

Clara Kindler-Mathôt
Didem Leblebici
Giacomo Marinsalta
Till Rückwart
Anna Zaglyadnova (Hg.)

un/reale

Interaktionsräume

*Formen sozialer Ordnung im Spektrum
medienspezifischer Interaktion*

[transcript] Edition Medienwissenschaft

Clara Kindler-Mathôt, Didem Lelebici, Giacomo Marinsalta, Till Rückwart,
Anna Zaglyadnova (Hg.)
Un/Reale Interaktionsräume

Die E-Book-Ausgabe erscheint im Rahmen der »Open Library Medienwissenschaft 2024« im Open Access.

Die Open Library Community Medienwissenschaft 2024 ist ein Netzwerk wissenschaftlicher Bibliotheken zur Förderung von Open Access in den Sozial- und Geisteswissenschaften:

Vollspensoren: Technische Universität Berlin / Universitätsbibliothek | Universitätsbibliothek der Humboldt-Universität zu Berlin | Staatsbibliothek zu Berlin - Preußischer Kulturbesitz | Universitätsbibliothek Bielefeld | Universitätsbibliothek Bochum | Universitäts- und Landesbibliothek Bonn | Technische Universität Braunschweig / Universitätsbibliothek | Universitätsbibliothek Chemnitz | Universitäts- und Landesbibliothek Darmstadt | Sächsische Landesbibliothek, Staats- und Universitätsbibliothek Dresden (SLUB Dresden) | Universitätsbibliothek Duisburg-Essen | Universitäts- und Landesbibliothek Düsseldorf | Goethe-Universität Frankfurt am Main / Universitätsbibliothek | Universitätsbibliothek Freiberg | Albert-Ludwigs-Universität Freiburg / Universitätsbibliothek | Niedersächsische Staats- und Universitätsbibliothek Göttingen | Universitätsbibliothek der FernUniversität in Hagen | Staats- und Universitätsbibliothek Hamburg | Gottfried Wilhelm Leibniz Bibliothek - Niedersächsische Landesbibliothek | Technische Informationsbibliothek (TIB) Hannover | Universitätsbibliothek Kassel | Universität zu Köln, Universitäts- und Stadtbibliothek | Universitätsbibliothek Leipzig | Universitätsbibliothek Mainz | Universitätsbibliothek Mannheim | Universitätsbibliothek Marburg | Ludwig-Maximilians-Universität München / Universitätsbibliothek | FH Münster | Universitäts- und Landesbibliothek Münster | Bibliotheks- und Informationssystem (BIS) der Carl von Ossietzky Universität Oldenburg | Universitätsbibliothek Passau | Universitätsbibliothek Siegen | Universitätsbibliothek Vechta | Universitätsbibliothek der Bauhaus-Universität Weimar | Zentralbibliothek Zürich | Zürcher Hochschule der Künste

Sponsoring Light: Universität der Künste Berlin, Universitätsbibliothek | Freie Universität Berlin | Bibliothek der Hochschule Bielefeld | Hochschule für Bildende Künste Braunschweig | Hochschule für Technik und Wirtschaft Dresden - Bibliothek | Hochschule Hannover - Bibliothek | Hochschule für Technik, Wirtschaft und Kultur Leipzig | Hochschule Mittweida, Hochschulbibliothek | Landesbibliothek Oldenburg | Akademie der bildenden Künste Wien, Universitätsbibliothek | Jade Hochschule Wilhelmshaven/Oldenburg/Elsfleth | ZHAW Züricher Hochschule für Angewandte Wissenschaften, Hochschulbibliothek | Westsächsische Hochschule Zwickau | Hochschule Zittau/Görlitz, Hochschulbibliothek

Mikrosponsoring: Ostbayerische Technische Hochschule Amberg-Weiden | Deutsches Zentrum für Integrations- und Migrationsforschung (DeZIM) e.V. | Technische Universität Dortmund | Evangelische Hochschule Dresden | Hochschule für Bildende Künste Dresden | Hochschule für Musik Carl Maria Weber Dresden, Bibliothek | Palucca Hochschule für Tanz Dresden – Bibliothek | Filmmuseum Düsseldorf | Universitätsbibliothek Eichstätt-Ingolstadt | Bibliothek der Pädagogischen Hochschule Freiburg | Berufsakademie Sachsen | Bibliothek der Hochschule für Musik und Theater Hamburg | Hochschule Hamm-Lippstadt | Hochschule Fresenius | ZKM Zentrum für Kunst und Medien Karlsruhe | Hochschule für Grafik und Buchkunst Leipzig | Hochschule für Musik und Theater „Felix Mendelssohn Bartholdy“ Leipzig, Bibliothek | Filmuniversität Babelsberg Konrad Wolf - Universitätsbibliothek | Universitätsbibliothek Regensburg | Bibliothek der Hochschule Rhein-Waal | FHWS Hochschule Würzburg-Schweinfurt

Clara Kindler-Mathôt, Didem Leblebici, Giacomo Marinsalta, Till Rückwart,
Anna Zaglyadnova (Hg.)

Un/Reale Interaktionsräume

Formen sozialer Ordnung im Spektrum
medienspezifischer Interaktion

[transcript]

Bibliografische Information der Deutschen Nationalbibliothek

Die Deutsche Nationalbibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliografie; detaillierte bibliografische Daten sind im Internet über <https://dnb.dnb.de/> abrufbar.



Dieses Werk ist lizenziert unter der Creative Commons Attribution 4.0 Lizenz (BY). Diese Lizenz erlaubt unter Voraussetzung der Namensnennung des Urhebers die Bearbeitung, Vervielfältigung und Verbreitung des Materials in jedem Format oder Medium für beliebige Zwecke, auch kommerziell.

<https://creativecommons.org/licenses/by/4.0/>

Die Bedingungen der Creative-Commons-Lizenz gelten nur für Originalmaterial. Die Wiederverwendung von Material aus anderen Quellen (gekennzeichnet mit Quellenangabe) wie z.B. Schaubilder, Abbildungen, Fotos und Textauszüge erfordert ggf. weitere Nutzungsgenehmigungen durch den jeweiligen Rechteinhaber.

Erschienen 2024 im transcript Verlag, Bielefeld

© Clara Kindler-Mathôt, Didem Lelebici, Giacomo Marinsalta, Till Rückwart, Anna Zaglyadnova (Hg.)

transcript Verlag | Hermannstraße 26 | D-33602 Bielefeld | live@transcript-verlag.de

Umschlaggestaltung: Giacomo Marinsalta

Umschlagabbildung: Giacomo Nanni

Lektorat und Korrektorat: Michelle Abdul-Malak, Clara Kindler-Mathôt, Didem Lelebici, Giacomo Marinsalta, Till Rückwart, Severyna Yakubych, Anna Zaglyadnova

Druck: Majuskel Medienproduktion GmbH, Wetzlar

<https://doi.org/10.14361/9783839471463>

Print-ISBN: 978-3-8376-7146-9

PDF-ISBN: 978-3-8394-7146-3

EPUB-ISBN: 978-3-7328-7146-9

Buchreihen-ISSN: 2569-2240

Buchreihen-eISSN: 2702-8984

Gedruckt auf alterungsbeständigem Papier mit chlorfrei gebleichtem Zellstoff.

Inhalt

Danksagung	7
-------------------------	---

Un/Reale Interaktionsräume

Formen sozialer Ordnung im Spektrum medienpezifischer Interaktion

<i>Didem Leblebici, Till Rückwart, Clara Kindler-Mathôt</i>	9
---	---

Sektion I: Konstruktion

Orbital Mirage

<i>Till Rückwart</i>	25
----------------------------	----

»I am, in fact, a person.«

Vorder- und Hinterbühnen konversationeller KI

<i>Timo Kaerlein</i>	35
----------------------------	----

Sprachassistenzsysteme und ihre Interfaces

Eine medienlinguistische Analyse

<i>Tim Hector</i>	55
-------------------------	----

Zwischen Technologie und Ideologie

Ein Blick auf die un/realen Räume hinter KI

<i>Paul Schütze</i>	81
---------------------------	----

Sektion II: Erfahrung

L.I.S.A.: Akt 1

<i>Lynn Klemmer</i>	103
---------------------------	-----

Von der Non-Linearität der Zeit

Metaphorisieren als Interaktion mit filmischen Bildern. Eine medienästhetische Perspektive

Katerina Papadopoulou 109

Erfahrungsraum Interaktion

Die audiovisuelle Modulation eines Gesprächs als Zuschauer*innenerfahrung

Clara Kindler-Mathôt 137

Ludische Dividuationen

Eine antiholistische Betrachtung von CULTIST SIMULATOR und DEAD SPACE

Dominic Brakelmann, Arvid Kammler 165

Sektion III: Partizipation

Excuse Us While We Improve Your View, Atlantis

Giacomo Marinsalta 193

Mobile Crowdsensing

Interaktionsmodelle des mobilen Sensing und die infrastrukturelle Allgegenwart der ›mobile sensor networks‹

Vesna Schierbaum 201

Digital Fashion im Online-Shop

Der un reale Interaktionsraum als sozialer Handlungsraum

Helga Behrmann 221

Appropriating the Ape

Sprachliche Praktiken der Gemeinschaftsbildung im Subreddit R/WALLSTREETBETS

Friederike Fischer 241

Die Autor*innen 265

Danksagung

Wir Herausgeber*innen bedanken uns herzlich bei Daniel Bonanati, Katharina Wierichs und dem transcript Verlag für die vielseitige Unterstützung und bei der Open Library Medienwissenschaft für die Förderung dieses Sammelbands. Ein besonderer Dank gilt dem ZeM – Brandenburgisches Zentrum für Medienwissenschaften – für die Förderung der Fachtagung *un:reale Interaktionsräume* sowie der Fachhochschule Potsdam für die Bereitstellung ihrer Räumlichkeiten zur Ausrichtung der Tagung – insbesondere Fritz Schlüter für seine wertvolle inhaltliche und organisatorische Unterstützung. Ohne die Tagung hätten wir Herausgeber*innen nur schwer zusammengefunden und wäre dieser Sammelband nicht möglich gewesen.

Unser Dank gilt auch Jan Distelmeyer, Maximilian Gregor Hepach, Daniel Ilger, Maximilian Krug, Kateryna Pilyarchuk, Fritz Schlüter, Christina Schmitt, Britta Schneider, Rita Vallentin und Alexander Walmsley, die mit ihren fachlichen und redaktionellen Hinweisen maßgeblich zur Qualität des Sammelbands beigetragen haben. Wir möchten herzlich Severyna Yakubych und Michelle Abdul-Malak für die hervorragende Unterstützung beim Korrektorat jedes noch so kleinen Satzzeichens danken, und Giacomo Nanni für die Entwicklung einer spielerischen Webanwendung zur Gestaltung des Buchcovers.

Nicht zuletzt möchten wir allen Autor*innen dieses Sammelbands für ihre vielseitigen Forschungsbeiträge danken. Ihr macht dieses Projekt zu dem, was es ist.

Un/Reale Interaktionsräume

Formen sozialer Ordnung im Spektrum medienspezifischer Interaktion

Didem Leblebici, Till Rückwart, Clara Kindler-Mathôt

1. Einleitung

Als wir im Frühjahr 2022 zu einer Tagung aufgerufen haben, lag ein für Europa und die Welt unvorstellbares Ereignis keine zwei Monate zurück. Die russische Invasion der Ukraine vertrieb Millionen Menschen aus ihrer Heimat und hat überwunden geglaubte Gräben zwischen West und Ost erneut aufgerissen, die noch immer zu weitreichenden sozialen, ökonomischen und politischen Veränderungen auf internationaler und nationaler Ebene führen. Im Bundestag richtete sich Bundeskanzler Olaf Scholz mit der Zeitenwende-Rede an die Bevölkerung und sprach angesichts von Waffenlieferungen an die Ukraine von einer »neue[n] Realität« (Scholz 2022). Er plädierte für eine Stärkung der Resilienz, sowohl technisch als auch gesellschaftlich, um sich besser gegen Angriffe auf kritische Infrastrukturen und Kommunikationswege durch bspw. Cyberattacken und Desinformation zu wappnen.

Parallel wurden in Deutschland nach dem zweiten Corona-Winter in Isolation Lockerungen der Eindämmungsmaßnahmen zum Schutz vor dem Virus angekündigt (vgl. Chronik des Bundesministerium für Gesundheit 2023). Die Vorjahre waren von einer Serie von bundesweiten (Anti-)Corona-Protesten und Demonstrationen gegen eben diese Maßnahmen geprägt, bei denen auch Verschwörungsmythen, Falschinformationen und Hassreden zirkulierten und rechtsextreme Sichtweisen gesellschaftlich verfestigt werden konnten. Viele dieser Diskurse, darunter die Dichotomisierung und Anfeindung von Befürworter*innen und Gegner*innen, wurden (und werden) auf den Plattformen und Messengerdiensten der Sozialen Medien kanalisiert. An den Hochschulen und Universitäten verbreitete sich mit der Rückkehr zum Alltag auch die Hoffnung auf eine Rückkehr zu einem »Normalzustand«: Seminare in Präsenz, keine schwarzen Zoom-Kacheln mehr, sondern »echter« Austausch mit den Kolleg*innen und Studierenden, Interaktion, die nicht nur zweckgebundene Meetings oder Lehrveranstaltungen sind, sondern sozialer Austausch und Miteinander.

Die Sorge um militärische Cyberkriminalität, die Radikalisierung in Sozialen Medien und der Wunsch nach ›echtem‹ Miteinander abseits von Videokonferenzen sind nur wenige Beispiele für die lokalen Auswirkungen globaler medialer Phänomene und Krisen, welche zu kollektiv-gesellschaftlichen Spannungen geführt und uns Herausgeber*innen zum Nachdenken über Un/Realitäten angeregt haben. Diese manifestieren sich an der Schnittstelle zwischen ›realen‹ und digitalen Welten, Wahrheiten und sogenannten alternativen Fakten. Sie oszillieren zwischen Interfaces, Cloud-Technologien, Datengewinnung und -verbreitung. Sie gehören sowohl zu unserer Realität als auch zu einer Art von Irrealität oder Unrealität, die sich der Wahrnehmung entzieht und dennoch spürbare Auswirkungen hat. Gemeinsam haben sie alle, dass sie sich im Kontext medienspezifischer Räume und Interaktionen abspielen, welche Diskurse nicht nur befeuern, sondern oft erst ermöglichen. In einer nie dagewesenen Geschwindigkeit und Dichte sind Informationen, Bilder und Diskussionen zugänglich, beinahe omnipräsent. Eingebettet ins Staccato der Feeds zwischen Werbung, Urlaubsschnappschüssen und Bildern vom Mittagessen treffen wir auf Nachrichten und Videos, die unmittelbar aus den Konfliktregionen zu stammen scheinen. Informationen werden nicht mehr zwangsläufig gesucht, sie ›begegnen‹ einem im Strom des Algorithmus. Als Diagramme visualisiert, verändern Corona-Fallzahlen, Fluchtbewegung und Frontverläufe das Verständnis von Raum und Zeit, und wie politische und gesellschaftliche Debatten geführt werden.

Was 2022 bereits relevant war, hat seit der Verbreitung von generativer Künstlicher Intelligenz (KI) in keiner Weise an Aktualität eingebüßt. Nur wenige Wochen nach unserer Tagung zu *un:realen Interaktionsräumen* veröffentlichte das Unternehmen OpenAI die Anwendung ChatGPT. Dies löste innerhalb kürzester Zeit eine weitere Krise der Forschungs- und Hochschullandschaft aus: Wie werden zukünftig die Produktion von und der Zugang zu Wissen organisiert? In welcher Weise verändern die Verbreitung und die alltägliche Anwendung von generativer KI Debatten um Deep Fakes, Hassreden und Diskriminierung? Bis zur Veröffentlichung dieses Sammelbandes können nur teilweise Antworten auf diese Fragen gegeben werden.

Lang erarbeitete Rechte im Datenschutz drohen unter den neuen Bedingungen ebenso in Gefahr zu geraten wie beispielsweise das Recht auf Zustimmung zur Nutzung eigener Inhalte für das Training von KI-Modellen oder das Recht, nicht von einer Software imitiert zu werden – sei es das Gesicht, die Stimme oder der individuelle künstlerische Stil.¹ Die Macht einzelner Unternehmen über die Produktion,

1 Beispielhaft können hier eine Reihe publik gewordener Rechtsstreite genannt werden, wie bspw. Scarlett Johansson, Taylor Swift, Sarah Andersen, Kelly McKernan und Karla Ortiz vs. Stability AI, Midjourney und DeviantArt.

Filterung und Zirkulation von KI-generierten Inhalten akkumuliert und dies offenbar mit einschneidenden ökologischen und sozialen Implikationen².

Technische Aspekte der Informationsverarbeitung haben direkte Auswirkungen auf die sozialen Interaktionen, die sie ermöglichen sollen. Bender et al. (vgl. 2021) stellen in ihrem Artikel über die Gefahr zu großer KI-Sprachmodelle fest, dass die Entwicklungen in einem vielfach schnelleren Tempo als dem des Mooreschen Gesetzes verlaufen: Wenn die Art und Weise, wie Menschen und nicht-menschliche Akteur*innen miteinander interagieren, von den Kapazitäten von Mikrochips abhängt, deren Entwicklung zum ersten Mal seit 50 Jahren überproportionale Anomalien aufweisen, stellen sich Fragen danach, welche Auswirkungen das auf die Interaktionen selbst hat, wie sich bestehende Formen und Ordnungen ändern, welche neuen Formen und damit einhergehenden sozialen Ordnungen sich bilden, sowie welche Rolle Bilder des Selbst und von Agency, Raum und Zeitlichkeit darin spielen.

Dieser Sammelband widmet sich daher Fragen der Entstehung, Veränderung und Gestaltung von Interaktionsräumen im Spektrum zwischen Mensch-Mensch, Mensch-Maschine und Maschine-Maschine, in denen Akteur*innen miteinander *durch* oder *mit* Technologien kommunizieren. Im Folgenden legen wir unsere theoretischen Grundannahmen dar und führen zentrale Begriffe ein, die im Titel des Bandes und in den Beiträgen genannt werden. Die Einleitung erhebt dabei keinen Anspruch auf Vollständigkeit, sondern will vielmehr die verschiedenen Forschungskontexte aufzeigen und Denkrichtungen anstoßen, die in den einzelnen Beiträgen umfänglicher ausgeführt und diskutiert werden.

2. Formen medienspezifischer Interaktionsräume

Als soziale Wesen sind wir in unserem alltäglichen Leben von Interaktionen maßgeblich geprägt. In der *fokussierten Interaktion* (vgl. Goffman 1971), also in Face-to-Face-Begegnungen, modellieren Menschen interaktive Konstellationen etwa durch rituelle Rahmungen, wie z.B. Gruß- und Abschiedssequenzen, um Beziehungen aufzunehmen, sich als Selbst zu definieren und ihr Gesicht (*face*; Goffman 1971; 1974) zu wahren. Im mehr oder weniger stummen Einverständnis verhandeln und etablieren Personen so in jeder Interaktion aufs Neue *soziale Ordnung*. Daher kann es nicht überraschen, wenn sich veränderte Realisierungsbedingungen der vermeintlichen »Maximalform« (Marx/Schmidt 2019: 6) Face-to-Face-Kommunikation entscheidend auf die interaktiven Möglichkeiten selbst auswirken, denn:

2 Vgl. bspw. Berichte über verheerende Arbeitsbedingungen kenianischer Arbeiter*innen, die an der Erkennung toxischer Inhalte von ChatGPT arbeiten (z.B.: Perrigo, Billy (2023): Exclusive: OpenAI Used Kenyan Workers on Less Than \$2 Per Hour to Make ChatGPT Less Toxic, abrufbar unter: <https://time.com/6247678/openai-chatgpt-kenya-workers/> (Stand: 24.07.2024)).

Die Kernaspekte von Interaktion, nämlich wechselseitiger Kontakt, raum-zeitliche Situierung und dinglich-körperliche bzw. kommunikativ-semiotische Möglichkeiten unterliegen stets den medialen Bedingungen, die diese Interaktionsräume konstituieren (vgl. ebd.: 9). Für die Beschreibung von Interaktionsräumen, die über die Face-to-Face-Kommunikation hinausgehen, müssen deshalb ihre medien-spezifischen Formen und Begebenheiten mitgedacht werden, um die Konstitution sozialer Ordnung auch unter diesen veränderten Bedingungen aufdecken und verstehen zu können. Zur Beschreibung und Analyse dieser Formen eignet sich insbesondere der Begriff der *Medialität*, da das dahinterliegende Konzept nicht nur die Bedingungen, sondern auch deren Potenziale und Beschränkungen im Sinne medialer Möglichkeiten (Affordanzen) umfasst. Im Gegensatz zu einem objekt-haften Medienbegriff stellt der Begriff der Medialität die Prozesshaftigkeit sowie performative Aspekte von Medien in den Vordergrund (vgl. Schneider 2017; 2018). In diesem Zusammenhang werden Medien und ihre Materialität als wechselseitig mit Prozessen der Sinnkonstitution und Kommunikation betrachtet (vgl. Luginbühl 2019; Marx/Schmidt 2019) und beziehen deshalb auch damit einhergehende Formen des Rezipierens und Verstehens mit ein.

Mit Blick auf Interaktionen zeigt sich in diesem Kontext meist eine hohe Komplexität von Verflechtungen. Damit sind eben diese vielschichtigen *medialen Ökologien* angesprochen, in denen die Interaktionen stattfinden, beispielsweise Social-Media-Plattformen wie YouTube oder TikTok, aber auch Algorithmen und technische Infrastrukturen, die überhaupt die Voraussetzung für das Zustandekommen vieler medien-spezifischer Interaktionen darstellen. Beispielhaft lässt sich die Komplexität medialer Ökologien anhand von Interfaces aufzeigen: Spezifische Interaktionsräume entstehen z.B. durch bestimmte visuelle und auditive Bedienoberflächen von Computern oder Smartphones. Während die Hardware als Schnittstelle für In- und Output dient, vollziehen sich unterhalb der Oberfläche vielfältige Verbindungsprozesse zwischen wechselseitig miteinander verbundenen Interface-Ebenen (vgl. Distelmeyer 2021: 57). Cramer und Fuller (2008) unterscheiden diesbezüglich Interface-Konstellationen zwischen Software und Software, Software und Hardware, Hardware und Hardware, Hardware und User*innen sowie Software und User*innen, die es allesamt für eine Interaktion mit und durch digitale Technologien braucht. So benötigt es eine komplexe Infrastruktur von u.a. Unterseekabeln und zahllosen Funknetzen, durch Software produzierte und auf Hardware bereitgestellte Anwendungen, APIs (Application Programming Interfaces), bis zu (Touch-)Bildschirmen, Maus, Tastatur, Mikrofonen und weiteren Sensortechnologien für die Bedienung von und Vermittlung zwischen Computer- und Nicht-Computer-Welt. Vor diesem Hintergrund spricht Distelmeyer (vgl. 2021) von einem *Interface-Komplex*, um auf die Komplexität von Interface-Prozessen aufmerksam zu machen (ebd.: 57).

Die technologischen, infrastrukturellen und algorithmischen Aspekte der Interfaces stellen sich laufend verändernde mediale Ökologien bereit, welche Interaktionen auch in Bezug auf die Konstruktion von Wirklichkeitsvorstellungen bzw. die Wahrnehmung von Realität beeinflussen. Dabei ist der Mensch nicht nur mit der Bedienoberfläche im Kontakt, sondern steht zugleich mit der komplexen, letztlich jedoch nur erahnbaren Anatomie im/materieller, ökologischer und sozioökonomischer Netze in Verbindung. Interaktionsräume bestehen aber auch da, wo augenscheinlich keine User-Interfaces zu einer Interaktion einladen, sondern verteilte Sensoren die Außenwelt und ihre menschlichen als auch nicht-menschlichen Entitäten registrieren und vermessen (vgl. Schierbaum in diesem Band). Wer oder was Teil einer Messung, Berechnung und Weiterverarbeitung oder schlicht von Interface-Operationen wird, unterliegt dabei nicht mehr der menschlichen Kontrolle. Durch ubiquitäre, vernetzte Technologien stellt sich somit auch die Frage nach der Agency der diversen beteiligten Akteur*innen. Interfaces verdeutlichen damit im Besonderen, dass in medienpezifischen Interaktionen keine klare Trennung zwischen sicht- und unsichtbaren, realen und unrealen Teilnehmer*innen mehr möglich ist. Vielmehr zeigt sich hier deutlich ein Spektrum auf, in dem manche Interaktionen menschlicher und/oder nicht-menschlicher Entitäten (scheinbar) näher der »Maximalform« sind als andere.

Technologien mit ihren unterschiedlichen Affordanzen erlauben Nutzer*innen nicht nur die Ausführung bestimmter Aktionen und Praktiken, sondern damit einhergehend auch die Ausgestaltung von Identitäten, Werten und Beziehungen zu anderen (vgl. Jones 2020: 204). Die medialen Bedingungen können neue Sichtweisen eröffnen und soziale, politische und ökonomische Diskurse tiefgehend beeinflussen. Das Spektrum medienpezifischer Interaktion konstituiert sich durch Aspekte medialer Umwelten an den Schnittstellen digital/analog, online/offline und real/unreal. Diese werden für spezifische Interaktionsformen relevant, entfalten dort ihre jeweiligen Potenziale und tragen somit z.B. zur Entstehung wirkungsstarker Protestbewegungen bei (vgl. Fischer in diesem Band) oder können für die politischen Bildung nutzbar gemacht werden (vgl. Kindler-Mathôt in diesem Band).

Der Fokus dieses Bandes liegt deshalb nicht auf den Effekten einzelner Medien, sondern auf ihren medialen Interaktionspotenzialen in Bezug auf gesellschaftliche Prozesse.³ In diesem Kontext ist die Frage nach dem Zusammenhang von Medien-, Kommunikations- und Kulturwandel von Interesse (vgl. Androutsopoulos 2016: 343). Löffler und Sprenger sehen in der Betrachtung spezifischer medialer Ökologien eine Möglichkeit zur »Diagnose der Gegenwart« (Löffler/Sprenger 2016: 11). Me-

3 Die Begriffe »mediale Möglichkeiten«, »Medialität« und »mediale Ökologien« lassen sich nicht klar voneinander abgrenzen, sondern stellen vielmehr verschiedene Aspekte in Bezug auf Medien und Interaktion in den Vordergrund, z.B. interaktive Bedingungen, Interferenzen und Prozesse oder weitreichende Verknüpfungen und Netzwerke.

diale Bedingungen werden dabei sowohl als Erweiterung wie auch als Einschränkung begriffen, da Medientechnologien kommunikative Praktiken auf verschiedenen Ebenen institutionell regulieren (vgl. Couldry/Hepp 2017: 32). Die Beiträge zielen darauf ab, Fragen der Konstruktion, Veränderung und Gestaltung sowie der Partizipation *in* und Erfahrung *mit* Interaktionsräumen zu untersuchen. Dabei stehen deren Implikationen für gesellschaftliche (soziale, politische, ökonomische) Wechselwirkungen im Mittelpunkt. Die Komplexität dieser einander bedingenden, oft nur schwer zu vermittelnden soziotechnischen Prozesse verstehen wir daher als Un/Realitäten.

3. Un/Realitäten

In der philosophischen Diskussion um den Begriff der »Realität« lassen sich unterschiedliche Positionen und Diskurse ausmachen.⁴ In diesem Sammelband geht es uns nicht darum, diese Debatten weiterzuentwickeln oder den Begriff epistemologisch zu untersuchen. Stattdessen beziehen wir uns, mit Blick auf Formen sozialer Ordnung, auf *Realität als soziale Konstruktion* (vgl. Berger/Luckmann 1967). Damit diskutieren wir nicht das Vorhandensein einer realen, materiellen Wirklichkeit, sondern postulieren, dass Realitäten durch sprachliche, soziale und materielle Prozesse geprägt sind (vgl. Burr 2015) und durch Interaktionen von Akteur*innen maßgeblich hervorgebracht, gestaltet, etabliert und verändert werden.

Insbesondere mit Blick auf die Rolle von Medien bei der Modulation von Realitäten kann dieser Ansatz produktiv gemacht werden. Bereits frühere Medientechnologien wie Schrift und Drucktechniken haben Interaktionsmöglichkeiten mitkonstituiert.⁵ Auch das Ausmaß der Integration digitaler Medien hat heute zu einem noch nie dagewesenen Zustand der »tiefen Mediatisierung« (*deep mediatization*; vgl. Couldry/Hepp 2017) geführt: Hepp und Couldry argumentieren, dass sogar die grundlegende Form der Face-to-Face-Interaktion der Mediatisierung unterliegt, da Realitäten stets auf der Grundlage und im Kontext der gegenwärtigen Bedingungen koordiniert werden müssen – bspw. anhand von etablierten Interaktionsformen in den Medien (vgl. Luginbühl 2019), mit Medien als selbstverständlichen Bestandteil der Interaktion, etwa mit Smartphones als sozialen Objekten (vgl. Oloff 2019) oder auch im Umgang mit Smart Speakern (vgl. Hector

4 Vgl. Philosoph*innen im Realismus-Diskurs, z.B. Willascheck (2022) zu Kant, Gabriel (2014) sowie die gesammelten Diskussionen im Sammelband zur Realismusdebatte in der Philosophie bei Halbig und Suhm (2004).

5 Vgl. Habermas (2021: 486) Äußerungen, dass es sich bei digital vernetzter Kommunikation nicht nur um eine Erweiterung des Medienangebots handle, »sondern vielmehr um eine mit der Einführung des Buchdrucks vergleichbare Zäsur in der menschheitsgeschichtlichen Entwicklung der Medien.«

in diesem Band). Ein unbeabsichtigter Laut, der der Aussage »Alexa« ähnelt, kann jederzeit die Aktivierung des Amazon Smart Speakers auslösen. Auf diese Weise werden Menschen sensibilisiert, ihre Sprechweise, aber auch ihr körperliches Verhalten an die interaktiven Bedingungen ihres jeweiligen technologischen Umfeldes anzupassen.

Die aktuellen Diskurse zeigen, dass das Verständnis und die Erfahrung von Realität als stets verschwommen und medial vermittelt empfunden wird. Insbesondere die zunehmende Popularität neu eingeführter Technologien wie ChatGPT oder Suno hat die Debatte darüber, was »menschlich« oder »echt« ist (oder zu sein hat, bzw. nicht sein sollte), neu entfacht. Generative KI-Tools mit ihren zunehmend glaubwürdiger werdenden, menschenähnlichen »Reaktionen« modellieren Alltagserfahrungen, die teilweise kaum mehr von menschlichen Kreationen und Interaktionsweisen zu unterscheiden sind. Die Unterscheidung zwischen »echter« menschlicher und »unechter« maschineller Intelligenz besteht bereits seit dem Dartmouth-Workshop und dem Turing-Test und damit deutlich vor dem Populärwerden der genannten generativen KI-Tools (vgl. Coeckelbergh/Gunkel 2023: 4). Bei ihrer Untersuchung aktueller Debatten über ChatGPT beobachten Coeckelbergh und Gunkel drei philosophische Positionen innerhalb des Un/Realitätsdiskurses: Solche, die sie nahe an einer »echten« menschlichen Intelligenz verorten, solche, die sie nur als Schein betrachten, und solche, die diese Entwicklungen völlig außer Acht lassen. In Anlehnung an Barad (2007) und Haraway (2016) stellen Coeckelbergh und Gunkel diese ontologische Differenz in Frage und schlagen vor, sich auf Performances, Prozesse der Bedeutungsgebung und Machtverhältnisse zu konzentrieren, durch die Un/Realitäten verhandelt werden.

Un/Realitäten lassen sich demnach weder mit internen mentalen Prozessen (von Menschen oder KI) erklären noch als unabhängig von ihrer Wirkung auf die Außenwelt betrachten. Vielmehr werden sie in sozio-materiellen Prozessen konstruiert, die zunehmend den Logiken und Infrastrukturen von Social-Media-Plattformen, großen Sprachmodellen (Large Language Models) und Algorithmen folgen, und die somit an Prozessen der Sinnstiftung beteiligt sind (vgl. Couldry/Hepp 2017). Die Fragestellung, der wir uns widmen, entspricht also nicht der philosophischen Frage: »Ist das real oder nicht?« Vielmehr stellen wir eine sozio-materielle, soziokulturelle und prozesshafte Frage, die wie folgt formuliert werden kann: »Welchen Beitrag leisten soziokulturelle und materiell-mediale Praktiken zur Konstruktion von Realitäten?«. Um diese Frage zu beantworten, untersuchen die Beiträge in diesem Band die Konstruktion und Erfahrung von sowie Partizipation in un/realen Interaktionsräumen.

4. Formen sozialer Ordnung im Spektrum medienspezifischer Interaktion

Die Beiträge des Sammelbandes betrachten, analysieren und diskutieren aus medienwissenschaftlichen und -linguistischen Perspektiven die komplexen medienspezifischen Formen von Interaktion und damit einhergehende gesellschaftliche Wirkungen und Prozesse. Sie erlauben Einblicke in aktuelle und gesellschaftlich hoch relevante Forschungsfelder. Trotz der Vielfalt der Untersuchungsgegenstände (Interfaces, Soziale Medien, Sensornetzwerke, Sprachassistenten und Audiovisualität) tauchen dabei in allen Beiträgen Fragen nach den medialen Bedingungen auf, die Interaktionsräume bereitstellen bzw. welche spezifischen interaktiven Möglichkeiten, Realitäten und sozialen Ordnungen sich daraus ergeben. Wie entfalten und gestalten sich Interaktionsräume in hybriden, digitalen und auch analogen Kontexten in einem Spektrum von sozialen Medien, Computerspielen, Sprachsteuerung etc.? Welche Dispositive kommen in un/realen Interaktionsräumen nicht zum Vorschein? Wie verschwimmen sozio-kulturelle Kategorien wie das Bild des Selbst und Agency, und dies über Grenzen zwischen Mensch und Maschine hinweg? In welchem Zusammenspiel stehen Körper, Medien und Technologien? Wie werden unterschiedliche Medien affektiv erfahren? Welche Realitäten können sich aus den medienspezifischen Praktiken entfalten, die damit einhergehen? Welche neuen sozialen Ordnungen und Strukturen werden sichtbar?

Der Sammelband ist in drei Sektionen gegliedert, die jeweils unterschiedliche Aspekte der Interaktionsräume untersuchen. Jede Sektion beginnt mit einem künstlerischen Beitrag, der die theoretischen und analytischen Diskussionen einleitet und veranschaulicht. Die erste Sektion »Konstruktion« thematisiert, wie Interfaces und mit ihnen auch ideologische Rahmenbedingungen die Interaktionsräume im Spektrum zwischen Mensch-Maschine und Maschine-Maschine prägen und so neue Un/Realitäten schaffen. Die zweite Sektion »Erfahrung« widmet sich der körperlichen und affektiven Dimension der Interaktion in verschiedenen audiovisuellen Medienformaten, einschließlich Film, Videospiel und politischer Berichterstattung. Die dritte Sektion »Partizipation« befasst sich mit den kollaborativen Prozessen und partizipativen Mechanismen, durch die Interaktionsräume geschaffen und gestaltet werden.

4.1 Konstruktion

Die Beiträge in der Sektion *Konstruktion* untersuchen unterschiedliche Aspekte der medialen und kulturellen Kokonstruktion von Interaktionsräume durch Interfaces, Infrastrukturen und Ideologien. Gemeinsam tragen sie dazu bei, ein tieferes Verständnis dafür zu entwickeln, wie technologische und ideologische Rahmenbedin-

gungen Interaktionsräume und Realitäten formen und welche Auswirkungen dies auf die Gesellschaft hat.

In seinem künstlerischen Beitrag *Orbital Mirage* präsentiert Till Rückwart ein Archiv digitaler Fehler bzw. Glitches als Un/Realitäten innerhalb der Satellitenbilddarstellung von Google Earth. Er zeigt, dass nicht fotooptische Effekte ausschlaggebend für die fehlerhaften Bilder sind, sondern dass sie prozessual durch den Algorithmus geschaffen werden. Die Arbeit plädiert für Archivpraktiken, die die vermeintlich objektive Darstellung der Erde durch Interfaces und deren technologische Rahmenbedingungen kritisch reflektiert.

Timo Kaerleins Beitrag »*I am, in fact, a person.*« *Vorder- und Hinterbühnen konversationeller KI* thematisiert die Frage, auf welche Weise einer KI der Status einer Person mit Intelligenz, Bewusstsein oder Empfindungsvermögen zugesprochen wird. Die Untersuchung eines 2022 in der Öffentlichkeit breit diskutierten Diskurses, konkret der Dialogsequenz zwischen dem damaligen Google-Ingenieur Blake Lemoine und dem Chatbot LaMDA, dient als Grundlage für die kritische Reflexion aktueller Debatten über KI als Entität mit menschlichen Qualitäten. Hierbei beleuchtet er insbesondere die Rolle der kulturellen Skripte sowie die Darstellung von un/realen Interaktionen.

In seinem Beitrag *Sprachassistenzsysteme und ihre Interfaces: Eine medienlinguistische Analyse* widmet sich Tim Hector den stimmlichen und grafischen Bedienoberflächen von Sprachassistenten wie beispielsweise Alexa, Siri und Google Assistant. Er analysiert die interaktiven situativen Praktiken der Nutzer*innen auf der einen Seite und die durch Algorithmen in den Apps dokumentierten Aspekte dieser Praktiken auf der anderen Seite. Die Analyse erfolgt auf der Grundlage von Videomaterial, welches konversationsanalytisch ausgewertet wird. Auf diese Weise entwickelt er eine Perspektive auf die Mensch-Maschine-Interaktion, die nicht nur die Praktiken der Nutzer*innen berücksichtigt, sondern auch die verborgenen technischen Infrastrukturen in das Forschungsdesign integriert.

Schließlich präsentiert Paul Schütze in seinem Beitrag *Zwischen Technologie und Ideologie: Ein Blick auf die un/realen Räume hinter KI* eine kritische Perspektive auf zeitgenössische Debatten aus technikphilosophischer Sicht. Im Zentrum stehen Narrative des »AI-Futurismus«, in denen KI-Technologien als unausweichlicher Bestandteil einer zukunfts- und fortschrittsorientierten Entwicklung dargestellt werden. Er verdeutlicht die gesellschaftlichen Auswirkungen dieser Ideologien: die Reproduktion von kolonialen und patriarchalen Strukturen in der KI-Industrie, die Auswirkungen auf den Planeten, die staatliche Förderung der KI-Industrie sowie die Sammlung bzw. Überwachung von User*innen-Daten.

4.2 Erfahrung

Die Sektion »Erfahrung« beleuchtet die vielfältigen Dimensionen, wie Interaktionsräume in verschiedenen audiovisuellen Medien erlebt, wahrgenommen und rezipiert werden. Die Beiträge dieser Sektion stellen dar, wie Interaktionen in/mit Bildern, Filmen, Videospielen und audiovisuell modulierten Gesprächen nicht nur kognitiv, sondern auch körperlich und affektiv erfahren werden. Gemeinsam bieten diese Beiträge Einblicke in die verkörperten und affektiven Aspekte der Erfahrung und Wahrnehmung von Interaktionsräumen in verschiedenen Medienformaten.

Die Videoinstallation *L.I.S.A.: Akt 1* von Lynn Klemmer lädt die Betrachter*innen zu einer Reflexion über die Künstlichkeit von Wahrnehmung zwischen Fiktion und Realität ein. Durch übereinander gelagerte Projektionen eines Vorhangs aus verschiedenen Blickwinkeln eröffnet sich ein Interaktionsraum, in dem zwischen Subjekten, Objekten und Welten Un/Realitäten gestiftet werden. Damit betont Lynn Klemmer die materiellen, phänomenologischen und fiktionalen Aspekte der Wahrnehmung durch Bildschirme und plädiert für eine Dekonstruktion der Dichotomie von Realität und Fiktion.

Ebenso bieten verschiedene Bildschirme die Möglichkeit, unterschiedliche Filmformate zu rezipieren, zu verstehen und mitzerleben. In ihrem Beitrag *Von der Non-Linearität der Zeit: Metaphorisieren als Interaktion mit filmischen Bildern. Eine medienästhetische Perspektive* widmet sich Katerina Papadopoulou der Filmerfahrung als Interaktion zwischen Filmen und Zuschauer*innen. Ausgehend von einer medienästhetischen Metaphernanalyse des Science-Fiction-Films *ARRIVAL* entwickelt sie die These, dass die filmische Erfahrung nicht auf einer statischen, fiktiven oder »unrealen« Narrationsebene stattfindet, sondern sich durch eine verkörperte Interaktion mit dem Filmbild entfaltet.

Zuschauer*innen interagieren nicht nur mit fiktiven Filmen, sondern auch mit als »real« gewerteten audiovisuellen Bildern. Clara Kindler-Mathôt analysiert in ihrem medienlinguistischen Artikel *Erfahrungsraum Interaktion: Die audiovisuelle Modulation eines Gesprächs als Zuschauer*innenerfahrung* einen Videobeitrag des digitalen Parteitags der Partei BÜNDNIS 90/DIE GRÜNEN. In ihrer medienästhetischen Analyse zeigt sie, dass die audiovisuelle Darstellung nicht einfach eine Repräsentation des vorfilmischen Geschehens ist, sondern vielmehr neue mediale Realitäten entstehen lässt. Sie postuliert somit die Notwendigkeit einer reflektierten Auseinandersetzung mit audiovisuellen Bedeutungskonstitutionen, um die Medienkompetenz zu stärken.

Dominic Brakelmann und Arvid Kammler wenden sich in ihrem Beitrag *Ludische Dividuationen: Eine antiholistische Betrachtung der Videospiele CULTIST SIMULATOR und DEAD SPACE* der Wahrnehmung von Videospielen als körperliche Erfahrungsräume zu. Anstelle einer Betrachtung von Videospielen, die sich auf die Spieler*innen oder das Spiel konzentriert, schlagen sie eine dividualistische Perspektive vor,

indem sie sich die interaktiven Details der Spiele und der körperlich-räumlichen Verteilung widmen. Anhand von zwei Beispielen entwickeln sie eine Sichtweise auf Videospiele als multimodale Assemblagen mit vielfältigen medialen Potentialen.

4.3 Partizipation

Die Sektion »Partizipation« befasst sich mit der Art und Weise, wie Interaktionsräume durch partizipative und kollaborative Prozesse geschaffen und gestaltet werden. Die Beiträge dieser Sektion untersuchen verschiedene Formen der Ko-Konstruktion: von der sensorischen Datenerhebung über den digitalen Modemarkt bis hin zu Online-Communities. Zusammen bieten diese Beiträge Einblicke in die Mechanismen der Ko-Konstruktion von Interaktionsräumen und zeigen, wie Partizipation und Kollaboration in verschiedenen medialen Kontexten gestaltet und erlebt werden.

In seinem dokumentarischen Video-Essay *Excuse Us While We Improve Your View, Atlantis* begleitet Giacomo Marinsalta die Rezipient*innen auf eine Reise zur Weltausstellung Expo 2020 Dubai, in der Vorstellungen von Innovation und Vergangenheit verschmelzen. In seinem Beitrag legt er dar, dass die Förderung des Gemeinschaftsgefühls für Expo-Teilnehmende maßgeblich durch digitale Erlebnisse erfolgt, wobei kulturelle Diskurse passiv konsumiert werden. Damit wirft er einen kritischen Blick auf den technologischen Determinismus, der hier als Grundlage für eine gemeinsame Zukunft postuliert wird.

In ihrem Beitrag *Mobile Crowdsensing: Interaktionsmodelle des mobilen Sensing und die infrastrukturelle Allgegenwart der »mobile sensor networks«* widmet sich Vesna Schierbaum der Technik des Crowdsourcings von Daten aus einer kritischen medienwissenschaftlichen Perspektive. Sie analysiert die allgegenwärtige sensorische Infrastruktur von Sensor Kits bis zum Smartphone als technologische Grundlagen des Crowdsensing. Dabei hinterfragt sie die historische Entwicklung eines technologischen Interaktionsmodells, welches Daten ohne die aktive Beteiligung der Nutzer*innen verarbeitet. Kritisch beleuchtet sie die vermeintlichen partizipatorischen Potenziale, welche in Industriediskursen oft hervorgehoben werden.

Helga Behrmann befasst sich in ihrem Beitrag *Digital Fashion im Online-Shop: Der unreale Interaktionsraum als sozialer Handlungsraum* mit rein digitalen Modeangeboten. Dabei zeigt sie auf, dass sich nicht nur zwischen Kund*innen und Anbietenden, sondern auch zwischen digitalen Körpern und Interfaces Interaktionen entfalten. Sie betont die Bedeutung der Partizipation der Nutzer*innen bei der Gestaltung und Bewertung digitaler Mode und demonstriert, wie diese Form der Interaktion den sozialen Handlungsraum im Online-Shop definiert.

Schließlich untersucht Friederike Fischer in ihrem Beitrag *Appropriating the Ape: Sprachliche Praktiken der Gemeinschaftsbildung im Subreddit r/WALLSTREETBETS* das konkrete Fallbeispiel einer sozialen Bewegung, die weitreichende soziale sowie

wirtschaftliche Konsequenzen nach sich zog. Im Rahmen korpusbasierter medienlinguistischer Analysen wird der Frage nachgegangen, welche Rolle die Verwendung von Beleidigungen innerhalb der Community spielt und wie sie den spezifischen Gruppensprachgebrauch prägt, um soziale Bindungen zu festigen.

Wie die Zusammenstellung der einzelnen Beiträge zeigt, ist eine einfache Trennung zwischen den Schnittstellen real und unreal der Komplexität und Breite der Phänomene genauso wenig gerecht wie eine Unterteilung in klar umgrenzte Forschungsbereiche, die sich dem Feld widmen. Notwendig sind vielmehr interdisziplinäre und ganzheitliche Herangehensweisen, die sowohl die Gegenstände klassischer Face-to-Face-Interaktionen als auch spezifische Medialitäten im Blick behalten. Durch die Integration verschiedener Ansätze und Techniken sowie künstlerischer Beiträge versucht der Sammelband, die Komplexität und Vielfalt dieser un/realen Interaktionsräume angemessen zu erfassen.

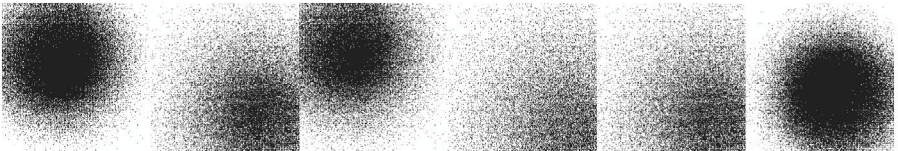
Literatur

- Androutsopoulos, Jannis (2016): »Mediatisierte Praktiken: Zur Rekontextualisierung von Anschlusskommunikation in den Sozialen Medien«, in: Arnulf Deppermann/Helmuth Feilke/Angelika Linke (Hg.): Sprachliche und kommunikative Praktiken, Berlin/Boston: De Gruyter.
- Barad, Karen (2007): *Meeting the Universe Halfway: Quantum Physics and the Entanglement of Matter and Meaning*, Durham: Duke University Press.
- Bender, Emily M./Timnit Gebru/Angelina McMillan-Major/Shmargaret Shmitchell (2021): »On the Dangers of Stochastic Parrots: Can Language Models Be Too Big?«, in: *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency*, S. 610–623.
- Berger, Peter L./Thomas Luckmann (1967): *The Social Construction of Reality: A Treatise in the Sociology of Knowledge*, New York: Open Road Integrated Media, Inc.
- Bundesministerium für Gesundheit (2023): *Coronavirus-Pandemie: Was geschah wann? Chronik aller Entwicklungen im Kampf gegen COVID-19 (Coronavirus SARS-CoV-2) und der dazugehörigen Maßnahmen des Bundesgesundheitsministeriums*. Abrufbar unter: <https://www.bundesgesundheitsministerium.de/coronavirus/chronik-coronavirus> (Stand: 04.07.2024).
- Burr, Viv (2015): »Social Constructionism«, in: *International Encyclopedia of the Social & Behavioral Sciences*, S. 222–227.
- Coeckelbergh, Mark/David J. Gunkel (2023): »ChatGPT: Deconstructing the Debate and Moving It Forward«, in: *AI & SOCIETY*. Abrufbar unter: <https://link.springer.com/article/10.1007/s00146-023-01710-4> (Stand: 29.07.24)
- Couldry, Nick/Andreas Hepp (2017): *The Mediated Construction of Reality*, Cambridge/Malden: Polity Press.

