zerlegt und anhand eines festgelegten Regelwerks oder Lexikons verarbeitet. Auf diese Weise werden Aussprache, Betonung und Satzmelodie (Prosodie) präzise bestimmt, um anschließend eine flüssige und natürlich klingende Lautfolge zu erzeugen (Stadler, 2024).

### **DSP-Komponente**

Die DSP-Komponente (Digital Signal Processing) ist für die Erzeugung der Lautfolge verantwortlich. Dabei können zwei Ansätze zum Einsatz kommen: Die artikulatorische Synthese, die darauf abzielt, die menschliche Lautbildung maschinell nachzuahmen, und die Signalmodellierung, bei der zuvor aufgezeichnete Sprachsignale angepasst und kombiniert werden. Beide Ansätze greifen auf Datenbanken zurück, die charakteristische Informationen zu Sprachsegmenten enthalten. Diese Segmente werden miteinander verknüpft, um die gewünschte Äußerung zu erzeugen. Je umfangreicher die Datenbank ist, desto mehr Laute können abgedeckt werden, wodurch der Bedarf an Nachbearbeitung sinkt und die Ausgabe natürlicher klingt (Stadler, 2024).

Da eine artikulatorische Synthese durch die hohe Komplexität einen enormen Aufwand erzeugen würde, werden überwiegend Signalmodellierungen verwendet. Im letzten Schritt werden die zusammengestellten Daten in eine Waveform (Audiodatei) umgewandelt, welche dann die synthetisierte Stimme darstellt.

## **Einsatzbereiche in der Synchronisation**

In der Synchronisation finden TTS-Technologien vielfältige Einsatzmöglichkeiten. Sie werden genutzt, um Dialoge in Echtzeit zu generieren, vorläufige Sprachfassungen

---

<sup>1</sup> Phonetisch beschreibt die Aussprache und Klangstruktur von Wörtern unabhängig von ihrer Bedeutung oder grammatischen Funktion.

für die Produktionsphase bereitzustellen oder fertige Synchronfassungen in verschiedenen Sprachen zu erstellen. Besonders in der Lokalisierung von Dokumentationen oder Bildungsinhalten hat sich TTS bewährt, da es erlaubt, Inhalte kosteneffizient und schnell in mehreren Sprachen bereitzustellen. Ein gutes Beispiel wäre hier die Dokumentationsreihe „Mythen und Monster“, welche von RTL Technology erstmals mit Hilfe einer KI-Technologie von ElevenLabs komplett mit synthetisierten Stimmen produziert wurde (Bertelsmann SE & Co. KGaA, 2024).

### Vorteile

Die Vorteile von TTS liegen in der Effizienz und Flexibilität der Technologie. Produktionszeiten von Voice Dubbing, Imagefilmen, Radio Spots und vielem mehr können erheblich verkürzt werden, da menschliche Sprecher in bestimmten Phasen ersetzt werden können (Hagen, 2023). Außerdem ist es möglich, konsistente und anpassbare Stimmen zu erstellen, die auf die Anforderungen eines Projekts zugeschnitten sind. Auch die Zugänglichkeit wird durch TTS verbessert, beispielsweise für sehbehinderte Menschen, die von audiovisuellen Inhalten profitieren (Stadler, 2024).

### Nachteile/Herausforderungen

TTS-Systeme stehen jedoch auch vor Herausforderungen. Eine der größten Hürden ist die emotionale Authentizität und die Natürlichkeit der generierten Stimmen. Obwohl die Technologie beeindruckende Fortschritte erzielt hat, wirken KI-generierte Stimmen oft noch monoton oder weniger ausdrucksstark als menschliche Sprecher. Das liegt daran, dass TTS-Modelle oft Schwierigkeiten damit haben, natürliche Rhythmen, Pausen und Betonungen der menschlichen Stimme nachzuahmen. Modelle müssen also häufig manuell programmiert und angepasst werden, wodurch ein hoher Zeitaufwand entsteht (Hagen, 2023). Zudem gibt es ethische Bedenken, etwa hinsichtlich des Missbrauchs synthetischer Stimmen oder der möglichen Reduzierung von Arbeitsmöglichkeiten für professionelle Sprecher.

## Konkrete Tool Beispiele

Einige konkrete Tools haben sich in diesem Bereich hervorgetan. ElevenLabs bietet eine Plattform für die Erstellung hochgradig realistischer synthetischer Stimmen und wird häufig in der Synchronisation genutzt. DeepDub hat sich auf die Lokalisierung von Filmen spezialisiert und ermöglicht eine emotionale Anpassung der KI-Stimmen an die Schauspieler.

Beide Systeme sind in der Lage den originalen Dialog in sämtliche Sprachen zu übersetzen, dabei verschiedene Sprecher zu erkennen, die Dialogspuren vom Hintergrundsound (Musik und Effekte) zu trennen und bei neu generierten Stimmen Timing, Tonfall und Emotionen zu übernehmen. Sowohl der übersetzte Text, als auch die generierten Tonspuren können im Nachhinein bei Bedarf manuell editiert und bearbeitet werden. Zudem werden verschiedenste Dateitypen unterstützt.

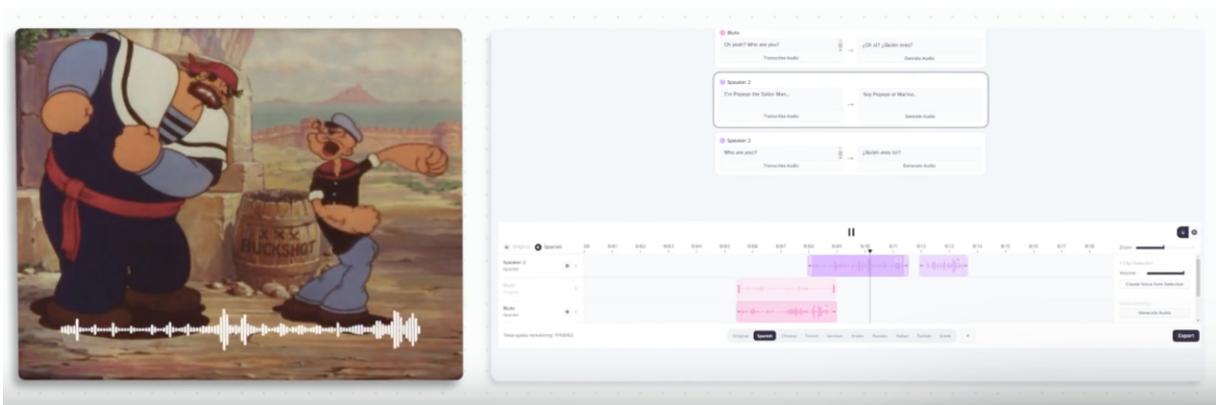


Abbildung 1 Dubbing Tool ElevenLabs (Quelle: ElevenLabs (2025))

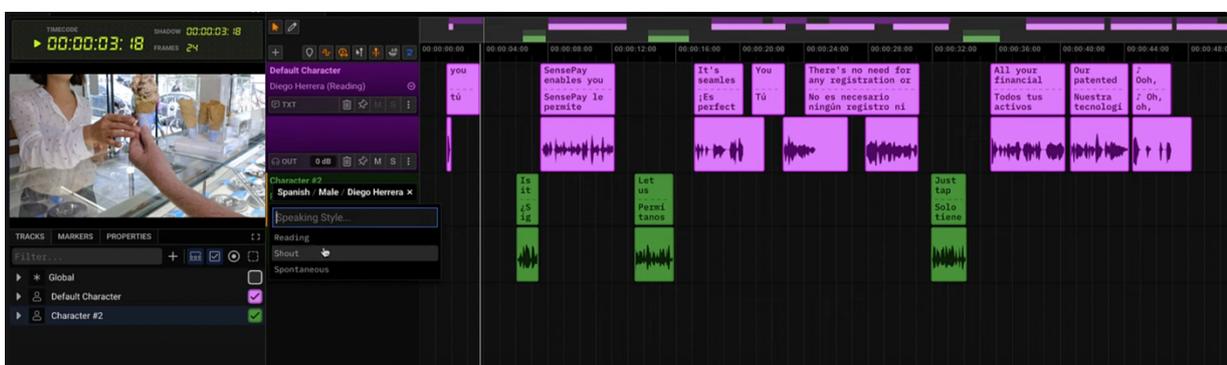


Abbildung 2 Dubbing Tool DeepDub (Quelle: DeepDub (2025))

## 4.2 Voice Cloning und Stimmklon-Technologien

Voice Cloning (auch Stimmklon-Technologie) bezeichnet die Fähigkeit, die Stimme einer Person künstlich zu reproduzieren. Dies wird durch den Einsatz fortschrittlicher Algorithmen und Sprachsynthese ermöglicht, die eine digitale Nachbildung der Stimme einer Person erstellen können. Diese Technologie verwendet Stimmdateien, die aus verschiedenen Quellen gesammelt wurden. Damit kann ein Stimmklon erzeugt werden, der jeden Text in einer Weise sprechen kann, die dem Original sehr nahekommt (SkimAI, 2024).

### Funktionsweise

Der Prozess besteht in der Regel aus zwei zentralen Komponenten: einem Text-to-Speech-(TTS)-Synthesensystem und einem Deep-Learning-Modell, das häufig auf einem generativen neuronalen Netzwerk basiert (Okeke, 2024).

Der erste und wohl wichtigste Schritt beim Klonen von Stimmen ist die Sammlung umfangreicher Sprachdaten. Dabei reicht es nicht aus, nur ein paar gesprochene Worte aufzunehmen. Um einen authentischen Stimmklon zu erzeugen, benötigt die Klon-Software eine große Vielfalt an Stimmproben. Diese sollten unterschiedliche Tonhöhen, Sprachmelodien und Muster enthalten, um die charakteristischen Eigenschaften der Stimme vollständig zu erfassen. Um die Qualität der Stimmdateien zu optimieren, werden bestenfalls (wenn möglich) hochwertige Mikrofone eingesetzt. Einige fortschrittliche Stimmklon-Programme verlangen sogar, dass bestimmte Sätze oder Phrasen gesprochen werden, damit die Software die individuellen Nuancen der Stimme präziser analysieren kann. Die gesammelten Sprachdaten dienen anschließend als Basis für die Erstellung eines detailgetreuen Stimmklons (SkimAI, 2024).

Nach der sorgfältigen Erfassung der Sprachdaten kommen hochentwickelte Algorithmen des maschinellen Lernens zum Einsatz, die das Kernstück der KI-Software für das Stimmenklonen bilden. Diese Algorithmen analysieren die aufgenommenen Stimmdateien und zerlegen sie in ihre einzelnen Bestandteile, wie Tonhöhe, Klangfarbe und die Sprechgeschwindigkeit.

Anschließend werden fortschrittliche Synthesetechniken angewendet, um diese Komponenten präzise zu reproduzieren. Das Ziel ist es, Stimmklone zu erstellen, die der Originalstimme nahezu identisch sind. Einige KI-Programme gehen noch weiter und integrieren emotionale Nuancen wie Freude, Trauer oder Begeisterung. Dadurch wird der Stimmklon nicht nur zu einer bloßen Nachahmung, sondern zu einer lebensechten und authentischen Reproduktion der Originalstimme (SkimAI, 2024). Genau diese Technologie wird bei der Synchronisation relevant.

Im nächsten Schritt kommt die Sprachsynthese zum Einsatz, welche ein essentieller Bestandteil des Klonprozesses darstellt. Diese Technologie nutzt die analysierten Sprachdaten, um neue Sätze synthetisch zu erstellen, sodass der Stimmklon eigenständig „sprechen“ kann. Fortschrittliche Algorithmen sind in der Lage, weit mehr als nur eine Stimme zu imitieren – sie können sich flexibel an unterschiedliche Akzente, Intonationen und sogar an verschiedene Sprachen anpassen (SkimAI, 2024).

Der finale und zugleich revolutionärste Schritt bei der Erstellung eines Stimmklons ist die eigentliche Generierung der synthetischen Stimme. In diesem Prozess werden alle zuvor gesammelten und analysierten Daten in einer Verbindung aus Technologie und Innovation zusammengeführt. Die KI-Software verarbeitet die detaillierten Stimmdateien (einschließlich ihrer Nuancen, Tonhöhen und Klangfarben) und kombiniert sie mit hochentwickelten Synthesealgorithmen. Das Ergebnis ist mehr als nur ein technisches Produkt. Es entsteht eine neue, künstliche Stimme, die die Originalstimme mit beeindruckender Präzision nachbildet. Die Fähigkeiten eines Stimmenklons beschränken sich nicht nur auf die reine Nachahmung. Dank der zugrunde liegenden Sprachsynthese ist der Klon in der Lage, vollkommen neue Sätze, Phrasen oder Dialoge in der Originalstimme zu generieren. Dadurch ergeben sich zahlreiche Anwendungsmöglichkeiten in unterschiedlichsten Bereichen (SkimAI, 2024).

## Einsatzbereiche in der Synchronisation

In der Film- und Fernsehbranche bietet Voice Cloning einige bedeutende Vorteile für die Synchronisation. Anstatt traditionell Synchronsprecher für jede Sprachversion eines Films oder einer Serie zu engagieren und somit für jede Sprache unterschiedliche Stimmen zu erzeugen, können Studios Voice-Cloning-Technologien nutzen, um Charakteren in verschiedenen Sprachen die „gleiche Stimme“ zu verleihen. Dies ermöglicht eine konsistente Darstellung von Charakteren über verschiedene Sprachversionen hinweg und reduziert gleichzeitig die Notwendigkeit, mehrere Synchronsprecher für verschiedene Sprachen zu engagieren (Video4net, o.D.). Auch tote Stimmen können mit Hilfe von Voice-Cloning-Technologien wieder zum Leben erweckt werden. So beispielsweise Pumuckls Originalstimme. Der bereits verstorbene Hans Clarin ist wieder Pumuckls Stimme in der RTL-Serie „Neue Geschichten von Pumuckl“. Gesprochen wird die Rolle des Pumuckl heute von Schauspieler Maximilian Schafroth. Eine KI-Software wird mit Sprachsampeln des alten Pumuckls trainiert und mit dem Tempo und der Modulation von Schafroth kombiniert (Urbe, 2023). Der Pumuckl kann nun in zwei verschiedenen Varianten geschaut werden. Die alte Stimme von Hans Clarin oder die neue Stimme von Maximilian Schafroth.

## Vorteile

Voice Cloning bietet in der Synchronisation von Filmen zahlreiche Vorteile, die die Produktionsprozesse effizienter gestalten und die Qualität der synchronisierten Inhalte erhöhen können. Einer der größten Vorzüge ist die Möglichkeit, einheitliche Charakterstimmen über alle Sprachversionen hinweg zu bewahren. Durch den Einsatz von Stimmklonen können Figuren in verschiedenen Sprachen mit der gleichen Stimmcharakteristik dargestellt werden, wodurch ein einheitliches und authentisches Erlebnis für Zuschauer weltweit geschaffen werden kann (Video4net, o.D.).

Ein weiterer bedeutender Vorteil ist die mögliche Kosten- und Zeitersparnis. Während die traditionelle Synchronisation oft mit einem hohen Aufwand an Sprecherkosten und Studiozeit verbunden ist, ermöglicht Voice Cloning eine

schnelle und kosteneffiziente Erstellung von Sprachversionen. Insbesondere bei Großproduktionen oder Serien ist dies ein erheblicher wirtschaftlicher Vorteil (Okeke, 2024).

Auch in kreativer Hinsicht eröffnet Voice Cloning neue Möglichkeiten. Beispielsweise können die Stimmen verstorbener Schauspieler/Sprecher wiederbelebt werden, um wieder in Filmen oder Serien zu erscheinen. Darüber hinaus können außergewöhnliche Stimmprofile für fiktive Charaktere entwickelt werden, die ohne diese Technologie schwer umsetzbar wären (SkimAI, 2024).

Ein weiterer Vorteil ist die Flexibilität, die mit Voice-Cloning-Technologien einhergeht. Beispielsweise bei kurzfristigen Drehbuchänderungen oder Nachbearbeitungen können durch die geklonten Stimmen problemlos neue Dialoge erzeugt werden, ohne dass der Sprecher erneut ins Studio kommen muss.

Dadurch kann der Produktionsprozess deutlich beschleunigt und eine größere Freiheit bei kreativen Entscheidungen ermöglicht werden (Speechify, 2024).

Voice Cloning bietet der Filmindustrie effiziente, einheitliche und kreative Möglichkeiten in der Synchronisation. Diese ergeben sowohl ökonomische Vorteile, als auch eine bessere Zuschauerbindung, da eine höhere Authentizität der Charakterdarstellung erzielt werden kann.

## Nachteile/Herausforderungen

Trotz der vielen Vorteile und Möglichkeiten des Voice Clonings ergeben sich auch einige Herausforderungen. Ähnlich wie TTS-Technologien können Voice-Cloning-Systeme die klanglichen Qualitäten und die emotionale Tiefe der Stimmen oft noch nicht ausreichend repräsentieren. Dadurch müssen oft Anpassungen und Korrekturen vorgenommen werden um ein passendes Ergebnis zu erzielen, was wiederum mit einem höheren Zeitaufwand verbunden ist (Hagen, 2023). Zudem können beim Klonen der Stimmen signifikante Fehler (Artefakte) auftreten. Dies kann dazu führen, dass das Tonmaterial neu aufgenommen werden muss, wodurch es zu Verzögerungen kommen kann (Hagen, 2023).

Neben den technischen Einschränkungen gibt es auch einige ethische Bedenken und rechtliche Aspekte, welche berücksichtigt werden müssen.

Durch die Möglichkeit, jede Stimme klonen zu können, besteht das Risiko von Betrug, Identitätsdiebstahl oder Missbrauch. Daher ist es dringend erforderlich, die Zustimmung der Person einzuholen, deren Stimme repliziert werden soll. Besonders wichtig wird dies, wenn es sich um Privatpersonen oder Persönlichkeiten des öffentlichen Lebens handelt. Zudem kann die Fähigkeit, jede Stimme klonen zu können, einige Fragen der Authentizität audiovisueller Inhalte aufwerfen (Video4Net, o.D.). Inhalte wie Nachrichtenbeiträge wichtiger Personen müssen möglicherweise stärker auf Echtheit geprüft werden.