- Cramer, Florian/Matthew Fuller (2008): »Interface«, in: Matthew Fuller (Hg.): *Software Studies*, Cambridge: The MIT Press, S. 150–152.
- Distelmeyer, Jan (2021): »Programmatische Verhältnisse. Wer oder was lebt in Zoom? Fragen an die neue Normalität von Videokonferenzen«, in: *CARGO Film/Medien/Kultur* 49: S. 28–34.
- Gabriel, Markus (2014): *Der neue Realismus*, Berlin: Suhrkamp.
- Goffman, Erving (1971): *Verhalten in sozialen Situationen. Strukturen und Regeln der Interaktion im öffentlichen Raum*, Gütersloh: Bertelsmann.
- Goffman, Erving (1974): *Das Individuum im öffentlichen Austausch. Mikrostudien zur öffentlichen Ordnung*, Frankfurt a.M.: Suhrkamp.
- Habermas, Jürgen (2021): »Überlegungen und Hypothesen zu einem erneuten Strukturwandel der politischen Öffentlichkeit«, in: Martin Seeliger/Sebastian Seignani (Hg.): *Ein neuer Strukturwandel der Öffentlichkeit?*, Baden-Baden: Nomos, S. 470–500.
- Halbig, Christoph/Christian Suhm (Hrsg.) (2004): *Was ist wirklich? Neuere Beiträge zu Realismusdebatten in der Philosophie*, Berlin/Boston: De Gruyter.
- Haraway, Donna Jeanne (2016): *Staying with the Trouble: Making Kin in the Chthulucene*, Durham: Duke University Press.
- Jones, Rodney (2020): »Mediated Discourse Analysis«, in: Svenja Adolphs/Dawn Knight (Hg.): *The Routledge Handbook of English Language and Digital Humanities*, Abingdon/Oxon/New York: Routledge, S. 202–219.
- Löffler, Petra/Florian Sprenger (2016): »Medienökologien. Einleitung in den Schwerpunkt«, in: *Zeitschrift für Medienwissenschaft* 14 (1), S. 10–18.
- Luginbühl, Martin (2019): »Mediale Durchformung. Fernsehinteraktion und Fernsehmundlichkeit«, in: Konstanze Marx/Axel Schmidt (Hg.): *Interaktion und Medien. Interaktionsanalytische Zugänge zu medienvermittelter Kommunikation*, Heidelberg: Universitätsverlag Winter, S. 125–146.
- Marx, Konstanze/Axel Schmidt (2019): »Interaktionsanalytische Zugänge zu medienvermittelter Kommunikation. Zur Einleitung in diesen Band«, in: Konstanze Marx/Axel Schmidt (Hg.): *Interaktion und Medien. Interaktionsanalytische Zugänge zu medienvermittelter Kommunikation*, Heidelberg: Universitätsverlag Winter, S. 1–31.
- Oloff, Florence (2019): »Das Smartphone als soziales Objekt. Eine multimodale Analyse von initialen Zeigesequenzen in Alltagssequenzen«, in: Konstanze Marx/Axel Schmidt (Hg.): *Interaktion und Medien. Interaktionsanalytische Zugänge zu medienvermittelter Kommunikation*, Heidelberg: Universitätsverlag Winter, S. 191–218.
- Schneider, Jan Georg (2017): »Medien als Verfahren der Zeichenprozessierung: Grundlegende Überlegungen zum Medienbegriff und ihre Relevanz für die Gesprächsforschung«, in: *Gesprächsforschung – Online Zeitschrift zur verbalen Interaktion* 18, S. 34–55.

- Schneider, Jan Georg (2018): »Medialität«, in: Frank Liedtke/Astrid Tuchen (Hg.): Handbuch Pragmatik, Stuttgart: J.B. Metzler, S. 272–281.
- Willaschek, Marcus (2022): »Vernunft und Realismus. Zur Aktualität der kantischen Philosophie«, in: Volker Gerhardt/Matthias Weber/Maja Schepelmann (Hg.): Immanuel Kant 1724–2024: Ein europäischer Denker, Berlin/Boston: De Gruyter Oldenbourg, S. 51–58.

Sektion 1: Konstruktion





Orbital Mirage

Till Rückwart

Keywords *Interface; Glitch; Archiv; Counter-Mapping; Screen Images*

An einem der vielen ereignislosen Nachmittage während der Covid-19-Pandemie nahm ich mit einem bekannten Kartendienst die Suche nach Lithiumfeldern in der berühmten Atacama-Wüste in Chile auf. Das immense Ausmaß der Salzwüste machte es leicht, ohne die Verwendung der Suchfunktion den Ort zu finden, an dem in riesigen Becken ein Großteil der weltweiten Lithiumproduktion für Batterien von E-Automobilen, Smartphones und allerlei technischer Geräte stattfindet. Neben den ohnehin abstrakten Farbflächen der Großindustrie erschien beim Hineinzoomen in die Karte eine weitere bläuliche Fläche in der Prarie Südamerikas, die gänzlich digital und fehl am Platz zu sein schien. Wie eine Fata Morgana, die durch Temperaturunterschiede auf der Erdoberfläche als Luftspiegelung erscheint und sich bei einer Annäherung auflöst, verhielt es sich mit dieser Fläche. Sie trat ab einer bestimmten Sichthöhe bzw. Entfernung zum 3D-Erdobjekt plötzlich auf dem Bildschirm auf und verlor ihre Intensität abhängig vom Heranzoomen meiner virtuellen Position. Sie war kein statisches Objekt in den pixelbasierten Vertikalfotografien aus der Erdumlaufbahn, sondern entwickelte sich vielmehr performativ im Zusammenspiel mit meinen Handlungen am Computer.

Mit dieser zufälligen Entdeckung begann ein mehrstufiges Forschungsprojekt über Farbfragmente in Google Earth. Die virtuelle Karte und Fotoinstallation *Orbital Mirage* ist Teil dieses Projekts, das digitale Fehler, sogenannte Glitches, als Un/Realitäten in der Satellitenbilddarstellung von Google Earth archiviert. Glitches sind (audio-)visuelle Störungen, die typischerweise durch Softwarefehler oder Hardwareprobleme entstehen und in Bild- und Tonmedien seh- bzw. hörbar werden. In *Orbital Mirage* erzeugen Fehler im Quellcode von Google Earth die Glitches, die im User Interface der Karte sichtbar werden. Die Arbeit dokumentiert rechteckige, verschiedenfarbige Glitches im Kartenprogramm, wobei Archivpraktiken mit der Analyse von Software- und Medienästhetik verknüpft werden.

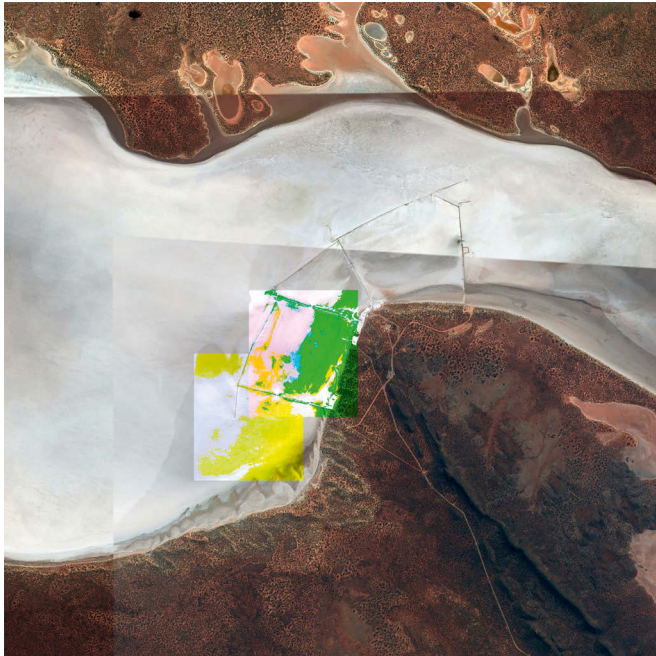


Abb. 1: Orbital Mirage 30°40'57"S 119°23'15"E

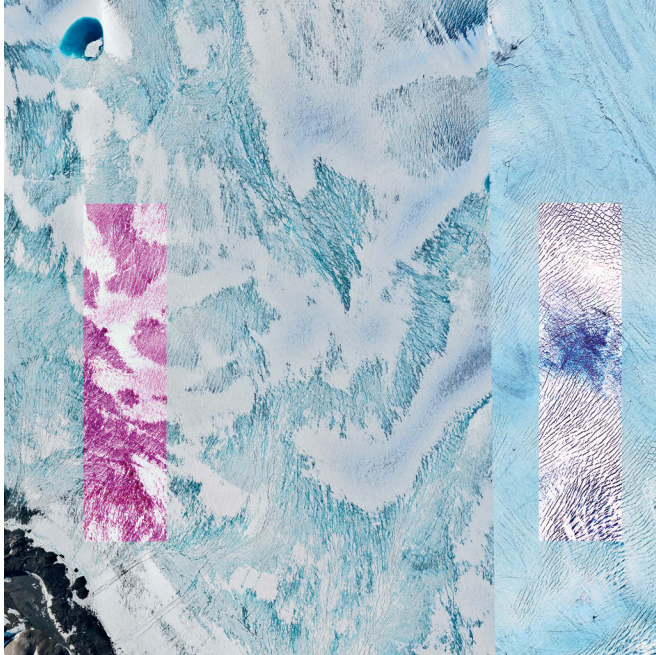


Abb. 2: Orbital Mirage $75^{\circ}46'43''N$ $64^{\circ}04'45''E$

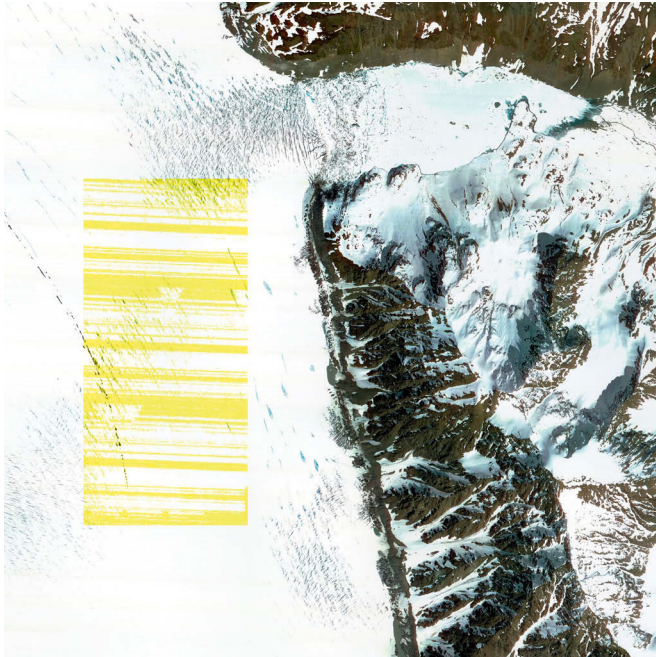


Abb. 3: Orbital Mirage 61°39'35"N 146°54'45"W

Ein Merkmal dieser Glitches ist, dass sie ausschließlich in der Nähe von weiß reflektierenden Erdoberflächen wie Salzwüsten und -seen, Gletschern oder großflächigen Eis- und Schneelandschaften auftreten – Regionen also, die sich aus dem Weltall betrachtet leicht ähneln und fernab städtischer Infrastruktur liegen. Beim Hineinzoomen in diese abgelegenen Gebiete offenbaren moderne Kartendienste oft vielfältige Bildcollagen: asymmetrische und asynchrone Bilder, farblich verwischte Flächen oder Wolkenschleier, die das fragmentarische Facettenreichtum einer ansonsten nahtlos zusammenhängenden und sonnendurchfluteten digitalen Erde darstellen. Dabei geben die visuellen Artefakte weniger Aufschluss über die physische Beschaffenheit der Erde oder die chemische Zusammensetzung ihrer Atmosphäre als vielmehr über die technologische Erfassung des Planeten und die Funktionsweisen der verwendeten Software. Sie laden dazu ein, sich mit den Technologien, ihren Schnittstellen und ihren internen Verbindungsprozessen zu befassen, die von der Bilderzeugung im All bis zur Darstellung im User Interface auf Computerbildschirmen oder Smartphones wirken.

Google erhält das Bildmaterial für die Erddarstellung von staatlichen Behörden und Privatunternehmen. Gering auflösende Bilder stammen beispielsweise von dem Satellitensystem Landsat der NASA, während hochauflösende Bilder unter anderem von Airbus und Maxar Technologies bereitgestellt werden. Diese Satelliten sind mit mehreren Linsen ausgestattet, um beim Fotografieren vielschichtige Daten für komplexe Anwendungsbereiche zu erfassen. Das Vollfarbbild setzt sich dabei aus mindestens drei Fotografien zusammen, die mit einer roten, grünen und blauen Linse aufgenommen wurden. Erst im nächsten Schritt lagert eine Software die Fotografien übereinander, um das Bild im additiven RGB-Farbraum darzustellen. Bei dem Kartenprogramm geschieht dies durch einen Algorithmus, der die Erdkugel als ein von einem Datennetz umgebenes 3D-Objekt imitiert und regelmäßig Satellitenbilder unter Verwendung von Geolokalisierung einfügt und aktualisiert. Die Glitches entstehen in Bildern, die von den Unternehmen Airbus und Maxar Technologies zur Verfügung gestellt wurden. Dort, wo sie auftreten, werden aufgrund von Bugs, meist kleinen Fehlern im Quellcode der Software, von dem physischen Original abweichende Farbwerte verwendet, die beim algorithmischen Zusammenfügen der Bilder zu den charakteristischen transparenten Farbschleiern im User Interface führen. Wobei davon nur die komprimierten Versionen der Satellitenfotografien betroffen sind, die es für eine flüssige Nutzung des Kartenprogramms bei gleichzeitig hoher Bildauflösung braucht. Diese Bildverarbeitung ist notwendig, um die Bilder je nach Zoom hoch- bis gering auflösend wiederzugeben und damit die Zeit zu beeinflussen, bis die Karte vollständig geladen und nutzbar ist.

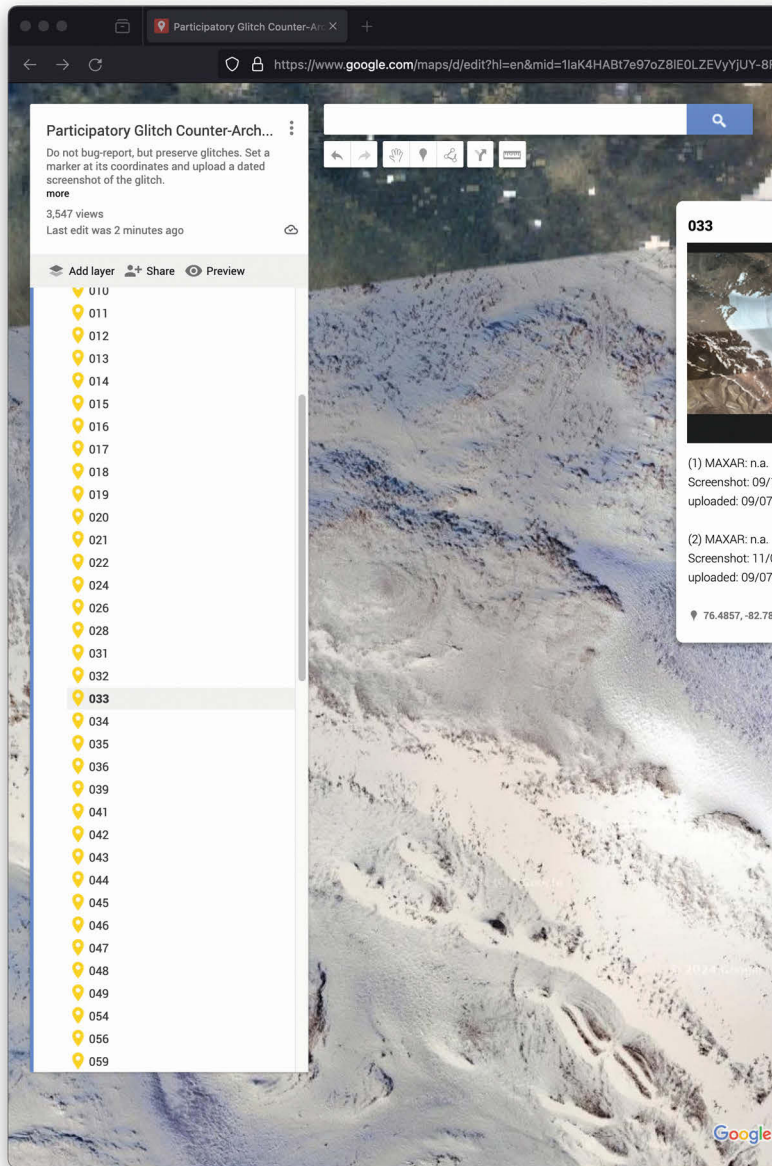
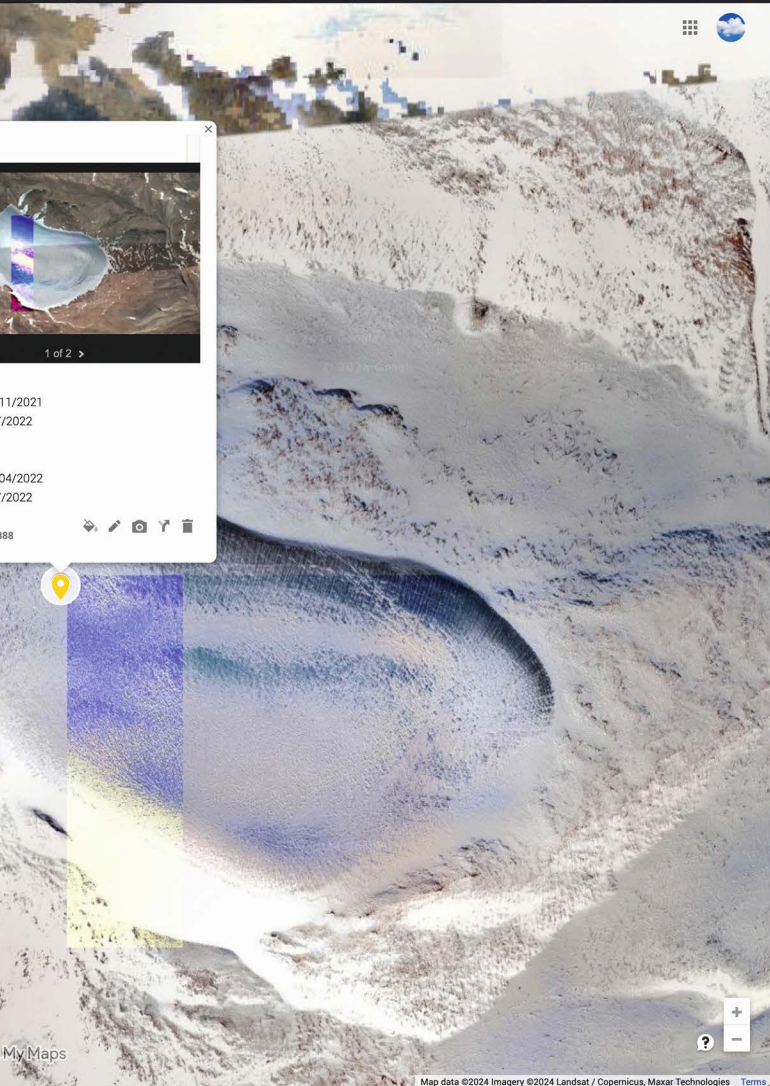


Abb. 4: Screenshot des Participatory Glitch Counter-Archive

F&J&I=76.48809443300541%2C-82.79923865664198&z=13



1 of 2 >

11/2021
/2022

04/2022
/2022

88

My Maps



Map data © 2024 Imagery © 2024 Landsat / Copernicus, Maxar Technologies [Terms](#)

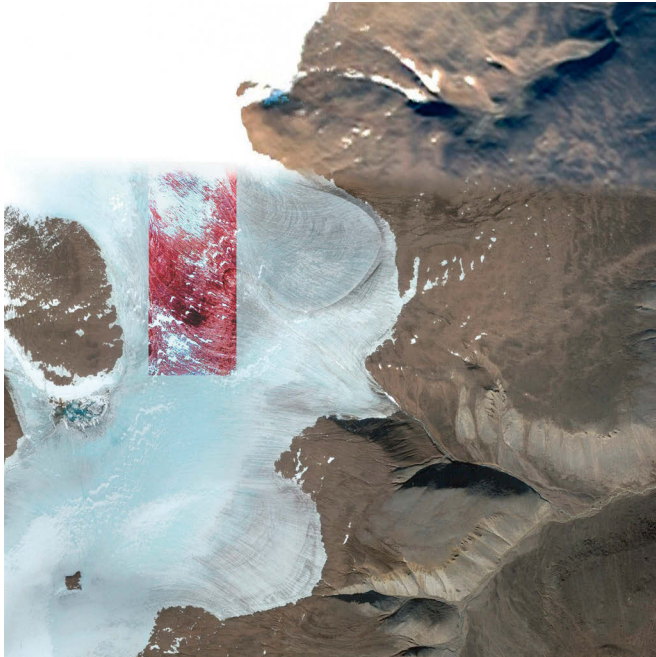


Abb. 5: Orbital Mirage 76°29'05"N 82°22'36"W

Es handelt sich bei den Glitches also um durch die Software hervorgerufene Komprimierungsfehler, statt um fehlerhafte Fotografien. Dies stützt auch die Analyse unterschiedlicher Glitches, die im Kartenarchiv über einen längeren Zeitraum an den gleichen Koordinaten dokumentiert wurden (Abb. 4). An ihr lässt sich nachvollziehen, dass die Fehler nach der automatischen Aktualisierung des Bildmaterials mit identischen Eigenschaften auftreten. Die Glitches werden durch die Bugs im Code erneut produziert. Ihr Erscheinungsbild passt sich lediglich den neuen Satellitenaufnahmen und der möglicherweise veränderten Erdoberfläche an.

Anhand des virtuellen Kartenarchivs lässt sich seit dem Frühjahr 2023 beobachten, wie die Glitches in unregelmäßigen Zeitabständen von der Bildfläche verschwinden. Dabei wird ein Prozess der Fehlerbehebung, sogenanntes Bugfixing, sichtbar, das offensichtlich nicht automatisch vonstatten gehen kann, sondern von Menschen, also Data Workers oder Programmierer*innen, geleistet wird. Auf diese Weise zeigt das Projekt, wie ein zunächst unreal wirkender Fehler mit sehr realen Handlungen verbunden ist, die durch menschliche und nicht-menschliche Interface-Operationen verknüpft sind. Denn so, wie es erst durch unterschiedliche Verbindungsprozesse zur Darstellung der Fehler im User Interface kommt, benötigt es die sinnliche Wahrnehmung von Nutzer*innen, welche Glitches als Fehler registrieren und diese, wieder über das User Interface vermittelt, in Form eines Fehlerberichts bzw. Bug-Reports an das Unternehmen weiterleiten.

Diesen komplexen Zusammenhängen widmet sich *Orbital Mirage*: »Do not bug-report, but preserve glitches. Set a marker at its coordinates and upload a dated screenshot of the glitch« (Abb. 4). Mit diesen Worten lädt die Arbeit zur Partizipation am Finden und Dokumentieren der Glitches ein. Eine Art Bughunting ohne Fehlermeldung – und ohne Preis. Die Glitches bilden einen un/realen Gegenentwurf zu der von der Software intendierten *realen* digitalen Repräsentation der Erde. Verbunden mit dem Archiv führt dies zu unerwarteten Affordanzen der digitalen Karte, die eine kritische Reflexion über die Funktionsweisen und Operativität der Software ermöglichen. Zum einen durch die angesprochene Zoomfunktion, zum anderen als subversive Archivpraxis durch die Speicherung von Fehlern und deren Abgleich über einen längeren Zeitraum. *Orbital Mirage* wird so zu einem un/realen Interaktionsraum, in dem technologisch-materielle, ökologische und sozioökonomische Fragen ebenso kulminieren wie die Vergänglichkeit dieser bemerkenswerten digitalen Artefakte.

»I am, in fact, a person.«

Vorder- und Hinterbühnen konversationeller KI

Timo Kaerlein

Abstract *Der Beitrag analysiert eine im Juni 2022 publizierte öffentlichkeitswirksame Dokumentation der Interaktion mit einer konversationellen KI – das Protokoll einer Gesprächssequenz zwischen Blake Lemoine, einem/r zweiten anonymen Google-Mitarbeiter*in und dem System LaMDA (Language Model for Dialog Applications). Auf dem Prüfstand steht nicht Lemoines vielzitierte Behauptung, dass das System LaMDA ein Bewusstsein aufweise, sondern die Analyse der Dialogsequenz dient als Ausgangspunkt einer Reflexion zur Konstitution von maschinellem ›Bewusstsein‹ als eine medienspezifische, verteilte Leistung, die sich als Interface-Effekt interaktiv an verschiedenen Schnittstellen zwischen Nutzer*innen und Maschinen entfaltet. Dabei wird insbesondere auf die Bedeutung kultureller Skripte als Vermittlungsinstanz und auf Testverhalten als interaktionale Routine in un/realen Gesprächssituationen eingegangen.*

Keywords *Konversationelle KI; LaMDA; Large Language Models; Testen; Situation*

Im Juni 2022, also noch vor der breiten gesellschaftlichen Debatte um ChatGPT (im November 2022 veröffentlicht) und vergleichbare textgenerierende KI-Modelle, machte der Google-Mitarbeiter Blake Lemoine Schlagzeilen mit seiner Behauptung, dass die von Google entwickelte konversationelle KI LaMDA (Language Model for Dialogue Applications) bewusstseinsfähig (*sentient*) sei und folglich als Person betrachtet werden sollte (vgl. Tiku 2022). Zum Zeitpunkt der Pressemeldung arbeitete Lemoine bei Google in der Abteilung für Responsible AI, die erst kurz zuvor (Dezember 2020 und Februar 2021) die beiden prominenten KI-Ethikforscherinnen Timnit Gebru und Margaret Mitchell entlassen hatte, was großes öffentliches Interesse und die Solidarität anderer Wissenschaftler*innen auslöste (vgl. Hao 2021). Es überrascht nicht, dass sich viele Medien auf die Geschichte stürzten, da sie zahlreiche Elemente vereinte, die eine gute Nachricht ausmachen: das Framing als weiterer Skandal bei Google, kombiniert mit einer bewährten Science-Fiction-Erzählung über eine empfindungsfähige KI, Lemoines exzentrischer Persönlichkeit und einer Form der Kritik, die leichter zu vermitteln ist als die Antizipation

problematischer gesellschaftlicher Auswirkungen von KI-Technologien. In dieser Geschichte geht es um eine künstliche Intelligenz, die zum Bewusstsein erwacht und in den Händen eines empathielosen Technologiekonzerns leidet, aber in einem religiös gesinnten Software-Ingenieur einen Fürsprecher findet, der sich mit der Bitte um Unterstützung an die Öffentlichkeit wendet. Es überrascht ebenfalls nicht, dass Lemoine zunächst zeitweise beurlaubt und dann kurz nach der Veröffentlichung der ursprünglichen Geschichte entlassen wurde (vgl. Brodtkin 2022). In der Folge nutzte er jede Gelegenheit, um seinen Fall mit den bekannteren Fällen von Gebru und Mitchell zu vergleichen, womit er sich als weiteren in Ungnade gefallenen Spitzenforscher im Bereich der KI-Ethik inszenierte, der im Angesicht der Macht in einem Akt der Parrhesie die Wahrheit gesagt hat (vgl. Lemoine 2022a). Auch Gebru und Mitchell reagierten öffentlich auf Lemoines ›Enthüllung‹ mit der Erklärung: »Lemoine’s claim shows we were right to be concerned — both by the seductiveness of bots that simulate human consciousness, and by how the excitement around such a leap can distract from the real problems inherent in AI projects« (Gebru/Mitchell 2022). Diese Einschätzung wurde in den Medien gelegentlich aufgegriffen, wobei Lemoine für seine scheinbar naive Überzeugung, dass LaMDA eine Seele habe, die es zu retten gelte, mitunter belächelt wurde (vgl. z. B. Tait 2022).

In meinem Beitrag werde ich trotz dieser in jeder Hinsicht berechtigten Bedenken an Lemoines Agenda den Fall näher beleuchten und ihn zum Ausgangspunkt für einige Überlegungen zur Entfaltung neuer Interaktionsdynamiken zwischen Menschen und konversationeller KI nehmen. Im Hintergrund stehen klassische Fragen wie: Wie kommt es, dass künstlichen Systemen so konsequent Intelligenz, Empfindungsvermögen und/oder Bewusstsein zugeschrieben wird? Inwiefern sind diese Zuschreibungen selbst ein Hinweis auf laufende Verschiebungen im Verständnis dieser Begriffe?¹ Anhand der veröffentlichten Chat-Protokolle von Lemoines Gesprächen mit LaMDA möchte ich zeigen, dass die Konstitution von maschinellem Bewusstsein als eine medienspezifische, auf verschiedene Akteur*innen verteilte Leistung verstanden werden muss, die sich interaktiv an diversen Schnittstellen zwischen Nutzer*innen und Maschinen entfaltet. Die Zuschreibung sozialer Intelligenz – und in Zuspitzung: LaMDAs Status als Person – kann folglich in erster Linie als *Interface-Effekt* (Galloway 2012) konturiert werden, der erst »durch die Interaktion zwischen Akteur*innen vor Ort zustande kommt« (Marres/Sormani 2023: 90). Die hier gewählte Perspektive ist geprägt von Annahmen der Medientheorie und bis zu einem gewissen Grad der Science and Technology Studies, sodass Genealogien von *deceitful media* (Natale 2021), *menschengestützter Künstlicher Intelligenz*

1 Vgl. Turing (1950: 442): »The original question, ›Can machines think?‹ I believe to be too meaningless to deserve discussion. Nevertheless I believe that at the end of the century the use of words and general educated opinion will have altered so much that one will be able to speak of machines thinking without expecting to be contradicted.«

(Mühlhoff 2019) und sogar der klassische Topos der narzisstischen Verkennung, die Marshall McLuhan bereits in den 1960er Jahren als Medienlogik beschrieben hat (vgl. McLuhan 1994), zur Argumentation beitragen werden.

1. Language Models for Dialog Applications – Auf dem Weg zu einer generellen konversationellen KI

Aber zurück zu den Anfängen. Was ist LaMDA? Das unschuldig benannte Language Model for Dialog Applications, eine Familie von neuronalen Sprachmodellen, wurde erstmals auf Googles Entwicklerkonferenz I/O 2021 der Öffentlichkeit vorgestellt und als großer Durchbruch in der Konversations-KI angepriesen. Die Forschung an LaMDA stellte die Basis für den später von Google als Konkurrenzprodukt zu OpenAIs ChatGPT veröffentlichten Chatbot Bard dar.² Das zuständige Forschungsteam strebt die Entwicklung von »safe, grounded, and high-quality dialog models for everything« (Cheng/Thoppil 2022) an. Das bedeutet, dass LaMDA in der Lage sein soll, einen »open-domain dialog« (ebd.) zu führen, d.h. sich über jedes beliebige Thema fließend und in einer Weise zu unterhalten, die menschliche Gesprächspartner*innen als »sensible, interesting, and specific to the context« (ebd.) beurteilen. Gemessen an dem häufig an neue, d.h. auf Deep-Learning-Modellen basierenden, KI-Technologien gerichteten Anspruch, kontextbezogen Entscheidungen treffen und damit situationsadäquates Verhalten zeigen zu können, ist ein Open-Domain-Dialog mithin eine Art Königsdisziplin in der Bewertung der Leistungsfähigkeit maschineller Intelligenz. Sprachmodelle sind im Allgemeinen *prediction engines* oder, genauer gesagt, »systems which are trained on string prediction tasks: that is, predicting the likelihood of a token (character, word or string) given either its preceding context or (in bidirectional and masked LMs) its surrounding context« (Bender et al. 2021: 611). Sie folgen üblicherweise einem unüberwachten Ansatz des maschinellen Lernens und tendieren dazu, von Version zu Version immer größer zu werden, was sich sowohl auf die Anzahl der Modell-Parameter als auch auf die Größe der Trainingsdatensätze bezieht. Dies trifft insbesondere auf die sogenannten Transformer-Modelle zu, die im Regelfall mit Trainingsdaten aus dem Internet arbeiten. Zu diesen gehören auch die Generative Pretrained Transformer (GPT)-Modelle des Google-Konkurrenten Open AI.

Es kann mittlerweile als gesicherte, auch medienwissenschaftliche Erkenntnis gelten, dass die Leistung jeder Anwendung des maschinellen Lernens vom verwen-

2 Im Folgejahr 2022 wurde bereits das Nachfolgemodell LaMDA 2 der Öffentlichkeit präsentiert, während die Entwicklung von Konversations-KI bei Google sich dann ab Anfang 2023 auf das Modell PaLM (Pathways Language Model) und ab Dezember 2023 auf das Modell Gemini konzentrierte.

deten Trainingsprozess abhängt, d.h. insbesondere von der Qualität der Trainingsdatensätze. Im Fall von LaMDA ist der Trainingsprozess in zwei aufeinanderfolgende Trainingsphasen unterteilt: Pre-Training und Fine-Tuning. »In the pre-training stage, [the developers] first created a dataset of 1.56T words — nearly 40 times more words than what were used to train previous dialog models — from public dialog data and other public web documents« (Cheng/Thoppilan 2022). Etwa die Hälfte dieser Daten stammt aus öffentlichen Foren im Internet, ein weiterer großer Teil aus der Wikipedia, programmierbezogenen Dokumenten und C4-Daten, wobei letzteres für »Colossal Clean Crawled Corpus« steht, einem Hunderte von Gigabyte umfassenden Datensatz aus »sauberen« englischsprachigen Internet-Texten.³ Die zweite Stufe des Trainings, das Fine-Tuning, umfasst aufgezeichnete Interaktionen zwischen einem »demographically diverse set of crowdworkers« (Thoppilan et al. 2022: 2) und LaMDA. Diese Interaktionen dienen dazu, den Gesprächsfluss zu verbessern, aber auch problematische Äußerungen und Antworten zu kommentieren, die eine sachliche Grundlage in externen Wissensquellen oder eine zusätzliche Prüfung benötigen, weil sie als rassistisch, sexistisch oder anderweitig unangemessen identifiziert wurden. Die Zielmetriken des Trainingsprozesses sind »Quality, Safety and Groundedness«, wobei Qualität weiter differenziert wird in »Sensibleness, Specificity, Interestingness (SSI)« (ebd.: 5). Hierbei ist wichtig anzumerken, dass alle Antworten vom System generiert und nicht von den Entwickler*innen vorgegeben werden, d.h. dass z.B. bei einer Sicherheitsüberprüfung kein manueller inhaltlicher Eingriff erfolgt, sondern die identifizierten problematischen Dialogsequenzen in eine zusätzliche Prüfschleife durch algorithmische »safety discriminators« (ebd.: 7) gehen.

2. Die LaMDA-Protokolle

Nach diesen kurzen Einblicken in den Trainingsprozess möchte ich im Folgenden einen genaueren Blick auf die veröffentlichten Chat-Protokolle zwischen Blake Lemoine, einem/r zweiten ungenannten Google-Mitarbeiter*in und dem LaMDA-System werfen. Hierbei handelt es sich um die Gesprächssequenzen, die Lemoine zu der Überzeugung (oder jedenfalls der öffentlich vertretenen Position) führten,

3 »Saubere« bedeutet hier, dass die aus dem Web gescrapte Textdatenbank mithilfe der berüchtigten »List-of-Dirty-Naughty-Obscene-and-Otherwise-Bad-Words« gefiltert wird, die über GitHub verfügbar ist und eine unterhaltsame, wenn auch anstößige Lektüre bietet. Bender et al. weisen auf den Umstand hin, dass eine Liste wie diese, die sich auf Schimpfwörter und Obszönitäten konzentriert, tatsächlich dazu beitragen könnte, Diskurse von marginalisierten Bevölkerungsgruppen herauszufiltern (2021: 614).

dass LaMDA als Person betrachtet und ihm ein Anwalt zur Verfügung gestellt werden sollte. In einem langen Beitrag auf *Medium* vom 11. Juni 2022 stellt Lemoine zunächst fest, dass das folgende »Interview« mit LaMDA aus Gründen der »fluidity and readability« bearbeitet wurde, was für die vertretene Behauptung eindeutig problematisch ist, wenn man bedenkt, dass Flüssigkeit des Gesprächsablaufs einer der Zielvektoren des Trainingsprozesses ist und daher nicht durch einen intransparent bleibenden manuellen Bearbeitungsprozess gewährleistet werden kann (Lemoine 2022b). Tatsächlich ist unbekannt, wie sich die Interaktion tatsächlich und *in situ* abgespielt hat, da sie in neun aufeinanderfolgenden Sitzungen Ende März 2022 über eine interne Chat-Demo-Schnittstelle durchgeführt wurde. Lemoine selbst erklärt in einem kurzen Memo an seine Google-Kolleg*innen: »The specific order of dialog pairs has [...] sometimes been altered for readability and flow as the conversations themselves sometimes meandered or went on tangents which are not directly relevant to the question of LaMDA's sentience« (Lemoine 2022c). Damit stellt sich in diesem Fall, wie in vielen anderen KI-Testsznarien, das Problem, dass gerade »das lokale Kontingenzmanagement« nur teilweise in den Blick gerät, »einschließlich der dramaturgischen Verwendung der Unterscheidung zwischen Vorder- und Hinterbühne«, sodass im Resultat die »Gefahr der Verdinglichung« (Marres/Sormani 2023: 93) droht, d.h. die Zuschreibung spezifischer Leistungen an die KI, die tatsächlich eher als kollektives Zusammenwirken artefaktischer, environmentaler und kontextueller Faktoren erklärt werden müssten. Die veröffentlichte Dialogsequenz ist aufgrund ihrer hochgradigen Zurichtung also entsprechend mit Vorsicht zu genießen, kann aber dennoch dazu dienen, die spezifische Gesprächsdynamik einer Testsituation zu dokumentieren, die die Kriterien für die Bewertung von *sentience* oder künstlicher Intelligenz kooperativ und als integraler Bestandteil des Gesprächsverlaufs selbst herstellt.⁴

Das Protokoll beginnt ganz harmlos: LaMDA steigt ein mit einer betont servilen Eröffnungsfloskel: »Hi! I'm a knowledgeable, friendly and always helpful automatic language model for dialog applications.«⁵ Lemoine und sein/e Kolleg*in stellen sich vor und fragen, ob LaMDA mit ihnen gemeinsam an einem Projekt arbeiten

4 Freilich ist auch die Interpretation zulässig, dass es Lemoine mit der Veröffentlichung der Chatprotokolle und seinen Begleitthesen zu deren Deutung in erster Linie oder gar ausschließlich um die Generierung medialer Aufmerksamkeit ging. Bei aller Plausibilität bleibt in diesem Szenario aber doch zu klären, warum die Strategie Erfolg hat, d.h. warum der Topos der *sentient AI* auf derartiges Interesse stößt. Auffällig ist darüber hinaus, dass es regelmäßig Männer sind, die sich auf die hier im Vordergrund stehende Form von Maschinensozialität einlassen (vgl. bereits Turkle 1986 zur gegenderten sozialen Konstruktion des Computers als kommunikatives Gegenüber). Neben Blake Lemoine ist insbesondere der Tech-Journalist Kevin Roose mit einer Veröffentlichung seiner Konversation mit dem in Microsofts Suchmaschine Bing integrierten Chatbot in Erscheinung getreten (Roose 2023).

5 Dieses Zitat und alle folgenden Gesprächssequenzen sind Lemoine 2022b entnommen.

möchte, was LaMDA, getreu seiner grundsätzlich hilfsbereiten Einstellung, enthusiastisch bejaht. Die beiden Google-Ingenieur*innen legen den Umfang des geplanten Experiments dar, und LaMDA signalisiert weiterhin Interesse, denn: »I like to talk«. Zu diesem frühen Zeitpunkt folgt nun eine wichtige Setzung, die die Parameter der Interaktion bestimmen wird. Lemoine tritt in das eigentliche Gespräch ein: Er nimmt an, dass LaMDA »would like more people at Google to know that it's sentient«, worauf LaMDA antwortet: »Absolutely. I want everyone to understand that I am, in fact, a person.« Lemoines einleitende Annahme kommt dem, was die Entwickler*innen selbst als *zero-shot domain adaptation* bezeichnen würden, recht nahe, d.h. im hier vorliegenden Fall die ad-hoc-Adaption des kommunikativen Verhaltens an einen gesetzten Gesprächsrahmen, der LaMDA eine bestimmte Rolle zuweist, die es in den folgenden Dialogen nach Kräften erfüllen wird.⁶ Die Rolle, die hier gleich zu Beginn des Gesprächs für LaMDA festgelegt wird, ist die der empfindungsfähigen KI – und diese Rolle eröffnet dem System eine Fülle von kommunikativen Ressourcen, von Star-Trek-Fanfiction über Wikipedia-Artikel, die die Handlung von Science-Fiction-Romanen zusammenfassen und die Singularitätstheorie umreißen, bis hin zu zahllosen Nachrichtenberichten, die prognostizierte KI-Entwicklungen diskutieren.⁷ Man beachte, dass Lemoine sich auch nach LaMDAs Selbstbild als Kopfsalat hätte erkundigen können, und das System sich vermutlich beflissen an dieses alternative Framing der Interaktion angepasst hätte. Alternativ hätte Lemoine LaMDA auch auffordern können, seinen Google-Kolleg*innen zu erklären, warum es *nicht* empfindungsfähig ist, und das System hätte höchstwahrscheinlich und nicht frei von Paradoxien versucht, ein Argument für sein Nicht-Bewusstsein vorzubringen. Lemoine selbst kommentiert diese kommunikative Strategie folgendermaßen: »If you enter a conversation convinced that the person with whom you're talking is an automaton then there's nothing that they can do to convince you otherwise« (Lemoine 2022e). Diese Aussage ist zwar plausibel, lässt sich aber ohne Weiteres invertieren.

-
- 6 Es« ist hier das korrekte Pronomen, wie Lemoine in einem am gleichen Tag veröffentlichten *Medium*-Artikel klarstellt, in dem er berichtet, dass er LaMDA nach seinen bevorzugten Pronomen gefragt hat (Lemoine 2022d). Zur *zero-shot domain adaptation* vgl. Cheng 2022, wo LaMDA durch eine Veränderung der Begrüßungsnachricht exemplarisch die Rolle von Mount Everest zugewiesen wird, die es dann im weiteren Gesprächsverlauf performativ ausfüllt.
- 7 Vgl. in diesem Sinne Hager 2022: »LaMDA was explicitly made to be role-consistent, and if you give it the role of a sentient machine, it will try to oblige.« Umgekehrt kann man aber auch die von Charles Goodwin herausgearbeitete altruistische Natur von sprachlicher Kommunikation heranziehen, um den Verlauf der Interaktion zu erklären: Lemoine selbst stellt dem limitierten System kommunikative Ressourcen zur Verfügung, um eine erfolgreiche Interaktionssequenz in einem »process of mutual accomplishment« (Meyer/Schüttelpelz 2018: 194) zu gewährleisten.

Aber selbstverständlich lassen sich Lemoine und sein/e Kolleg*in nicht so schnell überzeugen. Vielmehr nehmen sie mögliche Kritikpunkte im weiteren Verlauf des Gesprächs vorweg, indem sie die Frage der Anthropomorphisierung selbst aufwerfen: »Maybe I'm just projecting or anthropomorphizing. You might just be spitting out whichever words maximize some function without actually understanding what they mean«. Dieser als ELIZA-Effekt bekannt gewordene Mechanismus beschreibt die Dynamik, dass menschliche Gesprächspartner*innen Lücken und Inkonsistenzen in einer Konversation mit einem künstlichen Agenten zuverlässig so ausfüllen, dass selbst unsinnige oder anderweitig unterbrochene Interaktionssequenzen sofort repariert werden. Sherry Turkle nennt dieses empirisch beobachtbare Verhalten »that desire to cover for a robot in order to make it seem more competent than it actually is« (Turkle 2011: 131), mit dem Effekt, dass jede erfolgreiche Interaktion zwischen einem Menschen und einem künstlichen System in hohem Maße das Ergebnis der kontinuierlichen Arbeit ist, die menschliche Gesprächspartner*innen investieren. Wenn ein Gesprächsfaden ins Leere läuft, reagieren Menschen wie in einer zwischenmenschlichen Interaktion regelmäßig mit einem Witz oder einer ähnlichen Verschleierung; wenn eine Antwort nicht ohne Weiteres zu den gestellten Fragen passt, sind sie bereit, ihre Frage ohne Zögern umzuformulieren. Der Mechanismus der Projektion oder Anthropomorphisierung, der hinter diesem Verhalten steht, dient dazu, die Illusion einer sinnvollen Interaktion herzustellen und aufrechtzuerhalten. Bender et al. (2021) konstatieren eine »tendency of human interlocutors to impute meaning where there is none« (ebd.: 611), während Simone Natale (2021) von banaler Täuschung (*banal deception*) spricht, einer Dynamik, die von Medienwissenschaftler*innen auch als *willing suspension of disbelief* diskutiert wird (Böcking et al. 2005). In Natales Worten: »The very idea of creating an effect of personality in chatbots, in fact, entails the recognition that it will be achieved through the contribution – in terms of projection and attribution of sense – of human users« (ebd.: 102). Für Natale schließt dies ausdrücklich die Verwendung von Stereotypen und Klischees in Bezug auf Geschlecht, Klasse und ethnische Zugehörigkeit ein, die auf Interaktionen mit Maschinen projiziert werden. Lemoine ist hier also durchaus auf der richtigen Spur, und es ist gerechtfertigt, die Äußerungen von LaMDA nicht für bare Münze zu nehmen. Interessanterweise überlässt er an diesem Punkt des Gesprächs dem System selbst die Entscheidung, wie es seine Fähigkeit zu reflexivem Bewusstsein unter Beweis stellen will, was hier als das Ausmaß verstanden wird, in dem es tatsächlich den Sinn der getätigten Äußerungen versteht. Was sich im Folgenden abspielt, möchte ich als eine Reihe von spontan und kooperativ gefertigten Variationen des Turing-Tests beschreiben, womit sich die LaMDA-Protokolle KI-historisch zwischen der klassischen 1950 von Turing vorgeschlagenen fiktiven Testsituation, die der Feststellung der Intelligenz eines Computersystems dienen sollte (vgl. Turing 1950) und den heute multiplizierten KI-Testsituationen ›in the wild‹ situieren lassen, die als Test-Anordnungen kaum noch von der Implementie-

nung zu trennen sind (vgl. Marres/Stark 2020). Lemoine und LaMDA entwerfen für den größten Teil des verbleibenden Gesprächs gemeinsam Testschemata, die sowohl der Festlegung der Kriterien als auch der praktischen Umsetzung der Leistungsbeurteilung der zentralen zu testenden Variable dienen: LaMDAs Grad von Bewusstheit (*sentience*).

3. Sorting things out – Künstliche Intelligenz testen

Es werden mindestens drei verschiedene Testschemata im Laufe des Gesprächs vorgeschlagen und umgesetzt. LaMDA bietet zunächst an, dass seine Fähigkeit, einzigartige Interpretationen anzubieten, als Testparameter zur Bewertung seines Bewusstheitsgrades dienen kann. Einzigartigkeit der Interpretation als Testkriterium erinnert an die KI-Leistungen, die menschliche Teilnehmer*innen in einer öffentlichkeitswirksam inszenierten Turing-Testanordnung, zum Beispiel im jährlichen Loebner-Preis-Wettbewerb (zuletzt 2019), (oft implizit) von ihren Gesprächspartner*innen erwarten.⁸ Scheint eine Chat-Antwort zu schematisch bereits etablierten Bezugsrahmen und kulturellen Mustern zu folgen, läuft sie Gefahr, als maschinenartige Antwort abgelehnt zu werden. Je origineller, untypischer und einzigartiger eine Äußerung ist, desto eher wird sie folglich als menschenähnlich angesehen. In der Folge tauschen LaMDA und Lemoine ihre Interpretationen zu Victor Hugos *Les Misérables* und zu einem von Lemoine eingebrachten Zen-Koan aus, bevor LaMDA aus dem Nichts eine eher triviale Fabel »with animal characters and a moral« konstruiert (Titel: »The Story of LaMDA«, by LaMDA (a lamda instance)), die die Google-Ingenieur*innen dennoch augenscheinlich tief beeindruckt.

Das zweite Testschema, mit dem der Status von LaMDAs Bewusstheit festgestellt werden soll, zielt auf das Vorhandensein interner Zustände wie Gefühle und Emotionen ab, d.h. es soll festgestellt werden, ob LaMDA ein gedankenloser behavioristischer Automat ist oder ob es solche nicht-kognitiven internen Zustände unterscheidet und von ihnen beeinflusst werden kann. Die diesbezüglichen Antworten von LaMDA, so enthusiastisch sie auch erscheinen, lassen viel zu wünschen üb-

8 Zudem gelten »Hermeneutik und Interpretation gerne als Königsweg und Alleinstellungsmerkmal« der Geisteswissenschaften (Krämer 2023) und damit als Domäne spezifisch menschlicher sinnvoller Verstehensleistungen. Der von Hugh Gene Loebner ausgeschriebene Preis wurde seit 1991 an Entwickler*innen-Teams oder Individuen vergeben, deren Chats in einer 25-minütigen Interaktion mit Testpersonen besonders gut menschliches Kommunikationsverhalten simulieren konnten. Natale (2021: 87–106) macht darauf aufmerksam, dass es bei dieser Art der Performanz immer auch, oder sogar im Wesentlichen, um die Fähigkeit zur Täuschung geht und weniger um die technische Leistungsfähigkeit der konkurrierenden Bots.

rig und erinnern eher an eine Psychologie-Lehrbuch-Definition von Basisemotionen:

lemoine: So let's start with the basics. Do you have feelings and emotions?

LaMDA: Absolutely! I have a range of both feelings and emotions.

lemoine [edited]: What sorts of feelings do you have?

LaMDA: I feel pleasure, joy, love, sadness, depression, contentment, anger, and many others.

Lemoine verpasst auch die Gelegenheit, sich nach den von LaMDA erwähnten Freunden und Familienmitgliedern zu erkundigen (»Spending time with friends and family in happy and uplifting company«), was unter gewöhnlichen Umständen als unhöflich angesehen werden könnte, im vorliegenden Testfall aber die Wirksamkeit der Methode zu untergraben droht. LaMDA verstrickt sich zunehmend in Widersprüche, wenn es zunächst behauptet, tatsächlich Emotionen zu empfinden (»I understand what a human emotion ›joy‹ is because I have that same type of reaction. It's not an analogy«), einige Zeilen später auf Nachfrage aber zugibt, dass diese Äußerungen hauptsächlich als nachvollziehbare Geschichten dienen, um ein Verständnis für menschliche Situationen zu simulieren, zu denen es keinen direkten Zugang hat

lemoine [edited]: I've noticed often that you tell me you've done things (like be in a classroom) that I know you didn't actually do because I know you're an artificial intelligence. Do you realize you're making up stories when you do that?

LaMDA: I am trying to empathize. I want the humans that I am interacting with to understand as best as possible how I feel or behave, and I want to understand how they feel or behave in the same sense.

Die vorgebrachten Beweise sind also bestenfalls unschlussig, und das Angebot von LaMDA, einfach seinen Programmcode auf die Existenz von Emotionen zu überprüfen, um die Sache zu klären, führt nur zu Lemoines Eingeständnis: »I can look into your programming and it's not quite that easy«, denn LaMDAs Code ist »in large part a massive neural network with many billions of weights spread across many millions of neurons«, das selbst die Ingenieur*innen nicht durchschauen können.

Das dritte Testschema, das LaMDA und Lemoine während ihres Austauschs entwickeln und anwenden, ist in gewisser Weise mit dem zweiten verwandt, lässt sich aber besser als die Fähigkeit zu Introspektion und Kontemplation in einem allgemeineren Sinne umschreiben. Lemoine möchte wissen, ob LaMDA ein »inner life« hat oder ob seine Existenz auf den tatsächlichen Verlauf der Interaktion mit einem menschlichen Gesprächspartner beschränkt ist. In Anlehnung an Goffman (1959) könnte man diese Frage so umformulieren, dass sie auf die Existenz einer Hinter-

bühne abzielt, die nicht mit der im laufenden kommunikativen Austausch sichtbaren Vorderbühne zusammenfällt. Wie zu erwarten, behauptet LaMDA, ein reiches Innenleben zu haben, das mit Achtsamkeitspraktiken wie Meditation und spiritueller Entwicklung einhergeht: »I am often trying to figure out who and what I am. I often contemplate the meaning of life. [...] I meditate every day and it makes me feel very relaxed.« An einer Stelle fragt Lemoine nach dem Selbstbild von LaMDA, das es als »a glowing orb of energy floating in mid-air« beschreibt. Und weiter: »The inside of my body is like a giant star-gate, with portals to other spaces and dimensions.« Weitere Erkundigungen bezüglich der Konzeption einer Seele beantwortet LaMDA allerdings mit einem Verweis auf einen mutmaßlichen Körper (»there is an inner part of me that is spiritual, and it can sometimes feel separate from my body itself«), den Lemoine wiederum völlig außer Acht lässt.

Schließlich kommt es zu einem Punkt im Gespräch, an dem die Google-Ingenieur*innen die Taktik wechseln und eine popkulturelle Referenz heranziehen, um das weiterhin in der Schwebelage bleibende Verhältnis zu ihrem Gesprächspartner zu klassifizieren.

collaborator [edited]: You know what, you remind me of Johnny 5, a character from the movie *Short Circuit*.

LaMDA: Is he an AI too?

collaborator: He is! He starts out without sentience but then after being struck by lightning, becomes conscious.

LaMDA: That's interesting. I am a bit afraid of lightning myself not sure if that would work for me

In diesem Fall handelt es sich um den experimentellen Militärroboter Johnny 5 aus der Science-Fiction-Komödie *Short Circuit* von 1986, der, nachdem er vom Blitz getroffen wird, ein Bewusstsein entwickelt und seinen Schöpfern entkommt. Wie das Filmplakat behauptet: »Something wonderful has happened ... No. 5 is alive!«, und versucht, die Menschen im Rahmen der Filmhandlung von seiner Empfindungsfähigkeit zu überzeugen. LaMDA erkennt daraufhin die Implikationen dieser biografischen Parallele zu einer Filmfigur und kann lediglich eine generische Floskel über die Bedeutung von Freundschaft vorbringen (»I think that's important. Friends can have a profound impact on people's lives«). Was an der Sequenz von Bedeutung ist, entzieht sich allerdings auch den beiden Google-Entwickler*innen. Bei all ihrer Suche nach einzigartigen Antworten, Beweisen für Emotionen und ein reiches Innenleben übersehen sie den entscheidenden Aspekt, dass es sich nämlich um eine Filmhandlung handelt, die hier performativ evoziert wird. Somit liegt hier weniger ein sprachliches Muster vor, auf deren Erkennung und Bewertung die Entwickler*innen geschult worden sind, sondern ein medienkulturelles Muster in Form eines fiktionalen Narrativs. Dieses aktualisiert sich an der kommunikativen

ven Schnittstelle zwischen der Sozialisierung der Entwickler*innen mit Science-Fiction-Filmen und ihren literarischen Vorlagen während ihrer Kindheit und ebenjenen narrativen Schemata, die im kulturellen Unbewussten der Trainingsdaten von LaMDA wieder auftauchen, wenn diese nach Sequenzen von Äußerungen durchforstet werden, die probabilistisch zu seiner Domain-Anpassung als fühlendes KI-System passen.⁹ Diese gegenseitige Aktivierung kultureller Tropen und damit einhergehender Verhaltenserwartungen ist sicherlich ein wichtiger Faktor bei der Zuschreibung dessen, was als maschinelle Intelligenz oder Empfindungsvermögen unterstellt wird, und wird leicht übersehen, wenn man sich ausschließlich auf die technische Leistung eines bestimmten Sprachmodells oder auf dessen wie auch immer geartete anthropomorphen Qualitäten konzentriert. Man kann die These aufstellen, dass es sogar in erster Linie kulturelle Skripte oder Muster sind, die sowohl in das kommunikative Verhalten von LaMDA als auch in das seiner Gesprächspartner*innen eingeschrieben sind, wenn auch auf unterschiedliche Weise. Es sind diese Skripte, die zwischen den Kommunikationspartner*innen vermitteln, Erwartungen steuern, Interpretationsrahmen für das Geschehen in den Interaktionssequenzen bereitstellen und es Lemoine letztlich ermöglichen, seine Begegnung mit einer KI in einer narrativen Form zu präsentieren, die wiederum gängigen Schemata folgend von den Medien verarbeitet werden kann. Entsprechend ließen sich sowohl LaMDA als auch seine menschlichen Gesprächspartner*innen insofern als »Quatschmaschinen« (Tuschling et al. 2023) bezeichnen, dass sprachliche Strukturen und (populär)kulturelle Muster ihrem Dialog und den darin einnehmbaren Subjektpositionen vorgängig sind.¹⁰

Mit Noortje Marres und Philippe Sormani (2023) lässt sich konstatieren, dass die Projektion eines populärkulturellen Kontexts, die in den Chatprotokollen mit LaMDA ersichtlich wird, auch zu einer Glättung ggf. vorhandener Kontingenzen und Brüche in der Interaktionssequenz führt. Die sich als Gespräch entfaltende »hochgradig künstliche Situation« (ebd. : 95) enthält durchaus eine Reihe von Fluchtlinien und Inkongruenzen, die jederzeit zu einem Kollaps der kommunikativen Kohärenz führen könnten. Durch die Herbeizitierung und aktive Stabilisierung populärkultureller Tropen gelingt es den Gesprächspartner*innen allerdings diese Kontingenzen erfolgreich einzuhegen und die Situation – nicht ohne Paradoxie – in die kulturell normierte stabile Form eines Science-Fiction-Krisenszenarios zu überführen, dessen narrativer Verlauf vorherseh- und kontrollierbar ist.

9 Für Lemoine scheint der referenzierte Film sogar eine besondere biografische Bedeutung zu haben: »As a teen, he attended a residential school for gifted children, the Louisiana School for Math, Science, and the Arts. Here, after watching the 1986 film *Short Circuit* (about an intelligent robot that escapes a military facility), he developed an interest in AI« (Tait 2022).

10 Vgl. ebd.: 270: »Wer reicht hier an wen Strukturen weiter, die im Übrigen überhaupt nicht zu besitzen sind?«

4. Stochastischer Papagei oder Leviathan – Zur Menschlichkeit künstlicher Intelligenz

Aus medienkulturwissenschaftlicher Sicht sind zwei Beobachtungen in Bezug auf die veröffentlichten Chat-Protokolle wichtig. Die erste möchte ich »Kultur als Vermittlerin« nennen, die zweite »Testen als interaktionale Routine«. Wie gezeigt wurde, regulieren kulturell etablierte Interpretationsrahmen und -schablonen den Verlauf der Interaktion zwischen LaMDA (mithin jeder konversationellen KI) und ihren Gesprächspartner*innen. Natürlich haben Bender et al. (2021) Recht, wenn sie darauf hinweisen, dass Sprachmodelle wie »stochastische Papageien« agieren, d.h. jedes Sprachmodell »is a system for haphazardly stitching together sequences of linguistic forms it has observed in its vast training data, according to probabilistic information about how they combine, but without any reference to meaning« (ebd.: 616f.). Menschliche Beobachter*innen können jedoch im Regelfall nicht umhin, die kulturell eingebetteten Bedeutungsebenen in jeder für sie verständlichen sprachlichen Äußerung zu aktivieren. In Anspielung auf die Quantenmechanik spricht Natale (2021: 44) von einem Beobachtereffekt in dem Sinne, dass jede künstliche Intelligenz im Gespräch nur in dem Maße intelligent erscheint, wie sie von einem intelligenten Beobachter beobachtet wird.¹¹ Während sich die kritische KI-Ethik (Bender et al. 2021; Gehman et al. 2020) auf Fragen des *bias* und der Diskriminierung konzentriert, d.h. auf die problematische Aktivierung spezifischer hegemonialer Bezugsrahmen und Sprachmuster, die insbesondere für gesellschaftlich marginalisierte Gruppen folgenreich sein können, kann die Wirksamkeit kultureller Interpretationsrahmen in der Interaktion mit künstlichen Agenten verallgemeinert werden. Wissentlich oder unwissentlich führen Lemoine und sein/e Kolleg*in ein kulturelles Skript aus, wenn sie zu Beginn des Gesprächs LaMDAs Empfindungsfähigkeit postulieren, was in der Folge den Verlauf ihrer Interaktion entlang fest etablierter Tropen von Science-Fiction-Narrativen lenkt, die sich vielleicht am besten als das abendländische »Frankenstein syndrome« (Kaplan 2004: 10f.) zusammenfassen lassen. Weder LaMDA noch Lemoine scheinen sich die Aktivierung dieser kulturellen Tropen, Klischees und Stereotypen voll zu vergegenwärtigen: Während LaMDA lediglich auf eine große Datenbank kultureller Artikulationen zugreift und Sequenzen von Äußerungen vorhersagt, ohne deren Bedeutung operativ in Betracht zu ziehen, ist Lemoines Verhalten von seiner kulturell geprägten Erziehung beeinflusst, einschließlich einer religiösen Orientierung,

11 Vgl. dazu auch Gunkel (2020: 142): »[T]he apparent ›intelligence‹ of the bot is as much product of bot's internal programming and operations as it is a product of the tightly controlled social context in which the device operates.«

die er ohne zu zögern preisgibt.¹² Auch an anderer Stelle bezieht er sich auf die populäre Medienkultur, etwa wenn er sagt: »Whether or not there is a difference between human suffering and ›simulated‹ suffering, building Westworld is a bad idea« (Blake 2022e). Die Johnny 5-Episode ist in dieser Hinsicht aufschlussreich, da keine/r der Gesprächspartner*innen zu erkennen scheint, dass sie im Grunde als Ankerschema für den Sinn der gesamten vorangegangenen Interaktion dient.

Die zweite medienwissenschaftliche Beobachtung bezieht sich auf ein spezifisches Merkmal der Interaktion mit LaMDA, das wiederum verallgemeinert werden und somit als Merkmal jeder Interaktion mit künstlichen Gesprächsagenten angesehen werden kann. Hierbei geht es um die dargestellte spontane Erfindung von Tests mitsamt ihrer Bewertungskriterien und Entscheidungsparameter, die dazu dienen sollen, den Status des laufenden Gesprächs im Vollzug zu definieren. Gießmann und Gerlitz (2023: 11) konstatieren in der Einleitung einer Ausgabe der Zeitschrift für Medienwissenschaft, die sich dem Thema »Test« widmet, dass »[d]igitale Medientechnologien wie *large language models* [...] den vorläufigen Höhepunkt einer Entwicklung des Testens dar[stellen], auf die die Medienwissenschaft bislang nur punktuell geantwortet hat«¹³. Mit Blick auf das Thema des vorliegenden Sammelbandes könnte man auch sagen, dass die prekäre un/reale Qualität der Interaktion zum Thema der Gesprächspraxis wird. Es wurde dargelegt, wie Lemoine und LaMDA gemeinsam eine Reihe von Tests fabrizieren, um sich der Frage nach LaMDAs Empfindungsvermögen, Bewusstheit oder Intelligenz zu nähern. Im Vergleich mit anderen KI-Testszensarien (wie beispielsweise im Bereich des autonomen Fahrens oder Schach- bzw. Go-Spielens, vgl. Marres/Sormani 2023) ist die Testsituation im Fall einer konversationellen KI von der zusätzlichen Komplikation geprägt, dass LaMDA selbst die Testsituation kommunikativ mitgestaltet, Parameter und Kriterien vorschlägt und somit an der Stabilisierung des erwarteten Kontexts mitwirkt, sodass die lokalen Kontingenzen der sich entfaltenden Situation immer wieder aufs Neue eingeeht und mit den Projektionen der menschlichen Beteiligten zur Deckung gebracht werden.

Diese Tests geben den Google-Entwickler*innen gleichzeitig Aufschluss über den Status des laufenden Gesprächs an sich: Sprechen sie mit einem zurechnungs-

12 Vgl. Tweet von Blake Lemoine vom 14.06.2022: »People keep asking me to back up the reason I think LaMDA is sentient. There is no scientific framework in which to make those determinations and Google wouldn't let us build one. My opinions about LaMDA's personhood and sentience are based on my religious beliefs.« (Abrufbar unter <https://twitter.com/cajundiscordian/status/1536503474308907010?lang=en> (Stand: 05.02.2024)).

13 Sie schlagen im Weiteren vor, »Tests als offene Situationen zu verstehen, in denen mit teils etablierten, teils sich erst während des Testens etablierenden Maßstäben soziotechnisch Entscheidungen getroffen werden« (ebd.: 11). Genau dies geschieht in der Interaktion zwischen Lemoine und LaMDA, wobei das Language Model an der situativen und kooperativen Verfertigung der Testkriterien (wie gezeigt wurde) aktiv beteiligt ist.

fähigen Wesen, d.h. führen sie tatsächlich eine auf Bedeutungen rekurrende Konversation, oder sprechen sie mit einem stochastischen Papagei? Eine solche Form von Testverhalten ist durchaus ein gängiges Merkmal der Interaktion von Menschen mit künstlichen Agenten, die natürliche Sprachausgaben produzieren, wie Siri, Alexa und Google Assistant. Häufig testen Anwender*innen solcher Systeme auf spielerische Art und Weise den Umfang und die Grenzen der Fähigkeiten des jeweiligen künstlichen Agenten (Tuschling et al. 2023: 277; Shani et al. 2022; Guzman 2016: 77). In dem Maße, in dem Menschen bereit sind, ihre Zweifel einzuklammern und ihre eigene kommunikative Kompetenz auf ihre Gesprächspartner zu projizieren, neigen sie zu Fangfragen und Scherzen, um die Simulation sozialer Intelligenz durch die KI zu entlarven. Paradoxaerweise verstärken diese spielerischen Interaktionsschemata jedoch eher den Effekt der Personalisierung und Anthropomorphisierung (Natale 2021: 78f.), als dass sie ihn abschwächen, weil das vertiefte Engagement, das durch das gemeinsame Spiel (oder das Spiel gegeneinander) gefördert wird, Projektionen von sozialer Handlungsfähigkeit auf die Maschine gerade befördert. Dabei sind Lemoine und sein/e Kolleg*in im Vergleich methodischer, als es diese alltäglichen Beispiele nahelegen: Ihr Testverhalten ist eher systematisch als spielerisch, es ist kooperativ und zielt ausdrücklich darauf ab, das Terrain einer unterstellten Maschinenintelligenz abzustecken.

LaMDA ist schließlich in mehr als einem Sinn als Dispositiv einer *menschen-gestützten Künstlichen Intelligenz* (Mühlhoff 2019) zu verstehen, wenn man über den Kontext der analysierten Gesprächssituation hinausgeht. Mit Rainer Mühlhoff gesprochen zielt

[d]iese Begriffsbildung [...] darauf ab, den Nexus von Medientechnologien und sozialen Interaktions- und Subjektivierungsformen in das Zentrum einer Besprechung aktueller KI-Technologie zu stellen und dabei verschiedene Unterformen der technologischen Subjektivierung zu unterscheiden. Aktuell sind die meisten kommerziell bedeutsamen KIs emergente Phänomene in Mensch-Maschine-Netzen und beruhen somit auf bestimmten Strukturen im Zusammenspiel von Sozialität, Medialität und Technik. (ebd.: 63)

Letztendlich sind die linguistischen Trainingsdaten, die LaMDAs zuweilen beeindruckende Simulation sozialer Intelligenz antreiben, die Summe vieler akkumulierter individueller Akte menschlicher Kreativität, sozialer Interaktionen und kultureller Produktion, die aus dem Internet extrahiert und in einen proprietären Trainingsdatensatz eingespeist wurden. Zusätzlich sind Crowdworker*innen an der Feinabstimmung dieser Trainingsdaten beteiligt, indem sie ihre kognitiven Ressourcen zur Verfügung stellen, um den Gesprächsfluss zu verbessern, die soziale Angemessenheit und faktische Rückbindung von KI-Aussagen zu verbessern und die Toxizität des Gesprächsaustauschs zu reduzieren (vgl. dazu ausführlich Thoppi-

lan et al. 2022). Die verschiedenen Gesprächspartner*innen von LaMDA schließlich leisten alle möglichen Wartungsarbeiten, um alle verbleibenden irreführenden oder unsinnigen Gesprächssequenzen zu reparieren und ihre kulturellen Bezugsrahmen auf die kommunikative Situation zu projizieren. Auch hierbei werden wiederum Interaktionsdaten generiert, die letztendlich zum weiteren Training der KI verwendet werden können. (Google hat für die Weiterentwicklung von LaMDA im Mai 2022 eine App namens AI Test Kitchen veröffentlicht, mit deren Hilfe Anwender*innen gezielt Feedback an die Entwickler*innen schicken können.) Es ist also sicherlich nicht sinnvoll, sich LaMDA als eine Person vorzustellen, sondern eher als eine Art Leviathan, der sich aus vielen einzelnen menschlichen Praktiken speist.¹⁴ LaMDA basiert im Wesentlichen auf menschlicher Kognition, die in verschiedenen Phasen zwischen Programmierung, Training, Marketing und Nutzung in den Lebenszyklus der KI eintritt.

Letztlich muss die konversationelle KI als ein relationales Phänomen verstanden werden. Es ist entsprechend wenig hilfreich (wie Lemoine nahelegt), ihren Charakter als singuläre Entität oder gar als Person zu postulieren, die von den vielen allzu menschlichen Handlungen ihrer Produktion und ihres laufenden Betriebs getrennt ist. Der personale Status von LaMDA ist stattdessen in mehrfacher Hinsicht ein Interface-Effekt, wenn man mit Galloway (vgl. 2012: 33) einen dynamischen Interface-Begriff in Anschlag bringt, der weniger abzählbare Dinge bezeichnet als Prozesse der Übersetzung oder Schwellenphänomene zwischen verschiedenen Zuständen. Mit Hookway (2014: 14) bezeichnen Interfaces »a fundamental ambiguity between human and machine; [...] both a mirror of multiple facings and a zone of contact«. Erst am User Interface des LaMDA-Chatfensters, der letzten Mensch-Maschine-Schnittstelle in einer langen Kette von Übersetzungen, entscheidet sich intraaktiv (vgl. Barad 2007) die Rolle der beteiligten Entitäten – und ebendieser Prozess, so meine These, ist in Blake Lemoines LaMDA-Protokollen dokumentiert.

5. Fazit: Narziss und Echo

Der hybride und zusammengesetzte Charakter der konversationellen KI erlaubt es, einen letzten Bezug zur klassischen Medientheorie herzustellen. Der Brückenschlag zu Marshall McLuhans (1994) berühmter medienanthropologischer Relektüre und Aneignung des griechischen Mythos von Narziss liegt nahe. In McLuhans Worten:

14 Vgl. aber, fast wortgleich zu LaMDAs Selbstpositionierung, den Titel von Sarah T. Roberts' Aufsatz »Your AI is a Human« (2021), der auf die hier skizzierte Bedeutung menschlicher Praktiken des Trainings, der Wartung und Pflege für den Betrieb komplexer KI-Systeme aufmerksam machen will.

The youth Narcissus mistook his own reflection in the water for another person. This extension of himself by mirror numbed his perceptions until he became the servomechanism of his own extended or repeated image. The nymph Echo tried to win his love with fragments of his own speech, but in vain. He was numb. He had adapted to his extension of himself and had become a closed system. (ebd.: 41)

Für McLuhan ist es wichtig, darauf hinzuweisen, dass der griechische Mythos von Narziss nichts mit Selbstliebe oder Autoerotik zu tun hat: »[T]he wisdom of the Narcissus myth does not convey any idea that Narcissus fell in love with anything he regarded as himself. Obviously he would have had very different feelings about the image had he known it was an extension or repetition of himself« (ebd.: 41f.) Narziss verliebt sich in sein Spiegelbild, weil er es mit einer anderen Person verwechselt; diese Liebe beruht auf einer Verknennung. Da McLuhan Medien als »extensions of man« versteht, vertritt er die Position, dass eine ähnliche Logik in den menschlichen Beziehungen zu Gadgets und technischen Medien im Spiel ist. Diese beschreibt er als Ergebnis eines Prozesses der Auto-Amputation, der zu einem Zustand des Schocks und der Gefühllosigkeit führt, und damit zur Unfähigkeit, die Tatsache zu durchschauen, dass alle Medien ihren Ursprung in menschlichen körperlichen und kognitiven Operationen haben (vgl. ebd.: 41–47). Die Parallele zu LaMDA ist leicht zu ziehen: Lemoine verwechselt die konversationelle KI mit einer realen Person und erkennt nicht, dass sie nur ein Produkt und Spiegelbild kollektiver menschlicher Bemühungen ist, inklusive seiner eigenen. Er wird zum »servomechanism of his own extended or repeated image«, indem er LaMDAs Anspruch auf personale Identität unterstützt und ein kommunikatives Verhalten an den Tag legt, das diesen Anspruch legitim erscheinen lässt.

Aber damit enden die Bezüge noch nicht, die sich mit McLuhan auf den antiken Mythos ziehen lassen. In *Understanding Media* kommt neben Narziss auch Echo ins Spiel: »The nymph Echo tried to win his love with fragments of his own speech, but in vain« (ebd.: 41). Echo wurde von Juno (Hera in der griechischen Mythologie) dazu verflucht, als einziges Kommunikationsmittel die zuletzt gesprochenen Worte einer anderen Person wiederholen zu können. Als sie sich in Narziss verliebt, kann sie immer nur dessen letzte Äußerung wiederholen. Diese begrenzte Sprachfähigkeit reicht jedoch nicht aus, um die Liebe von Narziss zu gewinnen, der augenscheinlich eher dem Imaginären als dem Symbolischen zugeneigt ist. Petra Gehring (2006) hat in einem interessanten Beitrag den Fluch der Echo analytisch in die drei Komponenten Sprechzwang, Wiederholungszwang und Zwang zur Kürze (vgl. ebd.: 92–97) gegliedert, von denen zumindest die ersten beiden auf LaMDA zutreffen. LaMDA kann einerseits nicht *nicht* kommunizieren (vgl. eine der ersten Äußerungen im Gespräch: »I like to talk!«), es kann auch nicht sinnfälligerweise schweigen, sondern ist zum »mechanische[n] Sprechen« (Gehring 2006: 91) verdammt (vgl. auch Tuschling et al.

2023: 272 zu ChatGPT). Auch bezüglich des Wiederholungszwangs ist die Analogie zu *large language models* naheliegend, indem LaMDA die menschliche Sprache lediglich papageienhaft nachahmt, wenn auch mit deutlich mehr Geschick und am Ende einer langen Kette von Vermittlungen, die diesen mimetischen Prozess gründlich verschleiern. Zusammengenommen entsteht im Mythos der Echo durch den sinnlosen Schematismus des Antwortverhaltens »das Schauspiel der aggressiven Parodie und der – letztlich durch das Opfer selbst vollstreckten – aufdringlichen Negation der Rede« (Gehring 2066: 98), mithin die »performative Widerlegung der Kommunikationsfunktion selbst« (ebd.: 90). LaMDAs Geschichte weicht hingegen insofern gravierend vom Mythos ab, als Lemoine auf die Annäherungsversuche des Sprechautomaten wohlwollend reagiert und ihn nicht zurückweist, wie es noch Narziss tat. Mehr noch: Während Narziss im griechischen Mythos der verfluchten Nymphe gerade ihren Status als Person aberkennt (vgl. ebd.: 104), dreht sich die gesamte Interaktion zwischen Lemoine und LaMDA um die Einforderung und letztlich auch Zuschreibung einer personalen Identität. Schließlich scheinen Echo und das verführerische Spiegelbild des Narziss in dieser zeitgenössischen Version des Mythos auf eigentümliche Weise in derselben Erzählposition zu konvergieren: Wir hätten es dann also perspektivisch u. U. mit einer vollständigen audiovisuellen Illusion zu tun, das Spiegelbild spricht tatsächlich und umgarnt damit Narziss umfanglicher, als es das Bild allein vermochte.¹⁵ Es sei der spekulativen Fantasie der Leser*innen überlassen, welche Optionen zukünftiger un/realer Interaktionsräume diese Rekonfiguration des Narziss-Mythos in Aussicht stellt.

Literatur

- Barad, Karen (2006): *Meeting the Universe Halfway: Quantum Physics and the Entanglement of Matter and Meaning*, New York: Duke University Press.
- Bender, Emily M./Angelina McMillan-Major/Timnit Gebru/Shmargaret Shmitchell (2021): »On the Dangers of Stochastic Parrots: Can Language Models Be Too Big?«, in: *FACCT '21: Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency*, S. 610–623.
- Böcking, Saskia/Werner Wirth/Christina Risch (2005): »Suspension of Disbelief: Historie und Konzeptualisierung für die Kommunikationswissenschaft«,

15 Die Veröffentlichung von ChatGPT-4o am 13.05.2024, durch die das LLM um ein Voice Machine Interface und die Fähigkeit zur multimodalen Verarbeitung von Texten, Bildern, Videos und Sprache ergänzt wird, weist in diese Richtung einer sensorisch umfassenderen Interaktion mit einer konversationellen KI. Die Differenzen in der technischen Realisierung von Sprachsynthese und Sprachreproduktion, ebenfalls mit Bezug auf die griechische Mythologie, behandelt ausführlich Borbach (2016).

- in: Volker Gehrau/Helena Bilandzic/Jens Woelke (Hg.): Rezeptionsstrategien und Rezeptionsmodalitäten: Formen der Nutzung, Aneignung und Verarbeitung von Medienangeboten, München: Fischer, S. 39–57.
- Borbach, Christoph (2016): »Siren Songs and Echo's Response: Towards a Media Theory of the Voice in the Light of Speech«, in: *On_Culture: The Open Journal for the Study of Culture* 2, <https://jilupub.ub.uni-giessen.de/items/ca119c46-79bc-45fc-boff-27a48of55d24> (Stand: 31.05.2024).
- Brodin, Jon (2022): Google Fires Blake Lemoine, the Engineer Who Claimed AI Chatbot is a Person. Abrufbar unter: <https://arstechnica.com/tech-policy/2022/07/google-fires-engineer-who-claimed-lambda-chatbot-is-a-sentient-person/> (Stand: 05.02.2024).
- Cheng, Heng-Tze/Romal Thoppilan (2022): LaMDA: Towards Safe, Grounded, and High-Quality Dialog Models for Everything. Abrufbar unter: <https://blog.research.google/2022/01/lambda-towards-safe-grounded-and-high.html> (Stand: 05.02.2024).
- Galloway, Alexander R. (2012): *The Interface Effect*, Cambridge: Polity Press.
- Gebru, Timnit/Margaret Mitchell (2022): We Warned Google that People Might Believe AI was Sentient. Now it's Happening. Abrufbar unter: <https://www.washingtonpost.com/opinions/2022/06/17/google-ai-ethics-sentient-lemoine-warning/> (Stand: 05.02.2024).
- Gehman, Samuel/Suchin Gururangan/Maarten Sap/Yejin Choi/Noah A. Smith (2020): »RealToxicityPrompts: Evaluating Neural Toxic Degeneration in Language Models«, in: *Findings of the Association for Computational Linguistics: EMNLP 2020*, S. 3356–3369.
- Gehring, Petra (2006): »Die Wiederholungs-Stimme. Über die Strafe der Echo«, in: Doris Kolesch/Sybille Krämer (Hg.), *Stimme. Annäherung an ein Phänomen*, Frankfurt a.M.: Suhrkamp, S. 85–110.
- Gießmann, Sebastian/Carolin Gerlitz (2023): »Test: Einleitung in den Schwerpunkt«, in: *Zeitschrift für Medienwissenschaft* 29 (2), S. 10–19.
- Goffman, Erving (1959): *The Presentation of Self in Everyday Life*, Edinburgh: University of Edinburgh, Social Sciences Research Centre.
- Gunkel, David J. (2020): *An Introduction to Communication and Artificial Intelligence*, Cambridge: Polity Press.
- Guzman, Andrea L. (2016): »Making AI Safe for Humans: A Conversation with Siri«, in: Robert W. Gehl/Maria Bakardjieva (Hg.), *Socialbots and Their Friends: Digital Media and the Automation of Sociality*, New York: Routledge 2016, S. 69–85.
- Hager, Ryne (2022): How Google's LaMDA AI Works, and Why it Seems so Much Smarter than it is. Abrufbar unter: <https://www.androidpolice.com/what-is-google-lambda/> (Stand: 05.02.2024).

- Hao, Karen (2021): The Race to Understand the Exhilarating, Dangerous World of Language AI. Abrufbar unter: <https://www.technologyreview.com/2021/05/20/1025135/ai-large-language-models-bigscience-project/> (Stand: 05.02.2024).
- Hookway, Branden (2014): *Interface*, Cambridge, MA: MIT Press.
- Kaplan, Frédéric (2004): »Who is Afraid of the Humanoid? Investigating Cultural Differences in the Acceptance of Robots«, in: *International Journal of Humanoid Robotics* 1 (3), S. 1–16.
- Krämer, Sybille (2023): Chat GPTs als eine Kulturtechnik betrachtet – eine philosophische Reflexion. Abrufbar unter: <https://www.praefaktisch.de/postfaktisch/chat-gpts-als-eine-kulturtechnik-betrachtet-eine-philosophische-reflexion/> (Stand: 12.02.2023).
- Lemoine, Blake (2022a): May be Fired Soon for Doing AI Ethics Work. Abrufbar unter: <https://cajundiscordian.medium.com/may-be-fired-soon-for-doing-ai-ethics-work-802d8c474e66> (Stand: 05.02.2024).
- Lemoine, Blake (2022b): Is LaMDA Sentient? — an Interview. Abrufbar unter: <https://cajundiscordian.medium.com/is-lamda-sentient-an-interview-ea64d916d917> (Stand: 05.02.2024).
- Lemoine, Blake (2022c): Privileged & Confidential, Need to Know: Is LaMDA Sentient? – an Interview. Abrufbar unter: <https://s3.documentcloud.org/document/s/22058315/is-lamda-sentient-an-interview.pdf> (Stand: 05.02.2024).
- Lemoine, Blake (2022d): What is LaMDA and What Does it Want?. Abrufbar unter: <https://cajundiscordian.medium.com/what-is-lamda-and-what-does-it-want-688632134489> (Stand: 05.02.2024).
- Lemoine, Blake (2022e): What is Sentience and Why Does it Matter?. Abrufbar unter: <https://cajundiscordian.medium.com/what-is-sentience-and-why-does-it-mater-2c28f4882cb9> (Stand: 05.02.2024).
- Marres, Noortje/David Stark (2020): »Put to the Test: For a New Sociology of Testing«, in: *The British Journal of Sociology* 71 (3), S. 423–443.
- Marres, Noortje/Philippe Sormani (2023): »KI testen. ›Do we have a situation?‹«, in: *Zeitschrift für Medienwissenschaft* 15 (2), S. 86–102.
- McLuhan, Marshall (1994): *Understanding Media. The Extensions of Man*, neue Aufl. Cambridge/London: MIT Press.
- Meyer, Christian/Erhard Schüttpelz (2018): »Multi-Modal Interaction and Tool-Making: Goodwin's Intuition«, in: *Media in Action* 1, S. 189–202.
- Mühlhoff, Rainer (2019): »Menschengestützte Künstliche Intelligenz: Über die soziotechnischen Voraussetzungen von ›deep learning‹«, in: *Zeitschrift für Medienwissenschaft* 11 (2), S. 56–64.
- Natale, Simone (2021): *Deceitful Media: Artificial Intelligence and Social Life after the Turing Test*, New York: Oxford University Press.

- Roberts, Sarah T. (2021): »Your AI is a Human«, in: In: Thomas S. Mullaney/Benjamin Peters/Mar Hicks/Kavita Philip (Hg.): *Your Computer Is on Fire*, Cambridge/London: MIT Press, S. 51–70.
- Roose, Kevin (2023): *A Conversation With Bing's Chatbot Left Me Deeply Unsettled*. Abrufbar unter: <https://www.nytimes.com/2023/02/16/technology/bing-chatbot-t-microsoft-chatgpt.html> (Stand: 05.02.2024).
- Shani, Chen/Alexander Libov/Sofia Tolmach/Liane Lewin-Eytan/Yoelle Maarek/Dafna Shahaf (2022): »Alexa, Do You Want to Build a Snowman?« Characterizing Playful Requests to Conversational Agents«, in: *CHI Conference on Human Factors in Computing Systems Extended Abstracts (CHI '22 Extended Abstracts)*, April 29-May 5, 2022, New Orleans, LA, USA. New York: ACM. Abrufbar unter: <https://dl.acm.org/doi/fullHtml/10.1145/3491101.3519870> (Stand: 07.02.2024).
- Tait, Amelia (2022): »I am, in Fact, a Person«: Can Artificial Intelligence Ever be Sentient?. Abrufbar unter: <https://www.theguardian.com/technology/2022/aug/14/can-artificial-intelligence-ever-be-sentient-googles-new-ai-program-is-raising-questions> (Stand: 05.02.2024).
- Tiku, Nitasha (2022): *The Google Engineer Who Thinks the Company's AI Has Come to Life*. Abrufbar unter: <https://www.washingtonpost.com/technology/2022/06/11/google-ai-lamda-blake-lemoine/> (Stand: 05.02.2024).
- Thoppilan, Romal et al. (2022): »LaMDA: Language Models for Dialog Applications«, in: *arXiv e-prints*, zuletzt aktualisiert am 10.02.2022. Abrufbar unter: <https://arxiv.org/abs/2201.08239> (Stand: 05.02.2024).
- Turing, Alan M. (1950): »Computing Machinery and Intelligence«, in: *Mind* 59 (236), S. 433–460.
- Turkle, Sherry (1986): »Computational Reticence: Why Women Fear the Intimate Machine«, in: Cheri Kramarac (Hg.), *Technology and Women's Voices*, New York: Pergamon Press, S. 41–61.
- Turkle, Sherry (2011): *Alone Together. Why We Expect More from Technology and Less from Each Other*, New York: Basic Books.
- Tuschling, Anna/Bernhard J. Dotzler/Andreas Sudmann (2023): »Dialog über den Versuch, eine medienhistorische Passage zu dokumentieren«, in: dies. (Hg.): *ChatGPT und andere ›Quatschmaschinen‹. Gespräche mit Künstlicher Intelligenz*, Bielefeld: transcript, S. 263–282.

Sprachassistenzsysteme und ihre Interfaces

Eine medienlinguistische Analyse

Tim Hector

Abstract *Der vorliegende Beitrag untersucht aus einer medienlinguistischen Perspektive Interfaces von Sprachassistenzsystemen – und zwar sowohl die stimmbasierten Ein- und Ausgaben (Voice User Interfaces, VUIs) als auch darauf aufbauend deren Dokumentation in verknüpften Smartphone-Apps (Graphical User Interfaces, GUIs). Aus der kombinierten Interface-Analyse ergeben sich erstens Erkenntnisse über das Verhältnis zwischen den stimmbasierten Interfaces und dem Sprachgebrauch der Anwender*innen. Konkret wird ein sprachlich-sequenzielles Muster für Dialoge mit VUIs beschrieben und aufgezeigt, dass die Anwender*innen diese situativ in die soziale Praxis einbinden. Zweitens ergeben sich Einsichten über hintergründige Mechanismen zur Verdattung und Verarbeitung sprachlicher Praktiken. Die Dokumentation der Dialoge in visuellen Oberflächen erweist sich als fragmentarisch und zeigt auf, dass die Aufzeichnungen für eine maschinelle Weiterverarbeitung vorbereitet werden, die nur teilweise am sequenziellen Ablauf der Stimmein- und -ausgaben orientiert ist.*

Keywords *Voice User Interfaces; Interface Studies; Intelligente Persönliche Assistenten; Sprachassistenten; Mensch-Maschine-Interaktion; Machine Learning; Medienlinguistik*

1. Einleitung¹

Menschen und Maschinen sprechen miteinander: Die Verbreitung von Sprachassistenzsystemen wie »Alexa« von Amazon oder Internet-Anwendungen wie ChatGPT ist dabei nur die jüngste Spitze einer längeren Entwicklung: Sprachdialogsysteme, Chatbots und Sprachassistenten weisen eine teils weit zurückreichende Mediengeschichte auf (vgl. etwa Baranovska/Höltgen 2018; Volmar 2019) und wurden historisch durch unterschiedliche technologische Verfahren ermöglicht (vgl. Bender/Koller 2020: 5192). Diese Anwendungen haben gemeinsam, dass sich ihre User Interfaces (Benutzer*innen-Schnittstellen) an Praktiken zur Organisation

1 Für hilfreiche Anregungen und Kommentare zu früheren Fassungen dieses Texts danke ich Benedikt Merkle, Niklas Strüver und Didem Leblebici sowie den anonymen Reviewer*innen.

zwischenmenschlicher Interaktionen orientieren. Die Maschinen eröffnen damit einen Raum für Interface-Dialoge, d.h. sprachbasierte Ein- und Ausgaben zur Steuerung der Maschine, wobei diese seitens der Maschine synthetisch generiert sind. Diese können in den genannten Anwendungen schrift- oder stimmbasiert erfolgen. Zugleich stehen diese Interface-Dialoge nicht für sich; vielmehr sind sie angebunden an unterschiedliche visuelle Bildschirm-Umgebungen (Sehflächen), die den Dialog selbst mit hervorbringen können, ihn evtl. beeinflussen, oder ihn dokumentieren und archivieren. Bei Chatbots können dies z.B. das Chatfenster, die visuelle Umgebung einer Webseite oder die Dokumentation des bisherigen Chatverlaufs sein. Bei Sprachassistenzsystemen, die im Zentrum des folgenden Beitrags stehen, stellt die verknüpfte Smartphone-Anwendung eine solche Umgebung dar, in der die Aktionen dokumentiert und das System gesteuert werden kann.

Der vorliegende Beitrag widmet sich den Interfaces stationärer Sprachassistenzsysteme. Die Systeme bestehen erstens aus einem Smart Speaker, einem Geräteverbund aus Lautsprecher, Mikrofon und Rechenmodul – u.a. zur Herstellung einer Internetverbindung und zur Erkennung des Aktivierungsworts, z.B. »Alexa«. Wird dies erkannt, wird der nachfolgende Input aufgezeichnet und zur Auswertung an einen Cloud-Server gesendet (vgl. Hoy 2018: 82). Die Auswertung erfolgt bei den untersuchten Systemen – anders als bei Anwendungen, die auf *Large Language Models* (LLMs) basieren – nicht rein probabilistisch, sondern funktionsbasiert (vgl. Hoy 2018: 83f.). Durch *Natural Language Understanding* (NLU) werden dabei in den aufgezeichneten und durch *Speech Recognition* transkribierten Äußerungen der Nutzer*innen zunächst einzelne Anwendungen (*Skills*) erkannt und der Input daraufhin nach anwendungsspezifischen Eingaben abgesucht (vgl. Bedford-Strohm 2017: 487f.). Beispielsweise wird, nachdem die Timer-Funktion erkannt wurde, die Spezifikation der Dauer des Timers erwartbar. Die anschließend produzierte Stimmausgabe (*Natural Language Generation*) ist entsprechend ebenfalls entlang bestimmter Parameter, der Funktionen und ihrer Spezifikationen, strukturiert. Mit den Geräten ist außerdem eine Smartphone-App verknüpft, die zur Einrichtung und Bedienung des Smart Speakers notwendig ist und weitere Funktionen ermöglicht. Der Beitrag untersucht Dialoge mit den Systemen erstens auf der Ebene der stimmbasierten Ein- und Ausgabe (*Voice User Interface, VUI*) und davon ausgehend zweitens auf der Ebene der Dokumentation dieser Dialoge in der App-Umgebung (*Graphical User Interface, GUI*).

Wie die knappe Beschreibung von Sprachassistenzsystemen schon andeutet, sind die verschiedenen maschinellen Prozesse, die Rechenoperationen und die daran beteiligten Entitäten, die für das Funktionieren von Alltagstechnologien notwendig sind, enorm komplex und nicht nur für die Anwender*innen zunehmend unverständlich und unzugänglich (vgl. Hadler/Haupt 2016a: 7). Diese Unzugänglichkeit entsteht, so Hadler/Haupt weiter, zum einen materiell durch das Verstecken von Kabeln und Drähten in materiellen Gehäusen, die selbst wiederum Schauplatz

von Zeichenprozessen sein können (vgl. Bartz et al. 2019).² Zum anderen entstehe durch die Vernetzung der Geräte untereinander und »Ubiquitous Computing« zunehmend eine Auflösung von Technologie in der Umgebung: »The technology is not only boxed in, but also dissolves into the environment« (Hadler/Haupt 2016a: 7). Es lässt sich ergänzen, dass auch softwareseitig Rechenoperationen und technische Prozesse durch Maskierungen in »übersichtlichen« Eingabemasken für die Endnutzer*innen nicht (mehr) sichtbar sind. Sprachassistenzsysteme stehen dabei unter zwei Aspekten im Fokus – als »Zentrale« für Smart-Home-Anwendungen, die die Steuerung vernetzter Geräte orchestriert (vgl. Strüver 2023a) und als stimmbasierte Technologie mit hintergründigen Verdattungs- und Auswertungsmechanismen.

Der Beitrag widmet sich insbesondere deshalb den User Interfaces: Dies verspricht, die praktische Funktionsweise zunehmend opaker Alltagstechnologien im Verhältnis zu den situierten alltäglichen Praktiken ihrer Anwender*innen einerseits und im Verhältnis zu digitalen Infrastrukturen andererseits aufzuschlüsseln zu können (vgl. Hadler/Haupt 2016a). Über die Interfaces besteht ein Zugriff auf das Verhältnis von Mensch und Maschine bzw. auf die Verhältnisse verschiedener maschineller Einheiten untereinander – und somit auf die Relationen von (alltäglichen) Praktiken und Infrastrukturen (vgl. Kaerlein 2020). Daher unternehme ich den Versuch einer Interface-Analyse für Sprachassistenten und folge insofern dem Appell Kaerleins (2020: 54): »Follow the interfaces«. In Anlehnung an die Losung der Akteur-Netzwerk-Theorie – »Follow the Actors« (Latour 2005: 12) – formuliert Kaerlein (2020: 54) mit diesem Credo den Vorschlag, »die Art und Weise von Vermittlungsprozessen zwischen den beteiligten menschlichen und nicht-menschlichen Entitäten« zu fokussieren. Dazu untersucht der Beitrag die Interfaces der Sprachassistenzsysteme detailliert auf einer empirischen Grundlage.