Ein gutes Beispiel für den Missbrauch der Voice-Cloning-Technologie, ist der Fall der BBC-Moderatorin Liz Bonnin, deren Stimme ohne ihr Einverständnis für Werbung genutzt wurde (Webb, 2024). Aufgrund solcher Vorfälle, müssen dringend rechtliche Maßnahmen getroffen werden, um die Öffentlichkeit und ihre Einzelpersonen sowie Künstler zu schützen und einen verantwortungsvollen Umgang mit diesen mächtigen Technologien zu gewährleisten. Gesetzesvorschläge wie der NO FAKES Act in den USA (Webb, 2024) oder das KI-Gesetz, welches 2026 in Deutschland in Kraft treten soll, sollen KI-Technologien vertrauenswürdig machen und die Grundrechte der Menschen schützen (Europäische Kommission, 2024).

### Konkrete Tool Beispiele

Es gibt einige Anbieter, die Voice-Cloning-Systeme anbieten. Darunter beispielsweise ElevenLabs, Resemble AI, Murf.ai oder Speechify. Um zwei dieser Tools näher zu betrachten, werden die Voice-Cloning-Technologien von ElevenLabs und Resemble AI ausführlicher betrachtet.

#### **ElevenLabs**

ElevenLabs bietet zwei verschiedene Optionen zum Klonen einer Stimme an: das Instant Voice Cloning (IVC) und das Professional Voice Cloning (PVC). Die Voice-Cloning-Technologien von ElevenLabs unterstützen 32 Sprachen und versprechen eine perfekte Sprachkopie, welche den Tonfall, die Betonung und die emotionale Bandbreite bewahren. Eine präzise Steuerung soll durch ein trainierbares Modell ermöglicht werden, wodurch der gewünschte Sprechstil eingestellt werden kann. Die IVC-Variante ermöglicht eine schnelle und einfache

Replikation, während die PVC-Variante eine genauere und anpassbare Option ist. IVC nutzt kein benutzerdefiniertes KI-Modell, sondern greift auf bereits vorhandene Trainingsdaten zurück, wodurch ein Stimmklon aus nur wenigen Sprachsamples in einer sehr kurzen Zeit erstellt werden kann (ElevenLabs, 2025). Dieses Modell ist jedoch nicht für sehr einzigartige Stimmen geeignet, da möglicherweise keine ähnlichen Trainingsdaten zur Verfügung stehen. In diesem Fall sollte auf die PVC-Variante zurückgegriffen werden. Mit dieser Variante kann ein hyperrealistisches Modell anhand eines großen Datensatzes trainiert werden. Dadurch kann ein Modell erstellt werden, welches nicht von der Originalstimme zu unterscheiden ist. Dieses System ist jedoch mit einem deutlich höheren Zeitaufwand verbunden (ElevenLabs, 2025).

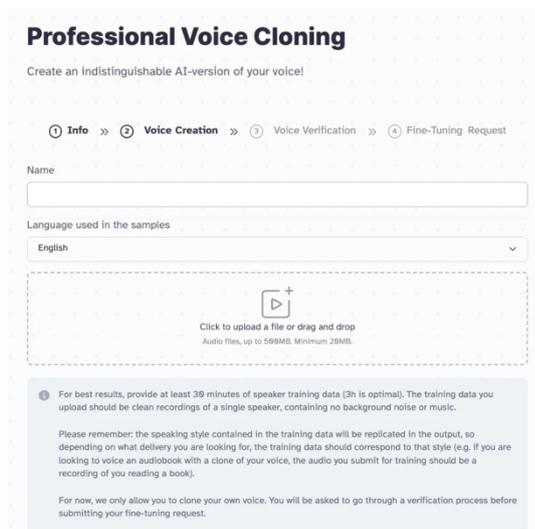


Abbildung 4 PVC Tool ElevenLabs (Quelle: Reddit (2025))

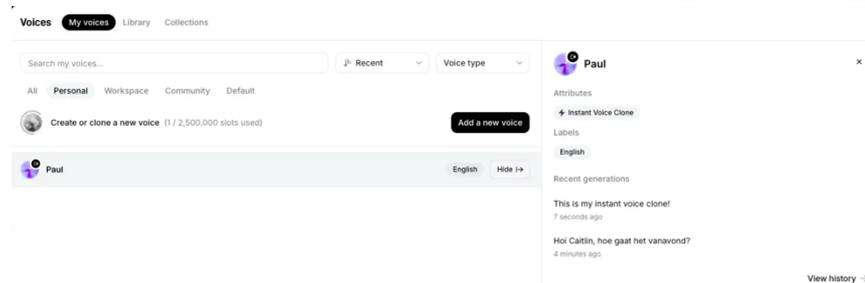


Abbildung 3 IVC Tool ElevenLabs (Quelle: ElevenLabs (2025))

## Resemble AI

Das Voice-Cloning-Tool von Resemble AI verspricht durch Advanced Voice Cloning (welches die Betonungen, Feinheiten und den Rhythmus ermittelt) eine Nachbildung der Stimme, die praktisch identisch mit dem Original ist. Anders als ElevenLabs unterstützt Resemble AI über 142 Sprachen. Resemble AI verspricht andere Anbieter in Sachen Schnelligkeit und Effektivität (Stimmen mit nur zehn

Sekunden Audiomaterial schnell zu klonen) weit zu übertreffen und damit Zeit und Ressourcen zu sparen. Dafür wird lediglich eine klare Stimmprobe der Zielstimme benötigt (Resemble AI, 2025).

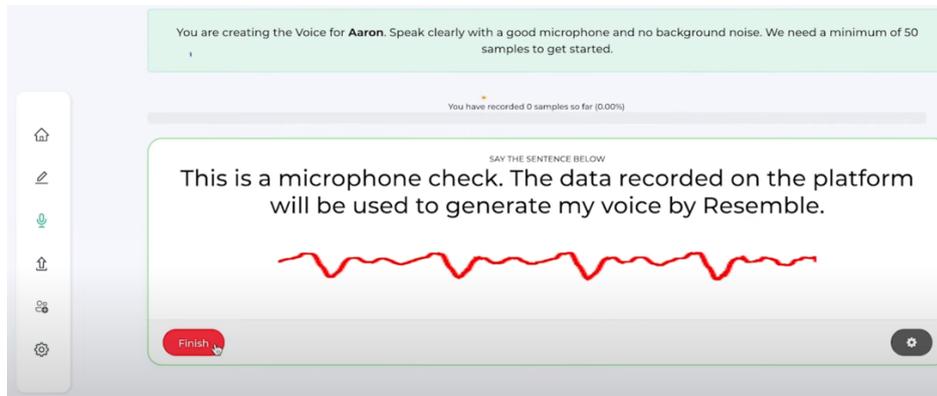


Abbildung 5 Voice Cloning Tool Resemble AI (Quelle: Resemble AI (2025))

Für die Anwendung in der Synchronisation wird eine hohe Qualität und eine präzise Replikation der Originalstimme mit all ihren Eigenschaften benötigt, um eine verwendbare (von den Konsumenten akzeptierbare) Übersetzung zu erhalten.

Voice-Cloning-Technologien finden also bereits ihre Anwendung in der Synchron-Branche und werden vor allem für die Replikation der Stimmen verstorbener/nicht zur Verfügung stehender Sprecher/Schauspieler verwendet. Voice Cloning kann ein sehr nützliches Tool sein, welches Prozesse in der Tonpostproduktion wesentlich beschleunigen und vereinfachen kann. Jedoch ist es ein sehr mächtiges Tool, welches für schädliche Zwecke missbraucht werden kann. Um dies zu verhindern und einen vertrauenswürdigen Umgang zu garantieren, sind dringliche rechtliche Maßnahmen erforderlich.

### 4.3 KI-gestützte Lippensynchronisation

Eine weitere Möglichkeit zur Revolutionierung der Synchronisation fremdsprachiger Produktionen ist die KI-gestützte Lippensynchronisation. Hierbei geht es darum, dass gesprochene Dialoge eindeutig auf die Lippenbewegungen der sprechenden Rolle passen. Bei der traditionellen Synchronisation von Filmen oder Serien ist eine perfekte Lippensynchronität kaum umsetzbar, da der übersetzte Text andere Lippenbewegungen erfordert. Diese Abweichungen können für manch aufmerksame

Zuschauer irritierend und in manchen Fällen sogar störend sein. Selbst gut produzierte Synchronfassungen können in der Regel kein makellooses Ergebnis erzielen. Dies soll durch KI-Technologien ermöglicht werden. Deep-Fake-Systeme sollen Lippenbewegungen automatisch und in viele verschiedene Sprachen synchronisieren können. Dabei wird das Bildmaterial nachträglich an die gewünschte Zielsprache angepasst. Theoretisch wäre dies auch durch klassische CGI-Technik umsetzbar, was jedoch mit einem beträchtlichen manuellem Animationsaufwand verbunden wäre (Bastian, 2021). Dies wäre sehr kostspielig und würde keinen wirtschaftlichen Vorteil schaffen.

### Funktionsweise

KI-gestützte Lippensynchronisations-Systeme (auch Lip-Sync-Systeme genannt) basieren in der Regel auf neuronalen Netzen, welche sowohl Audio- als auch Videosignale analysieren und die Lippenbewegungen dementsprechend generieren/anpassen. Die Eingaben bestehen üblicherweise aus drei Komponenten: das Bild, auf welchem der Mund verändert werden soll, ein Referenzbild (üblicherweise mehrere), welches zeigt, wie die Person aussieht und eine Audioeingabe. Durch die Verarbeitung dieser Daten mit drei Encodern, werden komprimierte Darstellungen erstellt, welche miteinander interagieren und somit realistische Mundformen erzeugen. Hierbei geht es darum, dort einen Mund zu generieren, wo er maskiert ist (beispielsweise bei animierten Produktionen) oder einen bestehenden Mund anzupassen. Dabei wird das Aussehen der Person und deren momentan gesprochenes Wort berücksichtigt. Dies ist ein sehr aufwendiger Prozess, welcher komplexe Modifikationen, sowie eine Vielzahl an Referenzbildern zur Erfassung des Aussehens der Person, verschiedene Gesichtsmodelle und unterschiedliche Audiokodierungsmethoden benötigt (Rezanov, 2024).

KI-basierte Systeme sind in der Lage Sprachspuren zu analysieren, dabei Phoneme<sup>2</sup> zu identifizieren und diese dann den entsprechenden Mundbewegungen zuzuordnen. Diese werden auch gleichzeitig an die richtige Stelle in der Zeitleiste

---

<sup>2</sup> Ein Phonem ist die kleinste Spracheinheit, welche Wortbedeutungen unterscheidet.

des Videos eingefügt. Dabei weiß die KI durch bestehende Datensätze oder neuen Analysen des Videomaterials, wie die Mundbewegungen der entsprechenden Phoneme auszusehen haben (Ames, 2023).



Abbildung 6 Lippenbewegungen für bestimmte Laute (Quelle: Vidnoz (2025))

Erzeugte Lippensynchronisationen können durch KI-Systeme zudem automatisch verbessert und optimiert werden, sodass die angepassten Inhalte beispielsweise mit Hilfe von hinzugefügten Antizipationsbewegungen<sup>3</sup> natürlicher aussehen.

Auch emotionale Anhaltspunkte können von KI- und ML-Algorithmen (ML für maschinelles Lernen) ermittelt und bei Mundbewegungen berücksichtigt werden. Die Mundbewegungen werden dann so angepasst, dass sie die jeweiligen Emotionen so genau wie möglich widerspiegeln. Viele KI-gestützte-Lippensynchronisations-Systeme übernehmen nicht nur die Lippenanpassung, sondern die gesamte Lokalisation der Produktion. Dadurch kann der Prozess beschleunigt und auf mehrere Sprachen angewandt werden (Ames, 2023).

### Einsatzbereiche und Möglichkeiten in der Synchronisation

Noch finden KI-gestützte-Lippensynchronisations-Tools in der Synchronbranche wenig Anwendung. Trotz der bisher oft noch mangelnden Qualität, bieten Lip-Sync-Technologien große Chancen und Möglichkeiten für die Synchronisation fremdsprachiger Produktionen. Filme, Fernsehsendungen oder Online-Videos können schnell und effizient in viele verschiedene Sprachen übersetzt und synchronisiert

---

<sup>3</sup> Vorrusschauende Bewegungen

werden. Dabei können jegliche Abweichungen der Lippenbewegungen korrigiert beziehungsweise vermieden werden. Störende Fehler, welche bei traditionellen Synchronisationen oft durch Asynchronität der Lippen auftreten, können dadurch mit Hilfe von künstlicher Intelligenz minimiert werden. Gleichzeitig kann durch die schnelle Lokalisierung ein größeres internationales Publikum erreicht werden, wodurch Einnahmequellen gesteigert werden können (Futuretrans, 2024). Auch bei animierten Produktionen können Zeichentrickfiguren oder CGI-Charaktere schneller mit einer Stimme versehen und zum Leben erweckt werden. Das ermöglicht eine realistischere und authentischere Darstellung der fiktiven Charaktere. Auch Bildungsinhalte können für Lernende auf der ganzen Welt, unabhängig ihrer Muttersprache, synchronisiert und zugänglich gemacht werden. Damit können globale Lernmöglichkeiten gefördert und ausgeweitet werden (Mindverse, 2024). Unternehmen können KI-basierte-Synchronisations-Techniken nutzen, um personalisierte Werbevideos zu produzieren, welche auf verschiedene Zielgruppen abgestimmt sind. Dabei können auch internationale Partner und Kunden problemlos erreicht werden. Zudem ermöglichen es bestimmte Tools, realistische Avatare zu erstellen, welche Produkte und Dienstleistungen in verschiedenen Sprachen präsentieren (Mindverse, 2024).

KI-gestützte-Lippensynchronisations-Technologien können also vielseitig eingesetzt werden und bieten die Möglichkeit präzisere und schnellere Prozesse zu erzielen. Außerdem kann durch große Sprachauswahl eine deutlich größere Reichweite erreicht werden.

## Vorteile

Wie auch die anderen bereits vorgestellten KI-Technologien können auch bei dieser Technologie Zeit und Kosten eingespart werden. Produktionszeiten können erheblich gekürzt und Inhalte schneller bereitgestellt werden. Prozesse, welche auf traditioneller Weise sehr kostspielig wären, können durch Automatisierung deutlich kosteneffizienter gestaltet werden (Futuretrans, 2024).

Durch die nahezu perfekt angepassten Lippenbewegungen an eine beliebige Sprache, wirkt die Synchronisation natürlicher und sorgt für ein besseres und glaubwürdigeres Zuschauererlebnis.

Zudem können große Mengen an Inhalten in mehreren Sprachen effizient produziert oder im Nachhinein übersetzt werden.

All diese Vorteile treten natürlich nur dann ein, wenn das angewandte System ein vom Nutzer akzeptierbares Ergebnis erzielt. Dies kann sich als sehr schwierig herausstellen, da insbesondere der deutsche Markt an eine sehr aufwendige und vor allem anspruchsvolle Adaption gewöhnt ist (Lückerath, 2025).

## Nachteile/Herausforderungen

Die Entwicklung und Implementierung von KI-gestützter Lippensynchronisation stehen vor einigen Herausforderungen, welche beträchtliche Nachteile für den Einsatz von Lip-Sync-Technologien hervorbringen. Diese sind sowohl von technischer als auch ethischer Natur.

### **Technische Herausforderungen**

Einer der größten technischen Hürden besteht darin, eine realistische visuelle Qualität zu erreichen und dabei eine präzise Synchronisation zwischen Audio und Video zu gewährleisten. Das Team von Rask AI beispielsweise musste mehrere Strategien anwenden, um diesen Schwächen entgegenzuwirken. Dazu gehören unter anderem die Modifizierung der neuronalen Netzwerkarchitektur, die Anpassung und Verbesserung des Trainingsverfahrens oder auch die Optimierung des Datensatzes (Rezanov, 2024).

Eine weitere große Herausforderung ist die Anpassung der Lippensynchronisation an verschiedene Sprachen. Entscheidend ist hierbei die Länge der übersetzten Texte, da diese in den jeweiligen Sprachen unterschiedlich lang sind. Deshalb ist es dringend erforderlich bei der Videosynchronisation die Geschwindigkeit der Aussprache zu kontrollieren und gleichzeitig die Länge der Originaldialoge und die Dauer der Ziltexte zu bestimmen. Zur Berücksichtigung dieses Aspekts, können SpeechTempo-Modelle eingesetzt werden, wie es bei Vidby der Fall ist. Hierfür

werden Algorithmen eingesetzt, welche die Lippenbewegungen analysieren und dabei die Tonspur entsprechend anpassen (The Decoder, 2023).

### **Ethische und gesellschaftliche Herausforderungen**

Zusätzlich besteht die Gefahr des Missbrauchs der Lip-Sync-Technologien. Durch die Erstellung manipulativer Deepfakes, welche falsche Informationen verbreiten oder Personen in unangebrachten Situationen darstellen, können erhebliche Auswirkungen auf die Privatsphäre oder den Ruf von Individuen entstehen (Bieß & Pawelec, 2020).

### **Datenschutz und Sicherheit**

KI-Systeme zur Lippensynchronisation verarbeiten oft riesige Mengen persönlicher Daten, wodurch Datenschutzbedenken auftreten. Die Einhaltung von Datenschutzstandards wie der DSGVO in Europa ist für die verantwortungsvolle Nutzung dieser Technologien und den Schutz jedes Individuums dringend erforderlich (Software-Aspekte, 2024).

### **Hohe Rechenleistung**

Ein weiterer Nachteil ist die hohe Rechenkapazität, welche KI-basierte Videobearbeitungstools benötigen, um die enormen Datenmengen zu verarbeiten und den gewünschten Inhalt zu generieren.

### **Konkrete Tool Beispiele**

Einer der führenden Unternehmen für KI-gestützte Lippensynchronisation ist Flawless, mit ihrem TrueSync Tool, welches eine einwandfreie visuelle Übersetzung in jegliche Sprachen verspricht. Eine perfekte visuelle Synchronisation soll erreicht werden, indem TrueSync jede Leistungsnuance eines Schauspielers erkennt und daraufhin neue Lippenbewegungen generiert.

Es soll die Reichweite des Publikums global erhöhen und eine schnellere, bessere Lokalisierung ohne Nachdrehs oder Kompromisse gewährleisten (Flawless, o.D.).

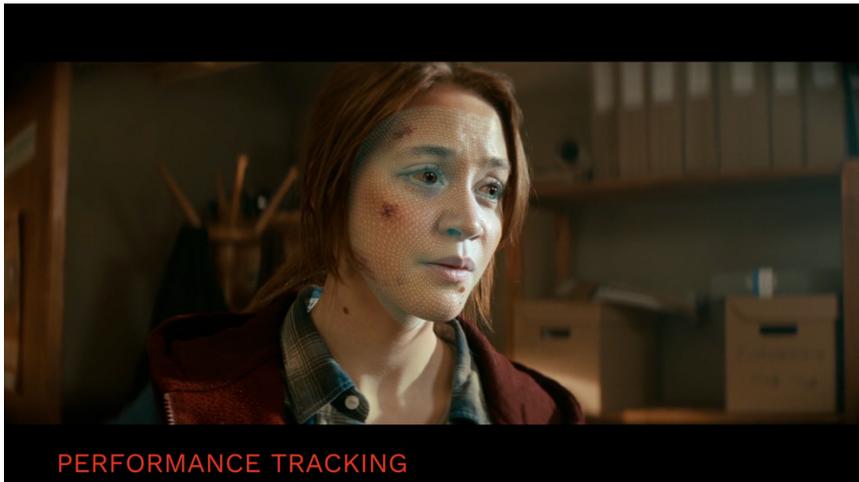


Abbildung 7 *Flawless TrueSync Tool* (Quelle: *Flawless* (2025))

Die KI-gestützte Lippensynchronisation bietet enorme Potenziale in der Medienproduktion und eröffnet neue Möglichkeiten für eine authentischere Synchronisation. Trotz der technischen und ethischen Herausforderungen wird die Technologie weiterhin optimiert und könnte in Zukunft eine höhere Einsatzquote, besonders in der Film-, Game- und E-Learning-Branche erzielen.

#### 4.4 Automatisierte Übersetzung und Kulturadaption durch KI

In einer zunehmend globalisierten Welt spielt die automatisierte Übersetzung eine große Rolle bei der Überwindung von Sprachbarrieren. Künstliche Intelligenz ermöglicht eine schnelle und präzise Übersetzung von Texten und gesprochenen Inhalten. Dabei können auch kulturelle Aspekte integriert werden. Moderne KI-Systeme können neben der Wortbedeutung auch kulturelle Kontexte, Redewendungen und sprachliche Nuancen analysieren. Anwendung findet diese Technologie hauptsächlich in Bereichen wie internationaler Geschäftskommunikation, Film- und Spielelokalisierung oder E-Learning. Trotz ihrer Vorteile stehen KI-gestützte Übersetzungen vor Herausforderungen. Besonders bei der Erfassung von Emotionen, Humor und kulturellen Einzelheiten stoßen diese Systeme oft auf Probleme.

## Funktionsweise

KI-gestützte Übersetzungssysteme sind in der Lage große Mengen an Textdaten zu analysieren und anhand dieser die Strukturen und Muster verschiedener Sprachen zu erkennen. Durch maschinelles Lernen können diese Systeme ihre Genauigkeit konstant optimieren. Dabei können die verschiedensten Eigenschaften und Bedeutungen einer Sprache gelernt und berücksichtigt werden. Ein bedeutender Aspekt bei der automatisierten Übersetzung von Texten ist die kulturelle Adaption. Hierbei handelt es sich um die kontextuelle Übertragung von kulturellen Elementen von der Ausgangssprache in die gewünschte Zielsprache. Um dies zu ermöglichen, wird ein tiefes Verständnis sowohl der Ausgangs- als auch der Zielkultur benötigt. Große Sprachmodelle (Large Language Models (LLMs)) sind in der Lage, diese kulturellen Elemente in verschiedenste Sprachen zu übertragen, da diese mit riesigen Datenmengen (deshalb Large) trainiert werden und somit über ein immenses kulturelles Wissen verfügen. Im Gegensatz zu anderen Übersetzungsmodellen, welche auf die reine Übersetzungsaufgabe trainiert werden, sind bei LLM-Systemen die kulturellen Aspekte in deren Parametern eingebettet (Singh et al., 2024).

LLMs basieren auf einem Transformator Modell, welches Eingaben in Token umwandelt und gleichzeitig mathematische Gleichungen ausführt, um Beziehungen zu ermitteln. Durch schnelle Lernmethoden, können LLMs verschiedene Teile oder sogar den gesamten Kontext eines Satzes berücksichtigen. Eingaben werden gelesen, interpretiert und anschließend verarbeitet. Um Texteingaben interpretieren zu können, müssen LLMs mit umfangreichen Datensätzen trainiert werden. Dabei sind die Menge und die Qualität der Daten essentiell für die Lernfähigkeit des Systems (Safar, o.D.).

## Einsatzbereiche in der Synchronisation

In der Medienbranche, besonders bei der Synchronisation von Filmen, Serien oder Videospiele, könnte die KI-gestützte Übersetzung eine bedeutende Rolle spielen. Sie ermöglicht eine schnelle und kosteneffiziente Lokalisierung von Inhalten auf verschiedenen Märkten. Dabei können nicht nur Dialoge übersetzt, sondern auch

kulturelle Nuancen angepasst werden, um die Authentizität zu bewahren und die Zuschauerbindung zu stärken. So können Drehbücher schnell in die unterschiedlichsten Sprachen übersetzt werden. Davon können besonders Autoren profitieren, welche in Minderheitssprachen schreiben, wodurch das Übersetzen ins Englische oder in andere Sprachen bisher sehr aufwendig war. Auch kleinere Unternehmen haben durch die Möglichkeit der Kostenersparnis, höhere Chancen ein größeres Publikum zu erreichen. Jedoch sind dennoch oft eine sorgfältige Prüfung und Bearbeitung notwendig, da LLM-Systeme oft nicht fehlerfrei arbeiten oder angemessene Texte generieren (MacDonald, 2024).

### Vorteile

Die Integration von Künstlicher Intelligenz in Übersetzungsprozesse bietet vielerlei Vorteile. Besonders Geschwindigkeit und Effizienz können durch die Möglichkeit, große Textmengen innerhalb kürzester Zeit zu übersetzen, optimiert werden. Dabei können besonders Unternehmen mit einem großen internationalen Markt profitieren. Zudem sind KI-basierte Übersetzungssysteme sofort einsetzbar und können dabei mehrere Projekte zeitgleich bearbeiten. Durch diese Effizienzsteigerung können auch Projektkosten reduziert werden, da geringere Bearbeitungsstunden oder Gebühren anfallen. So können auch kleinere Start-Ups ohne enorme Kosten eine internationale Kundenbasis schaffen (ReSartus, o.D.).

Außerdem können die Bedürfnisse und Erwartungen verschiedener Zielgruppen durch die Fähigkeit der kulturellen Adaption angepasst werden.

### Nachteile/Herausforderungen

Trotz der Fortschritte stehen KI-gestützte Übersetzungen vor einigen Herausforderungen. KI-Systeme können Schwierigkeiten haben den gesamten Kontext oder die Details eines Texts zu erfassen, wodurch Ungenauigkeiten oder Fehler entstehen können. Zudem können KI-geschriebene Texte teilweise noch „roboterhaft“ klingen, wodurch die Qualität der Texte stark beeinflusst werden kann. Die Nutzung KI-basierter Übersetzungs-Tools wirft außerdem datenschutzrechtliche

Bedenken auf, was besonders bei sensiblen und vertraulichen Inhalten relevant ist (Degenkolb, o.D.).

Ein weiterer Punkt sind die kulturellen Nuancen, welche nicht immer vollständig reproduziert werden können. Für die exakte Wiedergabe von Humor, Redewendungen oder emotionalen Untertönen erfordert es oft menschliches Feingefühl. Auch Ethische Überlegungen stellen KI-gestützte Übersetzungen vor große Herausforderungen. Besonders in kreativen Bereichen kann der Einsatz von KI zu Bedenken hinsichtlich der Qualität und Authentizität führen (MacDonald, 2024).

Ebenfalls fraglich ist, ob die KI inhaltliche Kontexte im Bezug zu der gespielten Zeit, des Jargon, des Slangs richtig einordnen kann. Versteht die KI, wenn der Film in den 80er Jahren spielt, dass es noch keinen Van gibt? Oder wenn er vor den 70er Jahren spielt, dass das Wort geil noch nicht seine Zweitbedeutung hat? In welchem Alter, welchem Stand, welchem Ausbildungsgrad würde wer was wie sagen? Welcher Slang gilt ab wann? Welche Person würde es verwenden? Welcher Humor, der sich durch welche Worte ausdrückt, passt zu welcher Person?

Ob etwas von einem schwarzen 16-Jährigen in L.A. gesagt wird oder von einem weißen britischen Professor, macht einen enormen Unterschied. Trotzdem gibt es viele Fälle, in denen beide die gleichen Sätze verwenden können (Frommann, 2025).

Auch unterschiedliche Zeitformen müssen berücksichtigt werden. Beispielsweise verwenden Engländer überwiegend das Präteritum, während die Deutschen häufig das Perfekt nutzen (Frommann, 2025). Bei Übersetzungen müssen dann kreative Kompromisse und Lösungen gefunden werden.

Eine weitere große Herausforderung bei der automatisierten Übersetzung der Dialogbücher, ist die Lippenasynchronität. Beispielsweise kann ein TH im Englischen mit herausgestreckter Zunge, im Deutschen nur mit einem L ausgeglichen werden. Eine andere Möglichkeit wäre, den Satz komplett umzustellen (Frommann, 2025). All diese Faktoren müssen bei Übersetzungen berücksichtigt werden. Jedoch ist es sehr fraglich, inwieweit KI-basierte Systeme dazu in der Lage sind.

Zusammenfassend bietet die automatisierte Übersetzung und Kulturadaption mit Hilfe von KI erhebliche Vorteile in Bezug auf Effizienz und Reichweite. Dennoch ist die menschliche Expertise weiterhin unverzichtbar um sicherzustellen, dass kulturelle Feinheiten und kontextuelle Bedeutungen korrekt und authentisch vermittelt werden.

## 5. Auswirkungen auf die Synchronbranche

Die wachsende Integration künstlicher Intelligenz verändert die Synchronbranche grundlegend. Automatisierte Lippensynchronisation, KI-generierte Stimmen und maschinelle Übersetzungstechnologien ermöglichen eine schnellere und kosteneffizientere Produktion. Dadurch können traditionelle Abläufe zunehmend ersetzt oder ergänzt werden. Durch diese Entwicklungen werden traditionelle Standards infrage gestellt und beeinflussen sowohl die Arbeitsabläufe in Synchronstudios, als auch die Rolle menschlicher Sprecher. Jedoch bietet KI nicht nur neue Möglichkeiten zur Lokalisierung von Inhalten, sondern wirft auch Fragen zu Qualität, Authentizität und den langfristigen Auswirkungen auf die Branche auf. Auch ethische und rechtliche Fragen beschäftigen die Synchronindustrie enorm.

### 5.1 Veränderungen in Arbeitsabläufen und Berufsbildern

Die Einführung künstlicher Intelligenz in die Tonpostproduktion, besonders im Bereich der Filmsynchronisation, hat tiefgreifende Auswirkungen auf bestehende Arbeitsabläufe und Berufsbilder. Die traditionelle Filmsynchronisation erfordert eine hohe Zusammenarbeit verschiedener Fachkräfte wie Synchrosprecher, Tontechniker, Dialogregisseure und Cutter. Zudem ist die klassische Synchronisation ein sehr zeit- und arbeitsintensiver Prozess. Von der Übersetzung der Dialoge über die Anpassung der Lippenbewegungen bis hin zur finalen Mischung, werden alle Schritte manuell und zeitaufwendig durchgeführt. Mit der Integration von KI-Technologien können viele dieser Prozesse automatisiert und beschleunigt werden. Durch KI-gestützte Tools wie automatische Spracherkennung, maschinelle Übersetzung und Voice-Cloning-Systeme, können Dialoge schneller transkribiert, übersetzt und sogar in der Stimme der Originaldarsteller nachgebildet werden. Dies reduziert nicht nur den Zeitaufwand, sondern auch die Kosten, da weniger

manuelle Eingriffe notwendig sind. KI-Systeme wie DeepL oder Google Translate können bereits heute Texte in hoher Qualität übersetzen, während Tools wie Respeecher oder ElevenLabs realistische Sprachsequenzen synthetisieren können (IT-P GmbH, 2024). Diese Technologien verändern die Rolle der Synchronsprecher, welche zunehmend als „Stimmgeber“ für KI-Systeme fungieren, anstatt jeden Dialog manuell einzusprechen.

Gleichzeitig entstehen neue Berufsbilder und Kompetenzanforderungen. Fachkräfte in der Tonpostproduktion müssen sich zunehmend mit KI-Tools und deren Anwendung auseinandersetzen. Kenntnisse in der Bedienung von KI-Software, der Analyse von Sprachdaten und der Feinabstimmung automatisierter Prozesse werden immer wichtiger. KI-Berufe gewinnen immer mehr an Bedeutung, da sie die Schnittstelle zwischen Technologie und kreativer Gestaltung bilden. Kontinuierliche Weiterbildung und Anpassungsfähigkeit seitens der Fachkräfte werden dabei ebenfalls immer relevanter (Burgemeister, 2023).

Allerdings birgt diese Transformation auch Herausforderungen. Die Automatisierung könnte dazu führen, dass bestimmte Tätigkeiten, wie das manuelle Schneiden von Dialogen oder die klassische Synchronisation an Bedeutung verlieren. Für manche mag das ein Vorteil sein, da sie somit schneller und mehr arbeiten können und damit auch mehr Geld verdienen, während sie sich nicht mit den Schauspielern auseinandersetzen müssen (Von Medvey, 2025). Dies wirft jedoch Fragen nach der Zukunft traditioneller Berufsbilder und der ethischen Verantwortung im Umgang mit KI auf. Wie können die kreativen Aspekte der Synchronisation bewahrt werden, wenn KI-Systeme zunehmend standardisierte Lösungen anbieten? Und wie lässt sich sicherstellen, dass die Qualität und künstlerische Integrität der Synchronisation nicht unter der Effizienzsteigerung leidet?

Auch Nachwuchstalente werden es sehr schwer haben, sich in der Branche etwas aufzubauen. Dabei stellt sich die Frage, wie lange es noch Sinn macht Sprecher auszubilden? Oder was sagt man den Leuten, die das wirklich gerne machen würden (Von Medvey, 2025)?

Insgesamt zeigt sich, dass die KI-Revolution in der Tonpostproduktion sowohl Chancen als auch Risiken mit sich bringt. Während sie die Arbeitsabläufe effizienter gestaltet und neue Berufsfelder schafft, erfordert sie gleichzeitig eine kritische Auseinandersetzung mit den Auswirkungen auf die kreative Branche und ihre traditionellen Werte.

## 5.2 Qualität und Konsistenz der KI-gestützten Synchronisation

### Qualität der bestehenden Tools

Die Qualität der KI-gestützten Synchronisation hat in den letzten Jahren erhebliche Fortschritte gemacht, ist jedoch noch nicht auf dem Niveau menschlicher Leistungen. Tools wie ElevenLabs, DeepDub, Respeecher oder Voicery sind bereits in der Lage Sprachsynthesen zu erstellen, die den Originalstimmen anhand ihrer Tonhöhe, Klangfarbe und Sprachmelodie sehr nahekommen. Durch Deep-Learning-Algorithmen können sogar emotionale Nuancen berücksichtigt und nachgeahmt werden. Trotz der im ersten Eindruck oft beeindruckenden Leistung einer KI-generierten Synchronisation, stößt diese bei komplexeren Dialogen, welche häufig mit starken emotionalen Ausdrücken oder individuellen sprachlichen Eigenheiten verbunden sind, zumeist an ihre Grenzen. Dennoch sind in den letzten Jahren sowohl gute als auch schlechte Ergebnisse durch die Anwendung KI-basierter Tools entstanden. Eines der gut gelungenen Beispiele, ist die bereits erwähnte zum Leben erweckte Stimme von Hans Clarin als Pumuckl in den neuen Folgen der RTL-Serie (Urbe, 2023). Die durch Voice-Cloning wiedererweckte Stimme des „originalen“ Pumuckls, kommt dem Original beeindruckend nah.

Ein verhältnismäßig schlechtes Beispiel ist die von Magenta TV veröffentlichte Serie „Murderesses“. Die bereits nicht mehr verfügbare Serie wurde bereits 48 Stunden nach Veröffentlichung wieder von der Plattform genommen. Grund dafür war die auffällig schlechte Synchronisation, welche sich durch sehr monotone Dialoge und mangelnder Betonungen auszeichnete. Synchronisiert wurde die Serie mit Hilfe von künstlicher Intelligenz des Unternehmens DeepDub (Lückerath, 2025).

Die Qualität KI-gestützter Tools hängt also sehr stark von der Komplexität und Art des Projekts, sowie von der Datenmenge und -qualität ab, mit der das System trainiert wird. Zudem spielt auch die Fähigkeit, kulturelle und sprachliche Feinheiten zu erfassen eine essentielle Rolle.

### Vergleich zwischen KI und menschlicher Synchronisation

Im direkten Vergleich zwischen KI und menschlicher Synchronisation zeigt sich, dass menschliche Synchronsprecher zum aktuellen Zeitpunkt nach wie vor überlegen sind, wenn es um die Interpretation von Dialogen und die Anpassung an kulturelle Kontexte geht. In den verschiedenen Produktionsphasen, in welchen KI-basierte Tools eingesetzt werden können, sind bisher menschliche Leistungen kaum wegzudenken, da diese subtile emotionale Aspekte, Humor oder Ironie besser interpretieren und vermitteln können (Habimex, 2024). Sei es bei der Übersetzung der Dialoge, der Erstellung eines Dialogbuchs, den eigentlichen Sprachaufnahmen oder der Postproduktion, KI-Systeme haben meist Probleme damit, starke emotionale Ausdrücke präzise zu reproduzieren. Dadurch entsteht häufig ein „roboterhafter“ Klang, welcher einen enormen Einfluss auf die Authentizität der Synchronisation hat (Degenkolb, o.D.). Obwohl KI-Systeme Fortschritte gemacht haben, erreichen sie nicht die emotionale Tiefe menschlicher Stimmen.