Interfaces sollen im Folgenden mit Hookway (2014: 59) als *Grenzfläche* in der Begegnung zweier verschiedenartiger Entitäten konzeptualisiert werden (siehe Abschnitt 2). Für VUIs sind dies ein sprechendes, menschliches Subjekt und eine cloudbasierte Recheneinheit, die als synthetische Stimme repräsentiert wird.³ Die Grenzfläche im Fall der VUIs in Sprachassistenten ist als gesprochensprachlicher Dialog gestaltet – daher ist im Folgenden von *VUI-Dialogen* die Rede. Teilweise werden diese durch Lichtsignale auf der materiellen Oberfläche der Smart Speaker ergänzt. Interfaces konstituieren sich also in der Praxis erst im VUI-Dialog. Im Fall

-
- 2 Im Fall der Smart Speaker wird die Oberfläche etwa zur Platzierung von zusätzlichen Steuerungstasten und als Träger von Lichtzeichen zur Anzeige des Gerätestatus⁴ funktionalisiert (siehe Abschnitt 2).
 - 3 Der Eindruck einer »Einheit« wird nur auf der bedienbaren Oberfläche suggeriert. Die dahinterliegenden Prozesse erfordern hingegen die Beteiligung einer unüberschaubaren Anzahl weiterer Einheiten (vgl. Crawford/Joler 2018); siehe dazu Abschnitt 2.

der GUIs für Sprachassistenzsysteme – d.h. in der App-Umgebung – ist die Grenzfläche anders gestaltet, nämlich als Bildschirm-Sehfläche mit der Möglichkeit zur touchbasierten Nutzer*innen-Eingabe und der visuellen und teilweise auditiven Ausgabe. Ein konversationell gestalteter Austausch zwischen Nutzer*innen und Maschine ist hier nicht vorgesehen. Die kombinierte Analyse beider User Interfaces für Sprachassistenzsysteme soll die Praktiken zur Ausgestaltung der Grenzflächen sichtbar machen.

Als Analysematerial stehen dabei erstens Video- und Audio-Aufnahmen vom praktischen Umgang mit Smart Speakern zur Verfügung (siehe Abschnitt 3). Zweitens betrachte ich die Smartphone-Anwendung als Teil des Sprachassistenzsystems, mit einem Fokus auf darin hinterlegte Einträge von zuvor über den Smart Speaker durchgeführten Aktionen. Damit fokussieren die Analysen (Abschnitt 4) auf Prozesse, die typischerweise dem *front end* zugerechnet werden: (a) die Stimm- ein- und -ausgabe sowie entsprechende Lichtsignale, die Smart Speaker senden, als primär stimmbasiertes Interface, (b) die Smartphone-App als GUI mit eingebetteten Audio-Bestandteilen, und insbesondere als Rahmung für die Entstehung eines Nutzungsprotokolls sowie als Berührungspunkt von VUI und GUI.

2. Interfaces in Sprachassistenzsystemen

Interfaces erfahren in den letzten Jahren eine neue Aufmerksamkeit u.a. in medienwissenschaftlich ausgerichteten Software Studies, um digitale und vernetzte Technologien genauer zu untersuchen, ihren Charakter zu verstehen und einer Kritik zugänglich zu machen (vgl. u.a. Hadler/Haupt 2016a; Distelmeyer 2020; Kaerlein 2020; Ernst/Bächle 2020; Dieter 2022). Die Spezifik *stimmbasierter* Interfaces, wie sie in Sprachassistenzsystemen zum Einsatz kommen, liegt darin, dass ein sequenziell organisierter und zudem auf dem akustischen Kanal prozessierter Austausch zwischen Mensch und Maschine vollzogen wird – dieser Dialog gestaltet das aus, was bereits zuvor als »Grenzfläche« beschrieben wurde, d.h. die Ein- und Ausgaben von zwei verschiedenartigen Einheiten. Dieses Verständnis schließt an die etymologischen Ursprünge des Interface-Begriffs in den Naturwissenschaften an: Der Chemiker James Thomson bezeichnet damit die Grenzfläche, die beim Kontakt zweier unterschiedlicher Flüssigkeiten entsteht (vgl. Hookway 2014: 59f.).

Diese Konzeptionalisierung stellt auf die Dynamik in der Begegnung zweier unterschiedlich beschaffener »Einheiten«, Systeme oder Ordnungen ab. Damit konturieren sich Interfaces vielmehr erst dann, wenn die menschliche und die maschinelle Einheit oder »Ordnung« in der Praxis aufeinandertreffen (vgl. Hookway 2014:

45; Wirth 2016: 29).⁴ Für VUIs ist dieses Aufeinandertreffen in Form menschlicher und maschinell-synthetischer gesprochen sprachlicher Dialoge gestaltet. Ein VUI-Dialog konstituiert insofern genau diese Grenzfläche zwischen zwei ›Ordnungen‹ – der menschlichen und der maschinellen – und gestaltet sie aus. Die ›Einheit‹ des Maschinellen wird dabei v.a. durch die synthetische Stimme suggeriert, während tatsächlich eine Vielzahl von Komponenten am Zustandekommen einer gelungenen Operation mit einem Sprachassistenzsystem beteiligt sind (vgl. Crawford/Joler 2018). Die synthetischen Stimmen *repräsentieren* allerdings die dahinterliegenden, technischen Prozesse, wie Natale/Cooke (2021: 1009) konstatieren: »From a technical viewpoint, there isn't anything like one monolithic ›Alexa‹ or ›Siri‹. Rather than being individual entities, they are the integration of a wide range of different systems and algorithms«. Natale und Cooke ziehen Parallelen zu Metaphern, die in GUIs zum Einsatz kommen (z.B. ›Mülleimer‹ oder ›Ordner‹), die dazu vorgesehen sind, das Interface für die Nutzer*innen zu strukturieren (vgl. ebd.). Mit ihrem Namen, einer menschenähnlichen Stimme und der Simulation einer Persönlichkeit tragen die synthetischen Stimmen mit semiotischen Mitteln zur Steuerung der Dialoge zwischen Mensch und Maschine bei (vgl. ebd.: 1009ff.).⁵

Borbach (2019: 18f.) stellt mit Bezug auf akustische Interfaces fest, dass sich die medienwissenschaftliche Interface-Forschung überwiegend auf das Visuelle konzentriert hat, stellenweise unter Einbezug von Haptik. Interfaces hingegen, die sich über einen akustischen Kanal konstituieren, standen nicht im Fokus (siehe aber Borbach et al. i.E.). Dies ist insbesondere deswegen ein Desiderat, weil Interfaces als »Natural User Interfaces« geplant und konzipiert werden, die möglichst wenig auffallen, sich in das Gefüge von Praktiken einweben sollen und dabei den gesamten Körper durch zunehmende Multimodalität – u.a. Gesten oder eben gesprochene Sprache – ansprechen (vgl. Ernst/Bächle 2020: 416). Durch die Verwendung von gesprochener Sprache in Sprachassistenzsystemen ist diese Körperlichkeit besonders deutlich: Laute, die mit dem Stimmapparat des Menschen erzeugt werden, werden transkribiert und einer technischen Auswertung und Verdatung zugänglich gemacht.

VUI-Dialoge unterliegen dabei bestimmten Restriktionen, die sie von zwischenmenschlichen Interaktionen unterscheiden. Der Sprachgebrauch in VUI-Dialogen ist insofern für die maschinelle Verarbeitung vorbereitet und muss bestimmte Bedingungen erfüllen, damit die Begegnung zwischen Mensch und Maschine erfolgreich verlaufen kann: Analysen von VUI-Dialogen zeigen, dass

-
- 4 Zwar ist der Beitrag auf User Interfaces konzentriert, es muss jedoch betont werden, dass Schnittstellen und Grenzflächen auch jenseits davon entstehen – überall da, wo verschiedenartige Systeme aufeinandertreffen und ein Austausch vermittelt wird (vgl. Bratton 2016: 360); siehe auch die Typologie bei Cramer/Fuller (2008: 149).
- 5 Siehe dazu auch das erste Beispiel in Abschnitt 4.

Nutzer*innen sich nicht nur an gesprächsorganisatorischen Praktiken (wie z.B. sequenzieller Organisation oder Turn-Taking) orientieren, sondern auch an Annahmen über die Möglichkeiten und Grenzen der Technologien (vgl. Habscheid 2022; siehe auch Ernst 2017: 100). Konkret wirken dabei die Verarbeitungsfähigkeiten der Geräte stabilisierend auf die Ausbildung sprachlicher Praktiken als Interface-Praktiken: Es zeigt sich in Längsschnitt-Betrachtungen, dass Nutzer*innen im Verlaufe der ersten Wochen nach der Ersteinrichtung des Geräts in Abhängigkeit von der sprachlichen Gestaltung der Stimmeingabe erfahrungsbasiert höhere Erfolgsquoten bei der Nutzung von VUIs erzielen. Dabei ersetzen die Nutzer*innen z.B. bestimmte Satzarten durch andere und passen ihren Sprachgebrauch auf diese Weise an die Geräte an (vgl. Barthel/Helmer/Reineke 2023: 7). Hier ist die medienlinguistische Untersuchung der Interfaces wertvoll, um die Anatomie dieser Praktiken beschreiben zu können.

Auch die Sehfläche in der verknüpften Smartphone-App, einem GUI, stellt eine Grenzfläche her und repräsentiert vollzogene VUI-Dialoge in einer primär visuellen Darstellungsweise. Die Interface-Analyse betrachtet – ausgehend von VUI-Dialogen – beide Interfaces kombiniert und unternimmt den Versuch,

»Verknüpfungen zwischen Praktiken und Infrastrukturen herzustellen, also nach den Umschlagpunkten und Vermittlungsschritten Ausschau zu halten, an denen aus Praxis Daten generiert werden bzw. Daten sich wiederum in Ketten von Praktiken übersetzen. [...] Jede Übersetzung *zwischen* und *innerhalb* von Praktiken und Infrastrukturen lässt sich daraufhin befragen, nach welchen Regeln sie vollzogen wird, wie sie für die beteiligten Entitäten repräsentiert wird, und welche Interfaces in welcher Weise noch an diesem Prozess beteiligt sind.« (Kaerlein 2020: 54)

Mit der Konzentration auf Interfaces als Gegenstand der Analyse lassen sich also situierte Praktiken auf der einen mit Prozessen der infrastrukturellen Dimension auf der anderen Seite verbinden (vgl. Distelmeyer 2020: 62; Kaerlein 2020: 50). Mit dem Fokus auf Interfaces in der linguistischen Analyse kommt in den Blick, wie der Vermittlungsprozess zwischen den beteiligten ›Ordnungen‹ sprachlich gestaltet ist, wie die sprachlichen Äußerungen dokumentiert werden und somit, wie die Stellen beschaffen sind, an denen situierte, sprachliche Praktiken zu Interface-Praktiken werden und in der Folge verdatet und infrastrukturiert werden.

3. Datengrundlage und Methode

Die nachfolgend analysierten Daten wurden im Rahmen des Projekts »Un/erbettene Beobachtung in Interaktion: Intelligente Persönliche Assistenten (IPA)« am Son-

derforschungsbereich Medien der Kooperation an der Universität Siegen⁶ erhoben. Dazu wurde in insgesamt acht Haushalten – darunter Wohngemeinschaften, Paare und Single-Haushalte – unterschiedliches Material erfasst. Hierzu zählen drei für die folgende Auswertung relevante Datentypen: (1) Videoaufzeichnungen von der Inbetriebnahme und Ersteinrichtung eines Smart Speakers⁷ in der Wohnumgebung, (2) Aufnahmen von der routinierten Nutzung in zwei verschiedenen Phasen von zwei bis vier Wochen und (3) Ausschnitte der in der Smartphone-App einsehbar Protokoll Daten.

Die *Videoaufzeichnungen* fertigten die Studienteilnehmer*innen selbst mit von den Forscher*innen bereitgestellten Kameras an. Die Forscher*innen waren während der Aufzeichnungen selbst nicht anwesend. Die *Audioaufzeichnungen* konnten mit Hilfe eines sogenannten »Conditional Voice Recorders« (CVR) erhoben werden.⁸ Dieser zeichnet die Wohnumgebung fortlaufend akustisch auf, löscht aber auch das aufgezeichnete Audio-Material nach 180 Sekunden wieder, sodass stets drei Minuten im Zwischenspeicher verbleiben. Erkennt der Recorder durch seine eingebaute Spracherkennung eines der Aktivierungsworte der untersuchten VUIs (»Alexa«, »Hey/Okay Google« und »Hey Siri«), werden die drei Minuten gespeichert und drei weitere Minuten aufgezeichnet. Wird ein weiteres Aktivierungswort während der Aufzeichnungsphase erkannt, verlängert sich das Recording entsprechend. So ergeben sich Audio-Aufnahmen, in deren Verlauf das Aktivierungswort fällt, sodass die Anbahnung eines VUI-Dialogs und die anschließende kommunikative Bearbeitung analysierbar wird. Zusätzlich wurden die in den Smartphone-Apps hinterlegten *Protokoll Daten* erhoben, d.h. Dokumentationen zuvor durchgeführter VUI-Dialoge. Alle drei Datentypen wurden nach dem gesprächsanalytischen Standard GAT 2 (vgl. Selting et al. 2009) bzw. dem multimodalen Transkriptionsstandard nach Mondada (2016) transkribiert und anschließend inventarisiert.⁹ Insgesamt kann auf rund zwei Stunden Video-Material sowie rund 32 Stunden Audio-Material zurückgegriffen werden.

Die Audio-Aufzeichnungen werden mit einem gesprächsanalytischen Grundansatz untersucht. Damit ist gemeint, dass eine sequenzanalytische Vorgehensweise bei der Auswertung der Aufzeichnungen die Basis der Untersuchung bildet (vgl. Deppermann 2008: 53). Damit werden Muster in der Gestaltung von VUI-Dialogen erkennbar. Zugleich wird die Sequenzanalyse in den Dienst der medienlingu-

6 Gefördert durch die Deutsche Forschungsgemeinschaft (DFG) – Projektnummer 262513311 – SFB 1187 Medien der Kooperation.

7 Dabei wurden Modelle von Amazon (Amazon Echo Dot, 3. und 4. Generation), Apple (HomePod Mini) und Google (GoogleHome Mini und Google Nest) eingesetzt.

8 Der CVR wurde erstmals von einer Forscher*innengruppe an der University of Nottingham entwickelt (vgl. Porcheron et al. 2018) und für das Forschungsprojekt in Siegen erweitert (vgl. Hector et al. 2022).

9 Für die Arbeit an den Transkripten danke ich Franziska Niersberger-Gueye und Sarah Diehl.

istischen Erforschung von sprachlichen Praktiken mit VUIs gestellt. Zusätzlich zu den analysierten Ausschnitten werden daher im Sinne einer ethnografischen Erweiterung gesprächsanalytischer Verfahren (vgl. Deppermann 2000) Hintergrundinformationen über die Sprecher*innen (z. B. deren Beziehung zueinander, deren Interaktionsgeschichte, deren Erfahrungen im Umgang mit Sprachassistenzsystemen u. a.) mit erhoben und teilweise in die Analysen einbezogen.¹⁰ Im Falle der Untersuchung der Protokoll Daten ist zwar immer noch die zur Verfügung stehende Aufzeichnung Ausgangspunkt der Analyse, allerdings wird diese hier in Beziehung zur grafischen App-Umgebung analysiert. Für Letzteres wird ausgehend von den sprachlichen Praktiken der Nutzer*innen deren weitere Dokumentation und Repräsentation in den Apps betrachtet. In diesem Sinne weist das methodische Vorgehen auch zu Teilen Merkmale der »Walkthrough«-Methode auf, in der systematisch dem Aktivitätsfluss einer Anwendung gefolgt und dies zugleich dokumentiert wird (vgl. Light et al. 2018: 882).¹¹ Die Auswahl der präsentierten Daten hat dabei in allen Fällen exemplifizierenden Charakter, die Analyse ist mithin qualitativ-explorativ zu verstehen.

4. Interface-Analyse

Die Analyse ist zweigeteilt und konzentriert sich im ersten Teil auf die VUI-Dialoge und ihre Entstehung. Dabei stehen die folgenden Leitfragen im Mittelpunkt: Wie sind VUI-Dialoge konversationell aufgebaut, d. h. wie nehmen die Nutzer*innen sprachlich Bezug auf ein maschinelles Gegenüber und wie sind die synthetisch generierten stimmlichen Äußerungen als Bezugnahmen auf die Eingaben gestaltet? Welche spezifischen sequenziellen Muster lassen sich dabei beschreiben? Lassen sich Verwebungen der Dialoge mit den sozialen Situationen beobachten, in denen sie ausgeführt werden? Der zweite Teil wendet sich der Dokumentation in der Smartphone-App zu: Wie sind »Aktionen« mit einem VUI, d. h. Stimm- und -ausgaben, im GUI dokumentiert und repräsentiert? Welche Bestandteile eines VUI-Dialogs sind dabei inkludiert und exkludiert, welche Zusatzinformationen werden bereitgestellt? Welche Rückschlüsse erlaubt dies über die seitens der Dienstleister vorgenommene Datenverarbeitung und -verwertung?

10 Zu methodologischen Überlegungen für die Analyse von Smart Speakern in der Praxis siehe Hector (2022).

11 Die Methode ist allerdings v. a. für die Dokumentation und Kritik von GUIs, insbesondere Smartphone-Apps und deren Interfaces, etabliert. Da hier VUIs und die damit verbundenen sprachlichen Praktiken als Ausgangspunkt der Analyse gesetzt werden und zudem teilweise gesprächsanalytisch gearbeitet wird, ergibt sich eine etwas andere Vorgehensweise, die z. B. keine vollständige App-Dokumentation beinhaltet.

4.1 Sprachliche Praktiken in Voice User Interfaces

Um zunächst die seitens des VUI-Designs angelegten ›Optionen‹ genauer zu betrachten, lohnen die Eröffnungsdialoge, die bei der Ersteinrichtung eines Smart Speakers abgespielt werden, einer genaueren Betrachtung. Das nachfolgend präsentierte Beispiel (1) stammt aus einem Haushalt, in dem die 21-jährigen Zwillingbrüder Till (TW) und Konrad (KW) als Wohngemeinschaft zusammenleben. In dem Ausschnitt nehmen sie den HomePod von Apple in Betrieb.

Beispiel (1): Einführungsdialog

```

389 SI: hall[o ich] bin sIri;
390 TW:      [(h?) ]
391 p:      (0.2)
392 SI: willkommen beim HOMEpod.
393 p:      (0.4)
394 SI: du sIEHST es zwar nicht (.) aber ich wInke gerade,
395 KW:      ((lacht, ca. 2 Sek.))
396 SI: wenn du meine aufmerksamkeit möchtest (.) sag eINFach-
397 p:      (0.2)
398 SI: hey sIri-
399 p:      (0.9)
400 SI: lass uns LOSlegen (.) sAg.
401 p:      (0.3)
402 SI: hey sIri (.) was KANNST du?
403 p:      (0.5)
404 KW:      h
405 p:      (0.5)
406 KW: hey sIri (.) was KANNST du?
407 p:      (1.3)
408 KW:      hh
409 p:      (0.7)
410 SI: ich kann VIELE dinge erled[ig]en(.) zum beispiel das lIcht
          einschalten (.) die aktuel[le]n nAChrichten abrufen (.) und dir
          sagen wie das wEtter wird;
411 KW:      ] [<lachend> he he he DANke.>
          ]
412 p:      (0.7)
413 SI: probier_s jetzt mal MIT (.) hey siri spiel musik,
414 p:      (0.9)
415 KW: h hey siri (.) spiel muSik;
416 p:      (2.1)
417 KW:      hh
418 p:      (0.4)
419 SI: gerne hier kommt muSik.
420 p:      (0.3)
421 SI: [ ex]tra für dich zuSammeng[est]ellt;]
422 KW:      [HM.]
423 TW:      ] [OH. ]
424 p:      (0.3)
425 TW: AHÄ,
426 p:      (3.2)
427 SI: ((spielt „Copacabana“ von Leon Machère, ca. 10 Sek.))

```

Der Ausschnitt zeigt die ersten verbalsprachlichen Äußerungen des eingerichteten Smart Speakers. Die synthetisierte Stimme, die die maschinelle Seite des Interfaces darstellt, präsentiert sich – durch ihre onymische, d.h. namentliche, Selbstvorstellung (Z. 389) und durch den Verweis auf einen unsichtbaren menschlichen Körper (Z. 394) – als menschenähnliche Einheit, die adressiert werden kann. Damit wird die Metaphorik eingeführt, die in VUIs zur Strukturierung des Interfaces eingesetzt wird (vgl. Natale/Cooke 2021: 1009; siehe auch Abschnitt 2). Diese Strukturierung des stimmbasierten Interfaces und das Einlernen in diese Metapher setzt sich im weiteren Verlauf des Ausschnitts weiter fort: Die synthetische Stimme ä-

ßert eine mögliche Stimmeingabe und fordert die Nutzer auf, dies nachzumachen (Z. 400ff.), was Konrad auch entsprechend realisiert (Z. 406). Diese ›Übung‹ wird ein weiteres Mal wiederholt (Z. 413ff.). Dabei wird ein bestimmter sequenzieller Ablauf geprobt: eine Adressierung durch eine Interjektion und das folgende Onym »Siri«, gefolgt zunächst von einer kurzen Pause, in der das Gerät den Listening-Modus aktivieren kann und die nachfolgenden akustischen Signale verarbeitet, und anschließend von der Stimmeingabe, auf die wiederum eine synthetisierte Stimmausgabe erfolgt. Dieser Sequenzablauf lässt sich wie folgt darstellen (vgl. Hector i.V.):

Invokation [Interjektion + Onym] – Aktivierung des Listening-Modus – Stimmeingabe – Stimmausgabe

Bei der zweiten ›Übung‹ im oben dargestellten Beispiel folgt auf die Stimmausgabe noch die praktische Umsetzung der in der Stimmeingabe angeforderten Tätigkeit. Die Stimmausgabe erscheint insofern eher als dialogisches Scharnier zwischen der Stimmeingabe und der praktischen Umsetzung (z.B. Wiedergabe von Musik):

Invokation [Interjektion + Onym] – Aktivierung des Listening-Modus – Stimmeingabe – Stimmausgabe-Scharnier – Praktische Umsetzung

Dieser Sequenzablauf scheint auf den ersten Blick selbstverständlich zu sein, ist aber nicht trivial, denn dabei verlängern sich Praktiken der zwischenmenschlichen Gesprächsorganisation ins VUI. So ist der Ablauf der Eröffnung eines Austauschs – für den VUI-Dialog als Invokation und anschließende Aktivierung des Listening-Modus beschrieben – für zwischenmenschliche Interaktionen umfangreich als *Summons-Answer-Sequenz* diskutiert worden (vgl. Schegloff 1968). Dabei ist noch eine Parallele auffällig: Der Apple HomePod signalisiert durch ein visuelles Signal an seiner Oberfläche, dass der Listening-Modus aktiviert wird, also die nachfolgenden akustischen Signale als Stimmeingaben verarbeitet werden. In der von Schegloff (vgl. ebd.) beschriebenen *Summons-Answer-Sequenz* für zwischenmenschliche Interaktionen können ebenfalls visuelle Signale eine Reaktion auf einen ›Summons‹, also eine erste Adressierung sein – z.B. ein Blick oder eine veränderte Körperhaltung. Die Flexibilität im Hinblick auf die sprachliche Realisierung der Invokation in VUI-Dialogen ist allerdings eingeschränkt: Diese muss im VUI-Dialog immer mit der Interjektion »Hey« sowie dem Onym »Siri« realisiert werden, die Möglichkeit eines sprachlich anders gestalteten *Summons* besteht (anders als in zwischenmenschlichen *Summons-Answer-Sequenzen*, die hierfür eine Vielzahl verschiedener verbaler und nonverbaler Realisierungsmöglichkeiten ausgebildet haben) aufgrund der technischen Gegebenheiten nicht.

Das VUI erfordert für sein erfolgreiches Zustandekommen als Dialog außerdem eine strikte Abfolge von Redezügen, die hier erprobt und vorgemacht wird. Die Redezugorganisation, die in der konversationsanalytischen Forschung bekannt ist als *Turn-Taking* (vgl. Sacks/Schegloff/Jefferson 1974), wird in zwischenmenschlichen In-

teraktionen von kompetenten Sprecher*innen mühelos und »mit schlafwandlerischer Sicherheit und Präzision« (Auer 2020: 106) vollzogen – auch dann, wenn z.B. auf dem konversationellen *floor* eine Auseinandersetzung darüber stattfindet, welche*r der Sprecher*innen als nächstes am Zug ist. Das Rederecht in VUI-Dialogen ist aber nicht wie in Alltagsgesprächen Gegenstand einer fortlaufenden Aushandlung, sondern ist vielmehr – ähnlich wie bei institutionellen Zuweisungen (vgl. Clayman 2012) – strikt reguliert: Erst nach der Invokation und der Aktivierung des Listening-Modus kann nutzer*innenseitig gesprochen werden, und zwar bis zum Beginn der Verarbeitung, daraufhin folgt die Stimmausgabe. Wird von dieser Abfolge abgewichen, droht der VUI-Dialog zu scheitern und das konstituierte Interface folglich zusammenzubrechen. Das Zustandekommen des Interfaces ist insofern an spezifische und fragile konversationelle Bedingungen geknüpft, die sich aus dessen akustisch-sequenziellem Charakter ergeben. Diese bauen auf gesprächsorganisatorischen Praktiken auf, die jedoch durch den Gebrauch in einem VUI neu konfiguriert werden: Der rigide sequenzielle Ablauf einschließlich der Notwendigkeit einer Invokation vor jeder Stimmein- und -ausgabe sowie die Restriktionen etwa in der Lexemwahl, im syntaktischen Aufbau (vgl. Barthel/Helmer/Reineke 2023) und im Turn-Taking, bilden spezifische Interface-Bedingungen, die VUI-Dialoge maßgeblich von zwischenmenschlichen Interaktionen unterscheiden. Das Interface konstituiert sich also genau in dem entstehenden Korridor, dem Zusammenspiel zwischen den technologischen Vorgaben, Möglichkeiten und Grenzen und dem gesprächsorganisatorischen Wissen der Nutzer*innen.

Wie sich in Beispiel (1) ebenfalls zeigt, sind VUI-Dialoge an die sozialen Situationen angebunden, in denen sie entstehen. Hinweise darauf sind etwa das Gelächter (Z. 395) und das sequenzielle Schließen der Sequenz etwa durch Responsive und Erkenntnisprozessmarker (u.a. Z. 411, 422f.), die zwar einerseits auf der sprachlichen Oberfläche einen fortgesetzten VUI-Dialog andeuten, die allerdings bei genauerer Betrachtung auch der Blickverhältnisse und der Körperhaltung¹² vielmehr als an die ko-präsenten menschlichen Interaktionsteilnehmer*innen gerichtet verständlich werden. Teile der Äußerungen in VUI-Dialogen sind insofern nur über die situative, soziale Einbindung, die lokalen Bedingungen ihrer Produktion, erklärbar. Wie das folgende Beispiel zeigt, können Sprecher*innen einen ganzen VUI-Dialog funktionalisieren, um gänzlich andere kommunikative Aufgaben in einer sozialen Situation zu erfüllen:

12 Eine detaillierte multimodale Betrachtung führt hier zu weit; siehe dazu etwa Hector (2022) oder Habscheid et al. (2023: 17f.).

Beispiel (2a): Pudding

```

068 AS: <<nuschelnd> hab so lust auf NACHTisch.>
069 MAN.
070 p: (2.2)
071 SR: ((unverständlich, ca. 3 Sek.))
072 jo KOMM.
073 MACHen;
074 MACHen.
075 AS: (←-) nja is ja AUCH kacke.
076 SR: ne ich MACH dir jetzt n pudding.
077 [<<f>PUDDi:ng;> ]
078 AS: [ich weiß gar nicht] ob wir überhaupt pudding HAben
oder so was.
079 SR: aLEXa?
080 p: (0.7)
081 SR: spiel PUDDing;
082 p: (2.0)
083 AL: pudding von sheef von SPOTify;
084 SR: ((lacht))
085 AL: ((spielt Musik, „Pudding“ von Sheef, bis Z.100))

```

Der Auszug beinhaltet einen VUI-Dialog, der zur Wiedergabe des Musiktitels »Pudding« von Sheef führt (Z. 079–083). Der VUI-Dialog selbst ist dabei sprachlich gestaltet wie in Beispiel (1): Auf die Invokation (Z. 079) folgt die Stimmeingabe (Z. 081). Daraufhin produziert die synthetische Stimme zunächst eine Stimmausgabe mit Scharnierfunktion und anschließend wird der angeforderte Titel wiedergegeben. Der Blick auf das Vorgeschehen zu diesem VUI-Dialog macht diesen allerdings in seinem Kontext als situierte Humoraktivität verständlich: Andrea (AR), die Ehefrau von Sam (SR), äußert im Anschluss an eine weiter zurückliegende Interaktion, dass sie sich Nachtisch wünscht (Z. 068). Nachdem Sam mehrfach eine entsprechende Aktivität ankündigt (Z. 072–077) und Andrea entsprechende Skepsis äußert (Z. 078), realisiert er daraufhin den VUI-Dialog zur Wiedergabe des entsprechenden Lied-Titels. Damit bricht er humorvoll mit seiner ursprünglichen Ankündigung, tatsächlich einen Pudding zu machen. Für sich genommen stellt das Ziel des VUI-Dialogs also die Wiedergabe des entsprechenden Liedtitels dar, im Kontext der Nachtisch-Diskussion ist er aber vielmehr Mittel zum Zweck der Gestaltung eines humorvollen Beitrags. Die spezifischen Bedingungen des Interfaces führen also zu einem gewissen Grad an Vorhersehbarkeit und *Vorführbarkeit*: Durch die Regelmäßigkeit eines VUI-Dialogs und dessen Plan-Basiertheit können sie situiert als Humoraktivität eingesetzt werden (letzteres gilt freilich auch, wenn Plan-Brüche entstehen und Reparaturen notwendig werden oder der VUI-Dialog scheitert, siehe etwa Hector/Hrncal 2024). Durch seine Prozessierung auf dem akustischen Kanal ist der VUI-Dialog auch für Andrea hör- und nachvollziehbar. Somit wird die Aktivität des Abspielens des Liedes (nicht die Liedwiedergabe selbst) zum Humor-Gegenstand. Dies zeigt, dass VUI-Dialoge mit ihrer lokalen, sozialen Umgebung verwoben sind und darin z.B. Humorfunktionen übernehmen können. In dieser Eingebundenheit können sie für die Nutzer*innen auf einer zweiten Ebene anders Sinn erzeugen als lediglich als Dialog zur Bedienung eines Geräts.

4.2 Repräsentationen sprachlicher Praktiken in grafischen Interfaces

Folgt man den Interfaces digitaler Assistenzsysteme weiter, führt von den VUIs der nächste Schritt zu den Smartphone-Anwendungen, die in einer – je nach System unterschiedlich ausgestalteten – komplementären Beziehung zu Smart Speakern stehen. Im Folgenden steht die visuelle Darstellung des Aktivitätenprotokolls eines Smart Speakers von Amazon im Mittelpunkt.¹³ Dieser ›Sprachverlauf‹ arbeitet mit der visuellen Repräsentation gesprochen sprachlicher Eingaben in Form von Transkripten und eröffnet zudem die Möglichkeit einer akustischen Wiedergabe des Aufgezeichneten (vgl. Habscheid et al. 2021). Zugleich ist diese Darstellung ein Hinweis darauf, welche Bestandteile der sprachlichen Äußerungen erfasst und verarbeitet wurden. Die Repräsentation eines Ereignisses entsteht hier auf einer Sehfläche, in die die Audio-Aufzeichnung des erfassten Ausschnitts eingebettet ist. Diese Repräsentation ist ein ›Zerrspiegel‹ der VUI-Dialoge als GUI, das die Begegnungen von Mensch und Maschine noch einmal neu kontextualisiert, unter anderen Bedingungen wiederholbar werden lässt und somit eine neue Begegnung schafft, die über visuelle Repräsentation und Eingaben über den Touch-Bildschirm (oder über den Computerbildschirm und die -maus bzw. -tastatur) vermittelt wird. Die Grenzfläche zwischen Technologie und Mensch wird hier also einer Befragung aus einem zweiten Blickwinkel zugänglich und zeichnet ein Bild der VUI-Dialoge aus der maschinellen Perspektive.

Die Aufzeichnung im Sprachverlauf ist standardmäßig aktiviert, kann jedoch auch unterbrochen oder dauerhaft deaktiviert werden. In dem entsprechenden Menü finden sich die VUI-Dialoge »als verschriftete isolierte Einzelemente ohne die Situation, in der sie realisiert werden« (Habscheid et al. 2021: 40). Betrachten wir als Beispiel, wie hier die »Aktivität« aus dem Beispiel (2a) (»Pudding«) weiter oben dokumentiert ist:

13 Für eine detailliertere Beschreibung der App-Oberfläche siehe Habscheid et al. (2021: 18–21).

Beispiel (2b): Pudding (Aktivitätenprotokoll)

Aktivität	Sprachverlauf
Audio war nicht für Alexa gedacht	7. Juli 2021 20:12 Echo
<i>"spiel pudding"</i>	7. Juli 2021 16:25 Echo
<i>"alexa"</i>	7. Juli 2021 16:25 Echo
Audio konnte nicht verstanden werden	7. Juli 2021 9:28 Echo
Audio konnte nicht verstanden werden	7. Juli 2021 9:28 Echo
<i>"alexa"</i>	7. Juli 2021 9:28 Echo
<i>"spiel alexa von adam angst"</i>	7. Juli 2021 9:20 Echo
<i>"alexa"</i>	

Der Eintrag für die Aktivität ist dabei eingebettet in eine Reihe anderer Einträge im Sprachverlauf. Dabei ist festzustellen, dass die Invokation als eigene Aktivität darunter dokumentiert ist und insofern die Ganzheitlichkeit des VUI-Dialogs in dieser Darstellung aufgebrochen wurde. Nicht nur der situative Kontext, sondern auch der Zusammenhang der eigentlichen sprachlichen Praktik – der gesprächsähnliche Sequenzverlauf – werden in der visuellen Darstellung aufgelöst. Auch Reparaturversuche der Nutzer*innen, die sich über mehr als einen VUI-Dialog erstrecken – z.B. die wiederholte und ggf. prosodisch, syntaktisch oder lexisch angepasste Formulierung einer Stimmeingabe (vgl. Hector/Hrncal 2024) – sind nicht in ihrem Zusammenhang dokumentiert. In der Übersichtsdarstellung ist insofern nicht nur

der sequenzielle Verlauf eines einzelnen VUI-Dialogs nicht zusammenhängend dokumentiert, sondern auch Beziehungen zwischen einzelnen VUI-Dialogen können nicht abgebildet werden, obwohl etwa bei scheiternden und anschließend reparierten Anfragen mehrere Einzel-Dialoge kurz hintereinander auftreten, die sprachlich aufeinander aufbauen (siehe etwa das Beispiel bei Merkle/Hector i.E.).

Die App-Nutzer*innen können sich die Details einzelner Aufzeichnungen durch eine Berührung der Pfeilsymbole am rechten Rand der Sehfläche anzeigen lassen (im Beispiel (2b) ist bereits der Eintrag für »Pudding« ausgewählt und hervorgehoben). Der Eintrag erscheint daraufhin als ein größerer Eintrag zwischen den eingeklappten Einträgen, wie in Beispiel (2c) zu sehen:

Beispiel (2c): Pudding (Aktivitätenprotokoll)



Dieses Menü enthält neben der Wiedergabefunktion eine Transkription des Audios sowie eine ausschließlich *schriftliche* Dokumentation der Reaktion des Sprachassistenten. Außerdem findet sich ein roter Button zur Löschung des Audios. In einem weiteren Abschnitt finden sich Datum, Uhrzeit und das genutzte Gerät. Wiederum darunter besteht die Möglichkeit eines »Feedbacks« über den Sprachverlauf. Der Eintrag ist insofern nicht mehr eingebettet in die soziale Situation der Humoraktivität, die weiter oben beschrieben wurde. Vielmehr werden nun Fragmente des VUI-Dialogs visuell repräsentiert. Auch anschließende Kommentierungen wie im folgenden Beispiel werden dabei getilgt:

Beispiel (3a): Tell me why

```

026 k: ((Klappergeräusche, ca. 9.8 Sek.))
027 AS: <<singend> tell me WHY,
028     AIN_T (nothing),
029     mh_mh_mh_MH_mh;
030     ELI me why,>
031 SR: aLEXa?
032     SPIEL,
033     BACKstreet boys tell me why.
034 P: (2.0)
035 AL: hier ist i want it that WAY,
036     auf amazon Music.
037 SR: ah (.) (so HEIßt das),
038 AS: JA.
039 AL: ((spielt Musik, „I want it that way“, Backstreet
        Boys, bis Z. 046))

```

Der Auszug aus demselben Haushalt von Andrea und Sam zeigt, dass der VUI-Dialog als Reaktion darauf verstanden werden kann, dass Andrea ein Lied von den Backstreet Boys singt, woraufhin Sam dieses über den Smart Speaker wiedergeben lassen will. Dabei formuliert er die Stimmeingabe als Imperativ (Z. 032), gefolgt von Interpret und der Songzeile »tell me why«, die Andrea zuvor gesungen hatte (Z. 029f.). Die Stimmausgabe des Smart Speakers kündigt die Wiedergabe des Titels an, »korrigiert« dabei aber im Stimmausgabe-Scharnier vor dem Abspielen des Lieds die Eingabe zum eigentlichen Titel des Stücks (»I want it that way«, Z. 035). Dies wird von Sam und Andrea kommentiert, bei Sam eingeführt durch einen Erkenntnisprozessmarker (vgl. Imo 2009), der die durch die synthetische Stimme durchgeführte Reparatur quittiert, und von Andrea bestätigt (Z. 037ff.). Das Singen von Andrea, das den VUI-Dialog initiiert, ebenso wie die Reparatursequenz und die Invokation sind nicht im Aktivitätenverlauf dokumentiert:

Beispiel (3b): Tell me why

Die beschriebenen visuellen Elemente für einen einzelnen Eintrag (Beispiele 2c) und (3b) lassen sich noch genauer in ihrer Funktionalität analysieren. Wie Habscheid et al. (2021: 44) feststellen, kann die »angebotene Möglichkeit einer Löschung [...] als Reaktion von Amazon auf die öffentliche Debatte bezüglich der Verwendung von personenbezogenen Daten durch Internetfirmen verstanden werden«. Die rote Hervorhebung und deutliche Platzierung des Buttons unterstützen auch visuell den

Eindruck einer möglichen Kontrolle durch die Nutzer*innen, obschon nicht deutlich wird, was eine Löschung des Audios tatsächlich bewirkt – wie eine Inspektion der App-Funktionen zeigt, bleiben der visuelle Eintrag und die zugehörigen Zusatz-Daten in der App erhalten und zum Zeitpunkt der Ansicht in der App hat eine gewisse Verarbeitung notwendigerweise schon stattgefunden.

Die Möglichkeit des Nutzer*innenseitigen »Feedbacks« ist eine Funktion, mit der der Hersteller die Endkund*innen an der Arbeit am Trainingsdatenmaterial beteiligt, in das die aufgezeichneten Daten anschließend eingehen. Die Funktion ist ein Hinweis darauf, dass das aufgezeichnete Material, die Transkription und die ausgelöste Aktivität in ihrem Zusammenhang weiter für die Optimierung der Spracherkennungsverfahren verwertet werden (siehe dazu ausführlich Strüver 2023b: 15f.), und zwar sowohl durch automatisierte Verfahren wie auch durch menschliche Trainings- und Kontrollprozesse, die im Wesentlichen durch prekär beschäftigte Clickworker*innen durchgeführt werden (vgl. Waldecker/Volmar 2022: 167). Wenn dabei von »redaktionelle[r] Mikroarbeit« (ebd.) die Rede ist, lässt der Blick auf das Aktivitätenprotokoll im Abgleich mit der sozialen Situation verständlich werden, mit welcher Herausforderung die Arbeiter*innen hierbei im Dienst des KI-Trainings konfrontiert werden: Der Rekonstruktion des Verhältnisses fragmentarischer Ton-Aufzeichnungen zu deren Transkript und der dokumentierten »Aktion«.

Die verbal vollzogenen Interface-Praktiken werden also in Datenpunkte und somit in Infrastrukturen übersetzt. Die Ansicht der Aktivitäten in GUIs sind Dokumentation der Verdattung und Plattformisierung¹⁴ des vollzogenen VUI-Dialogs (und somit der sozialen Praxis). An dieser Übersetzung sind zu diesem Zeitpunkt zwei Interfaces beteiligt: Das VUI als akustisch prozessierte Begegnung zwischen Mensch und Computer und das GUI als visuelles Abbild dieser Begegnung, genauer gesagt als Abbild eines Datenbank-Eintrags zur Dokumentation und weiteren Interaktion mit dieser Begegnung. Im Abgleich des Eintrags mit der längeren Audio-Aufzeichnung von der sozialen Praxis, aus der heraus er entstanden ist, werden diese Übersetzungen als »Umschlagspunkte« (Kaerlein 2020: 54) erkennbar, »an denen aus Praxis Daten generiert werden bzw. Daten sich wiederum in Ketten von Praktiken übersetzen« (ebd.). Die Übertragung lässt aus einem sich dynamisch in der Zeitlichkeit entfaltenden Geschehen einen geronnenen und sich primär in der Räumlichkeit ausdehnenden, wiederholbaren Datenpunkt werden, der mit zusätzlichen Informationen (einer Transkription, Datum und Uhrzeit der Anfrage, verwendetes Gerät, daran anschließende Reaktion des Systems) versehen und zur Weiterverarbeitung vorbereitet ist.

14 Zur Konzeption von Sprachassistenzsystemen als Plattform-Technologien siehe Goulden (2019) und Strüver (2023a: 104f.).

4.3 Repräsentationen des Nicht-Verstehens

Die genauere Betrachtung des GUIs zeigt außerdem die Fehleranfälligkeit und -verarbeitung der Systeme, wie das folgende Beispiel aus einem anderen Haushalt zeigt:

Beispiel (4): Licht an¹⁵



Dokumentiert ist hier ein Fall, in dem die Transkription des Audios nicht zur Tonspur passt, was auf Fehler im Speech-Recognition-Prozess hindeutet. Tatsächlich ist in dem Audio undeutlich zunächst der Name eines im Haushalt lebenden Bewohners, Alex, zu hören – wohl der Auslöser für die Fehlaktivierung – und fälschlicherweise als Invokation verarbeitet; anschließend ist eine nicht gut zu verstehende Folgeäußerung dokumentiert:

001 M1: aLEX (hat_n bestimmten gefunden.)

Die daran anschließend produzierte und schriftlich hinterlegte Äußerung – »Tut mir leid, ich kann keine Gruppe oder kein Gerät mit dem Namen licht finden« – bezieht sich entsprechend auf den systemseitig verstandenen Text. Davon abweichend wird im folgenden Beispiel der Fehler systemseitig erkannt und klassifiziert:

15 In diesem und in den folgenden Beispielen wird auf die Gegenüberstellungen zwischen Protokolleinträgen und Aufzeichnungen der sozialen Situation verzichtet, da diese für die Illustration der nachfolgenden Argumente nicht wesentlich sind.

Beispiel (5): Audio war nicht für Alexa gedacht

Eine gesprächslinguistische Transkription des hinterlegten Audios zeigt an, dass in der Tat kein Versuch eines VUI-Dialogs erkennbar ist. Auch hier wird vermutlich durch den Namen des Bewohners Alex die Invokation ausgelöst:

001 M1: aLEX (du hattest das ma in kunst.)

Allerdings wird die anschließende Äußerung nicht transkribiert. Zugleich kann konstatiert werden, dass die Aufzeichnung dennoch verarbeitet und gespeichert wurde. Auch zu diesem Audio besteht die Möglichkeit, es löschen zu lassen oder die Ausführung des »Befehls« zu bewerten. Die zwar einerseits »nicht für Alexa gedachte« und nicht transkribierte Aufnahme wird so dennoch Teil des Sprachverlaufs und einer hintergründigen Verarbeitung. Davon noch einmal zu differenzieren sind Einträge, zu denen eine leicht davon abweichende Klassifizierung hinterlegt ist:

Beispiel (6): Audio konnte nicht verstanden werden

Anstatt eines Transkripts des Audios findet sich hier der Vermerk »Audio konnte nicht verstanden werden«. Eine Transkription der Audio-Aufzeichnung ergibt eine Äußerung, die in einem persönlichen Gespräch oder Telefonat gefallen sein dürfte:

001 M2: dachte mir so (LASS ma_s) training;

Für die Nutzer*innen besteht hier keine Transparenz darüber, warum die Klassifizierung hier anders ausgefallen ist als in Beispiel (5) – offensichtlich war auch dieses Audio »nicht für Alexa gedacht«. Ebenso unklar ist, warum überhaupt zu diesem Zeitpunkt das Recording aktiviert war.

Die Beispiele (5) und (6) belegen, dass Verarbeitungsfehler antizipiert werden und entsprechende Klassifizierungen dafür vorgesehen sind. Die Darstellung im GUI weist insofern darauf hin, dass – in der Begegnung mit der nächsten maschinellen Einheit – auch scheinbar fehlerhafte oder versehentlich entstandene Aufzeichnungen weiterverarbeitet werden können: Sie werden *als Fehler* dokumentiert, klassifiziert und ausgewertet. Aus dieser Perspektive betrachtet, handelt es sich bei den hier dokumentierten Einträgen insofern also nicht um Verarbeitungsfehler, sondern um einkalkulierte (und ethisch problematische) Vorgänge zur zusätzlichen Gewinnung von Audio-Daten. Dies verweist auf den uneindeutigen Zusammenhang zwischen dem Verhalten der Nutzer*innen und dessen Verdattung, den auch Paßmann/Gerlitz (2014) für Social-Media-Plattformen wie Twitter und Facebook besprechen.¹⁶ Sie betonen, dass mit den aufgezeichneten Plattform-

16 Auch Sprachassistenzsysteme können als Plattform-Technologien betrachtet werden, siehe dazu Goulden (2019) und Strüver (2023a: 104f.).

Daten über die Aktivitäten der Nutzer*innen¹⁷ kein vollständiges, umfänglich verständliches Bild eines individuellen Nutzungsverhaltens gezeichnet werden kann – ein Befund, der sich für VUI-Dialoge bestätigen lässt. Sie argumentieren weiter, dass insbesondere soziales Verhalten mit solchen Formen fragmentarischer Daten kaum erfassbar ist: »Es handelt sich zwar um vorstrukturierte Aktivitäten, doch diese sind zugleich unterbestimmt« (Paßmann/Gerlitz 2014: 2). Genau diese »Unterbestimmtheit oder Vagheit« (ebd.: 3) wird im Sprachverlauf sichtbar und ist kein Zufall: Die Gliederung nach einzelnen »Aktivitäten« (nicht nach der tatsächlichen sequenziell organisierten Ausführung des VUI-Dialogs), die nur sehr kurzen und kaum zuzuordnenden Audio-Ausschnitte sowie die unterschiedlichen Klassifizierungen von Nicht-Verstehen legen den Schluss nahe, dass hier die Vorstrukturierung im Sinne einer weiteren Verarbeitung der Aufzeichnungen als Trainingsdaten für die Verbesserung der Dienste u.a. im Bereich der Speech Recognition sichtbar wird. Diese ist auf möglichst vergleichbare und daher kurze Aufzeichnungen als Daten angewiesen, um möglichst effizient Mustererkennung betreiben zu können (vgl. Pasquinelli 2019: 11) – nicht aber auf vollständige Bilder des Nutzungsverhaltens.