Zudem bringen menschliche Sprecher ein tieferes Verständnis für kulturelle und sprachliche Feinheiten mit, was zu einer authentischen und akkuraten Darstellung führt. KI-Systeme haben häufig Schwierigkeiten, solche Nuancen korrekt zu interpretieren und wiederzugeben, wodurch weniger überzeugende Ergebnisse entstehen (Habimex, 2024). Auch Sprichwörter bereiten KI-basierten Systemen zumeist Probleme, da diese oft nicht korrekt interpretiert werden können. So werden Sprichwörter wie „Ich glaub, mein Schwein pfeift“ in „I think my pig is whistling“ übersetzt (Stopa, 2023).

Ein weiteres Beispiel wäre hier: Der Footballspieler, Running Back im Film Jerry Maguire, sagt: „Come on man, show me the money“. Das kann jedoch nicht mit „Zeig mir das Geld“ übersetzt werden. Im Film wurde dieser Satz dann in „Führ

mich zum Schotter“ übersetzt (Frommann, 2025). Es wurde also eine neue Redewendung erfunden, um sowohl dem Inhalt als auch dem Stil gerecht zu werden.

Es können also auch Redewendungen erfunden werden, weil es sich im Film ergibt, damit man der Figur folgen kann, obwohl es kein gebräuchlicher Slang oder idiomatische Redewendung ist (Frommann, 2025). Ob eine KI solche schwierigen Situationen meistern kann, ist fraglich.

Ein wesentlicher Vorteil der KI-Synchronisation liegt wiederum in der Kosten- und Zeiteffizienz. Automatisierte Prozesse ermöglichen eine schnellere Produktion bei geringeren Kosten, was besonders für Projekte mit begrenztem Budget attraktiv ist. Traditionelle Synchronisationen erfordern hingegen umfangreiche Ressourcen, einschließlich professioneller Sprecher und Tonstudios, was den Prozess teuer und zeitaufwendiger macht (Verbalate, 2024).

Außerdem bieten KI-Systeme eine hohe Flexibilität, da sie schnell auf Änderungen reagieren und verschiedene Sprachvarianten erzeugen können. Menschliche Sprecher benötigen für solche Anpassungen mehr Zeit und Ressourcen, was die Produktionsdauer verlängern kann (Speechify, o.D.).

Durch den Druck, welcher durch die immer weiter steigenden Produktionen, auf die Studios und Sprecher ausgeübt wird, leidet die Qualität der Synchronfassungen enorm. Während menschengemachte Synchronisationen durch den hohen Zeitdruck drohen, immer schlechter zu werden, entwickelt sich die KI rasant und wird immer besser (Von Medvey, 2025). Susanne von Medvey (2025) ist der Ansicht, die einzige Chance, die Sprecher haben, ist, dass sie Menschen sind und somit diese auch am besten verkörpern können. Sprecher müssen nicht mit der KI konkurrieren, sondern einfach menschlicher klingen und sein (Von Medvey, 2025).

Die Wahl zwischen KI-gestützter und menschlicher Synchronisation hängt von den spezifischen Anforderungen des Projekts ab. Während KI-Lösungen in Bezug auf Effizienz und Kosten Vorteile bieten, punkten menschliche Sprecher mit Authentizität und kultureller Sensibilität. Eine hybride Herangehensweise, die beide Methoden kombiniert, könnte in vielen Fällen die optimale Lösung sein.

## Nutzerakzeptanz und Wahrnehmung von KI-generierten Stimmen

Die Wahrnehmung von KI-generierten Stimmen und die damit einhergehende Akzeptanz dieser, sind entscheidende Faktoren für die Integration dieser Technologie in verschiedene Anwendungsbereiche, darunter auch die Synchronisationsbranche. Studien zeigen, dass Menschen zunehmend Schwierigkeiten haben, KI-generierte Inhalte von menschlich erzeugten zu unterscheiden. Eine Online-Befragung mit circa 3000 Teilnehmenden aus Deutschland, China und den USA ergab, dass KI-generierte Medien über alle Medienarten hinweg, übermäßig oft als menschengemacht eingeschätzt werden. KI-generierte Inhalte können also immer weniger als solche identifiziert werden (Forschung & Lehre, 2024). Auch Deepfake-Stimmen können bei zwei Drittel der Versuche überzeugen. Trotz der mangelnden Perfektion der Deepfake-Stimmen, sind diese in der Lage, die Wahrnehmung der Menschen zu täuschen. Jedoch gibt es durchaus Unterschiede bei den biologischen Vorgängen zur Unterscheidung KI-generierter und natürlicher Stimmen. Diese Unterschiede bleiben jedoch häufig unter der Wahrnehmungsschwelle (Roswandowitz, 2024).

Trotz der bisher sehr starken Entwicklung, bleibt die Akzeptanz von KI-generierten Inhalten durchwachsen. Eine Studie des Bayerischen Forschungsinstituts für Digitale Transformation (bidt) ergab, dass die Mehrheit der Internetnutzer (81%) bereits von generativer KI gehört haben, jedoch nur ein Drittel (36%) der Nutzer diese mindestens einmal selbst genutzt haben (Schlude et al., o.D.). Dies deutet darauf hin, dass trotz der technischen Fortschritte Vorbehalte gegenüber der Nutzung von KI bestehen.

Zudem spielt die Qualität synthetischer Stimmen eine wesentliche Rolle für die Nutzerakzeptanz dieser. Untersuchungen zeigen, dass synthetische Stimmen mit höherwertiger Qualität deutlich positiver wahrgenommen werden als solche mit geringerer Qualität. In einer Studie wurden vier KI-generierte Stimmen hinsichtlich ihres Geschlechts und ihrer Qualität bewertet. Die Ergebnisse zeigten, dass die beiden hochwertigeren Stimmen von den Teilnehmern in Bezug auf pragmatische, hedonische und Attraktivitätsmerkmale besser beurteilt wurden als die Stimmen geringerer Qualität (Cuciniello et al., 2022).

Ein weiterer Aspekt ist die Fähigkeit der KI, emotionale Nuancen zu vermitteln. Obwohl KI-generierte Stimmen zunehmend realistischer klingen, haben sie oft Schwierigkeiten subtile emotionale Elemente, Humor oder Ironie authentisch wiederzugeben. Dies kann dazu führen, dass sie als weniger natürlich oder überzeugend wahrgenommen werden. Besonders bei anspruchsvollen Formaten wie Spielfilmen, Serien oder Werbung sind diese emotionale Ausdruckskraft und die Authentizität essentiell.

Auch das jeweilige Einsatzgebiet, in welchem die KI-Stimme eingesetzt wird, ist entscheidend für die Akzeptanz der Nutzer. In technischen oder informativen Anwendungen, wie beispielsweise Navigationssysteme, Sprachassistenten oder E-Learning-Elemente, werden KI-generierte Stimmen oft als effizient und kostengünstig wahrgenommen und daher eher positiv aufgenommen. In solchen Fällen steht die Funktionalität im Vordergrund und die emotionale Tiefe der Stimme spielt hierbei eine weniger relevante Rolle.

Zusammenfassend lässt sich sagen, dass die Nutzerakzeptanz von KI-generierten Stimmen stark vom Einsatzgebiet abhängt. Während sie in technischen und informativen Kontexten gut angenommen werden, bleibt die menschliche Synchronisation in kreativen und emotionalen Bereichen vorerst unersetzlich. Durch die Weiterentwicklung der KI-Systeme wird sich zeigen, ob diese eine breitere Akzeptanz erreichen können. Dabei sind besonders die emotionale Ausdruckskraft und kulturelle Anpassung entscheidend.

### 5.3 Ethische und rechtliche Implikationen

Die Einführung künstlicher Intelligenz in die Filmsynchronisation und Tonpostproduktion wirft eine Reihe ethischer und rechtlicher Fragen auf, die sowohl die Branche als auch die Gesellschaft insgesamt betreffen.

#### Urheberrechtsfragen

Die rasante Entwicklung der Künstlichen Intelligenz ermöglicht es, Stimmen synthetisch zu erzeugen oder nachzuahmen, wodurch erhebliche urheberrechtliche

Fragen aufkommen. Gemäß §2 Absatz 2 des Urheberrechtsgesetzes (UrhG) sind Werke nur dann urheberrechtlich geschützt, wenn sie als „persönliche geistige Schöpfung“ gelten. Da KI-generierte Inhalte ohne menschliche Kreativität entstehen, fehlt ihnen dieser Schutzstatus. In diesem Zusammenhang stellt sich auch die Frage der Haftung. Wenn eine KI-Technologie ein Werk erschafft, welches gegen das Urheberrecht verstößt, wer haftet in diesem Fall bzw. wer trägt die Verantwortung? Der Nutzer der KI, der KI-Entwickler oder der Eigentümer der Technologie (Röttger, 2023).

Ein weiteres zentrales Problem besteht darin, dass KI-Systeme häufig mit urheberrechtlich geschütztem Material trainiert werden. Dies kann im Sinne des Urheberrechts als eine Vervielfältigung von geschütztem Material gesehen werden, welche nicht immer gestattet ist. Sind ursprüngliche Werke, welche für das Training genutzt wurden, im Endprodukt erkennbar, muss vor der Veröffentlichung der KI-Erzeugnisse die Erlaubnis des eigentlichen Rechteinhabers eingeholt werden. Eine Veröffentlichung oder Vervielfältigung ist dann nicht erlaubt, wenn diese Erlaubnis nicht vorliegt (Oswald, 2024).

Um diese Rechte zu erlangen, ist es in der Synchronbranche üblich, dass Synchronsprecher Verträge unterschreiben, in denen sie die Rechte ihrer Sprachaufnahmen für jegliche Zwecke abtreten. Darunter zählen auch zukünftige Technologien und Innovationen, wie es die Künstliche Intelligenz ist. Um sich vor einem Ersatz durch KI zu schützen, versuchen Sprecher aktuell einen sogenannten KI-Ausschluss-Passus in die Verträge einzubinden. Dieser soll verhindern, dass Sprachaufnahmen für den Gebrauch für KI-Tools genutzt werden (NDR, 2024). In der Gaming-Branche hat sich dieser bereits durch den hohen Einsatz der Sprecher weitläufig durchgesetzt. Dies ist aber nur durch den starken Zusammenhalt und Widerstand der Sprecher möglich. Jedoch ist der größte Teil der Synchronbranche noch immer ein unregelter Raum (Von Medvey, 2025). Auch sowohl Persönlichkeitsrechte als auch Datenschutzrechte spielen eine wesentliche Rolle bei der Nutzung von Stimmmaterial für KI-Tools. Bei vielen derzeit bestehenden Voice-Models werden Stimmdaten illegal gesammelt und für deren Lernmaterial genutzt. Deshalb ist es hier ebenfalls essentiell,

Nutzungserlaubnis einzuholen, um jegliche Rechte nicht zu verletzen (Loschelder, 2024).

Zusammenfassend sind die urheberrechtlichen Fragen bei KI-generierten Stimmen komplex und erfordern eine Anpassung der bestehenden Gesetze, um den Herausforderungen der digitalen Welt gerecht zu werden.

Sprecherverbände setzen sich dafür ein, dass Maßnahmen gegen den steigenden Einsatz von KI-Technologien getroffen werden. Jedoch sind die Fortschritte hierbei sehr schleppend und haben noch zu nichts Handfestem geführt. Dies liegt unter anderem an den starken Meinungsverschiedenheiten der Branchenmitglieder. Firmen Auftraggeber, Sprecher, Autoren und Regisseure sind sich über die aktuelle Lage und wie man damit umzugehen hat, nicht einig. Der Umgang mit der aktuellen Situation und der Entwicklung möglicher Regelungen kann somit nicht mit der rasanten Entwicklung der KI mithalten (Von Medvey, 2025).

### Ethische Fragen

Die zunehmende Integration von KI in die Synchronbranche eröffnet zwar neue Möglichkeiten, wirft jedoch auch erhebliche ethische Fragen auf.

Ein zentraler Punkt ist die potenzielle Gefährdung der Qualität von Synchronfassungen. Laut dem Verband Deutscher Sprecher (VDS) mangelt es KI-generierten Stimmen an Empathie, emotionaler Tiefe und kulturellem Kontext, was die Authentizität und das emotionale Erlebnis beeinträchtigen kann (VDS, 2024). Ein weiterer ethischer Aspekt betrifft die Gefährdung von Arbeitsplätzen innerhalb der Branche. Der VDS warnt davor, dass die Verwendung von KI für Synchronisation traditionelle Arbeitsplätze in der deutschen Synchronbranche bedrohen könnte, besonders wenn KI-generierte Fassungen auf Basis bestehender menschlicher Synchronisationen erzeugt werden (VDS, 2024).

Durch die Angst von KI-Technologien ersetzt zu werden, lehnen einige Sprecher sogar große Projekte ab, bei denen keine KI-Klausel im Vertrag steht. Gleichzeitig versuchen Sprecher, so viele Projekte anzunehmen, wie es ihnen möglich ist, solange ihr Job noch so besteht wie bisher (Von Medvey, 2025). Susanne von

Medvey (Sprecherin mit über 30 Jahren Synchronerfahrung) beschreibt die Situation mit einer Endzeit-Stimmung.

Zusätzlich entstehen durch die Nutzung von KI in der Synchronisation Fragen der Authentizität und Integrität von Medieninhalten. Die Möglichkeit Stimmen realistisch zu imitieren, könnte zu einer Verbreitung von Deepfakes führen, die das Vertrauen der Zuschauer in audiovisuelle Medien schädigen (Bolder, 2024).

Der Einsatz von KI-generierten Stimmen erfordert also eine sorgfältige Abwägung ethischer Aspekte. Es ist essentiell, Richtlinien zu entwickeln, die sowohl die Qualität und Authentizität von Synchronfassungen als auch die Rechte und Arbeitsplätze der beteiligten Fachkräfte schützen.

## Konsequenzen

Durch die bereits erläuterten rechtlichen und ethischen Bedenken, steht die Synchronbranche im Zwiespalt. Um die Rechte der Sprecher und Schauspieler nicht zu verletzen und das Ansehen deutscher Synchronfassungen zu schädigen, können KI-gestützte Systeme bisher nur begrenzt eingesetzt werden. Zusätzlich erschwert die Gefährdung des Arbeitsmarktes den Einsatz KI-gestützter Synchronisationstools massiv, da sich Synchronsprecher stark für den Erhalt ihres Arbeitsplatzes einsetzen. Auch die bisher noch mangelnde Qualität der KI-generierten Stimmen, führt zu einem zögerlichen Einsatz von KI-Tools. Dies wird sich jedoch voraussichtlich in zukünftigen Entwicklungen deutlich verbessern und zu beeindruckenden Ergebnissen führen.

Wenn die rechtlichen und ethischen Herausforderungen nicht angemessen berücksichtigt werden, steht die Synchronbranche zwischen einer traditionellen und einer KI-gestützten Produktionsmethode.

Hybride Modelle könnten eine Lösung sein, bei der die KI als unterstützendes Werkzeug dient. Dabei werden menschliche Sprecher weiterhin eingesetzt.

Beispielsweise besetzt ein Studio in Köln nur noch die Hauptrollen mit menschlichen Sprechern, während weniger relevante Rollen von einer KI übernommen werden (Von Medvey, 2025).

Dabei sollten jedoch faire Regelungen zur Nutzung vorhandener Werke und eine angemessene Vergütung vereinbart werden (VDS, 2024).

Zusammenfassend zeigt sich, dass die Konsequenzen der rechtlichen und ethischen Bedenken tief in die Struktur der Synchronbranche eingreifen könnten. Eine nachhaltige und ausgewogene Nutzung von KI-Technologien erfordert daher klare gesetzliche Rahmenbedingungen sowie eine ethisch verantwortungsvolle Umsetzung.

## 6. Zukunftsperspektiven der KI in der Synchronisation

Die Zukunft der KI in der Synchronisation verspricht drastische Veränderungen in der Film- und Medienbranche. Während aktuelle Systeme noch mit Herausforderungen wie emotionaler Ausdruckskraft und kultureller Sensibilität kämpfen, zeigen technologische Fortschritte, dass KI-basierte Stimmen zunehmend natürlicher und flexibler werden. Dies könnte langfristig Produktionsprozesse optimieren, Kosten senken und die Synchronisation in mehrere Sprachen einfacher machen. Dennoch bleibt die Frage offen, inwieweit KI menschliche Sprecher und deren Kreativität vollständig ersetzen kann oder nur als unterstützendes Werkzeug fungiert. Die zukünftige Entwicklung wird maßgeblich davon abhängen, wie rechtliche, ethische und qualitative Aspekte geregelt und weiterentwickelt werden.

### 6.1 Technologische Trends und mögliche Weiterentwicklungen

Die Integration von Künstlicher Intelligenz in die Filmbranche und deren stetige Weiterentwicklung und Verbesserung, führt zu bemerkenswerten Trends und neuen Möglichkeiten.

Technologien wie die KI-gestützte Lippensynchronisation, Übersetzung oder automatisierte Vertonung werden stetig weiterentwickelt und verbessert. Dadurch werden die bisher noch mit mangelnder Qualität behafteten Ergebnisse KI-gestützter Synchronfassungen voraussichtlich in Zukunft immer besser. Firmen wie die Synchronfirma TNT Media kündigten an, zukünftig KI-generierte Synchronfassungen für Märkte bereitzustellen, die bisher hauptsächlich Untertitel nutzten. Dies soll jedoch nicht auf Kosten bestehender Arbeitsplätze geschehen, sondern lediglich als Ergänzung bestehender Projekte dienen. Dabei sollen

Synchronsprecher projektbezogen das Nutzungsrecht deren Stimme für eine faire Vergütung übertragen (TNT Media, o.D.).

Auch große Plattformen wie YouTube haben begonnen, KI-basierte Synchronisationsfunktionen zu integrieren. Ende 2024 führte YouTube für bestimmte Nutzer eine Funktion ein, die die automatisierte Synchronisation ihrer Videos ermöglicht. Trotz der Kritik dieses Tools hinsichtlich der Qualität, zeigt dies das Potential und die Richtung, in die sich die Technologie entwickelt (Böhm, 2024).

Zudem demonstrieren Projekte wie die geplante Verfilmung des Lebens von Roberto Blanco die fortschrittlichen Möglichkeiten der KI in der modernen Filmproduktion (Pittelkau 2024).

Die kontinuierliche Verbesserung von KI-Technologien deutet darauf hin, dass zukünftige Systeme in der Lage sein werden, Emotionen und Nuancen menschlicher Sprache besser zu erfassen und wiederzugeben. Weltweit arbeiten Wissenschaftler und Technologieunternehmen daran, KI-Systeme besser auf das Erfassen und die Wiedergabe von Emotionen und menschlicher Ausdruckskraft zu trainieren (Bart, 2024). Dies könnte zu einer breiteren Akzeptanz und Anwendung von KI in der Filmsynchronisation führen. Jedoch bleiben die Herausforderungen in Bezug auf ethische Fragen und die Sicherstellung der Qualität bestehen. Außerdem stellt sich weiterhin die Frage, inwieweit KI-synthetisierte Stimmen mit realen menschlichen Stimmen und deren Kreativität und Flexibilität mithalten können. Sind KI-Systeme in der Lage, jahrzehntelange Übung und Erfahrung eines Synchronsprechers nachzubilden? Oder doch nicht? Diese Fragen lassen sich zum jetzigen Zeitpunkt nur mutmaßlich beantworten und sind davon abhängig, wie sich die KI-Tools zukünftig weiterentwickeln.

## 6.2 Chancen und Risiken für die Industrie

KI in der Synchronbranche: Fluch oder Segen? Wie bereits näher erläutert bietet der Einsatz KI-basierter Tools bei der Synchronisation fremdsprachiger Produktionen erhebliche Potentiale und eröffnet neue kreative Möglichkeiten. Jedoch birgt dies auch enorme Risiken für die Branche.

### Chancen

Durch den Einsatz von KI können Produktionsprozesse effizienter gestaltet und Kosten reduziert werden. Laut einem Bericht des Global Voiceover and Dubbing Market Report betragen die Kosten für klassische Synchronisationsdienste im Durchschnitt etwa 150 US-Dollar pro Videominute. Für ein einstündiges Video kann somit mit ca. 9000 US-Dollar pro Sprache gerechnet werden. Dabei können die Kosten je nach Komplexität und Verfügbarkeit der Sprecher variieren. KI-Synchrondienste kosten im Durchschnitt etwa fünf bis zehn US-Dollar pro Videominute. Somit liegen die Gesamtkosten einer KI-Synchronisation pro Sprache bei 300 bis 600 US-Dollar (Storyshell, o.D.). Diese Kosten können natürlich je nach Projekt und dessen Anforderungen stark variieren. Ob die Nutzung einer KI-Technologie für ein Projekt in Frage kommt, hängt also sehr stark von dessen Anforderungen ab.

Außerdem ermöglichen KI-gestützte Systeme eine schnellere Bearbeitung großer Datenmengen. Auch die kreative Ebene profitiert durch den Einsatz von Künstlicher Intelligenz. Tote Stimmen können wiederbelebt und Charaktere verjüngt werden. Somit werden neue kreative Wege ermöglicht, welche zuvor nicht möglich waren. Zudem können störende Fehler durch Unstimmigkeiten der Lippsynchronität mit dem Einsatz KI-gestützter Lippsynchronisation verhindert werden.

Dank der Kosteneinsparungen und der hohen Effizienz können Produktionen auch deutlich schneller abgeschlossen werden, was die Umsetzung weiterer Projekte ermöglicht.

Dadurch können auch kleinere Firmen, welche ein geringeres Budget besitzen, Synchronisationen ihrer Projekte umsetzen.

Nicht nur die Synchronisationsindustrie wird dadurch revolutioniert, sondern auch andere Bereiche profitieren von neuen Möglichkeiten. Beispielsweise durch personalisierte digitale Assistenten, interaktive Spiele oder integrative Technologien für Menschen mit Sprachbehinderungen (Bart, 2024).

Marktanalysen prognostizieren ein starkes Wachstum für KI-Synchronisationstools. Laut dem Marktforschungsunternehmen Valuates Reports wird erwartet, dass dieser Markt von 783 Millionen US-Dollar im Jahr 2023 auf 1,88 Milliarden US-Dollar bis 2030 ansteigen wird. Dies entspricht einer durchschnittlichen jährlichen Wachstumsrate von 14,2% (Speech Technology, 2024).

## Risiken

Gleichzeitig birgt der Einsatz von KI Risiken, insbesondere in Bezug auf Arbeitsplatzverluste und ethische Fragen. Die Automatisierung von Synchronisationsprozessen könnte traditionelle Arbeitsplätze gefährden, wodurch Unsicherheiten in der Branche entstehen. Durch die stetige Verbesserung der KI-Tools, könnten diese eines Tages so gut werden, dass sie die Rolle der Synchronsprecher komplett ersetzen könnten. Auch andere Berufe der Synchronbranche stehen vor grundlegenden Veränderungen. KI-Technologien werden in Synchronisationsprozesse integriert, wodurch sich Abläufe ändern und die Fähigkeiten der Fachkräfte gegebenenfalls ergänzt werden müssen. Zudem bleiben die Fragen des Urheberrechts und der Authentizität, insbesondere wenn KI-generierte Stimmen ohne vorherige Zustimmung der ursprünglichen Sprecher verwendet werden. Im Allgemeinen gibt es noch einige offene Fragen in Bezug auf rechtliche Grundlagen, welche vor der Implementierung KI-gestützter Prozesse geklärt werden müssen.

Zudem stellt sich die Frage der Zuschauerakzeptanz von KI-basierter Synchronisation. Ein gutes Beispiel ist hierfür der im Dezember 2024 erschienene Film „Black Dog – Weggefährten“. Als erster vollständig KI-synchronisierter Film feierte der Film in den Kinos seine Premiere. Dieses Ereignis löste Diskussionen über die Qualität und Authentizität solcher Produktionen aus. Dabei waren die Reaktionen auf die KI-generierte Synchronfassung sehr durchwachsen. Ein Teil des

Publikums war durchaus beeindruckt, während der andere Teil starke Kritik übte (Schultze, 2024).

Zusammenfassend bietet die Integration von KI in die Synchronisation fremdsprachiger Produktionen erhebliche Potentiale für Effizienzsteigerungen und kreative Innovationen. Gleichzeitig müssen jedoch die damit verbundenen Risiken, insbesondere in Bezug auf Arbeitsplätze und ethische Standards, sorgfältig abgewogen werden, um eine ausgewogene und verantwortungsvolle Nutzung dieser Technologie zu gewährleisten.

### 6.3 Mögliche Veränderungen der Konsumentenwahrnehmung

Durch die Nutzung KI-gestützter Tools zur Synchronisation fremdsprachiger Produktionen, ist die Änderung der Konsumentenwahrnehmung keineswegs zu vernachlässigen. Besonders die Akzeptanz KI-basierter Synchronisation und die Erwartungshaltungen, welche Konsumenten über jahrelange Erfahrungen aufgebaut haben, spielen eine wesentliche Rolle.

#### Akzeptanz

Ein zentraler Faktor ist die Akzeptanz von KI-generierten Stimmen. Während Zuschauer in funktionalen Bereichen wie Sprachassistenten oder Navigationssystemen bereits an synthetische Stimmen gewöhnt sind, bestehen in der Synchronbranche höhere Anforderungen an Natürlichkeit und emotionale Ausdrucksstärke. Mehrere Studien, darunter eine Studie vom Bayerischen Forschungsinstituts ergaben, dass es durchaus Bedenken von Seiten der Zuschauer bezüglich KI-Technologien gibt. Hierbei kommt es maßgeblich auf den Anwendungsbereich an (Schlude et al., o.D.). Während informative Bereiche, bei denen qualitative Nuancen nicht so relevant sind, eine höhere Akzeptanz aufweisen, begegnen Bereiche wie die Synchronisation von Filmen oder Serien eher einer hohen Kritik.

Ein weiterer Faktor, welcher die Akzeptanz KI-gestützter Synchronisation stark beeinflusst, ist der berufliche Status der Konsumenten. Ein Großteil der Befragten

einer Studie des bidt (Bayerisches Forschungsinstitut für Digitale Transformation) fürchten um ihren Job durch den vermehrten Einsatz von künstlicher Intelligenz. Dabei wünschen sich 52% eine stärkere Regulierung der generativen KI, wobei sich rund ein Viertel der Befragten einen kompletten Stopp der Weiterentwicklung wünscht (Schlude et al., o.D.). Durch die Angst, den eigenen Job zu verlieren, könnte die Akzeptanz der KI-gestützten Synchronisation stark beeinflusst werden. Der wohl bedeutendste Aspekt zur Akzeptanz KI-generierter Synchronisationen ist jedoch die Qualität und Authentizität der synthetisierten Stimmen, welche die Emotionen und Stimmungen der Charaktere widerspiegeln. Wie bereits erwähnt, gab es hierfür schon ein gutes Beispiel für die bislang mangelnde Akzeptanz KI-gestützter Synchronisation mit der polnischen Serie „Murderesses“, welche durch ihre schwache Qualität und Gefühlslosigkeit heftige Kritik der Konsumenten erntete (Lückerath, 2025).

Jedoch gab es vor einigen Jahren auch einen vergleichbaren Fall einer menschlich synchronisierten Produktion, welche aufgrund ihrer mangelnden Qualität abgesetzt wurde. Bestehende und neue Synchronisationen beeinflussen mit ihrer Qualität die Hörgewohnheiten der Zuschauer, ob nun positiv oder negativ. Konsumenten könnten sich an mittel bis schlechte Synchronisationen gewöhnen (Von Medvey, 2025).

Cornelius Frommann sagt:

*„Die Frage ist nicht, wie viel die KI falsch macht, sondern ob es akzeptiert wird, ob es reicht für die Zuschauer.“*

Wenn sich die Konsumenten an bestimmte Sprachmuster gewöhnen (besonders junge und moderne), da sie selbst durch den Einfluss vom Internet oder anderen Medien in diese Muster fallen und somit ein neues Sprachniveau schaffen, könnte die KI in der Lage sein, die Zuschauer zufriedenzustellen. Die Zuschauer empfinden die Synchronfassung dann menschlich genug, da sie denken, dass Menschen wirklich so reden könnten (Frommann, 2025). Dies ist zumindest bei KI-übersetzten Inhalten denkbar.

## Erwartungshaltungen

Zuschauer haben klare Erwartungen an die Qualität von Synchronisationen. Jahrzehntlang haben sie sich an hochqualitative menschliche Synchronisationen gewöhnt, die kulturelle Feinheiten und emotionale Elemente vermitteln. Synchronregisseur Clemens Frohmann beleuchtet in einem Interview die Herausforderungen bei der Erstellung natürlicher und glaubwürdiger Synchronfassungen. Frohmann betont, dass trotz technischer Fortschritte die menschliche Nuance und das emotionale Feingefühl schwer durch KI zu ersetzen sind (Bernklau, 2024). Besonders KI-gestützte Übersetzungstools werden Probleme mit schwierig übersetzbaren Inhalten wie Dialekte oder kulturelle Nuancen haben. Oft können Dialekte oder sprachliche Feinheiten nicht direkt übersetzt werden. In diesem Fall müssen Kompromisse gefunden werden, welche nah an die gegebene Situation kommen, ohne inhaltlich zu stark abzuweichen. Ob eine KI dazu in der Lage wäre, ist fraglich.

Die Erwartungshaltungen und Akzeptanz KI-gestützter Synchronisation stehen in engem Zusammenhang. Wenn die Erwartungshaltungen der Zuschauer in Bezug auf die Authentizität und Qualität zufriedenstellend erfüllt werden können, ist die Wahrscheinlichkeit deutlich höher, dass KI-Synchronisationen vom Publikum akzeptiert werden. Jedoch gibt es ein breites Publikum, welches der Nutzung von KI kritisch gegenübersteht und eine menschengemachte Synchronisation bevorzugt.

## 7. Fazit

### 7.1 Zusammenfassung der Ergebnisse

Die Arbeit hat gezeigt, dass Künstliche Intelligenz die Standards und Prozesse der Synchronbranche maßgeblich verändert. Während klassische Synchronisationstechniken auf menschliche Sprecher, aufwendige Tonmischung und manuelle Nachbearbeitung angewiesen sind, ermöglichen KI-gestützte Herangehensweisen eine weitgehende Automatisierung. Fortschritte in der Sprachnachbildung, der automatischen Lippensynchronisation und der klanglichen

Anpassung von Stimmen sind in der Lage, die Effizienz in der Tonpostproduktion erheblich zu steigern. Besonders bemerkenswert ist die Fähigkeit moderner KI-Modelle, Stimmen realistisch zu imitieren (Voice-Cloning) und sogar verstorbene Schauspieler stimmlich zu rekonstruieren.

Gleichzeitig stellen sich ethische und rechtliche Fragen, insbesondere im Hinblick auf Persönlichkeitsrechte, Urheberrechte und die künstlerische Integrität. Die Branche steht vor einer immensen Veränderung, welche die Arbeitsabläufe und die traditionelle Weise der Synchronisation erheblich beeinflussen. Sprecher sind um die Zukunft ihrer Karriere besorgt und fürchten von KI-Systemen ersetzt zu werden. Jedoch sind zur jetzigen Zeit noch viele Fragen in Bezug auf rechtliche, künstlerische und ethische Aspekte ungeklärt.

## 7.2 Beantwortung der Forschungsfrage

Die Forschungsfrage „Wie verändert Künstliche Intelligenz die Standards und Prozesse der Synchronisation, und welche Entwicklungen sind weiter zu erwarten?“ wurde in der Arbeit umfassend untersucht. KI verändert die Standards der Synchronisation insbesondere durch Automatisierung und Effizienzsteigerung. Manuelle Prozesse können zunehmend durch KI-gestützte Algorithmen ersetzt werden, die in der Lage sind Stimmen zu synthetisieren und diese an die Lippenbewegungen des Schauspielers anzupassen. Zudem ermöglicht KI eine präzisere und kostengünstigere Lokalisierung von Inhalten für internationale Märkte. Auch die kreative Ebene profitiert von dem Einsatz Künstlicher Intelligenz. So können tote Stimmen wiederbelebt werden, wenn Schauspieler nicht mehr zur Verfügung stehen.

Allerdings sind aktuelle KI-Modelle noch nicht in der Lage, die volle emotionale Tiefe menschlicher Sprecher zu reproduzieren. Dadurch verliert die Synchronisation an Authentizität und Ausdrucksstärke. Daher kommen derzeit häufig hybride Lösungen zum Einsatz, bei denen KI-generierte Stimmen mit menschlicher Nachbearbeitung kombiniert werden. Langfristig könnte KI jedoch traditionelle

Synchronmethoden weitgehend verdrängen, besonders wenn rechtliche und ethische Rahmenbedingungen klarer definiert werden.

### 7.3 Ausblick auf zukünftige Entwicklungen

Die zukünftige Entwicklung der KI in der Tonpostproduktion wird sich voraussichtlich auf einige zentrale Bereiche konzentrieren. Ein besonders wichtiger Aspekt ist die Verbesserung der Emotionalität und Authentizität künstlich generierter Stimmen. Während KI-Modelle technisch immer ausgereifter werden, bleibt die natürliche Ausdruckskraft menschlicher Sprecher bisher unerreicht. Künftige Forschungen könnten sich daher verstärkt darauf konzentrieren, Modelle zu entwickeln, die Emotionen präziser nachbilden und eine noch realistischere Sprachsynthese ermöglichen.

Darüber hinaus erfordert der zunehmende Einsatz von KI in der Synchronisation eine klare Regulierung, um den Schutz der Persönlichkeitsrechte zu gewährleisten und Missbrauch zu verhindern. Dies gilt insbesondere im Zusammenhang mit Deepfake-Technologien, die Stimmen täuschend echt imitieren können und somit ethische sowie rechtliche Fragen aufwerfen. Ohne entsprechende Richtlinien wird es sehr wahrscheinlich zu unerwünschten Manipulationen oder einer unautorisierten Verwendung von Stimmen kommen. Das würde sowohl für die Künstler als auch für die gesamte Branche weitreichende Konsequenzen nach sich ziehen.

Ein weiteres zentrales Forschungsfeld ist die Optimierung der Produktionsabläufe. Besonders im Hinblick auf die effizientere Zusammenarbeit zwischen menschlichen Synchronsprechern und KI-gestützten Systemen, bietet die Integration von KI in bestehende Synchronisationsprozesse große Potenziale. Hier könnte die Weiterentwicklung hybrider Ansätze dazu beitragen, die Effizienz zu steigern, ohne dabei qualitative Einbußen in Kauf nehmen zu müssen. Die Zukunft KI-integrierter Prozesse ist sehr umstritten und ruft verschiedenste Meinungen auf. Cornelius Frommann (2025) ist beispielsweise der Meinung, dass die KI immer Fehler machen wird. Sprecher wie Susanne von Medvey (2025) haben Angst um die Zukunft ihrer Jobs und sind wegen der mangelnden Regelungen aufgebracht.