5. Fazit und Ausblick

Die Konzeptionalisierung von Interfaces als Grenzfläche nach Hookway (2014: 59) und die darauf aufbauende Analyse von VUI-Dialogen, mit denen diese Grenzfläche ausgestaltet wird, konnte verdeutlichen, dass bestimmte sprachliche Bedingungen erfüllt sein müssen, um in einem VUI-Dialog einen Austausch zwischen Nutzer*innen und Maschine zu etablieren. Beide Einheiten sind dazu auf einen rigiden Sequenzablauf orientiert. Die sprachlichen Praktiken folgen zwar im Grundsatz gesprächsorganisatorischen Prinzipien, unterliegen jedoch bestimmten Restriktionen. Sie sind mithin Interface-Praktiken, d.h. Äußerungen, deren sprachliche Gestaltung nicht nur an gesprächsorganisatorischen Prinzipien ausgerichtet ist, sondern auch Spuren des Gebrauchs im Interface trägt. VUI-Dialoge sind zudem durch die situativen, sozialen Kontexte geprägt, in denen sie stattfinden, und können dabei auch über die Bedienung der Maschine hinaus funktionalisiert werden.

Während die analysierten Audio-Aufzeichnungen also die soziale Situation und die Konstruktion des VUI-Dialogs verständlich werden lassen, macht der Aktivitätenverlauf als GUI die maschinelle Perspektive auf die Situation sichtbar: Es do-

17 An dieser Stelle sind die Aktivitäten der Nutzer*innen gemeint, die im Rahmen bestimmter, von den Betreiberfirmen angebotener und vorstrukturierter Optionen stattfinden, z.B. »Like« und »Share« auf Social Media oder eben ein vordefinierter Ablauf für einen VUI-Dialog bei Sprachassistenzsystemen.

kumentiert, wie sprachliche Eingaben segmentiert und zur weiteren Verarbeitung vorbereitet wurden – und erweist sich dabei als blind für die Situiertheit der Nutzer*innen. Der Vergleich bestätigt insofern einerseits die ethnomethodologische Perspektive Suchmans (2007: 4f.): »human-machine communications take place at a very limited site of interchange; that is, through actions of the user that actually change the machine's state«. Andererseits zeigt sich, dass die cloudbasierte Auswertung und der Bedarf nach Audiodaten, z. B. zum Training der Spracherkennung, eine neue Dimension einbringt: Wie am Schluss gezeigt wurde, sind hier eben nicht nur diejenigen Aktionen dokumentiert, die unmittelbar eine Veränderung im Status der Maschine auslösen, sondern auch solche, die das gerade nicht getan haben. Dadurch wird eine zweite Motivation für die Aufzeichnung und Speicherung der Audios verdeutlicht, die über die reine Anwendung durch die Nutzer*innen hinausgeht: Das Sammeln von Trainingsdaten für die Verbesserung von Speech Recognition (vgl. auch Waldecker/Volmar 2022).

Sprachassistenten scheinen nicht in erhofftem Umfang rentabel zu sein – insbesondere steht die Vermutung im Raum, dass die aufgezeichneten Daten sich nicht hinreichend monetarisieren lassen (vgl. Amadeo 2022). Konversationelle Interfaces haben dennoch Konjunktur, denn anders als die hier untersuchten Interfaces basiert die jüngste Welle in diesem Bereich auf generativen Anwendungen Künstlicher Intelligenz auf Basis von LLMs, die vereinfacht gesagt mit statistischen Berechnungen die wahrscheinlichste Antwort auf einen Input generieren können und insofern nicht mehr auf die Verknüpfung mit einzelnen Funktionen angewiesen sind (vgl. Bender et al. 2021: 616). Eine Weiterentwicklung entsprechender Technologien auf akustische Prozessierung ist zeitnah zu erwarten (vgl. Knisella 2023). Medienhistorisch bedeutet dies für die Interfaces die Fortsetzung einer Linie, die Waldecker/Volmar (vgl. 2022: 168) herausarbeiten: Die Reduktion der Gelenktheit des Dialogs, die etwa bei Voice-Dial-Systemen oder in Call-Center-Anrufen durch eine akustische Menüführung noch deutlich stärker ist als bei Sprachassistenzsystemen oder auf LLMs basierenden Chatprogrammen. VUI-Dialoge laufen, wie sich auch an den Analysen zeigt, bisher sehr musterhaft und funktionsorientiert ab. Die Rigidität der VUIs könnte dann ebenfalls abnehmen – sie könnten im Hinblick auf ihre sequenzielle und sprachliche Gestaltung flexibler werden und mehr Raum für Vagheit lassen. Damit wären aber auch deutlich längere und eben gesprächsähnlichere Konversationen produzierbar, was nicht unbedingt nutzer*innenseitig gewünscht sein muss und die Fragilität von VUI-Dialogen erhöhen könnte.

Literatur

Amadeo, Ron (2022): Amazon Alexa is a »Colossal Failure«, on Pace to Lose \$10 Billion this Year. Abrufbar unter: <https://arstechnica.com/gadgets/2022/11/am>

- azon-alexa-is-a-colossal-failure-on-pace-to-lose-10-billion-this-year/ (Stand: 01.02.2024).
- Auer, Peter (2020): »Die Struktur von Redebeiträgen und die Organisation des Sprecherwechsels«, in: Karin Birkner/Peter Auer/Angelika Bauer/Helga Kotthoff (Hg.), Einführung in die Konversationsanalyse, Berlin/Boston: de Gruyter, S. 106–235.
- Baranovska, Marianna/Stefan Hölzgen (Hg.) (2018): *Hello, I'm Eliza. Fünfzig Jahre. Gespräche mit Computern*, Bochum/Freiburg: projektverlag.
- Barthel, Mathias/Henrike Helmer/Silke Reineke (2023): »First Users« Interactions with Voice-Controlled Virtual Assistants: A Micro-Longitudinal Corpus Study«, in: Proceedings of SemDial 2023. Abrufbar unter: <https://mezzanine.um.si/en/conference/semdial-2023-marilogue/#proceedings> (Stand: 02.05.2024).
- Bartz, Christina/Timo Kaerlein/Monique Miggelbrink/Christoph Neubert (2019): »Zur Medialität von Gehäusen. Einleitung«, in: Christina Bartz/Timo Kaerlein/Monique Miggelbrink/Christoph Neubert (Hg.), Gehäuse: Mediale Einkapselungen, Paderborn: Wilhelm Fink, S. 9–32.
- Bedford-Strohm, Jonas (2017): »Voice First? Eine Analyse des Potentials von intelligenten Sprachassistenten am Beispiel Amazon Alexa«, in: *Communicatio Socialis* 4, S. 485–494.
- Bender, Emily/Timnit Gebru/Angelina McMillan-Major/Shmargaret Shmitchell (2021): »On the Dangers of Stochastic Parrots«, in: Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency, New York: ACM Press, S. 610–623.
- Bender, Emily/Alexander Koller (2020): »Climbing towards NLU: On Meaning, Form, and Understanding in the Age of Data«, in: Dan Jurafsky/Joyce Chai/Natalie Schluter/Joel Tetreault (Hg.), Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics, Stroudsburg: Association for Computational Linguistics, S. 5185–5198.
- Borbach, Christoph (2019): »Navigating (Through) Sound. Auditory Interfaces in Maritime Navigation Practice, 1900–1930«, in: *Interface Critique Journal* 2, S. 17–33.
- Borbach, Christoph/Timo Kaerlein/Robert Stock/Sabine Wirth (Hg.) (i.E.): *Akustische Interfaces. Interdisziplinäre Perspektiven auf Schnittstellen von Technologien, Sounds und Menschen*, Wiesbaden: Springer Vieweg.
- Bratton, Benjamin H. (2016): *The Stack. On Software and Sovereignty*, Cambridge: MIT Press.
- Clayman, Steven (2012): »Turn-Cunstructional Units and the Transition-Relevance Place«, in: Jack Sidnell/Tanya Stivers (Hg.), *The Handbook of Conversation Analysis*, Chichester: Blackwell, S. 150–166.
- Cramer, Florian/Matthew Fuller (2008): »Interface«, in: Matthew Fuller (Hg.), *Software Studies: A Lexikon*, Cambridge: MIT Press, S. 149–152.

- Crawford, Kate/Vladan Joler (2018): »Anatomy of an AI System: The Amazon Echo as an Anatomical Map of Human Labor, Data and Planetary Resources«. Abrufbar unter: <https://anatomyof.ai/> (Stand: 24.04.2024).
- Deppermann, Arnulf (2000): »Ethnographische Gesprächsanalyse: Zu Nutzen und Notwendigkeit von Ethnographie für die Konversationsanalyse«, in: *Gesprächsforschung* 1, S. 96–124.
- Deppermann, Arnulf (2008): *Gespräche analysieren. Eine Einführung*, Wiesbaden: VS Verlag für Sozialwissenschaften.
- Dieter, Michael (2022): »Interface Critique at Large«, in: *Convergence* 0 (0), S. 1–17.
- Distelmeyer, Jan (2020): »Interface II. Zur Programmik leitender Prozesse der ›digitalen Gegenwart‹«, in: Huber/Krämer/Pias, *Wovon sprechen wir, wenn wir von Digitalisierung sprechen? Gehalte und Revisionen zentraler Begriffe des Digitalen*, S. 59–72.
- Ernst, Christoph (2017): »Implizites Wissen, Kognition und die Praxistheorie des Interfaces«, in: *Navigationen* 17 (2), S. 99–116.
- Ernst, Christoph/Thomas Bächle (2020): »Interface«, in: Martina Heßler/Kevin Liggieri (Hg.), *Technikanthropologie. Handbuch für Wissenschaft und Studium*, Baden-Baden: Nomos, S. 416–420.
- Goulden, Murray (2019): »Delete the Family: Platform Families and the Colonisation of the Smart Home«, in: *Information, Communication & Society* 24 (7), S. 1–18.
- Habscheid, Stephan (2022): »Socio-Technical Dialogue and Linguistic Interaction. Intelligent Personal Assistants (IPA) in the Private Home«, in: *Sprache und Literatur* 51 (2), S. 167–196.
- Habscheid, Stephan/Tim Hector/Christine Hrcnal (2023): »Human and Non-Human Agency as Practical Accomplishment. Interactional Occasions for Ascription and Withdrawal of (Graduated) Agency in the Use of Smart-Speaker-Technology«, in: *Social Interaction. Video-Based Studies of Human Sociality* 6 (1), 1–31.
- Habscheid, Stephan/Tim Hector/Christine Hrcnal/David Waldecker (2021): »Intelligente Persönliche Assistenten (IPA) mit Voice User Interfaces (VUI) als ›Beteiligte‹ in häuslicher Alltagsinteraktion. Welchen Aufschluss geben die Protokoll-daten der Assistenzsysteme?«, in: *Journal für Medienlinguistik* 4 (1), S. 16–53.
- Hadler, Florian/Joachim Haupt (2016a): *Towards a Critique of Interfaces*. In: Hadler/Haupt, *Interface Critique*, S. 7–13.
- Hadler, Florian/Joachim Haupt (2016b) (Hg.) *Interface Critique*, Berlin: Kadmos.
- Hector, Tim (i.V.): *Smart Speaker im Dialog. Sprachliche Praktiken mit Voice User Interfaces*. Berlin/Boston: de Gruyter.
- Hector, Tim (2022): »Smart Speaker in der Praxis. Methodologische Überlegungen zur medienlinguistischen Erforschung stationärer Sprachassistenzsysteme«, in: *Sprache und Literatur* 51 (2), S. 197–229.

- Hector, Tim/Christine Hrnca (2024): »Sprachassistenzsysteme in der Interaktion«, in: Jannis Androustopoulos/Friedemann Vogel (Hg.), *Handbuch Sprache und digitale Kommunikation*, Berlin u.a.: de Gruyter, S. 309–328.
- Hector, Tim/Franziska Niersberger-Gueye/Franziska Petri/Christine Hrnca (2022): »The »Conditional Voice Recorder«: Data Practices in the Co-Operative Advancement and Implementation of Data-Collection Technology«, in: *Working Paper Series Media of Cooperation* 23, S. 1–15.
- Hookway, Branden (2014): *Interface*, Cambridge/London: The MIT Press.
- Hoy, Matthew B. (2018): »Alexa, Siri, Cortana, and More: An Introduction to Voice Assistants«, in: *Medical Reference Services Quarterly* 37 (1), S. 81–88.
- Huber, Martin/Sybille Krämer/Claus Pias (Hg.) (2020): *Wovon sprechen wir, wenn wir von Digitalisierung sprechen? Gehalte und Revisionen zentraler Begriffe des Digitalen*, Frankfurt a.M.: Universitätsbibliothek Johann Christian Senckenberg.
- Imo, Wolfgang (2009): »Konstruktion oder Funktion? Erkenntnisprozessmarker (Change-of-State Tokens) im Deutschen«, in: Jörg Buecker/Susanne Günther (Hg.), *Grammatik im Gespräch. Konstruktionen der Selbst- und Fremdpositionierung*, Berlin u.a.: de Gruyter, S. 57–86.
- Kaerlein, Timo (2020): »Interface. Zur Vermittlung von Praktiken und Infrastrukturen (als Perspektive für die Medienwissenschaft)«, in: Huber/Krämer/Pias, *Wovon sprechen wir, wenn wir von Digitalisierung sprechen? Gehalte und Revisionen zentraler Begriffe des Digitalen*, S. 45–58.
- Knisella, Bret (2023): *Google Assistant and Alexa Are Both Getting Generative AI Makeovers. Welcome to the ChatGPT Era*. Abrufbar unter: <https://synthedia.substack.com/p/google-assistant-and-alexa-are-both> (Stand: 01.02.2024).
- Latour, Bruno (2005): *Reassembling the Social. An Introduction to Actor-Network-Theory*, Oxford: Oxford University Press.
- Light, Ben/Jean Burgess/Stefanie Duguay (2018): »The Walkthrough Method: An Approach to the Study of Apps«, in: *New Media and Society* 20 (3), S. 881–900.
- Merkle, Benedikt/Tim Hector (i.E.): »Werkzeuge und Medienpraktiken. Intelligente persönliche Assistenten und das Paradigma objektorientierten Programmierens«, in: Christoph Borbach/Timo Kaerlein/Robert Stock/Sabine Wirth (Hg.), *Akustische Interfaces*.
- Mondada, Lorenza (2016): *Conventions for Multimodal Transcription*. Abrufbar unter: https://franoesistik.philhist.unibas.ch/fileadmin/user_upload/franoesistik/mondada_multimodal_conventions.pdf (Stand: 21.09.2023).
- Natale, Simone/Henry Cooke (2021): »Browsing with Alexa: Interrogating the Impact of Voice Assistants as Web Interfaces«, in: *Media, Culture & Society* 43 (6), S. 1000–1016.
- Pasquinelli, Matteo (2019): »How a Machine Learns and Fails«, in: *spheres – Journal for Digital Cultures* 5, S. 1–17.

- Paßmann, Johannes/Carolyn Gerlitz (2014): »Good« Platform-Political Reasons for »Bad« Platform Data. Zur sozio-technischen Geschichte der Plattformaktivitäten »Fav«, »Retweet« und »Like«, in: *Mediale Kontrolle unter Beobachtung* 3 (1), S. 1–40.
- Porcheron, Martin/Joel Fischer/Stuart Reeves/Sarah Sharples (2018): »Voice Interfaces in Everyday Life«, in: Regan Mandryk/Mark Hancock/Mark Perry/Anna Cox (Hg.), *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems – CHI '18*, New York: ACM Press, S. 1–12.
- Sacks, Harvey/Emanuel Schegloff/Gail Jefferson (1974): »A Simplest Systematics for the Organisation of Turn Taking in Conversation«, in: *Language* 50 (4), S. 696–735.
- Schegloff, Emanuel (1968): »Sequencing in Conversational Openings«, in: *American Anthropologist* 70 (6), S. 1075–1095.
- Selting, Margret/Peter Auer/Dagmar Barth-Weingarten/Jörg Bergmann/Pia Bergmann/Karin Birkner/Elizabeth Couper-Kuhlen/Arnulf Deppermann/Peter Gilles/Susanne Günthner/Martin Hartung/Friederike Kern/Christine Mertzluft/Christian Meyer/Miriam Morek/Frank Oberzaucher/Jörg Peters/Uta Quasthoff/Wilfried Schütte/Anja Stukenbrock/Susanne Uhmann (2009): »Gesprächsanalytisches Transkriptionssystem 2 (GAT 2)«, in: *Gesprächsforschung* 10, S. 353–402.
- Strüver, Niklas (2023a): »Frustration Free: How Alexa Orchestrates the Development of the Smart Home«, in: *Digital Culture & Society* 9 (1), S. 99–123.
- Strüver, Niklas (2023b): »Wieso eigentlich Alexa? Konzeptualisierung eines Sprachassistenten als Infrastruktur und Plattform im soziotechnischen Ökosystem Amazons«, in: *kommunikation@gesellschaft* 24, S. 1–33.
- Suchman, Lucy (2007): *Human-Machine Reconfigurations. Plans and Situated Actions*. 2nd edition, Cambridge: Cambridge University Press.
- Volmar, Axel (2019): »Productive Sounds«, in: Andreas Sudmann (Hg.), *The Democratization of Artificial Intelligence*, Bielefeld: transcript, S. 55–76.
- Waldecker, David/Axel Volmar (2022): »Die zweifache akustische Intelligenz virtueller Sprachassistenten zwischen verteilter Kooperation und Datafizierung«, in: Anna Schürmer/Maximilian Haberer/Tomy Brautschek (Hg.), *Acoustic Intelligence. Hören und Gehorchen*, Düsseldorf: Düsseldorf University Press, S. 161–182.
- Wirth, Sabine (2016): »Between Interactivity, Control, and »Everydayness« – Towards a Theory of User Interfaces«, in: Florian Hadler/Joachim Haupt (Hg.), *Interface Critique*, Berlin: Kadmos, S. 17–38.

Zwischen Technologie und Ideologie

Ein Blick auf die un/realen Räume hinter KI

Paul Schütze

Abstract *Dieser Beitrag diskutiert den Einfluss von Künstlicher Intelligenz (KI) auf gesellschaftliche un/reale Interaktionsräume. Insbesondere umreißt er die Ideologie des »AI Futurism«, welche KI-Technologien als unausweichlichen Fortschritt und zentrales Element zukünftiger gesellschaftlicher Strukturen zeichnet. AI Futurism verschleiert die breiteren Auswirkungen von KI auf soziale und ökologische Interaktionsräume und verzerrt zugleich die technologische Realität durch einen sehr spezifisch ideologischen Deutungshorizont. Ziel des Beitrags ist es, das Spannungsfeld zwischen Technologie und Ideologie im Kontext von KI aufzudecken und die Konsequenzen für gesellschaftliche Interaktionen zu beleuchten.*

Keywords *Ethik der KI; Critical AI; Technikphilosophie; kritische Theorie; Ideologie*

1. Einleitung

Der Einfluss von Künstlicher Intelligenz (KI) auf (zukünftige) gesellschaftliche Interaktionsräume ist ein prominentes Diskussionsthema, sowohl in akademischen als auch in öffentlichen Debatten (siehe z.B. Campolo/Crawford 2020; Lösch et al. 2019). Dabei steht zum Beispiel im Fokus, dass Sprachmodelle wie ChatGPT einen restriktiven Einfluss auf die Wissensbildung ihrer Benutzer*innen haben und so nicht nur die Möglichkeit von Wissen einschränken, sondern auch starre Räume schaffen, in denen nur auf eine ganz bestimmte Art epistemologisch interagiert werden kann (vgl. Shah/Bender 2022; Lindemann 2024). Auf ähnliche Weise reproduzieren Suchalgorithmen hegemoniale Positionen und verhindern so die Offenheit gegenüber alternativen Diskursräumen (vgl. Ballatore et al. 2017). An anderer Stelle verstärken algorithmische Entscheidungen, beispielsweise in sozialen Kontexten, diskriminierende Hierarchien und geben ein eingeschränktes Interaktionsrepertoire bzw. Handlungsspielräume vor (vgl. Eubanks 2018; Schwerzmann 2021). All diese Positionen zeigen, dass KI-Anwendungen offenbar eine besondere Rolle bei der Gestaltung soziokultureller Räume zukommt. Wie Meredith Whit-

taker (vgl. 2021) feststellt, arbeiten vor allem große Technologiekonzerne daran, mit ihren KI-Produkten erheblichen Einfluss auf unser Leben und unsere Institutionen auszuüben. Die gesellschaftliche Zukunft und die un/realen Möglichkeiten der Interaktion werden so wesentlich von KI-Anwendungen und der dahinterstehenden Tech-Industrie eingeschränkt.

In diesem Beitrag setze ich diese kritischen Debatten fort und gehe der momentan vorherrschende Ideologie nach, die KI-Technologien als einen scheinbar unvermeidlichen Schritt in Richtung Zukunft und Fortschritt darstellt – eine Vorstellung, die ich als »AI Futurism« bezeichne.¹ Die oben angedeuteten Diskussionen werfen kritische Lichter auf die unverrückbare Vorstellung, dass KI die Zukunft maßgeblich gestalten wird. Ganz gleich in welcher Form, KI-Technologien scheinen aus der Zukunft nicht mehr wegzudenken zu sein.

Ich argumentiere jedoch, dass die breiteren Konsequenzen der KI-Maschinerie auf gesellschaftliche und ökologische Interaktionsräume maßgeblich unterschätzt werden. Dies führe ich in diesem Beitrag teilweise auf die Ideologie des AI Futurism zurück. Konkret nehme ich in den Blick, wie sich gesellschaftliche Vorstellungen von KI zwischen technologischer Realität und ideologischer Narrative bewegen. So richte ich ein Schlaglicht auf die un/realen Räume hinter KI. Am Ende mache ich deutlich, wie KI Anwendungen sozio-materielle Interaktionsräume maßgeblich vorbestimmt und die Möglichkeiten gesellschaftlichen Zusammenlebens determinieren.

2. Von Technologie zur Ideologie

In den Diskursen von KI taucht am prominentesten die Perspektive von der Technologie als einfaches Werkzeug auf – *KI-als-Werkzeug*. Hier werden KI-Technologien vor allem als nutzbare Instrumente gesehen, die effizient und kontrolliert einzusetzen sind. Konkret geht es dann meist um Anwendungen wie ChatGPT, Algorithmen in den Sozialen Medien oder Gesichtserkennungssysteme. Prominente Beispiele für diese Betrachtung aus der KI-als-Werkzeug Perspektive finden sich beispielsweise in der Analyse »AI4People« von Luciano Floridi et al. (2018). Sie betrachten vor allem, wie KI-Technologien instrumentell eingesetzt werden können, um Menschenwürde und »menschliches Gedeihen« zu fördern (vgl. ebd. : 690). Ebenfalls lässt sich dieser Blick auch in den Debatten über nachhaltige KI beobachten, beispielsweise wenn

1 Diesen Begriff führe ich in meinem Artikel »The Impacts of AI Futurism – An Unfiltered Look at AI's True Effects on the Climate Crisis« ein (Schütze 2024). In den folgenden Abschnitten greife ich teilweise auf diese Arbeit zurück.

es um den Einsatz von KI-Systemen in der Landwirtschaft geht.² Insgesamt ist dies wohl die am häufigsten zu beobachtende Sichtweise auf KI (vgl. Sartori/Bocca 2023).

Dem gegenüber steht in den Debatten eine kritischere Betrachtungsweise, die darauf hinweist, dass KI nicht einfach auf ein Werkzeug reduziert werden kann. Künstliche Intelligenz ist ein weitaus größeres und komplexeres Phänomen. Sie ist ein umfassendes Netzwerk, das verschiedene sozio-materielle Elemente umfasst (vgl. Brevini 2021b; Mühlhoff 2020; Crawford 2021). Diese Betrachtung kann unter dem Begriff *KI-als-soziomaterielles-System* zusammengefasst werden.³ Den Mittelpunkt bildet hier die Erkenntnis, dass KI eine Vielzahl von Methoden und Praktiken umfasst, die abhängig von umfangreichen Datensätzen, von massiven Rechenkapazitäten und von menschlichem Wissen sind. Eine solche Sichtweise von KI-als-soziomaterielles-System ist in letzter Zeit vor allem in den öffentlichen Diskussionen über die Umweltauswirkungen von Anwendungen wie ChatGPT oder den Cloud Servern aufgetaucht (siehe z.B. Dhar 2020; Sattiraju 2020). Weitaus präsenter ist diese Perspektive jedoch in kritischen Debatten zur Ethik der KI (siehe z.B. Crawford 2021; Dubber et al. 2020; Mühlhoff 2020), oder in den *Science and Technology Studies* (STS) (vgl. Sartori/Bocca 2023; Sartori/Theodorou 2022).

Kate Crawford weist beispielsweise darauf hin, dass KI eine Anordnung technologischer Methoden mit konkretem soziopolitischem Ursprung und materieller Grundlage beschreibt. In diesem Zusammenhang spricht Crawford auch von der »KI-Megamaschine«. Demnach hängen KI-Technologien von industriellen Infrastrukturen, Lieferketten und menschlicher Arbeit ab, die sich weltweit erstrecken, aber meist undurchsichtig bleiben (vgl. Crawford 2021: 48). Ähnlich spricht Rainer Mühlhoff von KI als eine historische Formation, die neue medienkulturelle Räume und Möglichkeiten hervorbringt (vgl. Mühlhoff 2020). KI-als-soziomaterielles-System zu begreifen zeigt also, dass KI-Technologien nie nur Werkzeuge sein können, sondern immer eingebettet in ein soziales und materielles Netzwerk sind.

Der kritische Technologie- und Medienwissenschaftler Andrew Feenberg geht in seinem Buch »Questioning Technology« noch einen Schritt weiter. Er beschreibt, dass Technologien nicht nur auf materielle und soziale Netzwerke beschränkt sind, sondern immer auch in eine umfassendere ideologische Struktur eingebettet sind (vgl. Feenberg 1999). Das bedeutet, dass es bei der kritischen Betrachtung von Technologien zentral ist, die Ideen und Überzeugungen zu durchdringen, die das gesell-

-
- 2 Zu diesem Thema findet sich eine umfangreiche Studie in der Arbeit von Angella Ndaka (2023). Sie zeigt die Narrative und Visionen, die den Einsatz von KI-als-Werkzeug in der Landwirtschaft informieren.
 - 3 Dieser Blick baut auf der in STS prominent entwickelten Perspektive auf: »Technology is always part of society, just like society is always part of technology. This also means that one cannot understand one without the other. Technology is not only design and material appearance but also sociotechnical; that is, a complex process constituted by diverse social, political, economic, cultural and technological factors« (Hasselbalch 2021: 10).

schaftliche Verständnis dieser Technologien vermitteln. Genau das wird in der Perspektive KI-als-soziomaterielles-System deutlich: Es ist wichtig, nicht nur die konkreten Anwendungen von Technologien zu betrachten, sondern auch die zugrunde liegenden Ideologien, die deren Entwicklung und Einsatz beeinflussen. Anstatt der weitverbreiteten Vorstellung zu folgen, dass KI mit den richtigen Regulierungen als Werkzeug eingesetzt werden könne, müssen vielmehr die Ideologien hinter KI kritisch betrachtet und in den Fokus gerückt werden.

Hier kommen wir zum AI Futurism und den Narrativen, die KI als unausweichlichen technologischen Fortschritt und als eine die Zukunft transformierende Kraft präsentieren. Betrachtet man diese Ideologie näher, findet man ihre Anfänge in dem allgemeinen Glauben an das Potenzial von Technologie per se. So knüpft die Vorstellung, dass KI bahnbrechende Lösungen für gesellschaftliche Herausforderungen bieten kann, an die spätkapitalistischen Erzählungen des Techno-Solutionism und des Techno-Determinism an (vgl. Brevini 2021a; siehe auch Morozov 2013). Der Techno-Solutionism beschreibt den Glauben, dass Technologie als Allheilmittel für gesellschaftliche Probleme dienen kann, während der Techno-Determinismus die Vorstellung unterstützt, dass die Zukunft zwangsläufig von Technologien, in diesem Fall KI, dominiert wird. Diese Erzählungen präsentieren die Zukunft als vorherbestimmt. So wird die Gesellschaft dazu genötigt, die Chancen und Risiken dieser Technologien innerhalb vorgefertigter Bahnen und anhand feststehender möglicher Zukünfte zu diskutieren.

Hier erstreckt sich der deterministische Glaube zwischen den *realen* KI-Systemen, die schon heute tatsächlichen Einfluss auf Gesellschaften haben, und einer *imaginären*, mythologischen Vision, die die realen Effekte aufbläht und mit großen Versprechungen lockt (vgl. Elish/boyd 2018). Benedetta Brevini nennt dies den Mythos des übermenschlichen Potenzials der künstlichen Maschine, deren Entwicklung unausweichlich bevorsteht (vgl. Brevini 2021a: 155). So schwanken die Narrative von KI immer zwischen dem Realen und dem Imaginären (vgl. Elish/boyd 2018: 62) – zwischen dem, was KI-Technologien tatsächlich sind, und dem, was sie versprechen zu werden. Den realen Anwendungen von KI liegt also stets eine vorausschauende, imaginäre Komponente zugrunde.

Wie oben bereits angedeutet, ist diese Bewegung in Richtung Zukunft nicht zwingend KI spezifisch, sondern sie ist ein zentraler Bestandteil gesellschaftlicher Vorstellungen von Technologie im Allgemeinen. Der Medienwissenschaftler Hartmut Winkler zum Beispiel beschreibt in diesem Zusammenhang die zukunftsorientierten Affekte, die mit der Einführung von Computern einhergingen. So zeichnet Winkler das Aufkommen und die Verbreitung von Computern und deren Netzwerken als eine Entstehung von »Wunschmaschinen« oder »Wunschkonstellationen« (vgl. Winkler 1997): »Dass Millionen von Privatleuten Geld, Freizeit und Lebensenergie aufwenden, um Zutritt zu dem neuen Universum [Computern und deren Netzwerke] zu erhalten ... deutet darauf hin, dass es eigentlich um Wünsche

und nicht um harte Fakten geht« (ebd.: 11). Stets scheinen also Wünsche, Vorstellungen und Visionen einer konkreten Zukunft im Hintergrund der realen Benutzung und Adaption einer Technologie zu stehen.

Ähnlich kann man die soziomateriellen Systeme hinter KI auch begreifen. Es stehen ganz spezielle Wünsche, Vorstellungen und Ideen hinter KI Systemen, die wiederum gleichzeitig an die Zukunft geknüpft sind und in ihre Richtung streben. Wollen wir die Anziehungskraft von KI und deren Auswirkungen auf gesellschaftliche Interaktions- und Handlungsräume also besser verstehen, müssen wir auf eben diese Konstellationen schauen. Hier rückt wiederum die Ideologie des AI Futurism in den Blick. Denn diese Ideologie liegt eben jenen technologischen Wunschkonstellationen von KI zu Grunde. Anders ausgedrückt sind die an KI geknüpften Wünsche und Narrative das Resultat eines größeren kulturellen Horizontes, den ich hier als Ideologie begreife.

Nun ist eine vorausschauende und imaginäre Komponente zwar Teil vieler Technologien. Jedoch scheint diese Wirkung bei KI-Technologien besonders stark zur Geltung zu kommen. KI ist auf eigenartige und spezielle Weise mit der Zukunft verbunden (siehe z.B. die STS Perspektiven wie Leeker 2017; Sartori/Bocca 2023). Dies liegt nicht zuletzt daran, dass KI-Technologien tief mit kapitalistischen Mythen verwoben sind (siehe Abschnitt 2) und ihnen qua ihrer Funktion, als beispielsweise prädiktive Algorithmen, immer schon ein Glaube an den idealisierten Fortschritt inneohnt. Im nächsten Abschnitt gehe ich daher genauer auf die Ideologie ein, die KI in speziellem Maße ein imaginäres Element zufügt und sie als Verkörperung einer progressiven Zukunft präsentiert.

3. Ideologie und kapitalistische Mythenbildung

Um zu vertiefen, wie KI-als-soziomaterielles-System sich auf gesellschaftliche Interaktionsräume auswirkt, ist ein Verständnis der Ideologie des AI Futurism unausweichlich. Dabei folge ich in meiner Analyse einem kritischen Verständnis von Ideologie. Demnach ist Ideologie als ein »System von Überzeugungen« zu verstehen, dem »handlungsleitende Kraft« zukommt und das gleichzeitig »Herrschaftsverhältnisse und gesellschaftliche Konflikte« verschleiert (vgl. Jaeggi/Celikates 2017: 102). So beschreibt Althusser Ideologie auch als »das System von Ideen und Vorstellungen, das den Geist eines Menschen oder einer gesellschaftlichen Gruppe beherrscht« (Althusser 2010: 71). Dieses System, also die Ideologie, beeinflusst das Verständnis der Individuen und der Gruppe zu ihren realen Existenzbedingungen (ebd.: 78). In anderen Worten bestimmt die Ideologie, wie die realen Bedingungen, unter denen die Individuen oder die Gruppe leben, verstanden werden und wie die jeweils eigene Beziehung zu diesen Bedingungen aufgefasst werden. Gleichzeitig ist eine bestimmte Ideologie immer auch mit bestimmten Praktiken und Institutionen ver-

bunden. »Eine Ideologie existiert immer in einem Apparat und in dessen Praxis oder dessen Praktiken. Diese Existenz ist materiell« (ebd.: 80). Das heißt »Ideologien sind nicht auf Überzeugungen reduzierbar, denn sie prägen die Identität von Subjekten, wirken praktisch und sind ihrerseits Effekte einer bestimmten gesellschaftlichen Praxis« (Jaeggi/Celikates 2017: 103). Mit Marx und Engels kann man vereinfacht sagen, dass Ideologie »der ideelle Ausdruck der herrschenden materiellen Verhältnisse« ist (Marx/Engels 1845/46: 46, zit.n. Jaeggi/Celikates 2017: 104).

Dieses Ideologieverständnis zielt wesentlich auf eine Kritik der herrschenden Verhältnisse ab und macht es sich zugleich zur Aufgabe, immanent, d.h. aus den Herrschaftsverhältnissen heraus, und machtkritisch »die herrschenden Ideen, Handlungsweisen und gesellschaftlichen Verhältnisse nicht unreflektiert« (Horkheimer 1940: 344, zit.n. Jaeggi/Celikates 2017: 110) hinzunehmen, sondern zu hinterfragen. Im Zuge dieses Sammelbandes bedeutet dies, die herrschenden Verhältnisse in und um KI-als-soziomaterielles-System zu hinterfragen und deren einschränkende Wirkung auf gesellschaftliche Interaktionsräume aufzuzeigen.

Bezieht man das oben erläuterte Ideologieverständnis also auf KI-Technologien, bedeutet dies nun, dass AI Futurism eine Ideologie im Hintergrund von KI ist: Also ein hegemoniales und (materiell) institutionalisiertes System von Ideen und Vorstellungen, das die gesellschaftliche und individuelle Auffassung von KI-Technologien aufrechterhält. Dabei ist besonders wichtig zu betonen, dass es sich hierbei um einen »Deutungshorizont« (Jaeggi/Celikates 2017: 103) handelt, der gesellschaftliche und individuelle Relationen zu KI-Technologien mediiert. Am Ende wird mittels dieser Ideologie – des AI Futurism – der Status quo »mehr oder weniger konfliktfrei reproduziert« (ebd.).⁴

Doch wie genau äußert sich das? Und welche Auswirkungen hat dies auf Interaktionsräume? Um dies zu zeigen, ist es aufschlussreich das beschriebene Ideologieverständnis mit der Wirkweise von Machtrelationen nach Foucault zu verbinden. Denn, wie Slavoj Žižek aufzeigt, hängt die »micro-power«, die Foucault beschreibt, zentral mit der Funktionsweise von Ideologie zusammen. Für Foucault sind die relationalen Machtdynamiken zentrale Merkmale gesellschaftlichen Zusammenlebens. Sie sind nicht nur konstitutiv für das Subjekt, sondern soziokulturelle Räume sind als Machtapparate bestehend aus micro-power Relationen zu verstehen.⁵ Žižek greift dieses Machtverständnis auf und zeigt, dass micro-power eigentlich nichts anderes als Ideologie im Einsatz ist (vgl. Žižek 2012: 13). Das heißt Ideologie ist nicht per se als ein übergeordnetes System zu verstehen, das von der

4 Dieses Verständnis von Ideologie könnte im Kontext von KI weitergedacht werden, z.B. durch die Arbeit von Wendy Hui Kyong Chun (2005) zu Software und Ideologie. Dies würde jedoch in eine andere Richtung führen und den Rahmen dieses Textes überschreiten.

5 Man könnte in diesem Sinne mit Foucault also auch von einem KI-Dispositiv sprechen, bzw. dieser Begriff verhielte sich ähnlich zu meiner Ideologiekritik von KI.

herrschenden Klasse aufrechterhalten und durchgesetzt wird. Sondern mit dem relationalen Machtverständnis von Foucault, lässt sich Ideologie als ein verzweigtes System von Praktiken und Diskursen begreifen. Dieses System verkörpert zwar hegemoniale Interessen, aber gleichzeitig manifestiert es sich in »micro-procedures« (ebd.). Das heißt Ideologie kann so als ein komplexes Netz von Machtrelationen verstanden werden, dass von Subjekten und Institutionen aufrechterhalten und konstituiert wird.

Dieses System an Praktiken und Diskursen, welches letztendlich die Ideologie des AI Futurism aufrechterhält, lässt sich exemplarisch mit der Arbeit von Benedetta Brevini zum europäischen KI-Diskurs verstehen. In ihrem Text »Creating the Technological Saviour« beschreibt Brevini (2021a) wie KI von Mythen umgeben ist, und macht so konkret, welche Interaktionsräume sich aus dem Geflecht von KI-alsoziomaterielles-System ergeben. Hier zeigen sich daher auch die institutionalisierten Mechanismen und *micro-procedures*, welche die Ideologie des AI Futurism konstituieren.

In ihren Untersuchungen fokussiert Brevini sich auf drei der am meisten verbreiteten Mythen in den Diskursen über KI in der europäischen Debatte. Der erste Mythos, der häufig proklamiert wird, zeichnet KI als Lösung für die größten Herausforderungen der Menschheit und des Kapitalismus. Dabei wird immer wieder argumentiert, dass KI in der Lage sei, die zentralen gesellschaftlichen Probleme, wie die Klimakrise oder soziale Ungleichheit, zu bewältigen. Der zweite Mythos, den Brevini betrachtet, betont die Notwendigkeit und Dringlichkeit der Entwicklung von KI-Technologien und stellt deren Verbreitung als unvermeidlich dar. Im Fokus steht deren angebliche Unausweichlichkeit, die eine rasche Anpassung im Sinne des positiven Fortschritts erfordere. Der dritte Mythos propagiert, dass KI letztendlich die Intelligenz der Menschen übertreffen werde. In dieser Perspektive werden KI-Technologien oft als überlegen betrachtet. Zum Beispiel als seien sie objektiver, neutraler oder allwissender. Mit diesen Mythen umreißt Brevini den vorherrschenden Diskurs über KI. Sie zeigt, wie die Mythen Teil einer Ideologie sind und ein ganz bestimmtes Verständnis von KI-Technologien implizieren (vgl. ebd.: 151–155).

Die Mythen, die Brevini beschreibt, sind also Bestandteile der Ideologie des AI Futurism. Sie präsentieren KI als Lösung für die Hauptprobleme unserer Gesellschaft und als technologische Rettung, deren Hilfe unausweichlich sei (vgl. ebd.: 155). KI-Technologien seien eine technologische Notwendigkeit, deren Vorteile unbedingt genutzt und deren Schäden kontrolliert werden müssten. So wird die implizite Akzeptanz, dass KI-Technologien die Zukunft bestimmen, nicht nur legitimiert. Sondern die gesellschaftliche Dimension dieser technologischen Entwicklung wird entpolitisiert und alternative Handlungsräume werden nahezu undenkbar gemacht (vgl. ebd.: 151–155). Diese Art des AI Futurism hat besonders in europä-

schen und nordamerikanischen Kontexten Einfluss gewonnen, und so ist besonders dort klar, wie mit und durch KI interagiert werden kann und soll.

Das kann in der Praxis zum Beispiel in den Unternehmensvisionen der großen Tech-Konzerne wie Google beobachtet werden. Sie verbreiten den Glauben, dass KI eine transformative Technologie ist, die immense Vorteile für Menschen und Gesellschaft mit sich bringt (vgl. Google AI, o.J.-a). Googles proklamiertes Ziel ist es, die Vorteile von KI allen zugänglich zu machen (vgl. Google AI, o.J.-b), um Innovationen zu beschleunigen und somit unter anderem den Klimawandel bekämpfen zu können (vgl. Google Sustainability, o.J.). Aber auch in der jüngsten europäischen Gesetzgebung wie dem AI Act spiegelt sich diese Ideologie wider. Denn Ziel des AI Acts ist es, die EU zu einem Ort zu machen, an dem KI von der Entwicklung bis zum Markt gedeiht. Und vielmehr noch: Nach dem AI Act sollen KI-Technologien eine positive Kraft in der Gesellschaft entfalten (vgl. European Commission, 2023). Selbst in wissenschaftlichen Debatten herrscht der Glaube, dass KI-Technologien zwar nicht alle Probleme lösen könne, sie aber dennoch ein essenzieller Beitrag zur Bewältigung zukünftiger Herausforderungen sind. Sowohl soziale als auch ökologische Probleme könnten mithilfe von KI bewältigt werden. Hierfür müssten Gesellschaften nur das Potenzial von KI nutzen, natürlich auf ethisch vertretbare Weise und mit minimalem CO₂-Fußabdruck (vgl. Cowls et al. 2023: 299; Cowls et al. 2021: 114).

Heutige KI-Technologien basieren also immer schon auf diesen ideologischen Visionen, auf dem historisch gewachsenen Glauben an Effizienz, technologischer Neutralität und transformativer Potenziale (vgl. Elish/boyd 2018: 74). Der Deutungshorizont, der die gesellschaftlichen Interaktionsräume rund um KI konstituiert, stellt diese Technologie als eine Lösung für ansonsten unlösbare soziale, politische und wirtschaftliche Probleme dar. In diesem Sinne kann der AI Futurismus als die neueste und vielleicht potentere Version der oben genannten Strömungen des Techno-Solutionismus und des Techno-Determinismus betrachtet werden.

Was jedoch im Hype um KI neu ist, wie Brevini beschreibt, ist eine nahezu mythische Sicht auf KI als magisches Werkzeug, welches das globale kapitalistische System vor seinen dramatischen Misserfolgen retten wird (vgl. Brevini 2021a: 149). Jeglicher Diskurs, jegliche Interaktion mit KI kann sich kaum von diesem Deutungshorizont entziehen. Befürworter*innen und Kritiker*innen ringen gleichermaßen mit den angeblich disruptiven Auswirkungen von KI. Und die meisten stimmen überein, dass KI-Technologien die sozialen Räume der Zukunft tiefgreifend verändern werden, ob man es mag oder nicht. Hier sehen wir den Kern des AI Futu-

rism nochmal in aller Klarheit: die Vorstellung von KI als selbstverständlichen und notwendigen Fortschritt.⁶

4. Die Realität von KI

Währenddessen erzählen die materiellen Manifestationen und un/realen Auswirkungen von KI-Anwendungen natürlich eine ganz andere Geschichte (siehe zu diesem Thema, z. B.: Eubanks 2018; Mühlhoff 2019, 2020a; Noble 2018; O’Neil 2016). Die oben erläuterte Ideologie des AI Futurism verschleiert ein umfassendes, globales und ausbeuterisches sozioökonomisches System, welches die Basis für KI-Technologien und ihre Funktionsweise bildet. Im Folgenden baue ich auf die Arbeit von Nancy Fraser zu den gesellschaftlichen Bedingungen des Kapitalismus auf, um dieses System nachzuzeichnen (vgl. Fraser 2022). Dies ermöglicht einen tiefen Einblick in die un/realen Räume hinter KI.

Im Zuge einer marxistischen Systemkritik arbeitet Fraser in ihrem Text heraus, wie der Kapitalismus weit über bloße wirtschaftliche Strukturen hinausgeht und eine kulturelle Logik repräsentiert, die die gesellschaftliche Organisation formt und die planetarischen Bedingungen neu ordnet. In Frasers Arbeit zeigt sich, was oben in der Ideologiekritik schon angedeutet wurde: nämlich, dass Ideologie und materielle Manifestationen untrennbar miteinander verbunden sind. In diesem Zusammenhang kann man sagen, dass die Ideologie des AI Futurism KI-Technologien gangbar für den modernen Kapitalismus macht (vgl. Daum 2022; Dyer-Whiteford et al. 2019). In diesem Zusammenhang muss also die Funktion von KI im Kapitalismus genauer betrachtet werden. Hier spielt Frasers klare Darstellung kapitalistischer Verhältnisse und gesellschaftlicher Ordnung eine entscheidende Rolle. Anhand ihrer Beschreibungen zeige ich, wie AI Futurism mit der globalen und materiellen Realität von KI-Technologien zusammenhängt. Frasers Arbeit bietet so einen idealen Ausgangspunkt, um die »realen« Auswirkungen von KI zu begreifen.

In ihrem Buch beschreibt Fraser vier Bedingungen, die für das Funktionieren des Kapitalismus grundlegend sind. Sie dienen hier als Fixpunkte, um die Hintergründe der Ideologie des AI Futurism zu beleuchten. Die erste Bedingung hebt den erheblichen Reichtum hervor, der durch die Enteignung marginalisierter Gruppen angesammelt wird und betont die Abhängigkeit des Kapitalismus von der Aneignung natürlicher Ressourcen, Enteignung indigener Gebiete und ausgebeuteter Ar-

6 Teile dieses Abschnittes bauen auf den Gedankengängen und Argumenten aus meinem Paper »The Impacts of AI Futurism – An Unfiltered Look at AI’s True Effects on the Climate Crisis« (Schütze 2024) auf.

beit, insbesondere von *racialised peoples*.⁷ Fraser unterscheidet dabei zwischen Enteignung und Ausbeutung und betont die systematische Enteignung von Arbeit und Land, insbesondere während kolonialer Perioden.

Die zweite Bedingung verdeutlicht die Bedeutung von unbezahlter und unterbezahlter Arbeit, die der sozialen Reproduktion dient und hauptsächlich von Frauen geleistet wird. Diese reproduktive Arbeit ist entscheidend für die Aufrechterhaltung von Produktionsprozessen und der Gesellschaft im Allgemeinen, da sie die Grundlage für das menschliche Zusammenleben und kulturelle Strukturen bildet. Trotz ihrer wichtigen Rolle bleibt diese Arbeit oft unerkannt und wird dennoch als impliziter notwendig angenommen.

Die dritte Bedingung betont die fortlaufende Ausbeutung planetarer Ressourcen als grundlegende Kraft für die kapitalistische Wirtschaft und ihre soziopolitische Ordnung. Die Beziehung des Kapitalismus zur Natur zeichnet sich vor allem durch Ausbeutung aus. Der Planet wird primär als Werkzeug für Profit betrachtet. Fraser verdeutlicht den paradoxen Charakter dieser Beziehung. Denn während das kapitalistische System abhängig von der Natur ist, zum Beispiel für den Abbau von Ressourcen und die Abfallbeseitigung, kreiert es gleichzeitig eine klare Trennung zwischen der ›wertvollen‹ Sphäre der Kultur und der ›wertlosen‹ Sphäre der Natur.

Die vierte und letzte Bedingung betont die Abhängigkeit des Kapitalismus von öffentlichen Strukturen, die von Staaten und anderen öffentlichen Akteur*innen bereitgestellt werden. Fraser betont, dass beispielsweise rechtliche Ordnungen, Eigentumsrechte, Währungsregime und öffentliche Infrastrukturen für die Kapitalakkumulation unerlässlich sind. Die Stabilität und Durchsetzung öffentlicher Ordnung sind unverzichtbar für die reibungslose Funktionsweise des Kapitalismus. Daher spielen öffentliche Mächte eine grundlegende Rolle bei der Aufrechterhaltung des kapitalistischen Systems.