## Quellenverzeichnis

- Amberscript (2023). *Vor- und Nachteile der Synchronisation*. Aufgerufen von <https://www.amberscript.com/de/blog/vorteile-und-nachteile-synchronisation/> am [10.01.2025].
- Ames, S. (2023). *Lippen-Synchronisations-Apps: Die ultimative Lippensynchronisationslösung für KI-Videoübersetzung*. Rask. Aufgerufen von <https://de.rask.ai/blog/lip-sync-apps#was-ist-ein-ai-video-editor> am [02.02.2025].
- Arndt, O., Creutz, B., Dornberg, F., Schenten, C. (2024). „Künstliche Intelligenz in der Kultur und Kreativwirtschaft“. Kompetenzzentrum Kultur- und Kreativwirtschaft des Bundes. Aufgerufen von [https://kreativ-bund.de/wp-content/uploads/2024/04/240415\\_Themendossier\\_KI\\_final\\_1.pdf](https://kreativ-bund.de/wp-content/uploads/2024/04/240415_Themendossier_KI_final_1.pdf) am [03.01.2025].
- Bart, S. (2024). *KI-Stimmen: Bedrohung oder Chance?* Mimi Kama. Aufgerufen von <https://www.mimikama.org/ki-stimmen-bedrohung-oder-chance/> am [12.03.2025].
- Bastian, M. (2021). *Deepfake-Lippensynchronisation: Tech-Revolution für Filme und Serien?* The Decoder. Aufgerufen von <https://the-decoder.de/deepfake-lippensynchronisation-tech-revolution-fuer-filme-und-serien/> am [31.01.2025].
- Bernklau, J. (2024). *Wieso klingen Synchronstimmen oft so unnatürlich?* Über Medien. Aufgerufen von <https://uebermedien.de/101122/wieso-klingen-synchronstimmen-oft-so-unnatuerlich/> am [15.03.2025].
- Bertelsmann SE & Co. KGaA. (2024). *RTL Technology nutzt erstmals KI für Synchronisation*. Aufgerufen von <https://www.bertelsmann.de/news-und-media/nachrichten/rtl-technology-nutzt-erstmal-k-i-fuer-synchronisation.jsp> am [22.01.2025].
- Bieß, C., Pawelec, M. (2020). *Schaden Deepfakes (wirklich) der Demokratie?* Universität Tübingen. Aufgerufen von <https://uni-tuebingen.de/forschung/zentren-und-institute/internationales-zentrum-fuer->

- [ethik-in-den-wissenschaften/publikationen/bedenkzeiten-bibliothek/weitere-blog-artikel/schaden-deepfakes-wirklich-der-demokratie/](https://www.ethik-in-den-wissenschaften.de/publikationen/bedenkzeiten-bibliothek/weitere-blog-artikel/schaden-deepfakes-wirklich-der-demokratie/) am [20.02.2025].
- BILD. (2024, 15. September). *Roberto Blanco durch KI verjüngt*. Aufgerufen von <https://www.bild.de/unterhaltung/stars-und-leute/roberto-blanco-durch-ki-maschine-verjuengt-sein-leben-wird-verfilmt-66e4317acb180e1cf920765a> am [19.01.2025].
- Bippes, T. (o.D.). *Wie KI die Medienbranche verändert*. Fernhochschule The Mobile University. Aufgerufen von <https://www.mobile-university.de/kommunikation/kuenstliche-intelligenz-medienbranche/> am [17.01.2025].
- Böhm, A. (2024, Dezember 01). *YouTubes neuestes Feature ist echt SCHLECHT...* [Video]. YouTube. Aufgerufen von <https://www.youtube.com/watch?v=a5rXnUy0KpM&t=605s> am [12.03.2025].
- Bolder, N. (2024). *Sprach-KI: Ist es ethisch vertretbar, Stimmen zu klonen?* Digital Pioneers. Aufgerufen von <https://t3n.de/news/sprach-ki-stimme-klonen-ethisch-vertretbar-1653481/> am [08.03.2025].
- Burgemeister, S. (2023). *KI Berufe: Zukunft oder Hype?* Internationale Hochschule Akademie. Aufgerufen von <https://www.iu-akademie.de/blog/ki-berufe-zukunft/> am [22.02.2025].
- BVR Bundesverband Regie (o.D.). *Synchronregie und Synchronautor/in*. Aufgerufen von [https://www.regieverband.de/berufsbild/Synchronregie\\_Synchronautor\\_in#:~:text=Synchronregisseure%20oder%20Synchronregisseurinnen%20gestalten%20mit,Regisseur%3Ain%20in%20Personalunion%20aus](https://www.regieverband.de/berufsbild/Synchronregie_Synchronautor_in#:~:text=Synchronregisseure%20oder%20Synchronregisseurinnen%20gestalten%20mit,Regisseur%3Ain%20in%20Personalunion%20aus) am [10.01.2025].
- Chiappani, P. (2024). *Künstliche Intelligenz in der Filmindustrie: Chancen und Herausforderungen*. KI-Company. Aufgerufen von <https://www.ki-company.ai/blog-beitraege/kuenstliche-intelligenz-in-der-filmindustrie-chancen-und-herausforderungen#:~:text=Postproduktion%3A%20Bei%20oder%20Bearbeitung>

[%20von,sogar%20ganze%20Szenen%20zu%20generieren.](#) am  
[19.01.2025].

Cuciniello, M., Amorese, T., Cordasco, G., Marrone, S., Marulli, F., Cavallo, F., Gordeeva, O., Callejas Carrion, Z., Esposito, A. (2022). *Identifying synthetic voices qualities for conversational agents*. ArXiv.

<https://doi.org/10.48550/arXiv.2205.04149>

DeepDub. (2022). *Revolutionizing film dubbing with AI-powered technology*. Abgerufen von [deepdub.ai](https://deepdub.ai) am [22.01.2025].

Degenkolb, J. (o.D.). *KI in der Übersetzung Wie können Buchübersetzungen anhand KI automatisiert/vereinfacht werden?* Verlage der Zukunft.

Aufgerufen von <https://www.verlagederzukunft.de/ki-in-der-uebersetzung/>  
am [26.02.2025].

Díaz Cintas, J., & Remael, A. (2021). *Audiovisual Translation: Subtitling*.

Routledge. <https://doi.org/10.4324/9781315759678>

Europäische Kommission (2024). *KI-Gesetz tritt in Kraft*. Europäische Kommission. Aufgerufen von

[https://germany.representation.ec.europa.eu/news/ki-gesetz-tritt-kraft-2024-08-01\\_de](https://germany.representation.ec.europa.eu/news/ki-gesetz-tritt-kraft-2024-08-01_de) am [31.01.2025].

Flawless AI. (2023). *AI and the future of film localization*. Abgerufen von [flawlessai.com](https://flawlessai.com) am [22.01.2025].

Forschung & Lehre (2024). *Menschen erkennen KI-generierte Medien kaum*.

Forschung & Lehre. Aufgerufen von <https://www.forschung-und-lehre.de/zeitfragen/menschen-erkennen-ki-generierte-medien-kaum-6430>  
am [06.03.2025].

Frommann, C. (2025). *Stellungnahme am 26.03.2025*. Das Gespräch befindet sich im Anhang dieser Arbeit.

Futuretrans (2024). *KI-Videosynchronisierung: Sprachzugänglichkeit im Film neu definiert*. Futuretrans. Aufgerufen von <https://future-trans.com/de/ai-video-dubbing-redefining-language-accessibility/> am [04.02.2025].

- Glänzel, P. (o.D.). *Der Aufstieg generativer KI in Media und Entertainment*. Qvest. Aufgerufen von <https://www.qvest.com/de/spotlight/der-aufstieg-generativer-ki-in-media-und-entertainment> am [19.01.2025].
- Goethe-Institut. (2023). *Wie Deutschland zur Synchronnation wurde*. Abgerufen von <https://www.goethe.de> am [03.01.2025].
- Habimex (2024). *Texte von Menschen vs. KI: Stärken und Schwächen im Vergleich*. Habimex. Aufgerufen von <https://www.habimex.de/texte-von-menschen-vs-ki-staerken-und-schwaechen-im-vergleich/> am [27.02.2025].
- IT-P GmbH (2024). *Mit KI Sprachbarrieren überwinden: Mehr Effizienz bei der Übersetzung*. IT-P. Aufgerufen von <https://www.it-p.de/blog/ki-sprachuebersetzung/> am [22.02.2025].
- KI-Praxisbeispiele. (o.D.). *Wie KI die Film- und Medienindustrie revolutioniert*. Aufgerufen von <https://www.ki-praxisbeispiele.de/wie-ki-die-film-und-medienindustrie-revolutioniert/> am [19.01.2025].
- Klingbeil-Döring, W. (2023). *“Die Auswirkungen von Künstlicher Intelligenz auf den Arbeitsmarkt“*. Bundeszentrale für politische Bildung. Aufgerufen von <https://www.bpb.de/themen/arbeit/arbeitsmarktpolitik/522513/die-auswirkungen-von-kuenstlicher-intelligenz-auf-den-arbeitsmarkt/> am [03.01.2025].
- LeCun, Y., Bengio, Y., Hinton, G. (2015). *Deep learning*. *Nature*, 521(7553), 436–444. <https://doi.org/10.1038/nature14539>
- Lückerath, T. (2025). *Wegen KI-Synchro: Magenta TV nimmt Viaplay-Serie offline*. DWDL. Aufgerufen von [https://www.dwdl.de/nachrichten/101325/wegen\\_kisynchro\\_magenta\\_tv\\_nimmt\\_viaplayserie\\_offline/](https://www.dwdl.de/nachrichten/101325/wegen_kisynchro_magenta_tv_nimmt_viaplayserie_offline/) am [04.02.2025].
- Loschelder, D. (2024). *Recht an eigener Stimme – Stimmnachahmung durch KI*. Loschelder Leisenberg Rechtsanwälte. Aufgerufen von <https://ll-ip.com/aktuelles/recht-an-eigener-stimme-stimmnachahmung-durch-ki/> am [08.03.2025].

- MacDonald, K. (2024). „Es wird immer verwirrender“: Kann KI Übersetzer ersetzen? The Guardian. Aufgerufen von <https://www.theguardian.com/books/2024/nov/11/it-gets-more-and-more-confused-can-ai-replace-translators> am [22.02.2025].
- Marketingscout. (2024, 25. September). *Künstliche Intelligenz in der Medien-Branche – Die Zukunft hat erst begonnen*. Aufgerufen von <https://www.marketingscout.com/trends/kuenstliche-intelligenz-in-der-medien-branche-die-zukunft-hat-erst-begonnen/> am [19.01.2025].
- Morrison, B. (2020). „Wie ist der Ablauf bei einer Synchronisation? Wie werden Filme synchronisiert?“. Pionierfilm. Aufgerufen von <https://filmproduktion-werbefilm.de/wie-ist-der-ablauf-bei-einer-synchronisation/> am [15.01.2025].
- Motion Picture Association. (2020). Theatrical and Home Entertainment Market Environment Report. Aufgerufen von <https://www.motionpictures.org> am [03.01.2025].
- m-Voices CAMPUS (2014). *Professionelle Sprecherausbildung und Vocal Coaching*. Abgerufen von [https://www.casting-network.de/workshops/m-voices\\_CAMPUS\\_Workshop\\_Flyer.pdf](https://www.casting-network.de/workshops/m-voices_CAMPUS_Workshop_Flyer.pdf) am [10.01.2025].
- Naumann, G. (2015). *Filmsynchronisation in Deutschland bis 1955*. Peter Lang Verlag. 10.3726/978-3-653-04772-1
- NDR (2024, Februar 05). *Zukunft der Kino-Branche – Werden Synchronsprecher durch KI ersetzt?* [Videobeitrag]. NDR. Aufgerufen von <https://www.ndr.de/fernsehen/sendungen/kulturjournal/Zukunft-der-Kino-Branche-Werden-Synchronsprecher-durch-KI-ersetzt,kulturjournal9572.html> am [08.03.2025].
- Okeke, N. (2024). *AI Voice Cloning: Wie es funktioniert und wichtige Details*. TargetTrend. Aufgerufen von <https://targettrend.com/de/ai-voice-cloning/#how-ai-voice-cloning-works> am [25.01.2025].
- Oswald, F. (2024). *Was sagt das Urheberrecht zu KI?* Basecamp. Aufgerufen von <https://www.basecamp.digital/ki-verstehen-was-sagt-das-urheberrecht-zu-ki/> am [08.03.2025].

- Otte, R. (2021). *Allgemeinbildung Künstliche Intelligenz Risiko und Chance*. Wiley-VCH GmbH. Weinheim. ISBN: 978-3-527-82701-5.
- Pedersen, J. (2017). *Subtitling Norms for Television*. John Benjamins Publishing.  
<https://doi.org/10.1075/btl.98>
- ReSartus (o.D.). *Die Rolle der KI in der Übersetzungsindustrie*. ReSartus.  
 Aufgerufen von <https://resartus.de/die-rolle-der-ki-in-der-ubersetzungsindustrie/> am [26.02.2025].
- Resemble AI (o.D.). *Künstliche Intelligenz (KI) zum Klonen von Stimmen, die menschlich klingen*. Resemble AI. Aufgerufen von  
<https://www.resemble.ai/voice-cloning/> am [31.01.2025].
- Rezanov, A. (2024). *Der Aufstieg der KI-Lippensynchronisation: Vom Uncanny Valley zum Hyperrealismus*. Dataconomy. Aufgerufen von  
<https://de.dataconomy.com/2024/11/05/der-aufstieg-der-ki-lippensynchronisation-vom-uncanny-valley-zum-hyperrealismus/> am [02.02.2025].
- Röttger, T. (2023). *Künstliche Intelligenz und Urheberrecht*. Gulden Röttger.  
 Aufgerufen von [https://ggr-law.com/urheberrecht/faq/kuenstliche-intelligenz-ki-urheberrecht/?utm\\_source=chatgpt.com#c19160](https://ggr-law.com/urheberrecht/faq/kuenstliche-intelligenz-ki-urheberrecht/?utm_source=chatgpt.com#c19160) am [06.03.2025].
- Roswandowitz, C. (2024). *Menschen erkennen KI-generierte Medien kaum*. Forschung & Lehre. Aufgerufen von <https://www.forschung-und-lehre.de/zeitfragen/menschen-erkennen-ki-generierte-medien-kaum-6430> am [06.03.2025].
- Safar, M. (o.D.). *Large Language Models – Grundlagen KI-getriebener Kommunikation*. Weissenberg. Aufgerufen von <https://weissenberg-group.de/was-ist-ein-large-language-model/> am [22.02.2025].
- Schlude, A., Schwind, M., Mendel, U., Stürz, R., Harles, D., Fischer, M. (o.D.). *Verbreitung und Akzeptanz generativer KI in Deutschland und an deutschen Arbeitsplätzen*. Bidt. Aufgerufen von  
<https://www.bidt.digital/publikation/verbreitung-und-akzeptanz-generativer-ki-in-deutschland-und-an-deutschen-arbeitsplaetzen/> am [06.03.2025].

- Schultze, T. (2024). *Die KI-Synchro von „Black Dog“: „Das Original verpflichtet“*. The Spot Media & Film. Aufgerufen von <https://the-spot-mediafilm.com/news/interviews/die-ki-synchro-von-black-dog-das-original-verpflichtet/> am [13.03.2025].
- Singh, P., Patidar, M., Vig, L. (2024). *Translating Across Cultures: LLMs for Intralingual Cultural Adaptation*. ArXiv. <https://doi.org/10.48550/arXiv.2406.14504>
- SkimAI. (2024). *Was ist Voice Cloning und wie funktioniert es?* Abgerufen von <https://skimai.com/de/was-ist-stimmenklonen-und-wie-funktioniert-es/> am [25.01.2025].
- Software-Aspekte (2024). *Herausforderungen der künstlichen Intelligenz: Risiken und Lösungen*. Software-Aspekte. Aufgerufen von <https://software-aspekte.de/blog/herausforderungen-der-ki/> am [20.02.2025].
- Stadler, M. (2024). *Text-To-Speech (TTS)*. Mindsquare IT-Beratung und Entwicklung. Aufgerufen von <https://mindsquare.de/knowhow/text-to-speech/> am [22.01.2025].
- Stopa, J. (2023). *Meine Kopie spricht sieben Sprachen*. Zeit Online. Aufgerufen von <https://www.zeit.de/digital/2023-09/ki-tool-uebersetzer-heygen-videos-test> am [27.02.2025].
- Storyshell (o.D.), *The ROI of Using AI Dubbing: A Comparison of Storyshell's Services and Traditional Dubbing Agencies*. Storyshell. Aufgerufen von <https://www.storyshell.io/blog/roi-storyshell-dubbing> am [09.04.2025].
- Speechify (o.D.). Aufgerufen von [https://speechify.com/de/blog/how-does-ai-dubbing-work/?srsltid=AfmBOoqckcXKUnA6SAwDYF\\_B6KpGOzRjqokcvGYqxEO1FoSeKpkw1tbh](https://speechify.com/de/blog/how-does-ai-dubbing-work/?srsltid=AfmBOoqckcXKUnA6SAwDYF_B6KpGOzRjqokcvGYqxEO1FoSeKpkw1tbh) am [03.01.2025].
- Speechify (o.D.). *KI-Synchronisation vs. traditionelle Synchronisation: Kosten, Geschwindigkeit und Effizienz im Vergleich*. Speechify. Aufgerufen von <https://speechify.com/de/blog/ai-dubbing-vs-traditional-dubbing/> am [27.02.2025].

- Speech Technology (2024). *AI Dubbing Market to Grow to \$1.88 Billion by 2030*. Speech Technology. Aufgerufen von <https://www.speechtechmag.com/Articles/News/Speech-Technology-News/AI-Dubbing-Market-to-Grow-to-%241.88-Billion-by-2030-165350.aspx> am [09.04.2025].
- The Decoder (2023). *Wie KI in der Videosynchronisation eingesetzt werden kann*. The Decoder. Aufgerufen von <https://the-decoder.de/wie-ki-in-der-videosynchronisation-eingesetzt-werden-kann/> am [20.02.2025].
- TNT Media (o.D.). *TNT & KI*. TNT Media. Aufgerufen von <https://www.tnt-media.de/tnt-ki> am [12.03.2025].
- Turing, A. M. (1950). *Computing Machinery and Intelligence*. *Mind*, 59(236), 433–460. <https://doi.org/10.1093/mind/LIX.236.433>
- Urbe, W. (2023). *KI erweckt die Originalstimme des Pumuckl wieder zum Leben*. RedaktionsNetzwerk Deutschland. Aufgerufen von <https://www.rnd.de/medien/pumuckl-bei-rtl-ki-weckt-originalstimme-von-hans-clarin-wieder-zum-leben-C4OI4QYIHBEGNEHCYXDH5TW6CQ.html> am [28.01.2025].
- VDS (2024). *Statement zu KI-generierten Synchronisationen von VDS, BSD und UVA*. VDS. Aufgerufen von <https://www.sprecherverband.de/aktuelles/statement-zu-ki-generierten-synchronisationen-von-vds-bsd-und-uva/> am [08.03.2025].
- Verbalate (2024). *Eine vergleichende Analyse*. Verbalate. Aufgerufen von <https://verbalate.ai/de/ai-synchronisation-vs-traditionelle-synchronisation/> am [27.02.2025].
- Video4net. (o.D.). *Voice Cloning mit KI-Tools: Eine neue Ära der Sprachsynthese*. Abgerufen von <https://www.video4net.de/service/blog-inhalt/764-voice-cloning-mit-ki-tools-eine-neue-aera-der-sprachsynthese> am [28.01.2025].
- Von Medvey, S. (2025). *Stellungnahme am 28.03.2025*. Das Gespräch befindet sich im Anhang dieser Arbeit.

Webb, M. (2024). *AI Voice Cloning: beunruhigender Trend setzt sich fort*.  
Techopedia. Aufgerufen von <https://www.techopedia.com/de/ai-voice-cloning-beunruhigender-trend> am [31.01.2025].

Weitzman, C. (2023). *“Wie funktioniert KI-Synchronisation: Ein Leitfaden”*.  
Wikipedia. (2023). *Synchronisation (Film)*. Abgerufen von  
[https://de.wikipedia.org/wiki/Synchronisation\\_%28Film%29](https://de.wikipedia.org/wiki/Synchronisation_%28Film%29) am  
[03.01.2025].