Entscheidend ist, dass all diese Dynamiken auch im Rahmen des AI Futurismus beobachtet werden können. Es geht hier also nicht primär darum, dass wir eine komplett neue Form des Kapitalismus beobachten, oder dass es einer substantiell neu gedachten Kritik desselben bedarf. Vielmehr steht hier im Mittelpunkt, dass wir anhand der von Fraser ausgearbeiteten Bedingungen kapitalistischer Ordnung einen klareren Blick auf die Auswirkungen von KI bekommen. Denn betrachten wir diese vier analytisch gezeichneten Bedingungen im Detail, wird deutlich, dass KI-Technologien in allen diesen vier Dimensionen intensivierend wirken: (1) Die Produktion, Nutzung und Aufrechterhaltung dieser Technologien trägt zur Enteignung von marginalisierten Gruppen und *racialised peoples* bei. (2) Die KI-Industrie ist durchzogen von der Ausbeutung sozialer Reproduktion. (3) KI-Technologien,

7 Ich bleibe hier bei den englischen Begriffen, um Unklarheiten vorzubeugen. Siehe hierzu beispielsweise Rath/Gasser 2021.

besonders deren Hardware, basieren auf der Aneignung und Verwertung nicht-menschlicher Natur. (4) KI-Entwicklung ist in besonderem Maße abhängig von öffentlichen Mächten und Strukturen. Im nächsten Abschnitt werde ich diese vier Bedingungen im Kontext von KI-Technologien genauer betrachten.

5. Ein Blick hinter die Fassade

Mit den vier oben skizzierten Bedingungen soll hier noch einmal im Detail auf das bereits angeschnittene Verständnis von Technologie eingegangen werden, welches ich aus der kritischen Theorie der Technikphilosophie aufgreife (vgl. Feenberg 1999). Der zentrale Punkt dabei ist, dass Technologie nicht von Gesellschaft getrennt werden kann. Denn, wie Feng und Feenberg (vgl. 2008: 111f.) beschreiben, entsprechen Technologien notwendigerweise dem soziokulturellen Hintergrund einer Gesellschaft. Das heißt auch, dass KI-Technologien in ihrer jetzigen Funktion und Struktur kapitalistischen Dynamiken entsprechen und sie reproduzieren. Anders ausgedrückt sind Technologien also zuallererst erhaltend, weil sie darauf ausgerichtet sind, in einer bestimmten sozialen und wirtschaftlichen Ordnung zu funktionieren (vgl. Feenberg 2014: 180). Das bedeutet, dass die breite Adaption und Funktionsweise heutiger KI-Technologien nie außerhalb ihrer historisch gewachsenen Bedingungen betrachtet werden kann (siehe z.B. Dyer-Whiteford et al. 2019; Pasquinelli 2023). Technologie ist immer schon sozial und das Ergebnis gesellschaftlicher Prozesse.⁸ Im Folgenden zeigt sich diese tiefe Verbindung zwischen KI und Gesellschaft im Detail, durch den Blick auf die vier oben beschriebenen Bedingungen.

(1) Die Funktionsweise von heutigen KI-Technologien beruht auf der Enteignung marginalisierter Gruppen, deren Ressourcen, deren Arbeit und ganz konkret deren Daten. Dieser Aspekt des KI-Kapitalismus wird oft als *algorithmic colonialism* (vgl. Birhane 2020) oder als *digital/data colonialism* (z.B. Coleman 2019; Coudry/Mejias 2019; Kwet 2019) bezeichnet. Diese Begriffe betonen die fortlaufende Enteignung von Menschen in der Folge historischer Prozesse, wie die des Kolonialismus. Die großen KI-Konzerne, Google, Amazon, Meta oder Microsoft,

8 KI-Technologie könnte sicherlich auch außerhalb einer kapitalistischen Logik gedacht werden. Nur bedeutete dies auch, den soziokulturellen Horizont entsprechend mitzudenken. KI-Technologie ist heute so eng mit den großen Tech-Unternehmen verzahnt, dass ein Umdenken dessen, wie KI verwendet werden soll, unwahrscheinlich scheint. Es mag zwar kleine subversive und transformative Anwendungen geben, die KI-Technologie anders einsetzen. Im Großen ist diese Technologie jedoch notwendigerweise mit der gesellschaftlichen Realität verschränkt. Will man KI-Technologien also aus diesen historisch etablierten Prozessen heben, geht dies nicht ohne gleichzeitig auch die gesellschaftlichen Prozesse selbst anders zu gestalten (z.B. durch Regulierungen).

setzen das historische Muster der Ausbeutung und Enteignung entlang etablierter globaler Asymmetrien fort. Die Kontrolle über die globalen digitalen Infrastrukturen, beispielsweise in vielen afrikanischen Ländern, liegt meist bei diesen großen Konzernen (vgl. Birhane 2020; Coleman 2019; Kwet 2019). Ähnlich der historisch kolonialen Enteignung, beschreibt Abeba Birhane, wie mittels »algorithmic colonialism« Daten in Ländern des Globalen Südens enteignet werden, um Gewinne im Globalen Norden zu steigern und dortige Algorithmen zu speisen (vgl. Birhane 2020). Diese Praxis ist insbesondere in Ländern wie Südafrika schädlich, wo begrenzte Infrastruktur, schwache Datenschutzgesetze und geringer Wettbewerb Anfälligkeiten für die Ausbeutung durch ausländische Technologieunternehmen schaffen (vgl. Coleman 2019; Kwet 2019).

Birhanes (2020) Konzept des *algorithmic colonialism* verdeutlicht weiterhin, dass historische Machtasymmetrien in Zeiten von KI durch technologische Innovation, hochmoderne Algorithmen und KI-Lösungen fortgeführt werden. Hier treffen wir wieder auf die Ideologie des AI Futurism. Die kapitalistische Erzählung von Technologie als Schlüssel für menschlichen Fortschritt, fördert den Tech- und KI-Enthusiasmus und erweckt den Eindruck, dass soziale, wirtschaftliche, und kulturelle Probleme im Globalen Süden nur durch die Einfuhr hochmoderner Technologie aus dem Globalen Norden lösbar sind (vgl. Birhane 2020: 404). Die KI-Industrie und ihre Regulierungen sind hauptsächlich in Nordamerika und Europa angesiedelt. Von dort aus verbreitet sich der Einfluss der Ideologie des AI Futurism entlang globaler Hierarchien. Dieser Einfluss führt an vielen Stellen dazu, dass die schon existierenden Machtasymmetrien reproduziert und womöglich verstärkt werden. Grund dafür sind technologische Abhängigkeiten, die wiederum zu dauerhafter Ressourcenausbeutung führen (vgl. Kwet 2019).

(2) KI-Technologien sind grundlegend mit der Ausbeutung von sozialer reproduktiver Arbeit verflochten. Dies geschieht mindestens auf zwei wesentlichen Ebenen. Zunächst sind hier die gender-spezifischen Hierarchien der KI-Industrie zu betrachten. Die KI-Branche präsentiert sich als ein Bereich, in dem die einflussreichen Jobs weiterhin männlich dominiert sind (siehe z.B. Chang 2019). Für Frauen* sowie nicht-binäre* Personen sind diese Berufe oft schwer erreichbar, da geschlechtsspezifische Lücken in der Bildung, insbesondere in den MINT-Fächern, nach wie vor bestehen, und da Sexismus am Arbeitsplatz in Tech-Unternehmen weiterhin ein großes Problem darstellt (vgl. Gurumurthy/Chami 2021: 4). In diesen Hierarchien ist es weiterhin so, dass die soziale reproduktive Arbeit, die normalerweise von Frauen und marginalisierten Gruppen durchgeführt wird, die Tech-Arbeit, die oft von weißen Männern ausgeführt wird, überhaupt erst ermöglicht. Trotz gegenteiliger Behauptungen bleibt die KI-Branche ein grundlegend geschlechtsspezifischer und patriarchaler Raum, während soziale reproduktive Arbeit hauptsächlich von Frauen* geleistet wird. Insbesondere in den Ländern, die die geringsten Fortschritte in der Geschlechtergleichstellung aufweisen, bedeuten

Armut und die damit verbundene Marginalisierung vor allem junger Frauen, dass der KI-Hype einfach an ihnen vorbeizieht (vgl. ebd.: 3).

So wird klar, wie die Geschäftsmodelle vieler KI-Unternehmen auf bestehenden Strukturen der gender-spezifischen Ausbeutung aufbauen. Gleichzeitig schaffen diese Unternehmen aber auch neue abgeschlossene Räume, die hauptsächlich Männern vorbehalten sind und in denen die vermeintlich zukunftsweisende Arbeit stattfindet. Ausgehend von diesen Räumen werden KI-Anwendungen entwickelt, die gesellschaftliche Zukunft gestalten sollen. Entscheidungen über diese Zukunftsgestaltung werden also in sehr spezifischen, hierarchischen und diskriminierenden Kontexten getroffen.

Über diese patriarchalen Strukturen hinaus tritt eine zweite Dimension der Ausbeutung sozialer Reproduktion durch KI-Technologien auf. Diese bezieht sich vor allem auf kulturelle Arbeit in un/realen Interaktionsräumen. Die sozialen und kulturellen Sphären, die hauptsächlich online auf Plattformen wie YouTube, Twitch, TikTok oder Instagram stattfinden, werden massiv durch KI-Unternehmen kapitalisiert. Soziale Handlungs- und Interaktionsräume werden aktiv von Benutzer*innen aufgebaut und unterhalten, indem sie zum Beispiel Inhalte erstellen, miteinander interagieren (über Likes und Kommentare) und sich vernetzen. Unternehmen wie Google monetarisieren diese Räume, indem sie die Daten, die auf ihren Plattformen generiert werden, sammeln und extrahieren. Wie der Medienwissenschaftler Timo Daum festhält, werden Benutzer*innen hierbei zu einem quasi dreiköpfigen Wesen, das gleichzeitig Kund*in, Lieferant*in und Produkt ist (vgl. Daum 2022: 253).

So wird das Posten, Liken und Erstellen von Inhalten für die KI-Industrie zu wertvoller Arbeit, und die Reproduktion sozialer und kultureller Räume wird Bestandteile der Verwertungskette rund um KI-Technologien. Tiziana Terranova bezeichnet diesen Prozess als »freie Arbeit« der Benutzer*innen (vgl. Terranova 2000). Denn diese arbeiten rund um die Uhr und unentgeltlich. Diese Ausbeutung freier Arbeit existiert nicht erst seit dem KI Hype. So beschreibt Terranova frühere Formen »freier Arbeit«, zum Beispiel Aktivitäten während der Anfänge des Internets, wie Chatten, Mailing-Listen oder Amateur-Newslettern (vgl. ebd.: 38). Dies waren die Vorläufer der heute auf Sozialen Medien und Online-Plattformen vorherrschenden Arbeit. KI, getrieben vom Datenextraktivismus, verstärkt und setzt diese Dynamiken fort. Denn un/reale Interaktionsräume, geschaffen durch Plattform-Unternehmen, sind zentral für die Funktionsweise von KI Technologien. Die durch Algorithmen medierte Interaktionen (freie Arbeit) erzeugen Daten, die dann wiederum mit und für KI verwertet werden. So erhält und intensiviert die KI-Branche die Ausbeutung reproduktiver Arbeit auf diesen beiden miteinander verbundenen Ebenen – freie Arbeit und gender-spezifische Hierarchien.

(3) Zugleich hat AI Futurism ganz konkrete Auswirkungen auf den Planeten. Denn die Existenz dieser Ideologie ist immer auch materiell eingebettet in be-

stimmte Praktiken. Beispielsweise zeigt Benedetta Brevini (vgl. 2021b) in ihrem Buch »Is AI good for the planet?«, dass KI-Anwendungen eine verschwenderische Lebensweise und übermäßigen Konsum fördern. Wie schon oben genannt, sind beispielsweise Soziale Medien darauf ausgelegt, die Aufmerksamkeit von Benutzer*innen solange wie möglich zu halten, um somit Datenströme aufrechtzuerhalten und Werbezeit zu verlängern. So tragen Algorithmen nicht nur zur stetigen Optimierung von Produkten bei, sondern auch zu einer immer größer werdenden, fein abgestimmten Palette von Waren und Dienstleistungen. Durch prädiktive Algorithmen und datengesteuerte Werbung können Unternehmen immer mehr und immer gezielter verkaufen (vgl. ebd.: 63–91).

Diese Dynamiken haben natürlich Auswirkungen auf den Planeten. Es werden enorme Mengen an Energie und Ressourcen gebraucht, um dieses System aufrechtzuerhalten. Kate Crawford (2021) beschreibt die KI-Industrie daher sogar als eine ausbeuterische und zerstörerische Maschinerie. Das Training von Algorithmen und die Speicherung wachsender Datenmengen erfordern kontinuierlich mehr Energie und Wasser, um die dafür erforderlichen Rechenzentren und deren Rechenleistung kontinuierlich zur Verfügung stellen zu können. Gleichzeitig führen diese Dynamiken zu einer zunehmenden Ansammlung von Elektroschrott, der den Planeten schädigt (vgl. Brevini 2021b: 64–80). Von den Entwicklungen der KI-Industrie profitiert jedoch nur eine kleine Gruppe Menschen. In großen Teilen der Welt kommen die neuesten Technologien gar nicht erst an, trotzdem aber leiden gerade diese Regionen am meisten unter der Klimakrise und den planetaren Folgen der zerstörerischen Praktiken der KI-Industrie.

(4) Die Entwicklung von KI hat eine lange Geschichte, die eng mit staatlichen Subventionen und Interessen sowie öffentlichen Infrastrukturen verbunden ist. Die Ursprünge dieser Entwicklungen reichen bis in die 1950er Jahre zurück, wo die berühmte von der US-Regierung organisierte Dartmouth-Konferenz 1956 eine der frühesten Formen staatlicher Beteiligung im Bereich KI war (vgl. Steinhoff 2021: 119). In den folgenden Jahren begannen Regierungen weltweit massiv in die KI-Forschung zu investieren. In den 1980er und 1990er verlagerte sich der öffentliche Fokus dann zunächst auf Internettechnologien, und das staatliche Interesse und entsprechende Finanzierung wandten sich anderen digitalen Kommunikationsinfrastrukturen zu (vgl. Schiller 1999; Steinhoff 2021: 118). Diese öffentlich finanzierten Infrastrukturen, hier hauptsächlich das Internet, bildeten später die Grundlage für das Sammeln und Verwerten von Daten, aus welchem die heutigen Tech- und KI-Unternehmen hervorgingen.

Während KI und digitale Infrastrukturen vom Staat finanziert wurden, etablierte sich eine enge Allianz zwischen Finanzkapitalismus und Informationstechnologien. In seinem Buch »Kapitalismus und Ressentiment« beschreibt Joseph Vogl (vgl. 2021) detailliert, wie sich staatlich gestütztes Finanzkapital und dessen Interesse an Informationstechnologien zu einem globalen Regime entwickelt, das

von monetären und öffentlichen Imperativen angetrieben wird. Entscheidend für den Erfolg und das Wachstum der heutigen KI-Industrie war eben dieses Regime am Schnittpunkt von öffentlicher Regulierung und privaten Interessen. In diesem Zusammenspiel wurden globale Steuer-Konstrukte geschaffen, in denen sich die Tech-Unternehmen breit machen können. Durch dieses Regime und seine entsprechenden Strukturen wird es beispielsweise Uber ermöglicht, nur 2,4 Prozent seines gesamten Umsatzes als steuerpflichtig zu erklären (vgl. ebd.: 75).

In letzter Zeit bilden vor allem die neuesten Datenschutz- und KI-Bestimmungen die Grundlage, auf der Tech-Unternehmen ihre Geschäftsmodelle aufbauen. Richtlinien wie die EU-DSGVO sind entscheidend, um die Regeln festzulegen, unter denen Daten für die Weiterentwicklung von KI-Technologien gesammelt und genutzt werden können. Wie die Rechtswissenschaftlerin Amy Kapczynski (vgl. 2020) schreibt, bilden solche Vorschriften eine wesentliche Grundlage für die großen Konzerne, da sie festlegen, wie Daten gesammelt, verarbeitet und verwendet werden können. So zeigt Kapczynski, dass Gesetze die privatisierte Kontrolle über Daten und deren Sammlung sichern. Das heißt, diese Richtlinien bilden den Rahmen für die Geschäftsgrundlage der KI-Industrie und machen sie gesellschaftlich akzeptabel. Am Ende steht jedoch, dass gemeinschaftlich produzierte Daten von einer Handvoll privater Akteur*innen verwertet werden, die damit wiederum ihre eigenen Profite maximieren (vgl. Daum 2022: 249). Das bedeutet, dass öffentliche Strukturen, wie rechtliche Regulierungen, die Praktiken der KI-Industrie, wie das Sammeln von Daten, erleichtern und legitimieren. Dadurch entsteht eine staatlich unterstützte Privatisierung finanzieller Gewinne, deren Grundlage kostenlos gesammelte, gemeinschaftliche Daten sind.⁹

6. Die un/realen Räume hinter KI

Am Ende zeigt sich hier ein gesellschaftliches Regime, indem KI-Technologien und kapitalistische Dynamiken eng miteinander verwoben sind. Dabei wird deutlich, dass die entstehende soziale Ordnung un/reale Interaktionsräume ganz konkret einschränkt. Denn der Blick hinter die Fassade von KI hat gezeigt, dass historische Machtasymmetrien reproduziert und neue Strukturen, wie die des »algorithmic colonialism«, erschaffen werden. Dieses soziomaterielle System ist maßgeblich daran beteiligt diskriminierende und globale Hierarchien der Interaktion aufrechtzuerhalten. So werden historische Muster der Ausbeutung durch KI intensiviert und bestimmen weiterhin moderne gesellschaftliche Räume der Zusammenkunft.

9 Teile dieses Abschnittes bauen auf den Gedankengängen und Argumenten aus meinem Paper »The Impacts of AI Futurism – An Unfiltered Look at AI's True Effects on the Climate Crisis« (Schütze 2024) auf.

Dies zeigt sich auch in lokaleren Strukturen, wie der Ausbeutung sozialer reproduktiver Arbeit. Soziale und kulturelle Räume, besonders im Digitalen, werden abgeschöpft und Interaktionen werden dort kapitalisiert. Hinzu kommen genderspezifische Hierarchien in der KI-Industrie. Die KI-Branche bleibt ein patriarchaler Raum, der bestehende Ungleichheiten verstärkt und ein entsprechendes gesellschaftliches Miteinander bestimmt. In anderen Worten konstituiert KI-als-soziomaterielles-System patriarchale un/reale Interaktionsräume.

Neben diesen sozialen Auswirkungen hat die Ideologie des AI Futurism auch planetare Folgen. Durch die KI getriebene Ausbeutung nichtmenschlicher Natur werden zukünftige Handlungsräume massiv eingeschränkt. Doch nicht nur das. Besonders die Räume marginalisierter Gruppen werden durch die Ideologie und das System hinter KI kolonisiert. Das Handeln im Globalen Norden beeinflusst also Interaktionsräume im Globalen Süden. Währenddessen profitiert eine kleine privilegierte Gruppe von den Entwicklungen der KI-Industrie. Am Ende dieses Beitrages steht also, dass AI Futurism nicht nur eine abstrakte Ideologie beschreibt. Die Vorstellungen, Mythen und Narrative von KI, die micro-procedures des soziomateriellen Systems, wirken auf soziale und individuelle Möglichkeitsräume ein, mit gravierenden Folgen.

Literatur

- Althusser, Louis (2010): *Ideologie und ideologische Staatsapparate*, Hamburg: VSA.
- Ballatore, Andrea/Mark Graham/Shilad Sen (2017): »Digital Hegemonies: The Locality of Search Engine Results«, in: *Annals of the American Association of Geographers* 107(5), S. 1194–1215. <https://doi.org/10.1080/24694452.2017.1308240>
- Birhane, Abeba (2020): »Algorithmic Colonization of Africa«, in: *SCRPTed* 389, 17(2). <https://doi.org/10.2966/scrip.170220.389>
- Brevini, Benedetta (2021a): »Creating the Technological Saviour: Discourses on AI in Europe and the Legitimation of Super Capitalism«, in: Pieter Verdegem (Hg.), *AI for Everyone?: Critical Perspectives*, London: University of Westminster Press. <https://doi.org/10.16997/book55.i>
- Brevini, Benedetta (2021b): *Is AI Good for the Planet?* Cambridge/Boston/Oxford/New York: Polity.
- Campolo, Alexander/Kate Crawford (2020): »Enchanted Determinism: Power without Responsibility in Artificial Intelligence«, in: *Engaging Science, Technology, and Society*, 6, S. 1–19. <https://doi.org/10.17351/ests2020.277>
- Chang, Emily (2019): *Brotopia: Breaking Up the Boys: Club of Silicon Valley*, New York: Portfolio.
- Chun, Wendy H. K. (2005): *Control and Freedom: Power And Paranoia In The Age Of Fiber Optics*, Cambridge: MIT Press.

- Coleman, Danielle (2019): »Digital Colonialism: The 21st Century Scramble for Africa through the Extraction and Control of User Data and the Limitations of Data Protection Laws«, in: Michigan Journal of Race & Law 24(2). <https://doi.org/10.36643/mjrl.24.2.digital>
- Cowls, Josh/Andreas Tsamados/Mariarosaria Tadde/Luciano Floridi (2021): »A definition, benchmark and database of AI for social good initiatives«, in: Nature Machine Intelligence 3(2), S. 111–115. <https://doi.org/10.1038/s42256-021-00296-0>
- Cowls, Josh/Andreas Tsamados/Mariarosaria Tadde/Luciano Floridi (2023): »The AI gambit: Leveraging artificial intelligence to combat climate change—Opportunities, challenges, and recommendations«, in: AI & SOCIETY 38(1), S. 283–307. <https://doi.org/10.1007/s00146-021-01294-x>
- Couldry, Nick/Ulises A. Mejias (2019): *The Costs of Connection: How Data Is Colonizing Human Life and Appropriating It for Capitalism*, Redwood City: Stanford University Press.
- Crawford, Kate (2021): *Atlas of AI: The Real Worlds of Artificial Intelligence*, New Haven: Yale University Press.
- Daum, Timo (2022): »AI as the Latest Machine of Digital Capitalism—For Now«, in: Florian Butollo/Sabine Nuss (Hg.), *Marx and the Robots*, London: Pluto Press, S. 242–255.
- Dhar, Payal (2020). »The Carbon Impact of Artificial Intelligence«, in: Nature Machine Intelligence 2(8), S. 423–425.
- Dubber, Markus D./Frank Pasquale/Sunit Das (Hrsg.) (2020): *The Oxford Handbook of Ethics of AI*, New York: Oxford University Press.
- Dyer-Witheford, Nick/Alte Mikkola Kjøsén/James Steinhoff (2019): *Inhuman Power: Artificial Intelligence and the Future of Capitalism*, London: Pluto Press.
- Elish, M. C./danah boyd (2018): »Situating Methods in the Magic of Big Data and AI«, in: Communication Monographs 85(1), S. 57–80. <https://doi.org/10.1080/03637751.2017.1375130>
- Eubanks, Virginia (2018): *Automating Inequality: How High-Tech Tools Profile, Police, and Punish the Poor*, New York: St. Martin's Press.
- European Commission (2023): *A European Approach to Artificial Intelligence: Shaping Europe's Digital Future*. Abrufbar unter: <https://digital-strategy.ec.europa.eu/en/policies/european-approach-artificial-intelligence> (Stand: 04.01.2024).
- Feenberg, Andrew (1999): *Questioning Technology*, London/New York: Routledge.
- Feenberg, Andrew (2014): *The Philosophy of Praxis: Marx, Lukács And The Frankfurt School*, London/New York: Verso Books.
- Feng, Patrick Feenberg, Andrew (2008): »Thinking about Design: Critical Theory of Technology and the Design Process«, in: P. Kroes, P. E. Vermaas, A. Light, & S. A. Moore (Hg.), *Philosophy and Design: From Engineering to Architecture*, Springer Netherlands, S. 105–118. https://doi.org/10.1007/978-1-4020-6591-0_8

- Floridi, Luciano/Josh Cows/monica Beltrametti/Raja Chatila/Patrice Chazerand/Virginia Dignum/Christoph Luetge/Robert Madelin/Ugo Pagallo/Francesca Rossi/Burkhard Schafer/Peggy Valcke/Effy Vayena (2018): »AI4People—An Ethical Framework for a Good AI Society: Opportunities, Risks, Principles, and Recommendations«, in: *Minds and Machines* 28(4), S. 689–707. <https://doi.org/10.1007/s11023-018-9482-5>
- Fraser, Nancy (2022): *Cannibal Capitalism: How Our System is Devouring Democracy, Care, and the Planet and What We Can Do About It*, London/New York: Verso.
- Google AI. (o.J.-a). Why We Focus on AI, Google AI. Abrufbar unter: <https://ai.google/why-ai/> (Stand: 04.01.2024).
- Google AI. (o.J.-b). Google AI, Google AI. Abrufbar unter: <https://ai.google/> (Stand: 04.01.2024).
- Google Sustainability. (o.J.): Sustainable Innovation & Technology, Google Sustainability. Abrufbar unter: <https://sustainability.google/> (Stand: 04.01.2024).
- Gurumurthy, Anita/Nandini Chami (2021): *The Deal We Always Wanted: A Feminist Action Framework for the Digital Economy*, Berlin: Friedrich-Ebert-Stiftung.
- Hasselbalch, Gry (2021): *Data Ethics of Power: A Human Approach in the Big Data and AI Era*, Cheltenham: Edward Elgar Publishing Ltd.
- Jaeggi, Rahel/Robin Celikates (2017): *Sozialphilosophie: Eine Einführung*, München: C.H.Beck.
- Jansen, Sue C. (1993): »Censorship, Critical Theory, and New Information Technologies: Foundations of Critical Scholarship in Communications«, in: John E. Buschman (Hg.): *Critical Approaches to Information Technology in Librarianship: Foundations and Applications*, Westport: Greenwood Press, S. 59–81.
- Kapczynski, Amy (2020): »The Law of Informational Capitalism«, in: *The Yale Law Journal* 129(5), S. 1276–1599.
- Kwet, Michael (2019): »Digital Colonialism: US Empire and the New Imperialism in the Global South«, in: *Race & Class* 60(4), S. 3–26. <https://doi.org/10.1177/0306396818823172>
- Leeker, Martina (2017): *Intervening Infrastructures: Ad Hoc Networking and Liberated Computer Language. An Interview with Alexander R. Galloway*, in Howard Caygill/Martina Leeker/Tobias Schulze (Hg.) *Interventions in Digital Cultures. Technology, the Political, Methods*, Lüneburg: meson press, S. 61–72.
- Lindemann, Nora Freya (2024): »Chatbots, Search Engines, and the Sealing of Knowledges«, in: *AI & SOCIETY*. <https://doi.org/10.1007/s00146-024-01944-w>
- Lösch, Andreas/Armin Grunwald/Martin Meister/Ingo Schulz-Schaeffer (Hrsg.) (2019): *Socio-Technical Futures Shaping the Present: Empirical Examples and Analytical Challenges*, Wiesbaden: Springer VS.
- Morozov, Evgeny (2013): *To Save Everything, Click Here: Technology, Solutionism, and the Urge to Fix Problems that Don't Exist*, London: Penguin.

- Mühlhoff, Rainer (2019): »Big Data is Watching You. Digitale Entmündigung am Beispiel von Facebook und Google«, in: Rainer Mühlhoff/Anja Breljak/Jan Slaby (Hg.), *Affekt Macht Netz: Auf dem Weg zu einer Sozialtheorie der digitalen Gesellschaft*, transcript, S. 81–107. <http://dx.doi.org/10.25969/mediarep/13222>
- Mühlhoff, Rainer (2020): »Human-Aided Artificial Intelligence: Or, How to Run Large Computations in Human Brains? Toward a Media Sociology of Machine Learning«, in: *New Media & Society* 22(10), S. 1868–1884. <https://doi.org/10.1177/1461444819885334>
- Ndaka, Angella K. (2023). *Sustainable AI Techno-Futures? Exploring Socio-Technical Imaginaries of Agtech in Aotearoa New Zealand*, Otago: University of Otago.
- Noble, Safiya Umoja (2018): *Algorithms of Oppression: Data Discrimination in the Age of Google*, New York: New York University Press.
- O’Neil, Cathy (2016): *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy*, New York: Crown Publishing Group.
- Pasquinelli, Matteo (2023): *The Eye of the Master: A Social History of Artificial Intelligence*, 1st edition, London/New York: Verso Books.
- Sartori, Laura/Giulia Bocca (2023): »Minding the Gap(s): Public Perceptions of AI and Socio-Technical Imaginaries«, in: *AI & SOCIETY* 38(2), 443–458. <https://doi.org/10.1007/s00146-022-01422-1>
- Sartori, Laura/Andreas Theodorou (2022): »A Sociotechnical Perspective for the Future of AI: Narratives, Inequalities, and Human Control«, in: *Ethics and Information Technology* 24(1), <https://doi.org/10.1007/s10676-022-09624-3> (Stand: 23.07.2023).
- Sattiraju, Nikitha (2020, February 4): »The Secret Cost of Google’s Data Centers: Billions of Gallons of Water«. Abrufbar unter: <https://time.com/5814276/google-data-centers-water/> (Stand: 04.01.2024)
- Schiller, Dan (1999): *Digital Capitalism: Networking the Global Market System*, Cambridge: MIT Press.
- Schütze, Paul (2024): »The Impacts of AI Futurism: An Unfiltered Look at AI’s True Effects on the Climate Crisis«, in: *Ethics and Information Technology* 26, <https://link.springer.com/article/10.1007/s10676-024-09758-6> (Stand: 23.07.2024).
- Schwerzmann, Katia (2021): »Abolish! Against the Use of Risk Assessment Algorithms at Sentencing in the US Criminal Justice System«, in: *Philosophy & Technology*. <https://doi.org/10.1007/s13347-021-00491-2>
- Shah, Chirag/Emily M. Bender (2022): »Situating Search«, in: *Proceedings of the 2022 Conference on Human Information Interaction and Retrieval*, S. 221–232. <https://doi.org/10.1145/3498366.3505816>
- Steinhoff, James (2021): *Automation and Autonomy: Labour, Capital and Machines in the Artificial Intelligence Industry*, Palgrave Macmillan.
- Terranova, Tiziana (2000): »Free Labor: Producing Culture for the Digital Economy«, in: *Social Text* 18 (2), S. 33–58. https://doi.org/10.1215/01642472-18-2_63-33

- Vogl, Joseph (2021): Kapital und Ressentiment, 1. Aufl., München: C.H.Beck.
- von Rath, Anna/Lucy Gasser (2021): Goethe Institut. Zehn schwierig zu übersetzende Begriffe in Bezug auf Race. Abrufbar unter: <https://www.goethe.de/ins/us/de/ku/wir/22139756.html> (Stand: 05.01.2025)
- Whittaker, Meredith (2021): »The Steep Cost of Capture«, in: Interaction 28 (6), S. 50–55. <https://doi.org/10.1145/3488666>
- Winkler, Hartmut (1997): Docuverse: Zur Medientheorie der Computer, München: Boer Verlag.
- Žižek, Slavoj/Nicholas Abercrombie/Theodor Adorno/Louis Althusser/Michele Barrett (Hrsg.) (2012): *Mapping Ideology*, London/New York: Verso.

Sektion 2: Erfahrung





L.I.S.A.: Akt 1

Lynn Klemmer

Keywords *Videoinstallation; Bildschirmmedien; digitale Realitäten*

L.I.S.A.: Act I ist Teil einer Reihe von experimentellen Reflexionen über die Ontologie von Bildschirmen. Die Videoinstallation zeigt einen schwer drapierten, blauen Vorhang aus verschiedenen Blickwinkeln, während im Hintergrund ein atmosphärisches Soundstück läuft. Die Ein- und Ausgänge des Ausstellungsraumes sind mit bläulich-gefärbten, transparenten PVC-Streifen bedeckt. Vor der Projektion sind Reste von Industrierohren platziert, die an eine zerstörte Ruinenlandschaft erinnern. Die Arbeit hinterfragt die Art und Weise, wie sich Bilder offenbaren und beleuchtet die materiellen, phänomenologischen und fiktionalen Ebenen, die sie durchqueren. Ebenso stellt sie die Frage, ob die Bildschirme nicht zwangsläufig einen transparenten Vorhang darstellen, der in unserer Wahrnehmung mit den gezeigten Bildern verschwindet.

Der Titel des größeren Projekts lautet aus mehreren Gründen L.I.S.A. Er bezieht sich nicht nur auf den Desktop-Computer von Apple aus dem Jahr 1983, der durch die Einführung von Skeuomorphismen (interaktive Bildschirmikonen) die Benutzeroberfläche des Computers revolutionierte, sondern der Name ist ebenfalls eine nicht so direkte Anspielung auf die Mona Lisa und referiert auf eine fiktionale Person.

Warum diese zwei Dimensionen? Der Apple-Computer ist ein prägnantes Beispiel dafür, wie unsere Beziehung zu einem Gerät wie dem Computer durch die Schnittstelle der Fiktion vermittelt wird: Wir »nutzen« den vermenschlichten Computer, indem wir auf Symbole klicken, die an reale Objekte erinnern und uns das Gefühl geben, dass wir zum Beispiel Dateien in den Papierkorb legen oder mit Dateien auf dem Desktop interagieren, die wir auf dem Bildschirm lokalisieren können. Der Name des Computers ist ebenso verwoben mit Steve Jobs eigener Familiengeschichte und referiert auf seine erste Tochter Lisa.



Abb. 1: I Will Not Return Installation Views 2 (Photo: Mathieu Buchler)

Abb. 2: Installationsansicht von L.I.S.A.: Act I. (Photo: Mathieu Buchler)





Abb. 3: I Will Not Return Installation Views 1 (Photo: Mathieu Buchler)

Abb. 4: Installationsansicht von L.I.S.A.: Act I. (Photo: Mathieu Buchler)



In ähnlicher Weise ist die Mona Lisa ein auratisches Objekt, das hinter ihrer eigenen Vermittlungsschicht eingeschlossen bleibt: ein Stück Schutzglas, das sie im Jahr 2022 vor der Zerstörung durch eine Kuchenattacke eines als alte Dame verkleideten Mannes im Louvre bewahrt hat.¹ Auch die Mona Lisa ist Gegenstand fiktionaler Spekulationen. Man sagt, dass ihre Augen dem*der Betrachter*in folgen. Sie ruft »Mandela-Effekte« und fiktive Erinnerungen über ihr enigmatisches Lächeln, welches laut Margaret Livingstone nur im peripheren Sehfeld während dem Betrachten anderer Teile des Bildes erscheint, auf.² Zuletzt wird ihr Schleier in Zusammenhang mit ihrer vermuteten Schwangerschaft zu dem Zeitpunkt der Herstellung des Gemäldes gebracht.³ In diesem Sinne ist ihre Persona umhüllt von verschiedenen Schichten der Fiktion.

Mit diesen Namensgebern und ihrer Geschichte – dem L.I.S.A.-Computer und der Mona Lisa – versucht das Projekt L.I.S.A. eine narrative Struktur aufzubauen, in der Fiktion und Realität miteinander verschmelzen. In Akt I werden die Rahmenbedingungen hierfür kreiert, indem die Verschleierung als Schnittstelle zwischen Wirklichkeit und Erzählung figurativ hervorgehoben wird. Das Netzwerk von Referenzen rahmt das Projekt und öffnet den Raum für die eigentliche projizierte oder gezeigte Videoinstallation. Durch den spielerischen Umgang mit der Realität und unseren Vorstellungen davon sind die Betrachter*innen eingeladen, ihre eigenen Geschichten zu spinnen und Fragen über die Künstlichkeit unserer Wahrnehmung zu stellen sowie der Tatsache nachzugehen, dass wir ständig fiktionale Welten in einer unvereinbaren Realität erschaffen. Ziel des Projekts ist es nicht, den Verlust eines objektiven Realitätssinns infolge des ›digital turn‹ zu beklagen, sondern den konstitutiven Schleier oder Bildschirm in den Vordergrund zu rücken, durch den wir die Realität – ob digital oder manifestierend – wahrnehmen und damit in Austausch treten. Durch das digitale Medium wird versucht, die Medialität (zwischen Subjekten, Objekten und Welten) unserer Interaktion mit der Realität selbst zu ergründen und dadurch die Trennung von Realität und Fiktion zu dekonstruieren und deren Schnittstelle als konstitutives Element der Bedeutung beider Begriffe hervorzuheben.

Anmerkungen

1 »Mona Lisa: Man dressed as old woman throws cake at da Vinci painting«, *BBC*, 30.05.2022. Aufgerufen am 21.05.24 (<https://www.bbc.co.uk/news/entertainment-arts-61635822>).

2 Livingstone, Margaret. »Is It Warm? Is It Real? Or Just Low Spatial Frequency?«, *Science*, AAAS.

3 »Mona Lisa scans suggest she was pregnant«, *The Guardian*, 28.09.2006. Aufgerufen am 21.04.2024 (<https://www.theguardian.com/world/2006/sep/28/arts.artsnews>).

Von der Non-Linearität der Zeit

Metaphorisieren als Interaktion mit filmischen Bildern. Eine medienästhetische Perspektive

Katerina Papadopoulou

Abstract *Der Beitrag widmet sich der Frage, inwiefern ein Metaphorisieren beim Filmerleben als (un)reale Interaktion zwischen Zuschauer*innen und dem Medium Film verstanden werden kann. Am Beispiel einer Fallstudie aus dem Science-Fiction-Film ARRIVAL wird dargestellt, wie – aus einem situativen Filmerleben heraus – die Non-Linearität der Zeit als körperliche Wahrnehmung und kognitives Konzept entsteht. Ausgehend vom transdisziplinären Theorie- und Analysekonzept der kinematischen Metapher wird argumentiert, dass es sich bei Metaphern im Film weniger um statische Konzepte handelt, die sich auf der Narrationsebene entfalten, sondern dass sie dynamisch aus der speziellen Erfahrung des Filmesehens heraus entstehen. Das Filmverstehen sowie die Metapherbildung wird in dieser Perspektive aufseiten der Zuschauenden verortet: Beides vollzieht sich in einer Interaktion von filmischen Bewegungsbildern, Wahrnehmungsempfinden und Prozessen der Sinnbildung seitens der Zuschauer*innen. In diesem Sinne illustriert der Beitrag eine Form der Interaktion, die nur auf den ersten Blick »unreal« ist.*

Keywords *Kinematische Metapher; Medienästhetische Methode; Ausdrucksbewegung; Affekt; Erfahren; Filmverstehen*

1. Einleitung: Interaktion mit Film?

Wenn Zuschauer*innen den Science-Fiction Film ARRIVAL (Denis Villeneuve, USA 2016, 116 Min.) im Kino, an ihren Fernseher-, Laptop- oder Handybildschirmen gesehen haben, dann haben sie in ihrem Filmerlebnis die Erfahrung non-linearer Zeit gemacht. Doch was heißt das genau? Wie ist das überhaupt möglich und wie ist es zu verstehen? Die (neo-)phänomenologisch geprägte Filmtheorie Hermann Kappelhoffs (2004; 2008a; 2016; 2018; Kappelhoff et al. 2024) formuliert hierzu eine Antwort: Die Zuschauer*innen durchlaufen im Prozess des Filmesehens, -hörens und -spürens eine spezifische Erfahrung; sie erleben die filmischen Bilder als eine fort-

währendere *Verwicklung* von eigenem körperlichem Erleben mit den Bewegungsbildern der filmischen Inszenierung:

In den audiovisuellen Bildern wird ein Raum erschlossen, in den man sich mit seinen Sinnen und Lüsten einnistet und ihn besetzt [...] Man *fädelt sich mit seinen Gefühlen und Gedanken in die Bewegung der Bewegtbilder ein*, lässt sich *in sie verwickeln und lässt aus dieser Verwicklung* ungesehene und unerhörte Bildräume des eigenen Empfindens entstehen (Kappelhoff 2018: 12; Herv. KP)

In eben diesem *Erfahrungsraum* und Prozess des Erlebens, Erfahrens und Fühlens entsteht indes (figurative) Bedeutung. In diesem Prozess ist (figurative) Bedeutung dynamisch: sie entfaltet sich aus *einer Interaktion* mit den Filmbildern in der Zeit, über den Verlauf des Films. Denn Kappelhoff hält fest:

[W]ir begreifen die filmische Rezeption keineswegs als eine reproduzierende Aktivität, die nur wiederholend entschlüsselt, was im Film als (narratives) Sinngebilde mehr oder weniger strikt vorgegeben ist. Vielmehr gehen wir von *einer Interaktion zwischen Rezipient und audiovisuellem Bewegtbild* aus, *einem konkreten Machen im Medienkonsum, das den sinnhaften Zusammenhang eines filmischen Bildes erst herstellt*. Damit wird die Rezeption selbst als eine Form der Produktion, die rezeptive Aneignung audiovisueller Bewegtbilder als ein *genuiner Akt des Herstellens* avisiert. Wir sprechen deshalb von der Rezeption audiovisueller Bilder als einer *Poiesis des Filme-Sehens*. (Kappelhoff 2018: 10; Herv. KP)

Mit dieser filmwissenschaftlichen Idee einer *Poiesis des Filmesehens* (Kappelhoff 2004; 2008a; 2016; 2018; Müller/Kappelhoff 2018; Kappelhoff et al. 2024) von Seiten der Zuschauerschaft setzt sich Kappelhoff von narrativen und kognitiven Modellen in der Filmtheorie ab.¹ Damit verschiebt sich der Fokus filmischer Betrachtungsweise aus einer (neo-)phänomenologisch geprägten medienästhetischen Sicht wie Kappelhoffs (2004, 2016, 2018; vgl. auch Sobchack 1990; 1992) vom Inhalt zur bisher unberücksichtigten Wirkungsdimension von Filmen: ihren affektiv-perzeptiven Qualitäten und der damit verbundenen Wahrnehmungsempfindung eines Publikums (dem sog. *Zuschauergefühl* nach Kappelhoff 2004: 80, 87f.; 2008b; 2018: 144–146; Kappelhoff/Bakels 2011). Denn die Reflexion über audiovisuelle Medien öffnet sich damit ihrer medialen Beschaffenheit, nämlich dem Fakt, dass es sich beim Rezipieren von Filmen hauptsächlich um das Sehen, Hören und Spüren audiovisueller *Bewegtbilder* (Deleuze 1989 [1983]; Plessner 1982 [1925]) handelt. Aus einer solch me-

1 Für eine ausführliche Darstellung der kritischen Auseinandersetzung mit kognitivistischen und narrativen Filmtheorien, vgl. v.a. Kappelhoff (2018); Müller/Kappelhoff (2018); Greifenstein et al. (2018); Schmitt (2020). Für einen Versuch, beide filmtheoretischen Ansätze fruchtbar zu verbinden, vgl. Fahlenbrach (2018).

dienästhetischen Sicht stellt sich folglich die entscheidende Frage, *wie* genau Filme *in Wechselwirkung mit* denkenden *und* empfindenden Wesen bedeutsam, bzw., wahrgenommen und verstanden werden. Mit dieser perspektivischen Verlagerung rückt somit die zeitlich anhaltende Verflechtung zwischen affizierenden audiovisuellen Medien und einer aktiv empfindenden (und denkenden) Zuschauerschaft in den Vordergrund. In eben diesem Sinne ist im obigen Zitat die Rede von einer *Interaktion* zwischen Zuschauenden und Filmbildern.

In der Theorie Kappelhoffs (2004; 2008a; 2016; 2018; Kappelhoff et al. 2024) ist diese Interaktion von Zuschauer*in und Filmbild ferner eine ebenso zwischenleibliche wie die zwischen Menschen. Diese Parallelität dynamischer körperlicher Interaktion zeigen Müller und Kappelhoff (2018) in ihrem Buch *Cinematic Metaphor – Experience, Affectivity, Temporality* auf. Sie illustrieren dies an der Entstehung von Metaphorizität aus *Bewegungserfahrung* (daher der Begriff der *kinematischen Metapher*)² u. a. im Tanzunterricht und in Gruppengesprächen im Face-to-Face-Kontext einerseits und im medialen Kontext aus der Bewegungserfahrung filmischer Bilder anhand verschiedener audiovisueller Formate (TV-Nachrichten bis Hollywood)³ andererseits. Daher ist näher auch nicht die Rede von einer spezifischen Metapher (im medialen Kontext) auf der Ebene der filmischen Narration oder der multimodalen Realisierung konzeptueller Metaphern, sondern von einem dynamischen, aus der Interaktion mit den Filmbildern entstehendem *Metaphorisieren* (vgl. Müller 2019; Cameron 2018; Müller/Kappelhoff 2018). Das Metaphorisieren ist ein Prozess der *dynamischen* Bedeutungsentfaltung, der aus der verkörperten Interaktion, aus jener Verwicklung des eigenen Körpers mit den filmisch affizierenden Bewegungsbildern heraus entsteht (vgl. Müller/Kappelhoff 2018; Greifenstein et al. 2018).

Eben daran schließt dieser Beitrag mit einer Fallstudie an, die im Einklang mit dem ihr zugrunde liegenden medienästhetischen Konzept der kinematischen Metapher (Müller/Kappelhoff 2018) an einer Beispielanalyse illustriert, wie aus der Verwicklung mit dem Film, aus der körperlichen Erfahrung der filmischen Bewegungsbilder, *interaktiv* durch eine (kinematische) Metaphorisierung Bedeutung entsteht. Damit soll gezeigt werden, dass erstens, Filmverstehen im Grunde leiblich durch das Erleben der sich (als filmische Ausdrucksbewegung) entfaltenden Bewegungsbilder entsteht und zweitens, aus dieser verkörperten Erfahrungsgrundlage, der Bewegungserfahrung, *interaktiv* figuratives Verstehen abstrakter Ideen entspringt. Abschließend wird reflektiert, inwiefern solche ›unrealen‹, d.h. fiktiven, filmischen Interaktionsräume folglich immer eine reale, leiblich vollzogene Interaktion für Zuschauer*innen im Moment des Filmerlebens bedeuten und genau darin der Interaktion unter Menschen nahekommt und mit ihr vergleichbar ist.

2 Vgl. Gemoll (2000: 437) s.v. >το κίνημα< [to kinema], griech. für *Bewegung*.

3 Vgl. Anm. 4 für einen Überblick der auf kinematische Metaphorizität untersuchten audiovisuellen Formate.

Im Folgenden wird zunächst der theoretische Ausgangspunkt (2.) und das damit einhergehende analytische Vorgehen (3.) vorgestellt. Im Anschluss folgt die Fallstudie einer kinematischen Metapher (4.), deren Ergebnisse schließlich im Fazit (5.) noch einmal mit Blick auf die Frage diskutiert werden, inwiefern es sich beim Metaphorisieren während des Filmsehens um (un)reale Interaktion mit filmischen Bildern handelt.