## Bildquellen

<https://deepdub.ai/solution/post-production-dubbing-studios> (Abb. 1)

<https://elevenlabs.io/docs/capabilities/dubbing> (Abb. 2)

[https://www.reddit.com/r/ElevenLabs/comments/14gupz0/professional\\_voice\\_cloning\\_how\\_is\\_the\\_quality/](https://www.reddit.com/r/ElevenLabs/comments/14gupz0/professional_voice_cloning_how_is_the_quality/) (Abb. 3)

<https://elevenlabs.io/docs/product-guides/voices/voice-cloning> (Abb. 4)

<https://www.resemble.ai/voice-cloning/> (Abb. 5)

<https://de.vidnoz.com/kuenstliche-intelligenz/lip-sync-apps.html> (Abb. 6)

<https://www.flawlessai.com/truesync> (Abb. 7)

## Anhang

Gespräch mit Cornelius Frommann am 26.03.2025

Cornelius Frommann ist Synchronregisseur und hat 30 Jahre Erfahrung als Synchronautor und -regisseur. Er beschäftigt sich maßgeblich mit der deutschen Sprache und deren linguistischen Eigenschaften. Als Synchronautor und Regisseur sorgt er für lippen-synchrone Texte bei Synchronisationen fremdsprachiger Produktionen.

### Stellungnahme Cornelius Frommann

Jo, was habe ich dir aus meiner Warte zu erzählen? Ich weiß nicht, was Jochen vielleicht schon angeteasert hat an bestimmten Takes, einzelnen Sätzen, einzelnen Denkweisen in der Sprache, wo er sagte, ich wäre vielleicht der Richtige dazu, was zu sagen. Ich bin grundsätzlich der Meinung, die KI wird immer viel falsch machen. Aber die Frage ist ja nicht, wie viel sie falsch macht, sondern ob es akzeptiert wird, ob es reicht für die Zuschauer.

Denn wenn die sich damit zufriedengeben, weil ihr eigenes Sprachniveau so ist, weil sie zum Beispiel durch Internet über Hip-Hop über alles einfach in bestimmte Sprachmuster reingefallen sind, dass sie diese Synchro für menschlich genug halten, weil sie denken, Menschen könnten so reden.

Das hängt wieder ganz stark von der Epoche ab, von der Zeit. Wenn das heute und relativ junge, moderne oder mitteljunge bis junge modernredende Menschen sind, dann kann die KI das vielleicht erwischen. Aber schnallt die KI, wenn der Film in den 80er Jahren spielt, dass es noch keinen Van gibt?

Schnallt die KI, wenn es vor den 70er Jahren spielt, dass das Wort geil noch nicht seine Zweitbedeutung hat? Zeit, Jargon, Slang, Einordnung. Welche Person sagt das?

Die eine Person spricht ganz klar wegen mit einem Genitiv und die andere Person spricht ganz eindeutig wegen mit einem Dativ. Also hunderte von Punkten, die die

KI vielleicht nicht mit diesem dritten, vierten Register eines Lexikons in welcher Zeit hätte wer wie was gesagt.

In welchem Alter, welchen Stand, welchen Ausbildungsgrad würde wer was wie sagen. Sondern sie sieht einen Satz, der auf Englisch ganz schlicht ist und der sich auf Englisch zum Beispiel, wenn wir Englisch nehmen, aus Spaß, auch nicht groß verändert.

Ob das ein schwarzer 16-Jähriger in L.A. sagt oder ob das ein weißer britischer Professor sagt, denn die könnten in vielen Fällen den gleichen englischen Satz verwenden.

Bei uns wären es zwei verschiedene. Also Sprachniveau, Alter, Bildungsgrad, Epoche.

Welcher Slang gilt ab wann? Welche Person würde es verwenden? Die persönlichen Eigenschaften, Spracheigenschaften einer Figur. Kann eine KI das so einschätzen wie wir, wenn wir eine Figur sehen und sagen, hier ist eher Schnotter, hier ist Cracknutte in Moskau oder hier ist... Woher?

Zu den Möglichkeiten von noch treffsicherer sein in deutscher Sprache, was die Zielsprache von allen Synchronen ist. Ein Übersetzerfreund hat mir vor langen Jahren mal gesagt (der hat Literatur übersetzt), der hat mir gesagt: Wo was ist, muss was hin.

Das ist ein wunderbarer Satz, denn es heißt nur, wenn der englische Satz Diabolik beinhaltet, muss ich nicht diabolisch sagen, dann kann am Ende bei mir auch der Satan in dem Satz stehen.

Aber das was da ist, muss da hin, in irgendeiner ähnlichen oder wirksamen Form. Beispiele, die zitierfähig sind.

Einer erklärt, wie du es machen musst.

„Whatever it is, you do it like this or like that and Bob's your uncle.“

Jetzt, wörtlich kann ich es nicht nehmen. Die KI kann Englisch übersetzen und wird sagen ... und schon ist es fertig.

Ja, aber bei mir hieß es, du machst es so und so und so, Klappe zu, Affe tot. Woher soll die KI das kriegen, wenn ein Idiom wie Bob's your uncle auch eine bestimmte Schnodderigkeit (das spielte in den 60ern) hat, und den Ausdruck gibt es ewig und wir nehmen: Klappe zu, Affe tot. Auf Bayerisch würde man sagen, dann ist der Kas Bissen, auf Schwäbisch würde man sagen, dann ist der Kittel geflickt.

Also, woher soll eine KI das ziehen?

Ein Kollege von mir hat vor vielen Jahren in einem sehr lustigen Film mit Tom Cruise, Jerry Maguire heißt der, einen tragenden Satz erschaffen. Da gibt es einen tragenden Leuchtturmsatz, der sich durch den ganzen Film zieht. Der Sportagent hat einen großen Football Star und will dem einen Vertrag besorgen. Der schwarze Spieler, Running Back, sagt zu ihm, come on man, show me the money. Das heißt ja nicht, zeig mir das Geld.

Sondern, mach, dass ich an die Kohle rankomme. Und dann musst du noch das Lippen-Synchron bedenken.

Show me the money. Wenn du deinen Mund dabei beobachtest, wie show me the money aussieht, dann ist bei zeig mir das Geld hinten das Geld auf the money.

Ja. Das sieht furchtbar aus. Also, dieser Kollege hat vor vielen Jahren den genialen Satz geprägt: „Führ mich zum Schotter“.

Das ist gut.

Das entspricht dem schwarzen, schnodderigen.

Auch wieder ein Beispiel, ab wann gibt es Hipster? Und ab wann gab es sie im Englischen und ab wann gab es sie im Deutschen?

Was ein Unterschied von 15 Jahren sein dürfte.

Ja. Show me the money, führ mich zum Schotter heißt eben auch der Prozess von wegen, dass ich an die Kohle rankomme.

Mach, bring mich da hin. Besorg mir einen Hammer-Vertrag.

Also, wie kann eine KI-Idiomatik hier unterscheiden. Schicht, Hautfarbe, weil L.A., Downtown, arme, schwarze Herkunft spricht man anders als Ostküste weiß.

Ja, das sind alles Dinge, die ich für ganz entscheidend halte, was die KI nicht kann. Oder nur mühsam lernen wird, denn sie muss ja immer Lexika abgleichen.

Oh, was ist das? Ah, da müssen wir Sprache von da ziehen, was haben wir denn da? In Sekundenschnelle oder in Zehntel, Tausendstel Sekundenschnelle.

Aber wie soll das gehen? Welcher Humor, der sich durch welche Worte ausdrückt, passt zu welcher Person? Also Menschen basteln, glaubwürdige Menschen in der deutschen Sprache basteln.

Ja, das ist mein Hauptfeld. Ich hab noch ein paar andere Punkte, aber mit diesen Bereichen beschäftige ich mich hauptsächlich.

Also, Stand, Bildung, situationsbezogen, welche Art Humor, welche Epoche und so weiter.

Das gilt dann noch für besondere Sprachformen, das kommt nicht oft vor, aber es kommt vor. Was ist denn mit dem relativ pathosreich wirkenden historischen Blankvers eines Bühnenstücks?

Also, die kann das, die weiß, was die Regeln eines Blankverses sind, die KI. Die weiß, wie Hebung, Senkung geht.

Aber wie soll sie denn Hebung, Senkung dann noch mit sinnvollen Alliterationen, die im englischen Original vielleicht da sind, versehen. „Wo im Alter alles enden muss“. Das sind Sachen, hatte ich gerade in einer Wilhelm Tell-Verfilmung.

Dass du das bewusst machst.

Bestimmte, also Stabreimartige Alliterationen, die im englischen vielleicht an einer anderen Stelle im Satz sind und bei einem anderen Wort, aber drin sind. Klammer auf, wo was ist, muss was hin.

Dann jetzt noch den Punkt, einen sinnvollen deutschen Satz synchron zu kriegen, dass er auf dem Gesicht, vom Rhythmus, von der Mimik, auch von der Gestik, wenn z.B. eine Handbewegung auf einem bestimmten Impuls ist, wo das starke Verb ist.

Musst du ja mit beachten. Also Körpersynchron, nicht nur Lippensynchron.

Da kann man die ganze Zeit lernen, wo B's sind und U's sind und B's sind, die man sieht. Aber woher sieht die KI, dass ein TH mit rausgestreckter Zunge, mit sehr weit vorne liegender Zunge, im Lippensynchron nur durch ein L zu retten ist? Ein L sieht wenigstens halbwegs so aus, als könnte die Zunge da vorne sein.

Und zwar sprechen wir nicht L, das klingt dann ein bisschen doof, aber um das Synchron zu retten, machen wir das. Oder wir stellen einen Satz komplett um. Den zweiten Teil in den ersten Teil und den ersten Teil in den zweiten Teil, weil es auf Deutsch so griffiger ist.

Oder synchroner ist. Wie soll die KI wissen, dass ich bei den Nöten im Synchron, z.B. Engländer benutzen immer das Präteritum, he went, he did, he gave, they told und so weiter.

Und wir benutzen eigentlich immer das Perfekt. „Die haben uns gesagt, dann bin ich da hingegangen, er hat mir eine reingehauen“. Also wir wollen gerne Perfekt hören, weil es natürlich ist.

Wir sagen nicht, er haute mir eine rein.

Du wirst an vielen Synchros hören, die du schon studiert hast, dass viele faule Synchronleute ganz viel ins Imperfekt gehen.

Weil sie dieses Problem nicht gelöst kriegen.

Jetzt kann ich aber z.B. mich manchmal retten, auch nicht immer. Ich kann mich manchmal retten und so kurz werden wie das englische Imperfekt, indem ich ins Erzählen des Präsens gehe.

„Der kommt rein und haut mir eine rein und sagt: Datteldaddeldat“

Diese Verkürzungen, Erzählen des Präsens und Verben weglassen, woher soll die KI das wissen, dass das geht oder bei wem das geht? So, das sind meine Ansätze.

Ich kann auch ein Slangwort zum ersten Mal erfinden. Weil es sich in dem Film auf einmal so ergibt, dass ich der Figur folgen kann, obwohl das kein gebräuchlicher Slang oder eine gebräuchliche idiomatische Redewendung ist, aber es ergibt sich so schön.

Also ob die KI menschliche Sprachwitze oder Sprachverstand oder Sprachvielfalt selber kreieren kann.

Und indem sie von Regeln abweicht, weil vorher gab es das nicht.

Und es ist aber auch, was darüber hängt, wie eine Angst oder wie ein grauer Nebel ist, die Frage, ob diese ganzen Aspekte, die Qualität ausmachen oder ausmachen sollten, ob die noch verstanden oder gewollt werden. Ich weiß nicht, die Sprache ändert sich, das sagen immer alle.

Manches ändert sich von allein durch Einschleifen. Und jetzt gibt es das, jetzt sage ich das sogar aus Versehen selber manchmal, obwohl ich mich 25 Jahre gewehrt habe.

Oder das weiß die KI sicherlich, dass es bestimmte Anglizismen gibt, die im Deutschen akzeptiert sind. Sie hat so viele deutsche Texte gelesen, dass sie

weiß, das „Realisieren“, „das habe ich gar nicht realisiert“, verwendet wird im Sinne von merken, schnallen, wahr werden.

Und dass es aber vor 40, 50 Jahren nur verwirklichen hieß.

To realize heißt im Englischen merken, verstehen. Aber im Deutschen war Realisieren ausschließlich verwirklicht.

Das vermischt sich, das ist unvermeidlich. Dass Wörter aus einzelnen Sprachen in andere Sprachen gehen.

Und dann gibt es noch so Schichten, Schichten sind das nicht, Szenensprachen.

Es gibt also wer gendert und wer gendert nicht und wer benutzt welche Ausdrücke und wer benutzt sie nicht. Oder wer erklärt bestimmte Worte für verboten.

Weil mit dem Wort selbst das Böse mitschwingt wie Neger oder Nigger. Jetzt machen die das aber in dem amerikanischen Film plötzlich die ganze Zeit, weil die Schwarzen gehen sich untereinander alle mit „ey, Nigger, come on, Nigger, you can't tell me that, Nigger“.

Da musst du was machen. Und die KI wird dann was auswählen und sagen, ja, das ist gemeint, das ist klug, die KI kann sagen, das ist gemeint als Slang-Anrede. Das heißt dann immer, „Alter, ey, weißt du“, so.

Die wird das eliminieren.

Aber habe ich damit diese Schwarze-Ghetto-Gang richtig beschrieben? Oder muss ich lieber Menschen herbeibringen, die doch Nigger sagen lassen? Nur weil die das auch tun.

In der Synchro dürfen wir das nicht mehr. Wir dürfen das überhaupt nicht mehr.

Wenn sie diese Härten nicht mehr haben. Oder, einer ist fies, ja? Er ist fies und benutzt schwul nicht als homosexuelles Wort, sondern als Weichei-Verhalten.

„Das ist ja voll schwul“. Ja, das höre ich nun mal auf der Straße und eine Person kann unsympathisch sein oder ich kann grinsen müssen dabei, ist ja egal,

man benutzt es und wann weiß die KI, da gehört es hin? Weil der Typ würde es so sagen.

Also nochmal, Slang meets Political Correctness und Verbote von Begriffen, nicht Verbote, Verhöhnung von Begriffen.

Das ist so mein Feld. Ich komme von der Idiomatik, von der Sprache, wie sie aus Menschen rauskommt.

Und wenn uns dann, wenn deiner Arbeit dann widersprochen wird durch die Wirklichkeit, indem einer sagt, ich habe hier einen KI-Film, der hat das alles geschnallt. Ja, dann müssen wir uns genau anhören und gucken, was hat der Feind dazugelernt.

Gespräch Susanne von Medvey am 28.03.2025

Susanne von Medvey ist eine deutsche Film-, Theater- und Synchronschauspielerin und steht schon seit über 30 Jahren vor dem Mikrofon. Mit über 1000 Sprechrollen hat sie jede Menge Erfahrung in der Synchronbranche gesammelt. Außerdem unterrichtet sie neue Talente in der Synchronbranche.

**Stellungnahme Susanne von Medvey**

Überhaupt in der Kunst ist die KI sehr bedrohlich. Und es dauert so wahnsinnig lange.

Ich verstehe das irgendwie nicht, dass es so lange dauert, bis es Regelungen geben wird. Das kann ich einfach nicht nachvollziehen.

In der Zwischenzeit. Die sind so schnell unterwegs. Die Entwicklung ist so rasant. Und der Umgang damit kommt einfach nicht hinterher. Das finde ich eigentlich das einzige Schlimme.

Dass das schon ein Problem ist. Und die Verbände kümmern sich und es kommt nichts zustande. Irgendwann müssen wir dann auch nicht mehr drüber nachdenken. Sie haben ja eh alles, was sie brauchen.

Und sie können alles überall abgreifen. Manchmal denke ich mir, der ganze Kampf ist teilweise ein Kampf der Eitelkeiten.

Das ist auch so mein Gefühl. Ich finde es einerseits ehrenwert und ganz toll, dass die KollegInnen wirklich kämpfen und alles Mögliche auf die Beine stellen. Und auf der anderen Seite denke ich mir, ich kriege es teilweise persönlich mit von Leuten, wo ich auch weiß, es ist eben auch der Kampf, endlich werden wir mal gesehen. Und das stört mich auch so ein bisschen.

Obwohl ich wie gesagt immer dankbar sein muss, dass es Menschen gibt, die für Sachen kämpfen, für die ich halt nicht so viel Einsatz bringe. Und kämpfen dann auch für mich mit. Das finde ich irgendwie toll.

Aber ich sehe halt wirklich auch die Eitelkeit dabei. Und auch so eine gewisse Verblendung.

Es geht nicht um Leben und Tod. Natürlich existenziell für manche schon. Aber ich finde, es wird teilweise auch von der falschen Seite betrachtet.

Wenn wir in der Ausbildung nicht besser werden. Und wenn wir uns Druck machen lassen und die Firmen geben den Druck der Auftraggeber weiter an uns. Und es soll alles billiger werden.

Es werden ganze Werke weggelassen. Möglichst. Und es soll alles schneller gehen und billiger werden.

Und wenn wir so arbeiten, dann werden wir automatisch, und das passiert gerade, schlechter. Es sitzen teilweise Leute in der Regie, wo ich denke, das kann nicht sein, dass der mir jetzt was sagen will. Das heißt, die Qualität wird schlechter, die KI entwickelt sich und wird besser.

Und da kann man auch irgendwie nahtlos dazu übergehen. Und dann macht es auch nichts mehr. Mein eigentlich wesentlicher Aspekt ist, wir haben nur eine Chance, nämlich dass wir Menschen sind.

Aber wenn wir das nicht begreifen, das trifft ja auf viele Gewerbe zu, wenn wir das nicht begreifen, dass das die Chance ist. Wir müssen nicht konkurrieren mit der KI, wir müssen einfach nur besser werden, indem wir menschlicher klingen und sind.

Und das würde schon helfen. Die Leute sind ja auch nicht blöd, die Zuschauer zu sein. Man darf sie nur nicht überschätzen.

Sie werden jetzt auch nicht verzweifeln, wenn der Unterschied nachher nicht mehr groß ist, dann gucken sie halt original, okay. Aber diese eine Chance haben wir ja und ich habe einfach so das Gefühl, dass diese Chance nicht, also dass wir uns alle nicht einig genug sind.