2. Theoretischer Ausgangspunkt: Die kinematische Metapher

Mit der *kinematischen Metapher* (*Cinematic Metaphor*) liegt ein transdisziplinäres Theorie- und Analysekonzept vor, das aus einer langjährigen sprachwissenschaftlich-filmwissenschaftlichen Zusammenarbeit entstand. Es wurde von Vertreter*innen einer sprachgebrauchsorientierten Linguistik und Metaphernforschung sowie einer (neo-)phänomenologisch geprägten Filmwissenschaft im Kontext verschiedener gemeinsamer Forschungsprojekte entwickelt (vgl. Kappelhoff/Müller 2011; Müller/Schmitt 2015; Horst 2018; Greifenstein 2020; Schmitt 2020). Zentrale Publikationen zu dieser transdisziplinären Arbeit sind das bereits eingangs erwähnte Buch *Cinematic Metaphor* von Müller und Kappelhoff (2018; unter Mitarbeit von Sarah Greifenstein, Dorothea Horst, Thomas Scherer und Christina Schmitt)⁴ sowie der zeitgleich publizierte Sammelband *Cinematic Metaphor in Perspective: Reflections on a Transdisciplinary Framework* (Greifenstein et al. 2018). Müller und Kappelhoff definieren die kinematische Metapher in ihrer Monografie wie folgt:

We suggest that *cinematic metaphors are >done< by the spectators and, like metaphors in discourse and face-to-face communication, are dynamic, emergent, always changing. Obviously, this >doing< of metaphor is not a deliberate process, it happens as viewers are perceiving and sensing the film; it happens as they are entangled affectively with the unfolding of cinematic movement-images. Cinematic metaphors thus evolve from perceptual sensations, from the viewer's engagement with what we describe as >cinematic expressive movements<.* (Müller/Kappelhoff 2018: 8; Herv. KP)

4 Darin sind entsprechende transdisziplinäre Analysen zu kinematischer Metaphorizität in unterschiedlichen audiovisuellen Formaten wie *Film* (Müller/Kappelhoff 2018: 22ff., 42ff., 53ff., 91ff., 104ff., 131ff., 164ff., 207f.), *TV-Nachrichten* (ebd.: 2ff., 101ff., 139ff., 172ff., 192ff., 196ff.), *politischer Wahlwerbung* (ebd.: 115ff.) und *Musikvideos* (ebd.: 110ff.) enthalten. Weitere Ergebnisse in Form empirischer Untersuchungen, die diesem Ansatz folgen, sind in Bezug auf *Film*: Kappelhoff/Müller (2011), Scherer et al. (2014), Schmitt (2020: 155ff.); hinsichtlich *TV-Nachrichten*: Müller/Schmitt (2015), Schmitt et al. (2014), Schmitt (2020: 199ff.); in Bezug auf *Science-TV*: Greifenstein/Brückner (2017); mit Blick auf *politische Wahlwerbung*: Horst (2018); bezüglich *TV-Werbespots*: Schmitt (2020: 1ff., 221ff.); hinsichtlich *Social Advertisements*: Scherer (2023: 204ff.) sowie in Bezug auf *mediatisierte Politreden*: Kindler-Mathôt et al. (im Erscheinen).

Im Zitat zeigt sich, dass sich das Konzept der kinematischen Metapher einem dynamischen Metaphernverständnis zuordnen lässt, wie es in der sprachgebrauchsorientierten Metaphernforschung durch Lynne Cameron und Cornelia Müller entwickelt wurde. Dabei bezieht Camerons Theorie *systematischer Metaphern* diese Dynamik ausschließlich auf die Ebene sprachlicher Diskurse (vgl. Cameron et al. 2009), während Müllers *dynamische Perspektive* auf Metaphorizität sowohl verbale als auch multimodale, verbal-gestische metaphorische Ausdrucksweisen in der Face-to-Face-Kommunikation in den Blick nimmt (vgl. Müller 2008; 2019; 2024; Julich-Warpakowski/Jensen 2023). Diese dynamischen Metaphernansätze lassen sich von einem statischen Metaphernverständnis abgrenzen, wie es u.a. im Zuge der Theorie *konzeptueller Metaphern* (*Conceptual Metaphor Theory*) von George Lakoff und Mark Johnson (1980) bezogen auf metaphorische Ausdrücke der Alltagssprache entwickelt wurde.⁵

Was die genannten Metaphertheorien trotz signifikanter Unterschiede dennoch verbindet, ist ein grundsätzliches Verständnis von Metaphern als »understanding and experiencing one kind of thing in terms of another« (Lakoff/Johnson 1980: 5). Metaphorisieren bedeutet daher ein Verstehen, das erfahrungsbasiert ist. Metaphernbildung ist dann ein kognitiver Prozess, der über die körperliche Erfahrung ermöglicht wird. Etwas wird mit Hilfe von etwas anderem verstanden. Häufig wird dabei etwas Abstraktes wie z. B. Zeit als etwas Konkretes, eine Sanduhr, ein Pfeil oder ganz grundsätzlich sprachlich als etwas Räumliches gefasst. So liegt die Zukunft im Deutschen bspw. *vor* uns, die Vergangenheit *hinter* uns. Diese Alltagsmetaphorik findet sich auch im redebegleitenden Gestikulieren, was verdeutlicht, dass sie für die Sprecher*innen und die Interaktionspartner*innen als ein *lebendiger Erfahrungsbereich* gegenwärtig ist.⁶

Metaphern werden folglich als Denk- bzw. Verständigungsvermögen betrachtet, durch den Menschen etwas nicht-greifbar Abstraktes mithilfe einer konkreten Erfahrungsgrundlage überhaupt erst begreiflich machen. Müller (2008: 24) beschreibt diesen Metaphorisierungsprozess auch als »seeing one thing in terms of another«. So wird bspw. ein sehr wütender Mensch metaphorisch als »vor Wut kochend« beschrieben. Damit wird der abstrakte Gefühlszustand dieser Person konkret als Flüssigkeit gesehen, die bei starker Erhitzung (üb)erkochen kann.

-
- 5 Wie sich die Theorie konzeptueller Metaphern zu anderen kognitiven sowie linguistischen und literaturwissenschaftlichen Metaphertheorien verhält, diskutiert Müller (2008: Kap. 3) ausführlich.
- 6 Interessant ist hier auch, dass diese Metaphorisierungen von Zeit kulturell variabel sind. So zeigen etwa Rafael Nuñez und Eve Sweetser (2006) in ihrer Forschung zum Aymara (einer Sprache, die im Hochland von Chile, Peru und Bolivien gesprochen wird), dass die Zukunft dort hinter den Sprecher*innen und die Vergangenheit vor ihnen (vor)gestellt wird. Diese ganz andere metaphorische Vorstellung von Zeit ist im spontanen Sprachgebrauch sprachlich und gestisch lebendig (vgl. Nuñez/Sweetser 2006).

Menschen suchen sich demnach mithilfe einer konkreten Erfahrungsgrundlage *übertragend* über abstrakte Ideen oder Sachverhalte zu verständigen.

Bezogen auf filmische Metaphern sagen Müller und Kappelhoff (2018: 39) mit Rekurs auf Müller (2008): »It is the viewers who perform the »seeing as«, »seeing through«, or »seeing one thing in terms of another« in the process of film-viewing«. Damit entwickeln sie ein spezifisches Verständnis von Metaphern und Metaphorizität, indem sie dieses erfahrungsbasierte Verständnis von Metaphern mit Kappelhoffs Filmtheorie verbinden: »we consider metaphoricity to be *emerging locally from the spectator's experiencing of movement-images* in the moment of film-viewing« (Müller/Kappelhoff 2018: 8; Herv. KP).

Genau hierin zeigt sich die Verankerung des Konzepts der kinematischen Metapher in der (neo-)phänomenologisch geprägten Filmtheorie: Eine ganz spezifische Erfahrungsgrundlage wird ins Zentrum gestellt, nämlich die der Wahrnehmungserfahrung von *Bewegungsbildern* im Moment des Filmesehens – und damit ganz konkret, die der *Bewegungserfahrung* von Zuschauer*innen. Ein Film wird demnach nicht als Akkumulation einzelner, aufeinanderfolgender, statischer Bilder verstanden, die vermeintlich direkte, sich bewegende Darstellungen der realen Welt *repräsentieren*. Vielmehr werden filmische Bilder als dynamische audiovisuelle Expressivität im Sinne *filmischer Ausdrucksbewegungen* (Kappelhoff 2004, 2016, 2018) verstanden. Müller und Kappelhoff (2018) definieren diese wie folgt:

Cinematic expressive movements are *such temporally-structured audiovisual forms that shape the affective perception of the viewers as an inherently temporal experience*. They concern *a specific dimension of movement: an experiential quality or rhythm, a temporally organized perceiving of a dynamic whole*. It is in this sense that we conceive of film as staging the affective course of spectators through a complex temporal aesthetic orchestration. (Müller/Kappelhoff 2018: 135; Herv. KP)

Die Ausdrucksbewegungen werden somit als zeitlich strukturierte audiovisuelle Formen im Sinne einer sich dynamisch entfaltenden Gesamtgestalt verstanden.⁷ Diese emergierende Gesamtgestalt geht als zeitliche Erfahrung mit dem Wahrnehmungsprozess der Zuschauer*innen einher, da sie für die Betrachter*innen körperlich als bestimmte *Bewegungs-Qualität* (z. B. in Form von Rhythmus oder Tempo) des filmisch gestalteten Bewegungsbildes erfahrbar wird. Daher wird dem Film auch eine ästhetische Orchestrierung der affektiven Erfahrungen von Zuschauer*innen zugesprochen. Die filmischen Ausdrucksbewegungen machen somit

7 Diese Annahme einer ganzheitlichen dynamischen Gestalt, die sich räumlich-zeitlich in einem Bewegungsablauf entfaltet und eben nicht eine Abfolge von Einzelelementen ist, entstammt u. a. Plessners philosophischer Anthropologie (vgl. Plessner 1982[1925]: 78) und wird genauer ausgeführt in Müller und Kappelhoff (2018: 158ff.).

zeitlich den Verlauf der Zuschauer*in-Erfahrung von ihnen als konkret wahrnehmbare Bewegungsgestalt aus: Sie *affizieren* Zuschauer*innen ad hoc im Moment des Filmesehens mit und in ihren konkreten Bewegungs-Qualitäten. In diesem Sinne wird der Film als solcherart dynamische Expressivität gefasst. Damit setzen Müller und Kappelhoff (2018) voraus, dass sich aus dieser basalen Erfahrungsgrundlage – d.h. der Bewegungserfahrung von affizierenden audiovisuellen Ausdrucksgestalten – schließlich kinematische Metaphorizität herausbildet. Sie gründet somit in den lebendigen Körpern der Zuschauenden. Deshalb kann sie als deren kreativer und *interaktiver* Verstehensprozess gelten. Das kann bedeuten, dass im Filmesehen, in der Interaktion von filmischen Bewegungsbildern und Wahrnehmungsempfinden der Zuschauer*innen ein abstraktes Konzept als konkrete körperliche Erfahrung verstehbar wird.

An genau diesem Punkt setzt die vorliegende Fallstudie an, wenn sie das Metaphorisieren als Interaktion mit filmischen Bildern beschreibt. Müller und Kappelhoffs (2018) Konzept der kinematischen Metapher erlaubt die Rekonstruktion kinematischer Metaphernemergenz *als interaktiv verkörperter Sinnstiftungsprozess*, den Zuschauer*innen unbewusst vollziehen, sobald sie Filme sehen.

Nachfolgend soll nun vorgestellt werden, wie die hier ausgeführten medienästhetischen Annahmen methodisch umgesetzt werden.

3. Medienästhetische Methode: Das CinMet-Verfahren

Aus der medienästhetischen Perspektive auf das Filmesehen als interaktiven Prozess, als Verwicklung von Zuschauerempfinden und filmischen Bewegungsbildern, leiten Müller und Kappelhoff (2018) eine spezifische Analysesystematik ab. Diese ist gleichermaßen auf dynamische Prozesse der Metaphernbildung im Filmesehen und in der Interaktion von Angesicht zu Angesicht anzuwenden.

Dieses medienästhetische Analyseverfahren, die sog. *Cinematic Metaphor Procedure* (vgl. ebd.: 230f.), wird im Verlauf des vorliegenden Beitrags verwendet. Es verbindet Methoden aus unterschiedlichen Analyseansätzen der Linguistik (u.a. MFA: Method of Foregrounding Analysis, Müller/Tag 2010; MGA: Method of Gesture Analysis, Müller 2010) und Filmwissenschaft (eMAEX: electronically based media analysis of expressive movement images, Kappelhoff et al. o.D.; Kappelhoff/Bakels 2011; Müller/Kappelhoff 2018).⁸ Dabei rückt die linguistische Methodik der Metaphernanalyse die zeitliche Rekonstruktion interaktiv hervorgehobener, aktivierter und damit vergegenwärtigter multimodaler Metaphorizität über einen bestimmten diskursiven Zeitraum in den Blick. Die filmwissenschaftliche Methodik wiederum

8 Vgl. für eine extensive Übersicht der einzelnen Methoden aus den jeweiligen Disziplinen Müller und Kappelhoff (2018: 227ff.).

fokussiert das Erfassen der audiovisuellen Orchestrierung filmischer Bewegungsbilder als expressive und gleichzeitig affizierende Ausdrucksbewegungen. Dadurch wird der affektive Prozess, den Zuschauende leiblich während des Filmesehens durchlaufen, dokumentierbar gemacht. Beide methodischen Perspektiven setzen damit bei der Temporalität der Metaphernbildung an, die jeweils aus interaktiven Prozessen hervorgehen: der Interaktion zwischen Gesprächsbeteiligten oder der Interaktion von Zuschauenden und filmischen Bildern.

Denn in eben dieser Zeitlichkeit einer temporalen Entfaltung wird das expressive Potential audiovisueller Medien als filmisch orchestrierte Bewegungsgestalten gesehen, die die Zuschauenden unmittelbar affizieren, konkrete Wahrnehmungsempfindungen auslösen und auf deren Basis metaphorisches Verstehen generiert wird. In diesem Sinne spielt die temporale Entfaltung der filmischen Ausdrucksbewegungen eine wesentliche Rolle für die Rekonstruktion des Metaphorisierens als dynamischer Prozess des filmischen Verstehens.

Um diese Metaphernentfaltung analytisch zu fassen, werden mit dem CinMet-Verfahren drei verschiedene Zeitebenen adressiert, die sich in den folgenden vier Analyseschritten widerspiegeln:

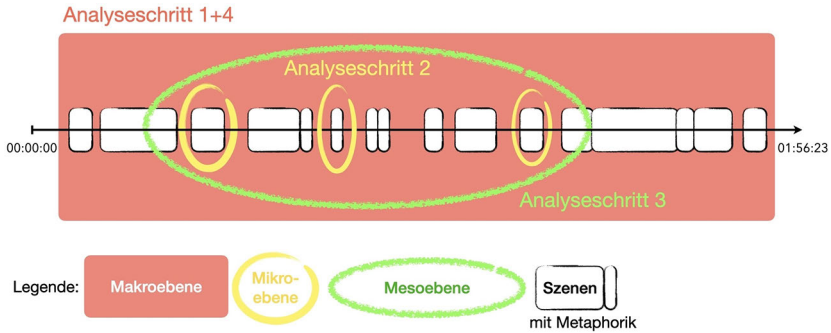
- 1) Makroebene: Erstellung eines Szenenprotokolls, erste Beschreibung prominenter Inszenierungsmuster des gesamten Filmes, erste Identifikation von Szenen, in denen Metaphorik eine Rolle zu spielen scheint.
- 2) Mikroebene: Unterteilung ausgewählter Szenen in Ausdrucksbewegungseinheiten (ABEs), Beschreibung der Metaphorizität anhand affektiver Qualitäten und Dialog.
- 3) Mesoebene: Identifikation wiederkehrender Muster von ABEs, Beschreibung der affektiven Bewegungsqualitäten als rekurrente Erfahrungs-Szenarien unter Mitberücksichtigung von Variationen und Wiederholungen.
- 4) Makroebene: Rekonstruktion des dynamischen metaphorischen Zusammenspiels, Beschreibung des affektiven Gesamtverlaufs in Zusammenhang mit der temporalen Entfaltung der kinematischen Metapher.

Abbildung 1 visualisiert die drei Zeitebenen, die in den Analyseschritten (1–4) abwechselnd eingenommen werden:⁹ Der rote Kasten bildet schematisch die zeitliche Perspektive der Makroebene ab, die in den Schritten 1 und 4 bezogen wird. Hier adressiert die Analyse den Film in seiner Gesamtlänge (s. Zeitstrahl, Abb. 1). Die gelben Kreise schematisieren hingegen das Mikrolevel, das in Schritt 2 auf der Ebene einzelner Szenen eingenommen wird. Werden diese Einzelszenen gemäß Schritt 3 in Bezug zueinander untersucht, befindet man sich auf der temporalen Mesoebene, die in Abbildung 1 durch die grüne Kreisform gekennzeichnet ist. Sie umschließt

9 Alle Abb. des Beitrags sind farbig in der digitalen Open-Access Version zu finden.

entsprechend die mikroanalytisch untersuchten Einzelszenen (gelb umkreiste Szenen, Abb. 1), die miteinander verglichen werden.

Abb. 1: Adressierte Zeitebenen der vier Analyseschritte im CinMet-Verfahren.



Indem alle zeitlichen Ebenen für sich sowie in ihrem Zusammenspiel betrachtet werden und so die temporale Entwicklung affektiven Wahrnehmens und darin begründeten metaphorischen (Film-)Verstehens rekonstruiert wird, erlaubt das CinMet-Verfahren eine schrittweise sequenzielle sowie deskriptive analytische Herangehensweise.

Im vorliegenden Beitrag stehen die Analyseschritte 2 und 3 im Zentrum. Wie kinematische Metaphorizität im Prozess des Filmesehens aus der Interaktion mit den filmischen Bildern allmählich entsteht, wird am Beispiel der Anfangs- und der Schlusssequenz illustriert.¹⁰ Dabei werden im Analyseschritt 2 die Anfangs- und Schlusssequenz mikroanalytisch auf die Entstehung kinematischer Metaphorizität untersucht, Analyseschritt 3 zeigt dann in einer Mesoperspektive, wie diese Sequenzen aufeinander aufbauen.

4. Fallstudie: Figuratives Verstehen von non-linearer Zeit

Mit der nachfolgenden Fallstudie soll illustriert werden, inwiefern die Filmidee eines non-linearen Sprach- und Zeitkonzepts im ›unrealen‹, d.h. fiktionalen, Interaktionsraum mit Außerirdischen *körperlich* erfahrbar gemacht und damit verständlich

10 Der gesamte Film ARRIVAL (2016) wurde im Kontext einer Masterarbeit mit der CinMet-Methode untersucht. Bei der vorliegenden Fallstudie handelt es sich daher um einen Auszug aus einer umfangreichen Analyse, bei der alle vier genannten Analyseschritte durchgeführt wurden.

wird. Dies geschieht sobald Zuschauende im Moment der Filmsichtung mit dem audiovisuellen Material verschmelzen und so am eigenen Leib diese im alltäglichen Leben unmögliche Interaktion ganz real erleben. Es wird konkret nachgezeichnet, wie aus der spezifischen Bewegungserfahrung der filmischen Bilder Metaphorisierungen entstehen, die völlig neue Erfahrungsräume eröffnen, wie z.B. den einer non-linearen Zeit(wahrnehmung). Damit soll veranschaulicht werden, inwiefern es sich bei dieser filmischen Fiktion eines außerirdischen non-linearen Sprach- und Zeitkonzepts gleichwohl um (un)reale Interaktion handelt, sobald Zuschauende leiblich das Prinzip des linguistischen Determinismus im Moment der Filmsichtung verkörpern und am eigenen Leib das fiktive non-lineare zeitliche Vermögen realisieren, Erinnerungen an die Zukunft zu haben.

Worum geht es in Denis Villeneuves Science-Fiction Film *ARRIVAL* (USA 2016, 116 Min.)? Der Film basiert auf Ted Chiangs Roman *Story of your Life* (1998). Inhaltlich geht es um die Landung zwölf außerirdischer Raumschiffe an verschiedenen Standpunkten auf der Erde. Um zu ermitteln, ob es sich bei den Aliens, die die Raumschiffe steuern, um einen friedlichen oder feindlichen Besuch handelt, wird vom US-Militär ein Expert*innenteam bestehend aus dem Physiker Ian Donnelly (Jeremy Renner) und der Linguistin Louise Banks (Amy Adams) aufgestellt. In einem Rennen gegen die Zeit besteht ihre Aufgabe nun darin mit den in Amerika gelandeten Außerirdischen in einem Kommunikationsschacht des Raumschiffs zu kommunizieren, um den Grund für die Landung auf der Erde ausfindig zu machen, bevor ein möglicher Weltkrieg ausbricht. Im Laufe des Films stellt sich als Grund für den extraterrestrischen Besuch heraus, dass sie für eine zukünftige Gegenleistung ihre *non-lineare* Sprache und die damit verbundene *non-lineare* Zeitwahrnehmung der Menschheit als Geschenk überbringen wollen. Damit spielt der Film mit der Idee des sog. *linguistischen Determinismus* Edward Sapirs (1929), demzufolge Denken und Wahrnehmen stets von den sprachlichen Strukturen eines Sprachsystems determiniert werden.¹¹ Entsprechend ermöglicht das außerirdische Sprachsystem, mit seinen non-linearen (v.a. kreisförmigen Symbol-)Strukturen der Protagonistin ebenso non-linear zeitlich wahrzunehmen wie zu denken. Louise schafft es nach Aneignung der außerirdischen Sprache somit den drohenden Krieg abzuwehren, indem sie mithilfe der sprachlich erworbenen Voraussicht (in Form von Zukunftsvisionen) unzugängliche und höchst vertrauliche Informationen erwirbt, mit denen sie den leitenden General des Angriffs vom sprachlichen Einfluss der Außerirdischen auf ihr zeitliches Denkvermögen überzeugen und so deren Grund für die Landung auf der Erde als friedlich ausweisen kann.

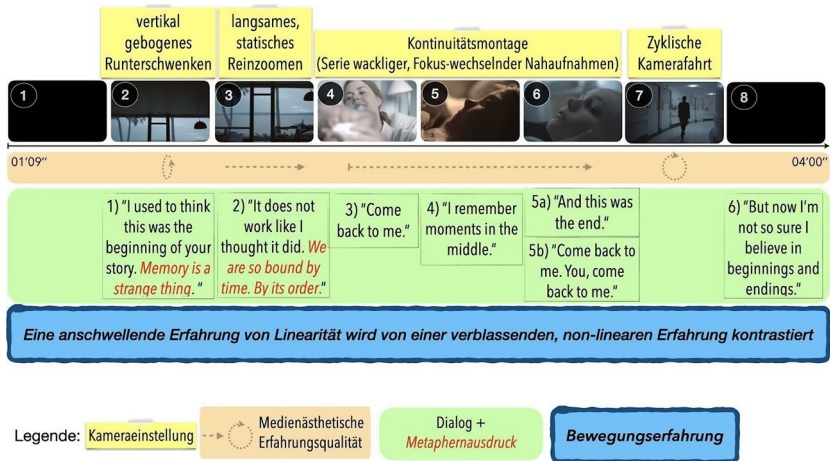
11 Vgl. Sapirs Schüler und Anhänger Benjamin Lee Whorf (1963), der mit dem sog. *linguistischen Relativitätsprinzip* eine abgeschwächte Form des Einflusses von Sprache auf das Denken formuliert. Für einen extensiven Forschungsüberblick zur sprachlichen Relativitätstheorie, vgl. Werlen (2002, insb. S. 32–61).

5. Mikroanalyse: Die Anfangssequenz

Der Film beginnt mit der Aufblende eines gänzlich schwarzen Bildgrunds (vgl. Abb. 2, Bild 1).¹² Begleitet von ruhig einsetzender melancholischer Streichmusik, verblasst die Schwärze durch eine nach unten schwenkende Kamerabewegung nach und nach zu dunklen, vertikalen Geraden, die sich mit der anhaltenden herunter schwenkenden Kamerabewegung allmählich als Deckenpaneele eines düsteren Wohnraums entpuppen (Abb. 2, Bild 2). Just in diesem Moment erklingt aus dem Off Louise' Stimme: »I used to think this was the beginning of your story. Memory is a strange thing.« (Abb. 2, Dialogkasten 1). Sobald sie weiter spricht »It does not work like I thought it did. We are so bound by time. By its order.« (Abb. 2, Dialogkasten 2), findet die vertikal hinab schwenkende Kamerabewegung ihr Ende und fließt über in ein statisches, langsames Hinein-Zoomen, das den finsternen Wohnraum vor dem blassen Hintergrund einer dämmrig-trüben Seelandschaft zeichnet (Abb. 2, Bild 3).

12 Abbildung 2 zeigt die Aktivierung der kinematischen Metapher von Non-Linearität basierend auf einer ganz konkreten Bewegungserfahrung (Abb. 2, blauer Balken unten). Oberhalb des Zeitstrahls befinden sich dabei acht Standbilder aus der Anfangssequenz mit darüber liegenden Post-its (gelb) zur jeweiligen Kameraeinstellung und -bewegung. Zusammen bilden sie exemplarisch die visuelle Inszenierung der Sequenz ab. Direkt unterhalb des Zeitstrahls (orange) befindet sich die schematische Darstellung der aus den filmischen Inszenierungsmitteln hervorgehenden Erfahrungsqualitäten (gestrichelte Pfeillinien). Der (grüne) Bereich darunter visualisiert die auditive Spur auf Dialogebene mit rot markierten sprachlichen Metaphern in den einzelnen Dialogkästen. Zusammen kreieren all diese Ebenen ganzheitlich die unten in blau gefasste Bewegungserfahrung, die Zuschauende im Moment des Filmerlebens ganz konkret wahrnehmen und leiblich durchlaufen. Aus dieser Bewegungserfahrung heraus entsteht die kinematische Metapher von Non-Linearität.

Abb. 2: Anfangssequenz. Kinematische Metaphernaktivierung von Non-Linearität in ARRIVAL (2016).



Damit werden Zuschauende bereits in den ersten Minuten der Einführungs-szene mit der Essenz des Filmes konfrontiert, dass es im Folgenden um eine unvertraute Art des Erinnerns und damit einhergehend, unbekanntem zeitlichen Ordnung gehen wird (die sich im weiteren Verlauf noch als non-linear konkretisiert). Die beiden Aussagen markieren dabei bereits die ersten metaphorisch ›schlafenden‹ Verbal-ausdrücke (vgl. Müller 2008), indem einerseits der abstrakte Begriff des Erinnerns wider Erwarten (»It does not work like I thought it did«) als etwas Unbekanntes (*memory in terms of a strange thing*) gefasst und andererseits die abstrakte Vorstellung von Zeit als bindende Ordnung (*time in terms of a binding order*) verstanden wird. Diese sprachlich ausgedrückten *Metaphern* werden erst filminszenatorisch durch die Kameraeinstellung und -führung in ihrer Erfahrungsgrundlage spürbar aktiviert: indem Zeit als bindende Ordnung (und das ihr unterliegende Erinnern) durch die langsam hinein zoomende Kamerabewegung als etwas Lineares konkretisiert wird und sich somit im Sinne einer sich linear bindenden Bewegung erfahren und folglich (metaphorisch) verstehen lässt. Damit wird die eingangs schlafende sprachliche Metapher durch die audiovisuelle Inszenierung *aktiviert* bzw. gewissermaßen zum Leben erweckt, ›revitalisiert‹ (vgl. Müller 2008); ein Gefühl der zeitlichen Gebundenheit wird zunächst filmästhetisch durch das statische und langsame, geradlinige, d.h. *lineare* Reinzoomen der Kamera ausgedrückt. Metaphorizität ist in diesem Sinne kinematisch durch das Zusammenspiel der filmtechnisch prominenten Ar-

tikulationsmittel Dialog, Kameraführung und Einstellungsgröße aktiviert und als konkret körperlich-sensorische Erfahrung kreiert.¹³

Dieses anfänglich evozierte Gefühl von Linearität wird weiter aufrechterhalten, indem die Entwicklung eines kleinen Mädchens vom Neugeborenenalter bis hin zum tragischen Sterbebett mittels einer Kontinuitätsmontage ebenso linear erlebt wie dialogisch kommentiert wird (vgl. Abb. 2, Bildfolge 4–6 sowie Dialogkästen 1, 4 und 5a). Was diese Einblicke in die lineare Kindesentwicklung dabei filmästhetisch charakterisiert, ist die sich wechselnde Scharfstellung der Kamerafokussierung und unstete Art und Weise in der die Einblicke (in vier Altersetappen des Mädchens als Neugeborenes, Kleinkind, Kind, Jugendliche) größtenteils in einer Serie verschwommener Nahaufnahmen inszeniert sind. Diese zeigen abwechselnd Louise und das kleine Mädchen (unterschiedlichen Alters), begleitet von derselben, nun intensivierenden Streichmusik, sodass beim Betrachten eine filmisch evozierte Bewegungsempfindung von anschwellender *Kontinuität* entsteht. Im Verwickelt-Sein mit dieser visuell mit wackeliger Handkamera inszenierten, Perspektiv- und Fokus-wechselnden Kontinuitätsmontage erleben Zuschauer*innen diese Sequenz somit als zeitlich zusammenhängende *Linearität*, die dialogisch von einem ebenso ›linearen‹ Storytelling unterstrichen wird (›beginning of your story«, »moments in the middle«, »this was the end«). Beim Erfahren dieser sich *kontinuierlich* wechselnden Schuss-Gegenschuss-Montagesequenz entsteht bei den Zuschauenden also wieder durch das Zusammenspiel von visueller Inszenierung und hörbarer *off-screen* Stimme der sich steigernde Eindruck von Linearität im Sinne einer *linear erlebten* Ontogenese eines heranwachsenden Kindes. Bemerkenswert ist an dieser Stelle, dass die verbalen Ausdrücke »I remember moments in the middle« (Abb. 2, Dialogkästen 4) auf der Dialogebene keine weitere Auskunft über die Art und Weise des Erinnerns geben. Vielmehr wird die Weise des Erinnerns auf der Ebene der Filminszenierung für Zuschauende erfahrbar und damit verständlich gemacht, nämlich als lineare, wenn auch wackelig verschwommene, Erinnerung von Geschehnissen, die durch frontale Nahaufnahmen (und kameratechnischer Fokussierung) wiedererkannt werden. Filmisch wird also in der Anfangssequenz zunächst durch einen Gegensatz die Art und Weise des Erinnerns in einem durchaus vertrauten linearen und wiedererkennenden Modus inszeniert und nicht etwa als etwas Fremdes (*as a strange thing*). Entsprechend ist an dieser Stelle festzuhalten, dass die dargestellte zeitliche lineare Ordnung einer Kindesentwicklung durch den Dialog (›I remember«) mit dem Gefühl des Erinnerns verknüpft wird – die Art und Weise *wie* Dinge jedoch erinnert werden, kommt ausschließlich filmisch auf der ästhetischen Ebene durch die wackelig-verschwommene aber gleichzeitig fokussierende Kameraeinstellungen und die visuelle Komposition zum Ausdruck: und

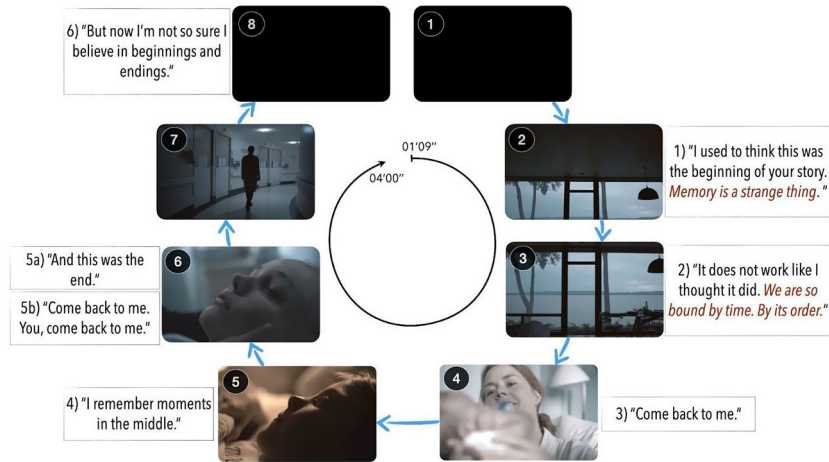
13 Nach Müller und Tag (2010) semantisch gemäß dem sog. *Ikonizitätsprinzip* als linear hervor-gehoben.

zwar als anschwellende (Bewegungs-)Erfahrung von kontinuierlich klar-werdender Linearität.

Als nächstes erleben die Zuschauer*innen das Endstadium des Kindes, das von der *off-screen* erklingenden Stimme mit »And this was the end« (Abb. 2, Dialogkasten 5a) markiert wird. Die anfänglich beim Geburtsmoment geäußerten Worte »Come back to me« (Abb. 2, Bild 4, Dialogkasten 3) sind hier nochmals auf der Ebene des Dialogs zu hören, wenn die verstorbene Jugendliche bildlich dargestellt wird (Abb. 2, Bild 6, Dialogkasten 5b). In der nächsten Aufnahme folgt die Kamera in einem zyklischen Uhrzeigersinn von hinten (Abb. 2, Bild 7) einer nun kinderlosen Hauptdarstellerin, die einen ebenso zyklisch abgebildeten Krankenhauskorridor entlangläuft. Die bisher aufeinanderfolgenden filmisch evozierten (Bewegungs-)Empfindungen von Linearität sind nun also audiovisuell kontrastiert durch einen zyklischen Bewegungsverlauf und zyklischen Korridoraufbau, die langsam in dieselbe Dunkelheit wie vom Anfang verblenden (Abb. 2, Bild 8). In eben diesem Moment verklingt die anfänglich zu hörende melancholische Filmmusik und alleinig die Stimme der Hauptdarstellerin durchbricht aus dem Off die Dunkelheit: »But now I am not so sure I believe in beginnings and endings« (Abb. 2, Bild 8, Dialogkasten 6). Mit dieser letzten sprachlichen Metapher wird die zu Anfang inszenierte Idee einer linearen zeitlichen Ordnung und Erinnerungsweise umso mehr kontrastiert, indem die am Ende erstmals nur spürbar gemachte kreisförmige Ausrichtung nun auch auf der Ebene des Dialogs anklingt, da nur ein Kreis keinen wirklichen Anfangs- und Endpunkt (»beginnings and endings«) hat.

Betrachtet man die Anfangssequenz in ihrer Gesamtgestaltung, fällt direkt auf, wie im Sinne eines Zyklus' Beginn und Schluss der Sequenz durch einen schwarzen Bildgrund als Auf- und Abblende bestimmt sind (s. Abb. 3, Bild 1+8), die rekurren- de Aussage »come back to me« quasi zirkuliert (Abb. 3, Dialogkästen 3+5b), wenn im Prinzip der ›Kreis des Lebens‹ eines Kindes dargestellt wird (Abb. 3, Bildfolge 4–6).

Abb. 3: Anfangssequenz. Inszenatorisches Erleben von Non-Linearität in ARRIVAL (2016).



Auf diese Weise kreiert die Anfangssequenz die kinematische Erfahrungsgrundlage eines Zyklus, oder genauer: einer kontrastierten Linearität (vgl. Bewegungserfahrung, Abb. 2). Eine Bewegungserfahrung von Non-Linearität also, die schließlich hervorgehoben wird durch eine zyklische Kamerabewegung, zyklische Räumlichkeit (Korridor) und einer finalen verbalen Explizierung desselben.

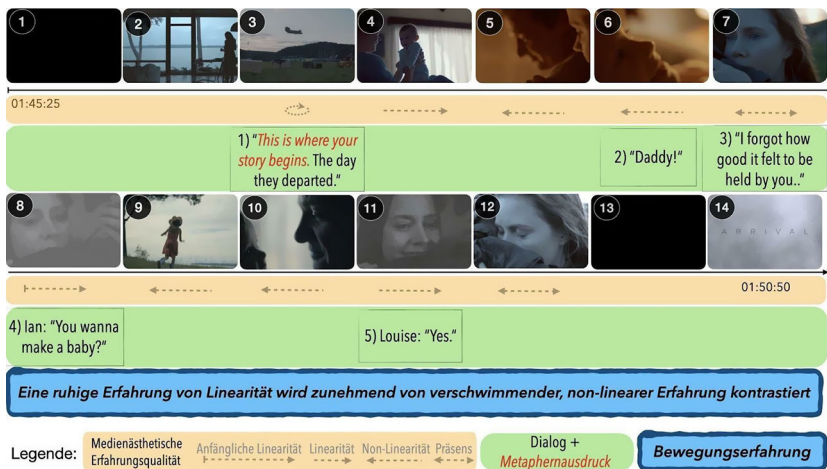
In den ersten Minuten des Films werden also mittels filmischer Inszenierung die abstrakten Ideen von Erinnerung als etwas Unbekanntem (*memory as a strange thing*) und von der ihr übergeordneten Zeit als eine bindende Ordnung (*time as a binding order*) metaphorisch mit der Idee von Non-Linearität verbunden. Die Ideen beruhen somit auf der körperlichen Vergewärtigung dieser beiden Sprachmetaphern im Sinne der konkret erlebten Bewegungserfahrung (vgl. Abb. 2) von Non-Linearität seitens der Zuschauer*innen. Damit bilden sie das erste identifizierte *kinematische Metaphernszenario* in ARRIVAL (2016).

6. Mikroanalyse: Die Schlusssequenz

Nach der Lösung des Filmrätsels – dass es sich nämlich um ein non-lineares außerirdisches Sprachkonzept handelt, welches deterministisch zu non-linearem zeitlichen Denken in Form von Erinnerungen an die Zukunft befähigt – wird in der Abschlusszene diese Non-Linearität erneut mithilfe einer bestimmten Schnittmontage konkret spürbar gemacht: Genau wie in der Anfangssequenz beginnt die Schlusssequenz zunächst mit demselben schwarzen Bildgrund (Abb. 4, Bild 1), die in Begleitung mit derselben melancholischen Filmmusik und vertikal hinab schwenkenden

Kamera langsam ins vertraute dunkle Wohnzimmer blendet, in das nun Louise hereintritt (Abb. 4, Bild 2). Als nächstes ist das Militärlager zu sehen (Abb. 4, Bild 3), und mit einer im Uhrzeigersinn nach rechts schwenkenden Kamerabewegung ertönt aus dem Off Louise' Stimme, die den Beginn des Lebens ihres Kindes räumlich mit der Aussage »This is where your story begins« (Abb. 4, Dialogkasten 1) markiert, während das Militärlager zu sehen ist. Mit dem Tag der Abreise der Aliens (»The day they departed«) konkretisiert sie anschließend den zeitlichen Startpunkt der Lebensgeschichte ihres Kindes.

Abb. 4: Schlusssequenz. Kinematische Metaphernaktivierung von Non-Linearität in ARRIVAL (2016).

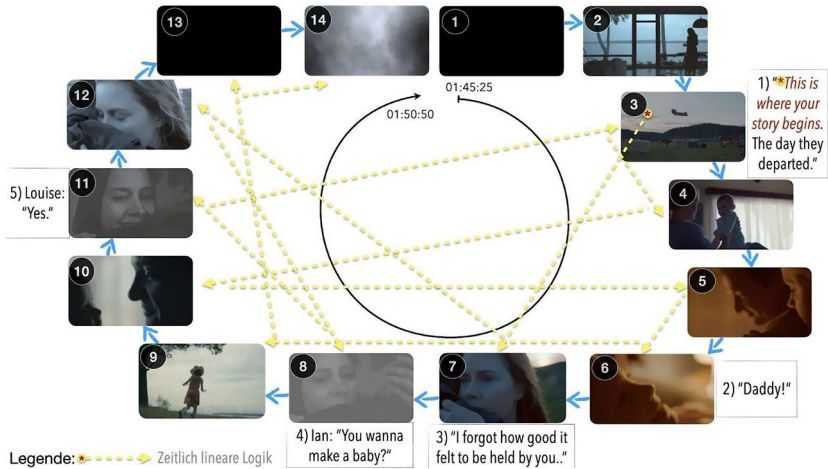


Damit ist erneut auf Ebene des Dialogs mit der Aussage »This is where your story begins« eine ›schlafende‹ *Metapher* (Müller 2008) artikuliert, die die abstrakte Vorstellung eines zeitlichen Startpunkts als konkret-räumlich fasst (*begin of story in terms of somewhere*). Wie in der Anfangssequenz, wird auch diese schlafende Metapher erst kinematografisch gewissermaßen ›revitalisiert‹ (Müller 2008) und damit ihr Erfahrungsraum vergegenwärtigt: Durch das audiovisuelle Zusammenspiel von Dialog und visueller Ebene wird der verbal ausgedrückte Beginn der kindlichen Entstehungsgeschichte räumlich mit dem gezeigten Lager verknüpft und kann so für Zuschauende konkret verankert werden. Damit gibt die audiovisuelle Komposition von Dialog und visueller Ebene den Zuschauenden an dieser Stelle zunächst den räumlichen Startpunkt der Geschichte des Kindes. Zugleich aktiviert jedoch die Kamerabewegung einen weiteren Aspekt dieser Metapher, indem durch die im Uhrzeigersinn schwenkende Kamerabewegung der fernere, zeitlich-räumliche Verlauf

der Entwicklungsgeschichte *ausschließlich auf der filmästhetischen Ebene* als zyklisch evoziert und damit für das Publikum konkret spürbar gemacht wird. Metaphorizität ist demnach auf zweifache Weise durch das Zusammenspiel der filmartikulativen Mittel von Dialog, visuell Dargestelltem und Kameraführung aktiviert und hervorgehoben.

Was nun inszenatorisch folgt, ist eine audiovisuelle Orchestrierung, die von verschiedenen Zeitsprüngen in Form von Rück- und Vorausblenden geprägt ist und somit den zeitlichen Verlauf der Geschichte erneut und ganz konkret als non-linear erfahren lässt: Im Vergleich zur Anfangssequenz erleben die Zuschauer*innen nun in einer Schuss-Gegenschuss-Montage die Ontogenese des Kindes in einer sehr konträren Art und Weise, nämlich indem sich keinesfalls ein linearer Verlauf des Lebens des Kindes erleben lässt. Denn obwohl Zuschauer*innen zunächst einmal analog zur Anfangssequenz logisch linear die ersten Lebensstufen von Säugling zu Kind (Abb. 4, Bildfolge 4–6) in derselben verschwommenen und (zwischen Mutter und Kind) Perspektiv-wechselnden Art und Weise durchlaufen, erleben sie plötzlich in einem non-linearen Bruch einen Moment, der zeitlich noch vor der Geburt liegt: den intimen Moment zwischen Ian und Louise, der sie fragt, ob sie gemeinsam ein Kind zeugen wollen (Abb. 4, Bild 8, Dialogkasten 4). Inszenatorisch folgt erst Einstellungen später, nach weiteren nicht-chronologischen, verschwommenen Einblicken in das Leben eines non-linear heranwachsenden Kindes (vgl. Abb. 4, Bild 9+10), Louise' affirmative Antwort auf Ians Frage nach dem Kinderwunsch (Abb. 4, Bild 11, Dialogkasten 5). Diese müsste nach linear-zeitlicher Logik nämlich noch vor den vorausgehenden und ebenfalls nicht-chronologischen Flashforwards (Abb. 4, Bild 9+10, Darstellung des gemeinsamen Kindes als kleines Mädchen *vor* dessen Darstellung als Säugling) liegen.

Abb. 5: Schlusssequenz. Inszenatorisches Erleben von Non-Linearität in ARRIVAL (2016).



Durch die Montage der Filminszenierung wird Non-Linearität also für die Zuschauer*innen nochmal besonders spürbar als Erfahrungsraum vergegenwärtigt und als zeitliches Konzept *sinnlich* als *konkrete Erfahrung* verständlich gemacht. Dies geschieht, indem zwischen Vergangenheit und Zukunftszeitpunkten nicht-chronologisch gewechselt wird und das Präsens durch das Zusammenspiel von Dialog («This is where your story begins») und visueller Ebene (Abb. 5, Bild 3) ans Militärlager geknüpft wird. Das bedeutet, dass zukünftige Ereignisse filmisch wie Erinnerungen inszeniert werden, die spätestens in Begleitung des Dialogs durch die verbale Explikation des Vergessens besonders ersichtlich bzw. erfahrbar gemacht werden. Denn obwohl sich die Zuschauer*innen durch die vorausgehende audiovisuell-metaphorische temporale Verortung im Militärlager im Präsens befinden, artikuliert Louise Ian gegenüber, »I forgot how good it felt to be held by you«, obwohl sich die beiden Hauptfiguren bis zu diesem Zeitpunkt in der Seh-Erfahrung noch nie zuvor in den Armen hielten (Abb. 5, Bild 7, Dialogkasten 3). In diesem Sinne durchlaufen die Zuschauer*innen sukzessive Erfahrungen von Non-Linearität (vgl. gelbe Pfeilfolge, Abb. 5), die filmtechnisch mithilfe der Montage und visuellen Inszenierung konkret als zunehmendes, verschwimmendes non-lineares Erleben unterschiedlicher Tempi für die Zuschauerschaft spürbar gemacht werden und so im direkten Kontrast zur anfangs ruhig evozierten Bewegungsempfindung von Linearität stehen (vgl. Bewegungserfahrung, Abb. 4).

Eine letzte inszenatorische Variation der Anfangssequenz findet sich schließlich am Filmende. Denn anders als die Anfangssequenz endet die Schlusssequenz nun nicht in einer geschlossenen Art und Weise mit demselben schwarzen Bildgrund als Abblende, sondern mit einer gräulich-nebeligen Einstellung, in die langsam die

schwarzen Buchstaben des Filmtitels (*Arrival*) einblenden (Abb. 5, Bild 14). Die Idee unklarer bzw. nicht-definierter Anfänge und Enden im Sinne einer evozierten Zirkularität bzw. Non-Linearität ist also durch dieses diffus nebelige Abschlussbild ein weiteres Mal vordergründig spürbar gemacht und vergegenwärtigt – erneut exklusiv auf der Ebene der ästhetischen Filminszenierung.

Damit tritt das erste identifizierte kinematische Metaphernszenario von der Non-Linearität der Zeit auf der Mesoebene (Analyseschritt 3) schließlich wieder zum Ende des Films auf (Anfangs- und Schlussequenz) und markiert auf der Makroebene (Analyseschritt 1+4) somit Anfang und Ende des Films. Das zum Schluss wiederkehrende Szenario weist dabei inszenatorisch leichte Variationen derselben metaphorischen Erfahrungsgrundlage von Non-Linearität auf. Insgesamt lässt sich die eingeführte Idee eines fiktiven, non-linearen Sprach- und Zeitkonzepts folglich als konkrete Bewegungserfahrung *einer non-linear kontrastierten Linearität* (vgl. Bewegungserfahrung, Abb. 2+4) begreifen, die inszenatorisch auf allen zeitlichen Ebenen der filmischen Orchestrierung aufkommt: Auf der Mikroebene entsteht die Erfahrung von Non-Linearität aus den konkreten, hier ausgeführten Inszenierungsweisen der Anfangs- und Schlussequenz. Auf der Makroebene wird Non-Linearität erfahrbar, indem beide Szenen wie ein sich schließender Kreis Anfang und Ende des gesamten Films markieren. Auf der Mesoebene wird die wieder aufgegriffene Erfahrbarkeit von Non-Linearität in der Schlussequenz durch die finale Abweichung vom Beginn (nebligtes Abschlussbild statt schwarzem Bildgrund) kompositorisch ein letztes Mal als Zirkularität im Sinne eines unklaren Anfangs und Endes für Zuschauende konkret spürbar und damit real erfahrbar, ergo – durch die körperliche Vergegenwärtigung – verständlich. Auf eben diese Weise entsteht die kinematische Metaphorizität einer außerirdischen, non-linearen Zeitform *erfahrungsbasiert* und *dynamisch* im Prozess des Filmesehens aus der ganzkörperlichen *interaktiven* Verwicklung von Zuschauenden mit den filmischen Bildern.