Mit allen meine ich die Firmen, die Auftraggeber, die SprecherInnen, die Autoren, die Regisseure, alle die, sind sich nicht einig genug in diesem Punkt.

Das ist für mich eigentlich das Wesentliche. Und mir tut es halt wahnsinnig leid für die vielen Jüngeren, die mit Elan irgendwie in den Job rein wollen und die Sachen mit denen viele anfangen, oder ich meine viele kommen halt auch nicht vom Schauspiel und sprechen aber trotzdem und machen halt so Audioguides und ein bisschen Doku und ein bisschen Werbung und solche Sachen. Und das fällt ja alles weg.

Jetzt so hochwertige Dokus für Arte oder so und synchron hochwertige Produktionen, die wird es noch eine Weile geben. Aber es ist halt schon dünn. Je weiter das dann fortschreitet, je besser die KI wird, desto weniger wird es dann. Ich kriege halt auch mit, die Kolleginnen teilweise wirklich, neulich eine wieder, die ein Projekt für 7000 Euro im Game nicht gemacht hat, weil die KI-Klausel nicht drin war. Und ich bin da immer so zwischen Hochachtung und Blauäugigkeit irgendwie, weil ich so denke, ja, einerseits natürlich super, weil so die große Freude, die dann in den Gruppen geteilt wird, auch dass die dann irgendwie schwer jemanden finden, der es dann macht, weil es doch jetzt einen größeren Zusammenhalt gibt, was ja sehr schön ist.

Aber ich frage mich halt wirklich, ob es das echt bringt oder ob die nicht einfach eh schon alles überall nehmen können, wo sie wollen. Ja, und ich verstehe es nicht ganz.

Also es wird immer wieder gesagt, nein, das ist nicht so und sie brauchen irgendwie aktuelle Sachen, um gezielt zu trainieren und ich bin mir da halt nicht so sicher.

Der VDS hat jetzt irgendwie eine Klausel, ich weiß nicht, ob du die schon kennst, die in die Verträge übernommen werden soll, da kann man irgendwie drauf bestehen.

Es ist ein Video rausgekommen, was prominente Feststimmen gemacht haben, was ich ganz schön finde. Also die einfach die Öffentlichkeit aufmerksam macht.

Das kann ich dir nachher mal schicken. Da ist dann die Stimme von Angelina Jolie und Cameron Diaz und so. Das haben sie echt, finde ich, gut gemacht.

So ein Schnellschuss, aber richtig gut gemacht. Okay, ich finde, mit solchen Aktionen kann man vielleicht einfach Leute aufmerksam machen.

Und dass das nicht ganz egal ist. Ich finde auch, es kommt halt auch viel von den Zuschauern eben, ob die das halt akzeptieren, dass das KI macht oder nicht.

Kann ich mir gut vorstellen, dass es welche gibt, die, wenn sie darauf aufmerksam gemacht werden, dann auch hinhören. Aber es gab ja auch jetzt diese eine Serie, ich glaube, es war eine polnische, die auf diesem Telekom-Sender komplett von KI gemacht wurde. Und da haben sich die Leute beschwert und die wurde abgesetzt.

Also auch das gibt's. Aber es gab auch, letztes Jahr oder vor zwei Jahren, gab's auch eine extrem schlecht synchronisierte Serie, die auch abgesetzt wurde, von Menschen synchronisiert, weil es so schlecht war.

Also es ist halt ganz schön, dass die Leute ja irgendwie offenbar das Ding wissen, noch nicht völlig wurscht ist. Aber ich weiß nicht, wie verlässlich das ist.

Und wenn dieses Deep-Dub sich durchsetzt, dann ist halt echt schwierig.

Vor allem, es entwickelt sich ja auch immer weiter. Klar, man sagt jetzt, okay, es ist jetzt noch nicht so der Burner, dass man das jetzt so gut benutzen kann, aber

es wird halt auch immer besser so. Und ich glaube, Deep-Dub wird wahrscheinlich ziemlich teuer.

Also wenn sich Brad Pitt das bezahlen lässt, ich glaube, dann sind wir mit Synchronisationen immer noch billiger. Also da kann ich mir vorstellen, dass das erstmal mehr kosten würde.

Ja, es ist ein schwieriges Thema. Weil es so viel beeinflusst.

Also die Angst, die das macht, ist halt auch nicht gut. Weil die Angst halt irgendwie auch so auslöst, dass ganz viele einfach nochmal schnell abgraben wollen, was geht. Es kommt so in eine Endzeit-Stimmung und das ist auch gar nicht gut.

Anstatt besser zu werden, sorgfältiger zu werden und einfach noch feiner abzunehmen vom Original und so, passiert teilweise wirklich genau das Gegenteil.

Und es ist auch so eine Grundstimmung, die halt einfach so bedrohlich ist.

Und wie gesagt, das Schöne ist, es ist irgendwie so eine, ja, vielleicht ist es auch eine Illusion, aber ein Stück weit ein Zusammenhalt.

Das ist nicht unbedingt üblich unter SchauspielerInnen. Und das finde ich eine ganz gute Entwicklung, denn, ich weiß noch nicht, wo das hinführt.

Also ich glaube, dass bestimmte Bereiche werden einfach, ich kann nicht so genau sagen, was ich glaube, wie lang das dauert, aber bestimmte Bereiche, da wird es relativ schnell gehen und andere Bereiche, glaube ich, hochwertige Kinofilme und eben hochwertige Dokus und sowas wird sich noch eine Weile halten, nehme ich an. Aber es gibt auch Mischformen. In Köln gibt es eine Firma, die hat, ich weiß nicht, ob sie es noch machen, aber die haben Filme gemacht, wo sie nur noch die Hauptrollen mit Menschen besetzen.

Ja, ist halt auch maßgeblich eben noch an der Qualität verschuldet, die ja jetzt einfach noch nicht so gut ist und die Leute ja einfach was Gutes wollen.

Und deswegen kann man das jetzt einfach auch noch gar nicht ersetzen. Aber die Frage ist halt auch wirklich, was gut ist.

Wir beeinflussen ja die Hörgewohnheiten. Wenn dieser Einfluss kein guter ist, also ich meine, mir geht es jetzt schon so, immer schon, dass ich manchmal einfach was nicht anschauen kann, weil es so schlecht synchronisiert ist, weil es so

wahnsinnig nervt. Und wenn die Leute halt so künstlich sprechen, ja dann ist es halt scheiße.

Und dann können wir auch nichts mehr bewirken. Und das meine ich. Die Leute haben sich daran gewöhnt. Die Leute gewöhnen sich an mittel bis schlechte Synchronisation.

Das Blöde beim Synchronen ist halt, wenn es wirklich gut ist, fällt es nicht auf. Das heißt, es ist schwierig, dann mit Lorbeeren zu ernten, weil es halt unauffällig ist.

Und das ist auch das Ziel. Deswegen habe ich auch so ein bisschen ein Problem mit der Eitelkeit. Aber ich sehe es im Moment ein, weil man ein Forum schaffen muss.

Aber ursprünglich habe ich immer gedacht, wenn ich die Eitelkeit bedient haben will, oder wenn ich gesehen werden will, dann muss ich halt was anderes machen. Dann gehe ich halt auf die Bühne oder dann mache ich ein Liederabend. Aber ich finde, Synchron ist dienen.

Aber jetzt hat sich das natürlich ein bisschen verändert und es ist vielleicht gar nicht schlecht, wenn man sichtbarer wird.

Ich glaube, dass sowieso von den jüngeren Leuten, die ja auch eh immer im Fokus sind als Zielgruppe, wird es eher so sein, dass die dann sagen, ach, dann gucke ich doch im Original. Und dann gibt es halt Original mit Untertiteln für nicht amerikanische Filme.

Das funktioniert ja auch. Es geht nicht um Leben und Tod in dem Sinne, aber es geht natürlich teilweise schon wirklich ums Überleben der Leute, die das machen. Wenn CutterInnen zum Beispiel ausgespart werden, damit es billiger wird, dann geht es da auch ums Überleben.

Der künstlerische Aspekt von deinem Beruf, der fällt dann natürlich auch ziemlich weg. Das ist dann halt auch nicht mehr so das, was es mal war. Aber für manche ist es vielleicht irgendwie eher, dass sie ganz laut Ja sagen, weil alles schneller geht, man kann in letzter Zeit mehr Geld verdienen, weil man eben

mehr Zeit zur Verfügung hat, und man muss sich nicht rumschlagen mit den Zickereien der Schauspieler.

Und muss sich mit denen nicht auseinandersetzen, was ja einigen auch nicht so viel Spaß macht. Also das ist auch eine Sichtweise, die ich von der technischen Seite her auch kenne.

Und ein wichtiger Aspekt ist halt auch, ich meine, der mich sehr trifft, da ich unterrichte, wie lange macht es noch Sinn, Leute auszubilden dafür? Ich unterrichte das an der Bayerischen Theaterakademie bei August Eberding und ich mache ja auch für die Stuttgarter Workshops mit Jochen und ich mache Einzelcoaching mit Leuten und da ist auch wirklich, da habe ich jetzt mal zwei Tage, einmal einen Tag in Berlin und bin eingesprungen und einen in München für Yellow Dub Marine gemacht. Das finde ich sehr fragwürdig. Aber ich denke, wenn wir ausbilden, dann müssen wir auch da extrem gewissenhaft sein.

Und aus menschlichen Gründen, man darf halt nicht mit den Hoffnungen der Leute spielen, man muss ihnen ganz klar sagen, wie die Lage ist. Und auf der anderen Seite, so wie das Yellow Dub Marine macht, kann man halt auch nicht qualitativ hochwertig ausbilden. Und das ist halt auch schwierig.

Die Fragen stellen sich einfach auch, was sage ich den jungen Leuten, die das machen wollen?

Der nächste entscheidende Punkt sind für mich diese ganzen halb professionellen QuereinsteigerInnen, die irgendwie auf LinkedIn sich gegenseitig hypen und liken und sich auch gegenseitig Workshops geben.

Und ich denke, ey, der Blinde führt den Namen, das geht gar nicht. Und die dann auch so ihre Demos reinstellen und ich kriege echt die Krise, was die in ihren Homestudios da aufnehmen für wenig Geld. Und die haben natürlich auch den Markt überschwemmt.

Das heißt, für Profis, die das wirklich ernst nehmen und vom Schauspiel kommen, ist das auch nochmal ein großer Marktanteil, der wegbricht. Ich finde es wirklich eine Plage. Und die meisten, diese ganzen Workshops, wo wirklich mit den Hoffnungen gespielt wird und die sich gegenseitig, wie gesagt, da stützen, aber die verdienen ja ihr Geld.

Also weniger natürlich, aber die Masse macht's. Die sitzen dann in ihrem Homestudio und für den Preis von allem machen die halt 3 Hörbücher und sind happy, weil sie sich ihren Kindertraum erfüllen.

Dass das möglich ist. Das war mal ein Beruf.

Das ist echt sehr schwierig alles. Also ich finde, dass die gesamte Branche eh relativ verkommt und jetzt kommt die KI noch dazu.

Und es trifft ja nicht nur uns. Wenn es um Technik geht oder was, es trifft ja dann auch Leute, die Musik aufnehmen wollen.

Es ist einfach eine andere Dimension, weil manche sagen, ja, das gab es ja früher auch. Es gab immer Veränderungen und dann kam irgendwie der Tonfilm und so.

Ja, okay. Das ist natürlich auch traurig, aber da sind halt ein paar Pianisten arbeitslos geworden. Aber es sind schon andere Dimensionen jetzt.

Ganz abgesehen von diesen ganzen Fake-Geschichten. Das ist auch ein großes Problem.

Und du kannst es nicht kontrollieren. Ich kann zum Beispiel, jemand kann ja meine Stimme nehmen, von was auch immer, was es alles schon gibt, und verändert die minimal, tut quasi ein bisschen was anderes dazu. Und keiner kann ja nachweisen, dass es irgendwas mit mir zu tun hat.

Die Möglichkeiten sind ja fast grenzenlos. Es gibt unfassbar viel, was man damit machen kann.

Es wird einfach richtig verpennt gerade. Und irgendwann kommen dann so Reaktionen. Wo diese ganzen Fake-Sachen herkommen.

Irgendwann weißt du halt sowieso nicht mehr, was echt ist und was nicht. Das ist das Allerschlimmste.

Da kann sich dann auch keiner mehr wehren. Dann ist es eigentlich schon fast zu spät.

Was wirst du da noch machen?

Es ist echt keine gute Zeit.

Insgesamt wird es einfach weniger. Auch die Firmen. Immer mehr Firmen haben immer weniger zu tun.

München ist ziemlich trostlos geworden. Mir geht es auch so gerade, dass ich mich sehr nach Berlin orientiere und hauptsächlich in Berlin arbeite. Es ist auch super schwer einzuschätzen, was da noch alles kommt.

Und das ist echt ein toller Beruf eigentlich. Und diese Freiheit, die das hat.

Und du bist immer frei. Und du kannst auch immer Nein sagen, keiner kann dich wirklich irgendwie binden.

Und man hat wirklich fast immer mit netten Leuten zu tun.

Und wenn nicht, dann muss man sich nur sagen, naja, ich bin ja gleich wieder weg.

Und deswegen finde ich es auch so unfassbar schlimm, was da gerade passiert, weil eben dieser Bereich da so ein bisschen kaputt gemacht wird dadurch.

Und das ist so schade. Aber wie gesagt, wir arbeiten also teilweise auch selber mit daran, dass der Bereich kaputt geht.

Das ist nicht nur die KI. Im Moment kommt viel zusammen.

Also es gibt so Entwicklungen, die sind nicht so gut. Auch dass Influencer Hauptrollen oder zumindest in großen Kinofilmen sprechen, ist auch jetzt nicht so der Hit.

Aber es geht halt um Verkauf. Und es geht wirklich immer überall darum, schnell noch was abzugraben.

Und mir geht es echt nicht, ich bin jetzt nicht so, wie soll ich es sagen, mir geht es nicht darum, dass es früher besser war. Im Gegenteil. Ich finde viele Entwicklungen zum Beispiel wirklich super.

Früher stand man gemeinsam vor dem Mikro und heute alleine. Was ich super finde. Ich liebe das, weil man einfach konzentrierter arbeitet und meine Philosophie im Synchronen ist sowieso das Dienen.

Ich diene der Schauspielerin, die das schon gespielt hat. Also ich muss das dafür jetzt nicht meinen Kollegen irgendwie neben mir haben, damit ich das Gefühl habe, ich habe einen Dialog. Wenn ich meiner Schauspielerin diene, reicht der Fokus.

Und das finde ich ganz toll, weil die meisten Kollegen dann sagen, ja, es war viel schöner, man sieht sich ja nicht mehr. Ich habe ja einen Freundeskreis und ich habe eine Familie. Man sieht sich trotzdem noch genug.

Diese konzentrierte Arbeit ist einfach schön, finde ich. Mir fehlt es überhaupt nicht.

Ich finde jetzt nicht, dass früher alles besser war, aber im Moment ist so ein Wandel, der echt gravierend ist.

Und was die Games betrifft, da scheinen sie ja wohl weiterzukommen.

Da haben alle wohl sehr konsequent sich geweigert und da werden wohl immer mehr so Klauseln eingebaut. Aber beim Synchron ist immer noch so ein komplett rechtsfreier Raum. Der BFFS sagt irgendwie, es braucht jetzt keine Klausel, das muss erst entwickelt werden, dann muss man sich Zeit lassen und das Ganze in Ruhe machen.

Das braucht es gar nicht, weil wenn es nicht explizit erlaubt ist, also explizit genau angesprochen, das KI-Thema im Vertrag, dann ist es sowieso verboten. Und alle anderen sagen aber, und auch der VDS, dass das nicht stimmt, sondern es muss explizit angesprochen sein, dass es nicht erlaubt ist, damit es verboten ist. Das ist ganz komisch, die widersprechen sich so.

Wie gesagt, der Synchron ist im Moment immer noch ein völlig unregelter Raum. Wenn man sich da auch nicht einig ist und jeder was anderes sagt, dann ist es noch schwieriger, was festzumachen oder was einzuführen.

Dann kämpfen teilweise schon, irgendwie gibt es Anhänger von dem Verband und dann treten die aus dem aus und was weiß ich.

Und ich merke auch, dass ich auch nicht so richtig Lust habe, mich dauernd damit zu beschäftigen. Ich will arbeiten. Mich nervt es auch.

Ich gucke mir auch von allem, was rumgeschickt wird, höchstens jedes fünfte Video an. Man kann sich darauf fixieren und das gefällt mir auch nicht.

Ja, das wird so zum Mittelpunkt der Welt und dann gibt es immer mehr und irgendwie verliert sich so der Blick dafür, wie wichtig das im Gesamtleben wirklich ist und für andere Menschen auch. Ich weiß nicht, ich habe das Gefühl, dass man sich da sehr leicht verschätzen kann und verrennen kann.

Ja, und es ist auch nicht richtig, die Ängste zu nähren, finde ich. Wenn einem was einfällt, kreativ, was man dagegensetzen kann, dann ist gut, aber dieser Hype von, hast du schon gesehen, oh Gott, das ist ja noch besser geworden, das finde ich blöd.

Aber dieses Video, was jetzt da überall rumgeht, was die prominenten Stimmen gemacht haben, finde ich eine gute Idee.

Das haben sie auch sympathisch gemacht. Ich kann dir das gleich mal schicken.

Bei mir ist ja auch vieles so sehr gefühlsmäßig, weil ich mich über manche Sachen auch nicht so konkret informieren will.

Also die Zeit, die ich damit verbringe, die Zeit vergeht eh so wahnsinnig schnell.

Und ich meine, das ist ja auch interessant, die haben ja auch diesen Pumuckl gemacht. Also das ist, finde ich, auch jetzt eher mal eine positive Geschichte zur KI-Stimme.

Das finde ich einfach eine gute Idee. Und das ist ja wohl auch daraus entstanden, dass der Redakteur so ein Pumuckl-Fan ist und einfach sich gewünscht hat, dass der Pumuckl klingt wie immer. Und das kann man ja auch wirklich machen.

Ja, das finde ich cool. Das ist wieder irgendwie eine coole Sache. Es gibt ja schon auch so ein paar neue kreative Sachen, die daraus entstehen, aber die Frage ist, zu welchem Preis.

Was fällt dafür weg, ja.