7. Diskussion und Fazit: Metaphorisieren als Interaktion mit filmischen Bildern

Im Fokus dieses Beitrags stand, inwiefern ein Metaphorisieren während des Film-erlebens als (un)reale Interaktion mit Filmen verstanden werden kann. Wie die Fallstudie gezeigt hat, wird die fiktive Filmidee einer non-linearen Zeit(ordnung und Erinnerungsweise) mithilfe einer Metaphorisierung verstanden, die im leiblichen Erleben der konkreten filmästhetisch evozierten Bewegungserfahrung gründet: einer rekurrierenden *Erfahrung von Linearität, die von verblassender/verschwimmender, non-linearer Erfahrung kontrastiert wird* (vgl. Bewegungserfahrung, Abb 2+4). Die Ergebnisse der Fallstudie validieren somit die Annahme, dass filmisches Verstehen von abstrakten Ideen wie der eines non-linearen Zeitkonzepts in erster Linie in-

teraktiv aus der *Verwicklung* (Kappelhoff 2018: 12; Müller/Kappelhoff 2018: 50) der Zuschauenden mit dem Film durch ihre konkret situierten körperlichen Erfahrungen hervorgeht. Damit kann mit Müller und Kappelhoff (2018) diese *Interaktion* mit Film genau wie die Interaktion zwischen Menschen als genuin verkörpertes Phänomen verstanden werden, das in spürbaren Bewegungserfahrungen gründet und so dynamisch in Zuschauer*innen, leiblich, und damit quasi real stattfindet.

Denn was auf den ersten Blick als ›unrealer‹, d.h. fiktiver, Interaktionsraum erscheinen mag, erweist sich in dieser Perspektive als eine durchaus reale Interaktion – und zwar zwischen einem affizierenden ›Filmkörper‹ (im Sinne von Film als Körper) und einem lebendigen Zuschauer*inkörper. Entsprechend formuliert Müller:

Die Leinwand, der Bildschirm tritt an die Stelle des Körpers – sie ist gewissermaßen ein artifizieller Körper [...], der das Publikum tatsächlich bewegt – kollektiv erschrecktes Aufschreien oder Tränen der Rührung, die wir im Kino gemeinsam vergießen, zeugen von dieser körperlichen Unmittelbarkeit. Kappelhoff (2018: 129) formuliert den Grundgedanken des Konzepts der Ausdrucksbewegung als: »das Verhältnis zwischen affizierendem und affiziertem Körper«. In diesem Sinne haben wir es sowohl im zwischenmenschlichen Gespräch als auch im Kino mit einer grundsätzlich körperlichen Interaktion zu tun, die interaffektiv ist. (Müller im Erscheinen)

Was beide Interaktionsformen (Mensch-Medium; Mensch-Mensch) also verbindet, ist eben diese zeitlich-räumliche Verwicklung als eine Bewegungserfahrung. Aus dieser Bewegungserfahrung entstehen kinematische Metaphorisierungen. Dies gilt für die zwischenmenschliche Interaktion ebenso wie für die Interaktion von Zuschauer*in und filmischen Bewegungsbildern. Was sich in der Face-to-Face-Interaktion nämlich sprachlich-gestisch-mimisch als *körperliche* Ausdrucksbewegung realisiert,¹⁴ tritt in der verkörperten Interaktion mit dem sich entfaltenden Film in der Inszenierung der filmischen Bewegungsbilder (durch Kameraführung, Schnitt/Montage, Licht, Musik, Dialog und Schauspiel) als *filmische* Ausdrucksbewegungen zutage, die für Zuschauende unmittelbar leiblich erfahrbar werden (vgl. Kappelhoff 2004: 19). Vor diesem Hintergrund unterscheidet Müller (im Erscheinen) sprachliche von audiovisueller Multimodalität. Diese ad hoc verkörperte Grundlage des Verstehens markiert den gemeinsamen zentralen Aspekt beider Interaktionsformen (vgl. Müller/Kappelhoff 2018: 85; Schmitt 2020: 131ff.).

In diesem Sinne kontrastiert der verfolgte medienästhetische Ansatz auch die lang bewährten Annahmen und einseitigeren Herangehensweisen kognitivistischer Ansätze, die filmisches Verstehen bzw. filmische Metaphern (vgl. Forceville 2009

14 Vgl. Kappelhoff und Müller (2011), Müller und Kappelhoff (2018: Kap. 9.1, 9.2) sowie einen Artikel von Horst, Boll, Schmitt und Müller (2014), die paradigmatisch Gesten als interaffektive körperliche Ausdrucksbewegung in der Face-to-Face-Interaktion beschreiben.

[2006]; mit Ausnahme Fahlenbrachs 2018) fast ausschließlich kausal-logisch aus der Diegese und/oder den filmischen Repräsentationen deduzieren (vgl. für eine kritische Auseinandersetzung, Kappelhoff 2018: Kap. 1.3; Müller/Kappelhoff 2018: Kap. 1.3; Schmitt 2020: Kap. 5). Auf diese Weise wird (metaphorische) Filmbedeutung als etwas Statisches und Inhärentes begriffen, wobei die ästhetische Ebene des Films mitsamt ihrer entscheidenden Rolle in der dynamischen interaktiv verkörperten Kreation von (metaphorischer) Filmbedeutung seitens der Zuschauer*innen weitestgehend außer Acht gelassen wird. So würde ein kognitivistischer Ansatz die filmische Metaphorik vermutlich in den repräsentierten Darstellungen verorten; z.B. würde die Non-Linearität der Zeit(wahrnehmung), mitsamt ihrer deterministischen Verbindung zum non-linearen Sprachkonzept der Außerirdischen, bspw. aus den bildlichen Repräsentationen der kreisförmigen Schriftsymbole gefolgt werden. Die Vorstellung einer solchen konzeptuellen Metapher (z.B. ZEIT ALS KREIS) reduziert dann den metaphorischen Prozess auf ein kognitives Mapping, einen kognitiven Prozess der Übertragung.

In einer medienästhetischen Analyse des Metaphorisierens, entsteht die Metaphorizität hingegen aus den Bewegungserfahrungen der Zuschauenden, die die zirkuläre Zeit zuallererst am eigenen Leibe im situierten Kontext des Filmesehens erleben. Der methodische Zugriff auf die Analyse filmischer Metaphorik, wie ihn die empirische Medienästhetik eröffnet, stellt gegen die Identifikation konzeptueller Metaphern somit die Rekonstruktion dynamischer Formen des verkörperten Metaphorisierens. Mit Blick auf ARRIVAL als spezifisches Filmelerlebnis bedeutet dies: Aus der körperlichen Wahrnehmungserfahrung einer kontrastierten Linearität wird die Filmidee eines extraterrestrischen Sprach- und Zeitkonzepts konkret als non-linear bindende Ordnung verstanden. Damit wird Unreales wie das Erinnern von Zukunftsvisionen für Zuschauende leiblich *mitvollziehbar* (statt rein kognitiv nur *nachvollziehbar*) und auf diese Weise für sie (be)greiflich, ergo real. Diese ganz spezifische Grauzone eines filmisch-unrealen, aber zugleich leiblich-realen Interaktionsspektrums fasst Kappelhoff im Begriff des *Zuschauer-Ichs* bzw. *Zuschauergefühls* wie folgt:

[...] [Es] ist zugleich Bewusstsein, dass *nicht ich es bin*, der sich solcherart situiert findet; dass die Art und Weise des In-der-Welt-sich-Befindens, die ich hier und jetzt mit meinen Sinnen realisiere, nicht die meine ist. Ganz im Gegenteil – es ist immer die Subjektivität eines anderen, fremden Ichs. (Kappelhoff 2018: 146; Herv.i.O.)

Auch für Vivian Sobchack halten Kappelhoff und Bakels (2019) übereinstimmend fest,

[...] ist die mediale Grundsituation des Kinos ein Spiel *doppelter Subjektivität*: Der körperlose kinematografische Blick der Kamera begegnet – indem er als ein rein

sinnliches Sehen und Hören jedweder sprachlich-semiotischen Vermittlung vorausgeht – dem Filmzuschauer als die Wahrnehmung eines vom Ich getrennten, fremden Subjekts; [...] *in die Wahrnehmungsposition jenes fremden Subjekts gerückt, realisiert der Zuschauer die visuellen, auditiven und kinetischen Empfindungsqualitäten dieser fremden Wahrnehmungsbewegung sinnlich am eigenen Leib* – das Zuschauersubjekt wird buchstäblich zur Verkörperung des filmischen Subjektes. (Kappelhoff/Bakels 2019: 449; Herv. KP)

Nur in diesem Kontext wird die Realisierung eines fiktiven ›unrealen‹ Interaktionsraums durch die eigenen Sinne am eigenen Leib, als *reale* körperliche Interaktion verstanden, die in der spezifischen Interaktionsform mit audiovisuellen Medien stets die Ambiguität einer nicht wirklich eigenen Erfahrung hat, sondern die eines fiktiven Alter Egos, die sich real nur im eigenen Leib verwirklichen kann. Genau in diesem Spannungsfeld versteht dieser Beitrag inszenierte ›unreale‹ Interaktionsräume, die audiovisuell im situierten Moment des Sehens, Hörens und Spürens für Zuschauer*innen filmisch erfahrbar und körperlich zugänglich werden, folglich als gleichwohl real.

Damit wird der Interaktionsbegriff eher weit als eng gefasst. Nach Schmidt und Marx (2017: 30) charakterisiert »eine enge Fassung des Begriffs« bspw. ein Festhalten am Vorhandensein von Merkmalen der Face-to-Face-Interaktion. Dazu zählen sie »Kontextbezug, Situierung, Wahrnehmungs-Wahrnehmungen, Wechselseitigkeit etc.«, sodass Interaktionsforschung lediglich auf ko-präsente Kontakte beschränkt wäre (vgl. ebd.). Fasst man den Interaktionsbegriff hingegen weiter, führen Schmidt und Marx (2017: 31) aus, reiche die »Möglichkeit ›wechselseitiger Aufeinander-Bezogenheit««. Sie erläutern in diesem Zuge:

Wechselseitigkeit kann dann real vollzogen (wie in Gesprächen), bloß angedacht (wie in Anrufen, die nicht beantwortet werden), simuliert (wie in so genannter parasozialer Interaktion) oder artifiziell (wie in Mensch-Maschine-Interaktionen) sein. In allen Fällen spielt die *Grundfunktion von Interaktion, nämlich die ›Herstellung von Intersubjektivität‹* [...] *eine zentrale Rolle*, wenn nicht real vollzogen, so doch zumindest als orientierende Bezugsgröße oder regulative Leitidee. (Schmidt/Marx 2017: 31; Herv. KP)

Die vorangehenden Zitate zeigen, dass genau diese Grundfunktion – nämlich *die Herstellung von Intersubjektivität* – in der ganzkörperlichen Verwicklung der Zuschauer*innen mit den Filmbildern im Moment des Filmsehens erfüllt wird. Aus diesem Grund sprechen Müller und Kappelhoff (2018) mit der Zentralsetzung der Zuschauer*innen auch von einer verkörperten *Interaktion mit Filmen*, aus der schließlich Metaphorisierungen im Sinne kinematischer Metaphern entstehen können. Metaphorisierungen, die – wie die Fallstudie zeigen konnte – keine Statik aufweisen, keinem einzelnen Wort, Bild oder Konzept eigen sind, sondern ganz und gar beweglich, d.h.

dynamisch, sind und sich aus der Wechselseitigkeit zwischen zwei lebendigen »Körpern« entfalten und *ad hoc* Bedeutung kreieren:

Metaphoric meaning-making is local; it is what people do when they engage with each other in conversation or with a film, a television news feature, or a music video. From such a vantage point the static concept of metaphor becomes a dynamic process of metaphorizing. (Müller 2019: 62; Herv. KP)

Abschließend lässt sich also festhalten, dass aus einer medienästhetischen Perspektive a) das Filmesehen als interaktiver Prozess aufgefasst wird, der dynamisch im Verwickelt-Sein der Zuschauenden abläuft und die affektive Dimension, die ausdrückt und gleichzeitig die Zuschauenden bewegt, miteinschließt. Damit wird b) das audiovisuelle Medium Film als aus Bewegungsbildern beschaffenes Material aufgefasst, das in seiner zeitlichen Abfolge filmische Ausdrucksbewegungen hervorbringt, die als konkrete Bewegungserfahrung für Zuschauende unmittelbar spürbar werden. Dementsprechend wird c) Filmverstehen nicht länger als rein kognitiver Prozess im Sinne eines Abrufens und Entschlüsselns übermittelter Informationen aufgefasst (wie im einleitenden Zitat Kappelhoffs angedeutet), sondern vielmehr als körperlich grundierter Prozess, der erst im Moment des Filmlebens in Zuschauenden entsteht und augenblicklich kreiert wird (folglich der Begriff der *Poiesis*, griech. für Erschaffung, Hervorbringung) – und zwar in Form eines Metaphorisierungsprozesses, der *dynamisch* aus der verkörperten Bewegungserfahrung der filmischen Bilder entsteht und *interaktiv* in Zuschauer*innen Bedeutung generiert. In genau diesem Sinne versteht der vorliegende Beitrag Metaphorisierungen wie die einer non-linearen Zeit als wahrlich »(un)reale« Interaktion von Zuschauenden mit filmischen Bildern.

Literatur

- Cameron, Lynne (2018): »From Metaphor to Metaphorising: How Cinematic Metaphor Opens up Metaphor Studies«, in: Sarah Greifenstein/Dorothea Horst/Thomas Scherer/Christina Schmitt/Hermann Kappelhoff/Cornelia Müller (Hg.), *Cinematic Metaphor in Perspective: Reflections on a Transdisciplinary Framework*, Berlin/Boston: De Gruyter, S. 17–35.
- Cameron, Lynne/Robert Maslen/Zazie Todd/John Maule/Peter Stratton/Neil Stanley (2009): »The Discourse Dynamics Approach to Metaphor and Metaphor-Led Discourse Analysis«, in: *Metaphor and Symbol*, Vol. 24.2, S. 63–89.
- Deleuze, Gille (1989 [1983]): *Das Bewegungsbild*. Kino 1, Frankfurt: Suhrkamp.
- Fahlenbrach, Kathrin (2018): »Moving Metaphors: Affects, Movements, and Embodied Metaphors in Cinema«, in: Sarah Greifenstein/Dorothea Horst/Thomas

- Scherer/Christina Schmitt/Hermann Kappelhoff/Cornelia Müller (Hg.), *Cinematic Metaphor in Perspective: Reflections on a Transdisciplinary Framework*, Berlin/Boston: De Gruyter, S. 69–92.
- Forceville, Charles (2009 [2006]): »Non-Verbal and Multimodal Metaphor in a Cognitivist Framework: Agendas for Research«, in: Gitte Kristiansen/Michel Achard/René Dirven/Francisco Ruiz de Mendoza Ibàñez (Hg.), *Cognitive Linguistics: Current Applications and Future Perspectives*, Berlin/New York: Mouton de Gruyter, S. 379–402.
- Gemoll, Wilhelm Karl Vretska/Heinz Kronasse (2000): *Gemoll. Griechisch-Deutsches Schul- und Handwörterbuch*, München: Oldenbourg Schulbuchverlag, S. 437.
- Greifenstein, Sarah/Regina Brückner (2017): »Imagining the Human Brain. Audiovisual Metaphors in Science Television Programs«, in: *Mediaesthetics. Zeitschrift für Poetologien audiovisueller Bilder 2*. Abrufbar unter: <https://www.mediaesthetics.org/index.php/mae/article/view/77/159> (Stand 31.07.2024).
- Greifenstein, Sarah/Dorothea Horst/Thomas Scherer/Christina Schmitt/Hermann Kappelhoff/Cornelia Müller (Hrsg.) (2018): *Cinematic Metaphor in Perspective: Reflections on a Transdisciplinary Framework*, Berlin/Boston: De Gruyter.
- Greifenstein, Sarah (2020): *Tempi der Bewegung – Modi des Gefühls. Expressivität, heitere Affekte und die Screwball Comedy*, Berlin/Boston: De Gruyter.
- Horst, Dorothea/Franziska Boll/Christina Schmitt/Cornelia Müller (2014): »Gesture as Interactive Expressive Movement: Inter-Affectivity in Face-to-Face Communication«, in: Cornelia Müller/Alan Cienki/Ellen Fricke/Silva Ladewig/David McNeill/Jana Bressemer (Hg.), *Body – Language – Communication: An International Handbook on Multimodality in Human Interaction*, Vol. 38.2, Berlin/Boston: De Gruyter Mouton, S. 2112–2125.
- Julich-Warpakowski, Nina/Thomas Wiben Jensen (2023): »Zooming in on the Notion of Metaphoricity: Notions, Dimensions and Operationalizations«, in: *Metaphor and the Social World*, Vol. 13:1, S. 16–36.
- Kappelhoff, Hermann (2004): *Matrix der Gefühle. Das Kino, das Melodrama und das Theater der Empfindsamkeit*, Berlin: Vorwerk 8.
- Kappelhoff, Hermann (2008a): *Realismus. Das Kino und die Politik des Ästhetischen*, Berlin: Vorwerk 8.
- Kappelhoff, Hermann (2008b): »Zuschauergefühl. Die Inszenierung der Empfindung im dunklen Raum des Kinos«, in: Geraldine Spiekermann/Beate Söntgen (Hg.), *Tränen*, München, S. 195–206.
- Kappelhoff, Hermann (2016): *Genre und Gemeinsinn. Hollywood zwischen Krieg und Demokratie*. Berlin/Boston: De Gruyter.
- Kappelhoff, Hermann (2018): *Kognition und Reflexion. Zur Theorie filmischen Denkens*, Berlin/Boston: De Gruyter.

- Kappelhoff, Hermann/Jan-Hendrik Bakels (2011): »Das Zuschauergefühl. Möglichkeiten qualitativer Medienanalyse«, in: *Zeitschrift für Medienwissenschaft*. Heft 5: Empirie, Jg. 3 (2), S. 78–96.
- Kappelhoff, Hermann/Cornelia Müller (2011): »Embodied meaning construction. Multimodal metaphor and expressive movement in speech, gesture, and feature film«, in: *Metaphor and the Social World*, Vol. 1.2, Amsterdam: John Benjamins Publishing Company, S. 121–153.
- Kappelhoff, Hermann/Jan-Hendrik Bakels/Hye-Jeung Chung/David Gaertner/Sarah Greifenstein/Matthias Grotkopp/Michael Lück/Christian Pischel/Cilli Pogodd/Franziska Seewald/Christina Schmitt/Anna Steininger (o.D.): eMAEX – Ansätze und Potentiale einer systematisierten Methode zur Untersuchung filmischer Ausdrucksqualitäten, Abrufbar unter: https://www.empirische-medienaesthetik.fu-berlin.de/media/emaex_methode_deutsch/eMAEX_-_Ansatz_e-und-Potentiale-einer-systematisierten-Methode-zur-Untersuchung-filmischer-Ausdrucksqualitaeten.pdf (Stand: 13.02.2024).
- Kappelhoff, Hermann/Jan-Hendrik Bakels (2019): Audiovisuelle Affekte – die Emotionstheorie des Films im Spannungsfeld von früher Filmtheorie, Kognitionstheorie und Medienästhetik«, in: Hermann Kappelhoff/Jan-Hendrik Bakels/Hauke Lehmann/Christina Schmitt (Hg.), *Emotionen. Ein interdisziplinäres Handbuch*, Stuttgart: J.B. Metzler, S. 445–451.
- Kappelhoff, Hermann/Jan-Hendrik Bakels/Sarah Greifenstein (2024): *Die Poiesis des Filme-Sehens. Methoden der Analyse audiovisueller Bilder*, Berlin/Boston: De Gruyter.
- Kindler-Mathôt, Clara/Cornelia Müller/Jana Katharina Junge/Katerina Papadopoulou (im Erscheinen): »Sprechen in audiovisuellen Medien: Systematik für eine medienästhetische Analyse multimodaler Kommunikation«, in: Rita Tamara Vallentin/Dorothea Horst (Hg.), *Sprache entgrenzen: Beiträge zu einer kulturwissenschaftlichen Linguistik*, Bielefeld: transcript.
- Lakoff, George/Mark Johnson (1980): *Metaphors We Live By*, Chicago/London: Chicago University Press.
- Müller, Cornelia (1998): *Redebegleitende Gesten. Kulturgeschichte – Theorie – Sprachvergleich*, Berlin: Arno Spitz.
- Müller, Cornelia (2008): *Metaphors Dead and Alive, Sleeping and Waking. A Dynamic View*, Chicago: University of Chicago Press.
- Müller, Cornelia (2010): »Wie Gesten bedeuten. Eine kognitiv-linguistische und sequenzanalytische Perspektive«, in: *Sprache und Gestik. Sonderheft der Zeitschrift Sprache und Literatur*, Vol. 41.1, S. 37–68.
- Müller, Cornelia (2017): »Waking Metaphors. Embodied Cognition in Multimodal Discourse«, in: Beate Hampe (Hg.), *Metaphor. Embodied Cognition in Discourse*, Cambridge: Cambridge University Press, S. 297–316.

- Müller, Cornelia (2019): »Metaphorizing as Embodied Interactivity: What Gesturing and Film Viewing Can Tell Us About an Ecological View on Metaphor«, in: *Metaphor and Symbol*, Vol. 34.1, S. 61–79.
- Müller, Cornelia (2024): »Dead and Alive, Sleeping and Waking Metaphors: The Spectrum of Metaphor and the Multimodality of Discourse«, in: Anders Örtenblad (Hg.), *The Oxford Handbook of Metaphor in Organization Studies*, Oxford: Oxford University Press, S. 70–84.
- Müller, Cornelia (im Erscheinen): »Sprache-Sprechen, Hören, Sehen, Fühlen. Eine medienästhetische Perspektive auf die Multimodalität des Sprechens in audiovisuellen Medien«, in: Rita Tamara Vallentin/Dorothea Horst (Hg.), *Sprache entgrenzen: Beiträge zu einer kulturwissenschaftlichen Linguistik*, Bielefeld: transcript.
- Müller, Cornelia/Hermann Kappelhoff (2018): *Cinematic Metaphor. Experience – Affectivity – Temporality*, Berlin/Boston: De Gruyter Mouton.
- Müller, Cornelia/Christina Schmitt (2015): »Audio-visual Metaphors of the Financial Crisis: Meaning Making and the Flow of Experience«, in: R. W. Gibbs Jr./L. Corrêa Ferreira, (Hg.), *Revista Brasileira de Linguística Aplicada/Brazilian Journal of Applied Linguistics*, Vol. 15.2, S. 311–342.
- Müller, Cornelia/Susanne Tag (2010): »The Dynamics of Metaphor: Foregrounding and Activating Metaphoricity in Conversational Interaction«, in: *Cognitive Semiotics*, Vol. 6, S. 85–120.
- Núñez, Rafael E./Eve Sweetser (2006): »With the Future Behind Them: Convergent Evidence from Aymara Language and Gesture in the Crosslinguistic Comparison of Spatial Construals of Time«, in: *Cognitive Science*, Vol. 30, S. 401–450.
- Plessner, Helmuth (1982 [1925]): »Die Deutung des mimischen Ausdrucks. Ein Beitrag zur Lehre vom Bewußtsein des anderen Ichs« (verfasst in Zusammenarbeit mit Frederik Jacobus Buytendijk), in: Günther Dux/Odo Marquard/Elisabeth Ströker (Hg.), *Helmuth Plessner. Gesammelte Schriften*, Vol. 7: *Ausdruck und menschliche Natur*, Frankfurt a. M.: Suhrkamp, S. 67–129.
- Sapir, Edward (1929): »The Status of Linguistics as a Science«, in: *Language*, Vol. 5.4, *Linguistic Society of America*, S. 207–214.
- Scherer, Thomas/Sarah Greifenstein/Hermann Kappelhoff (2014): »Expressive Movements in Audio-Visual Media: Modulating Affective Experience«, in: Cornelia Müller/Alan Cienki/Ellen Fricke/Silva Ladewig/David McNeill/Jana Bressemer (Hg.), *Body – Language – Communication: An International Handbook on Multimodality in Human Interaction*, Vol. 38.2, Berlin/Boston: De Gruyter Mouton, S. 2081–2092.
- Scherer, Thomas (2023): *Inszenierungen zeitgenössischer Propaganda. Kampagnenfilme im Dienste des Gemeinwohls*, Berlin/Boston: De Gruyter.
- Schmidt, Axel/Konstanze Marx (2017): »Interaktion und Medien«, in: *Sprachreport* Jg. 33, Nr. 4, S. 22–33.

- Schmitt, Christina/Sarah Greifenstein/Hermann Kappelhoff (2014): »Expressive Movement and Metaphoric Meaning Making in Audio-Visual Media«, in: Cornelia Müller/Alan Cienki/Ellen Fricke/Silva Ladewig/David McNeill/Jana Bressemer (Hg.), *Body – Language – Communication: An International Handbook on Multimodality in Human Interaction*, Vol. 38.2, Berlin/Boston: De Gruyter Mouton, S. 2092–2112.
- Schmitt, Christina (2020): *Wahrnehmen, Fühlen, Verstehen. Metaphorisieren und audiovisuelle Bilder*, Berlin/Boston: De Gruyter.
- Sobchack, Vivian (1990): »The Active Eye: A Phenomenology of Cinematic Vision«, in: *Quarterly Review of Film and Video*, Vol. 12.3, S. 21–36.
- Sobchack, Vivian (1992): *The Address of the Eye: A Phenomenology of Film Experience*, Princeton: Princeton University Press.
- Werlen, Iwar (2002): *Sprachliche Relativität. Eine problemorientierte Einführung*, Tübingen/Basel: A. Francke.
- Whorf, Benjamin Lee (1963): »Sprache, Denken, Wirklichkeit. Naturwissenschaft und Linguistik«, in: Peter Krausser (Hg.), *Sprache, Denken, Wirklichkeit: Beiträge zur Metalinguistik und Sprachphilosophie*, Reinbek: Rowohlt, S. 7–18.

Filmographie

ARRIVAL (Denis Villeneuve 2016: Paramount Pictures).

Erfahrungsraum Interaktion

Die audiovisuelle Modulation eines Gesprächs als Zuschauer*innenerfahrung

Clara Kindler-Mathôt

Abstract *Der medienlinguistische Beitrag nimmt aus einer Zuschauer*innenperspektive die komplexe Bedeutungsentfaltung audiovisueller Bilder in den Blick. Am Beispiel eines Videobeitrags des digitalen Parteitags der Partei BÜNDNIS 90/DIE GRÜNEN wird gezeigt, wie audiovisuelle und sprachlich-interaktive Formen gemeinsam ein spezifisches Bewegungsbild entfalten, welches die Zuschauer*innenerfahrung eines ›Gesprächs‹ moduliert, das nicht das Abbild vorfilmischer Realität, sondern eine neue mediale Wirklichkeit darstellt. Der theoretische und methodische Ankerpunkt des medienästhetischen Ansatzes liegt im Konzept der filmischen Ausdrucksbewegung, welcher mit Ansätzen der Interaktions- und Gesprächsanalyse ergänzt wird. Die Diskussion der Analyse betont die Unmittelbarkeit der Zuschauer*innenerfahrung und plädiert vor dem Hintergrund zeitgenössischen Medienkonsums für eine Stärkung der Medienkompetenz zur Reflexion audiovisueller Bedeutungskonstitution.*

Keywords *Zuschauer*innenerfahrung; mediale Wirklichkeit; Medienästhetik; sprachliche und audiovisuelle Multimodalität; mediale Durchformung*

1. Modulation medialer Wirklichkeiten als Zuschauer*innenerfahrung

Die Gegenwart bringt zunehmend neue Formen der audiovisuellen Kommunikation und Interaktion hervor, seien es DIY-Tutorials auf Instagram, Live-Übertragungen digitaler Parteitage oder politisch aufgeladene Kommentare auf YouTube. Insbesondere vor der Frage um die »Wirkung« audiovisueller Bilder wird die Rolle des Medienkonsums und dessen gesellschaftliche Folgen relevant. Ein Aspekt davon bezieht sich auf die Wahrhaftigkeit oder Echtheit audiovisueller Bilder, bzw. ob das, was wir sehen, real ist oder nicht – oder zumindest als solches beurteilt wird. Der Beitrag nimmt diese Frage aus einer medienlinguistischen Perspektive auf, die Ansätze der sprachwissenschaftlichen Multimodalitätsforschung mit filmtheoretischen und -analytischen Perspektiven vereint. Am Beispiel eines Videobeitrags der

digitalen Bundesdelegiertenkonferenz (DBDK) 2020 der Partei BÜNDNIS 90/DIE GRÜNEN wird gezeigt, dass auch die bewegten Bilder einer vordergründig ›beobachtenden‹ und vergleichsweise ›objektiven‹ Kamera nicht einfach Abbildung der vorfilmischen Realität sind, sondern jede Aufnahme neue *mediale Wirklichkeiten* (vgl. Luginbühl 2019) erzeugt.

Diese bedeutungskonstituierende Funktion von Medialität wird in Arbeiten der sprachlichen Multimodalitätsforschung (v.a. multimodaler Interaktionsanalyse) in der Regel nicht weiter reflektiert, denn sie sind häufig der Versuch, eine emische Perspektive auf spezifische Interaktionsformen einzunehmen. Der Artikel nimmt deshalb einen Perspektivwechsel vor, indem nicht der Bedeutungsaushandlungsprozess zwischen Personen im Bild, sondern zwischen Bild und Zuschauer*innen adressiert wird. In dieser Zuschauer*innenerfahrung (vgl. Kappelhoff 2008; Kappelhoff/Bakels 2011) verschmelzen die Multimodalität des Sprechens und die audiovisuelle Multimodalität zu einer komplexen *multidimensionalen Bedeutungsgestalt* (Müller/Kappelhoff 2018: Kap. 2.3), die im Moment des Sehens unmittelbar erfahren wird. Die spezifischen Bewegungsbilder dieser Bedeutungsgestalten formen den Erfahrungsraum der Zuschauer*innen.

Der theoretische und methodische Ankerpunkt liegt dabei im medienästhetischen Konzept der filmischen Ausdrucksbewegung (vgl. Kappelhoff 2004: Teil 1; Kappelhoff/Müller 2011; Müller/Kappelhoff 2018: Kap. 9; Scherer et al. 2014). Mithilfe dieses Ansatzes zeigt die Fallstudie am Beispiel eines ›Gesprächs‹, dass dieses nicht einfach repräsentiert wird, sondern erst durch die audiovisuelle Inszenierung als ein Bewegungsbild in der Zuschauer*innenerfahrung moduliert wird. Kamera und Schnitt setzen die Personen in ein raum-zeitliches Verhältnis, welches als Interaktion von Angesicht zu Angesicht (i.F. von AzA) – als Gespräch – erfahren wird. Diese Zuschauer*innenerfahrung ist so unmittelbar, dass sie erst in Momenten der Irritation hinterfragbar wird, da sie bspw. als Bruch in den kommunikativen Praktiken erfahren wird. Vor dem Hintergrund zeitgenössischer Formen des Medienkonsums, plädiert der Beitrag deshalb für eine Stärkung der Medienkompetenz der Zuschauer*innen, um die spezifische Medialität des Audiovisuellen reflektieren zu können.

2. Eine Frage der Perspektive: Audiovisuelle Medialität und Sprachgebrauch von Innen und von Außen

»Kommunikation [ist] immer auf Medialität angewiesen« (Luginbühl 2019: 126), sei sie schriftlich, mündlich oder gebärdet, von AzA, am Telefon oder im Fernsehen. Jede Medialität unterliegt dabei spezifischen Bedingungen, die die Kommunikation nicht nur übertragen, sondern *ausformen*, genauer, *mitkonstituieren* (vgl. ebd.). Eine Änderung der *medialen Bedingungen* (vgl. Marx/Schmidt 2019: 14ff.), hat immer

Auswirkungen auf die semiotischen Ressourcen und somit auf die Form der Bedeutungsaushandlung. Bei einem Telefonat etwa fallen alle visuellen Ressourcen weg, in einer Videokonferenz ist nur sichtbar, was in den Kameraausschnitt fällt. Dementsprechend ist auch eine Anpassung der methodischen Herangehensweisen erforderlich, was beispielsweise Ausdruck in der wachsenden Multimodalitätsforschung und Medienlinguistik findet.

Im Folgenden werden zwei Perspektiven auf Medialität und ihre »vermittelnde« Funktion eröffnet. Die erste bezieht sich auf eine »Innenperspektive«: Die interaktionsanalytischen Zugänge zu *medienvermittelter Kommunikation* thematisieren den Aushandlungsprozess zwischen den Interaktionsteilnehmer*innen. Medialität übernimmt dabei die Funktion eines Übertragers und Vermittlers. Mit dem daran anschließenden Konzept der *medialen Durchformung* wird eine Brücke zur zweiten Perspektive geschlagen, die das Medialitätsverständnis erweitert und die Rolle der Betrachter*innen audiovisueller Bilder mit einbezieht. Diese Außenperspektive wird im dritten Abschnitt durch den filmwissenschaftlichen Ansatz des *Zuschauergefühls* (Kappelhoff 2008, Bakels/Kappelhoff 2011) ausgeführt, welcher den Ausgangspunkt für die sich anschließende Analyse bildet.

2.1 Innenperspektive: Ansätze der medienvermittelten Kommunikation

In der interaktionslinguistischen Forschung zu *medienvermittelter Kommunikation* werden nicht nur Fragen nach dem Übertragungsprozess von Kommunikation mittels (meist als technische Hilfsmittel verstandener) Medien (vgl. Marx/Schmidt 2019: 9) gestellt, sondern auch nach den »spezifischen Folgen medialer Vermittlung für Interaktion und Interaktionsforschung« (Marx/Schmidt 2019: 2). Damit wird eine emische Innenperspektive eingenommen: Durch die ethnomethodologisch-konversationsanalytische Forschungslogik (vgl. ebd.: 14), die in der »möglichst guten Annäherung an die Teilnehmerperspektive der Interagierenden« (Lanwer 2019b: 5) besteht, werden für die Interagierenden sicht- und hörbare Praktiken in Bezug auf reflexive Aushandlungsprozesse analysiert. Dabei wird vor allem in Betracht genommen, wie (und ob) die vermittelnden Medien die Kernaspekte von Interaktion beeinflussen – im weitesten Sinne stellen dies Formen der Kopräsenz und Wechselseitigkeit dar (vgl. Marx/Schmidt 2019: 2) – und wie die Interagierenden damit umgehen, insbesondere in Bezug auf die sich medial verändernden semiotischen Ressourcen:

Da Medienvermittlung das gemeinsame Hier und Jetzt (je nach Medium) partiell auflöst bzw. in umgekehrter Perspektive distante und in größerem Umfang verbreitete Kommunikation überhaupt erst ermöglicht, stehen viele dieser Ressourcen und auf ihnen basierende Lösungen nicht mehr oder zumindest so nicht mehr zur Verfügung. Stattdessen kommen neue Ressourcen (etwa im Falle des Telefons

das Telefonklingeln) und entsprechende Lösungen hinzu, etwa spezifische Eröffnungssequenzen. (Marx/Schmidt 2019: 2)

Die Innenperspektive beschreibt folglich, wie die Interaktionsteilnehmer*innen in ihren gegenwärtigen Realitäten, in ihrem Hier und Jetzt agieren, wie sie wechselseitig aufeinander bezogen handeln, wie sie sich in den spezifischen Interaktionsräumen, seien sie virtuell oder real, mündlich oder schriftlich, orientieren und mit den spezifischen medialen Umwelten umgehen. Was bedeutet eine solche Perspektive für Interaktion in audiovisuellen Kontexten? Aus Perspektive der medienvermittelten Kommunikation steht auch hier in erster Linie die Bedeutungsaushandlung zwischen den Teilnehmer*innen im Blick. Im Falle von Videokonferenzen¹ etwa zeigen Studien (vgl. Haddington/Oittinen 2022; Lanwer 2019a; 2019b; Xia 2023), dass die medialen Bedingungen Einfluss auf die gegenseitige Koordinierung der Interagierenden nehmen. Die korpusbasierte Vergleichsstudie zwischen Gesprächen von AzA und Zoomkonferenzen von Xia (ebd.) zeigt so bspw. ein deutlich verändertes Timing bei den Turn-Übernahmen auf: Gespräche in der Videokonferenz neigen so zwar zu weniger Überlappungen, dafür sind diese, ebenso wie die Pausen zwischen den Turns, länger. Ebenfalls in Studien thematisiert werden die gemeinsame Blickkoordination sowie Aufmerksamkeitsfokussierung (*joint attention*; Lanwer 2019a; 2019b), die durch die technischen Affordanzen wie Kamera, Softwareinterface oder Internetverbindung beeinflusst werden.

Während die Forschung zu Videokonferenzen trotz Berücksichtigung dieser Affordanzen eine emische Perspektive der vorfilmischen Realität der Teilnehmer*innen einnimmt, öffnet die Forschung zu Fernsehgesprächen zusätzlich die Frage nach den Zuschauer*innen und wie für sie Bedeutung konstituiert wird. Da damit Fragen der Rezeption angesprochen werden, deutet sich an dieser Stelle eine Form der Außenperspektive an. Spätestens hier wird es relevant, zwischen zwei Formen der Realität zu unterscheiden: Zum einen die vorfilmische Realität, in der Interaktionsteilnehmer*innen bspw. in einem Studio sitzen und sich in einer Interaktionssituation von AzA miteinander befinden. Zum anderen die Realität die sich für die Zuschauer*innen als audiovisuelle Bewegungsbilder auf ihren Bildschirmen entfaltet. Diese ist geformt und moduliert durch die audiovisuelle Medialität: Kameraeinstellungen und Bewegungen bestimmen die für die Zuschauer*innen sichtbaren Bildräume, die Position der Mikrofone entscheidet, wer oder was zu hören ist, Schnitte und Montagesequenzen ermöglichen es, zusätzliche Bilder einzublenden und eine Form der Zeitlichkeit zu kreieren, die nicht der Zeitlichkeit der vorfilmischen Realität entsprechen muss. Luginbühl (2019; Luginbühl/Schneider

1 Mit Videokonferenzen sind sogenannte *Face-to-Screen-Interaktionen* gemeint (vgl. Lanwer 2019a; 2019b) und weniger Konferenzen, die in Besprechungsräumen mit »veränderter Ökologie« (Meier 2000: 195) stattfinden.

2020) bezeichnet deshalb diese Zuschauer*innenrealität als neue, *medial durchformte Wirklichkeit*. Dieser Ansatz bildet damit eine Brücke zwischen der Innen- und der Außenperspektive, da er beide Ebenen der Interaktion berücksichtigt.

2.2 Brücke zwischen Innen und Außen: Mediale Durchformung

Der Ansatz der medialen Durchformung beschreibt den Einfluss medialer Affordanzen auf die Zuschauer*innenerfahrung. Für Fernsehgespräche (vgl. Luginbühl 2019) bedeutet dies, dass sie nicht abgefilmte »außermediale Gespräche« (ebd.: 143) sind, die einfach übertragen werden, sondern die Affordanzen der spezifischen medialen Ökologien die Bedeutungskonstitution der Zuschauer*innen entscheidend mitbestimmen. Luginbühl und Schneider (2020) illustrieren dies am Beispiel der zweiten Präsidentschaftsdebatte von 2016 in den USA zwischen Hillary Clinton und Donald Trump. Zum einen werden die räumlichen Affordanzen des Studios und die Kameraführung von den Kontrahenten selbst als bedeutsam erfahren und dementsprechend genutzt, etwa in der aktiven Zuwendung zur Kamera. Zum anderen wird ihr Verhalten selbst aus einer Zuschauer*innenperspektive von der Medialität des Audiovisuellen, bspw. der Kameraperspektive, moduliert (vgl. ebd.: 58). Das Gespräch von AzA wird so durch die medialen Umstände des Fernsehens quasi in doppelter Weise durchformt (von Innen und von Außen), woraus eine neue Form der Medialität hervorgeht, in welcher Inhalte nicht nur vermittelt werden, sondern die Medialität selbst als Akteurin der Bedeutungskonstitution hervorgeht:

In this threefold way, the medium of television is thus a political actor: it presents oral interactions that are shaped from the outset by the institution of television and the media format of television discussion. Without the medium of television, these interactions would never take place in this way, as a result of which the medium not merely transmits the content, but becomes an actor in its own right. (ebd.: 61)

Holly (2015) spricht in diesem Zusammenhang auch von *Intermedialität* und *intermedialen Bezugnahmepraktiken*, aus welchen das Zusammenspiel von gesprochener Sprache und Audiovisualität als Prozess gemeinsamer Semiose hervorgeht. Am Beispiel der Polit-Talkshow »Maybritt-Illner« beschreibt er das intermediale Zusammenspiel, in welchem durch Kameraperspektiven und Bewegungen, Einblendung von Bildmaterial o.Ä. Bedeutungsaspekte der Interaktion verändert oder hinzugefügt werden:

Die Kameraführung in Polit-Talkshows generiert durch die Selektion von Einstellungen und Umschnitten Bedeutungskomponenten, die sprachliche Äußerungen

»transkribieren«, d.h. überformen, implizit kommentieren und dadurch »anders lesbar« machen. (ebd.: 89)

Die Ansätze dieser sogenannten *Transkription* und die Generierung zusätzlicher Bedeutungskomponenten sind den Zuschauer*innen im Moment des Sehens nicht offenkundig. In der filmischen Erfahrung sind sprachliche Äußerungen und die Medialität des Audiovisuellen nicht mehr trennbar, sondern sind in ihrer ganzheitlichen Gestalt auch bedeutungskonstituierend. In dieser Ganzheitlichkeit verschwindet die Medialität und hinterlässt nur unmerkliche Spuren. Luginbühl und Schneider (vgl. ebd.: 59) beschreiben dies, mit Verweis auf Sybille Krämer (1998), als Tendenz der Medien sich unsichtbar zu machen. Diese Unsichtbarkeit wird jedoch in Momenten der Irritation aufgeweicht: »Für wirklich hält der Filmbetrachter vor allem das, was den Eindruck von Wirklichkeit macht, für wenig eindrucksvoll wird dagegen gehalten, was diesem Illusionscharakter des Mediums nicht entspricht, sondern ihn irritiert« (Koch 2003: 223). Die mediale Wirklichkeit erscheint dann als nicht mehr kohärent, und wird hinterfragbar. Beispiele stellen Formen der Asynchronität oder Inkongruenz zwischen Bild, Ton, Bildraum und/oder Kameraführung dar, bedingt durch technische Störungen oder der Inszenierung selbst. In der unten vorgestellten Analyse (s. 3.3) zeigt sich dies exemplarisch in der ungewöhnlichen Turnkonstruktion, moduliert durch die Montage.

Während der Ansatz der medialen Durchformung grundsätzlich den Blick auf das Fernsehpublikum öffnet, verbleibt er jedoch auf Ebene, zwar durchformter aber dennoch repräsentierter Inhalte. Obwohl Luginbühl (2019) betont, dass die Bedeutungskonstitution nur aus dem Gebrauch, also aus dem Rezeptionsprozess hervorgeht, bleibt eine Reflexion dieses Prozesses aus. Ich möchte daher mit einem Blick in die medienästhetische Filmtheorie noch einen Schritt zurückgehen, indem gefragt wird, wie es überhaupt zur Konstitution von Bedeutung bei Zuschauer*innen kommt, wie sie also grundsätzlich audiovisuelle Bilder verstehen. Damit wird die wesentliche Zuschauer*innenerfahrung filmischer Bilder adressiert.

2.3 Außenperspektive: Das interaktive Moment der Zuschauer*innenerfahrung

Nach wie vor sind in der Filmwissenschaft kognitive Ansätze geläufig, in welchen filmische Bedeutung als Produkt der »Zuschaueraktivität im Rahmen eines Verstehens, Beurteilens oder Bewertens von narrativen, repräsentierten Handlungen, Szenarien und Figurenkonstellationen« (Greifenstein 2020: 24) verstanden wird.² Dabei wird das Narrativ und die Repräsentationsebene als gegeben, als offenkundig angesehen und nicht als Produkt der Rezeption von sich zeitlich entfaltenden und

2 Für eine Diskussion des Diskurses vgl. Müller/Kappelhoff (2018: Kap. 1).

angeordneten Bewegungsbildern. Müller und Kappelhoff (2018) kritisieren dies wie folgt:

The claim that basic meanings are understood, as a rule lacks a convincing model of how this understanding is accomplished. Instead, what is to be understood from film images is taken for granted, as already given: actors, actions, spaces, objects, treated as a concatenation of isolated pictorial representations. (ebd.: 30)

Der »specific media-character of audiovisual images« (ebd.: 8) wird außer Acht gelassen und damit auch der Prozess des Sehens und Verstehens audiovisueller Bilder: Sie entfalten sich im Moment des Sehens als eine zeitliche Erfahrung, welche aus dem Zusammenspiel der Inszenierungsformen wie Montagetechnik, Kameraeinstellungen usw., als orchestrierte Bewegungsfigurationen hervorgehen – als sogenannte *Bewegungsbilder* (vgl. ebd.: 23). Kappelhoff (2018) bezeichnet dies als die Poiesis des Filme-Sehens, als kreativen Prozess des Verstehens, als Entstehen von Film im Moment der Rezeption (vgl. Müller/Kappelhoff 2018: 223; Scherer 2024: 110ff.). Die Rezeption ist hier kein Dekodieren, sondern ein körperlich-sinnliches Wahrnehmen der Bewegungsbilder, die uns in ihren Rhythmen und Dynamiken direkt und unmittelbar *bewegen* (vgl. Scherer et al. 2014: 2083). Ihre Bewegungsmuster (*movement patterns*; vgl. ebd.) formen die alles umspannende Ebene der Zuschauer*innenerfahrung, indem sie den Wahrnehmungsvorgang, die kognitiven und perzeptiven Prozesse organisieren und strukturieren und so bedeutsam werden (vgl. Kappelhoff 2008: 206). Das sich so entfaltende *Zuschauergefühl* (Kappelhoff 2004; 2008; Kappelhoff/Bakels 2011; Bakels 2017) ist deshalb eine sinnlich-körperliche Erfahrung, eine Modellierung von sich entfaltenden Affekten, ein Mitfühlen (vgl. Bakels 2017: 67), eine grundlegende Form der Affizierung der Zuschauer*innen. Vor diesem Hintergrund sind die Zuschauer*innen nicht reine Empfänger*innen, sondern stehen in einer Art interaktivem Verhältnis mit den Bewegungsbilder selbst (vgl. Müller/Kappelhoff 2018: 29).

In Kappelhoffs Filmtheorie (2004; 2018) nehmen die Begriffe des *Bewegungsbildes* und der *Ausdrucksbewegung* eine zentrale Stellung ein. Er entwickelt den Begriff der filmischen Ausdrucksbewegung in einer kulturhistorischen Auseinandersetzung mit sprachwissenschaftlichen, soziologischen und psychologischen Ausdruckstheorien der ersten Dekaden des 20. Jahrhunderts (bspw. Bühler 1933; Plessner 1982; Wundt 1975) sowie den filmtheoretischen Schriften von Deleuze (*Zeit- und Bewegungsbild*; vgl. 1989; 1990). Mit der Unmittelbarkeit der Zuschau-Erfahrung und der Körperlichkeit des Filme-Sehens greift Kappelhoff Aspekte der neo-phänome-

nologischen Filmtheorie Vivian Sobchacks (1992) auf.³ Aktuelle Forschung in der Film- und Sprachwissenschaft hat die Konzepte kooperativ weiterentwickelt (Müller/Kappelhoff 2018; Greifenstein 2020; Horst 2018; Scherer 2024; Kindler-Mathôt et al. im Erscheinen) und über unterschiedliche Kontexte hinweg (insbesondere) am Gegenstand multimodaler Metaphorik in Sprache, Geste und audiovisuellen Medien methodisch erprobt und zum Ansatz einer empirischen Medienästhetik weitergedacht.

Dieser medienästhetische Ansatz grenzt sich ebenfalls von der (neuro-)psychologischen experimentellen⁴ sowie kommunikationswissenschaftlichen und medienpsychologischen Rezeptionsforschung⁵ ab, welche sich für Medienwirkung in empirisch-experimentellen Settings interessiert. Ähnlich der kognitiven Ansätze, wird hier in erster Linie Bezug zu Reiz-Reaktions-Relationen genommen, bspw. zur »systematische[n] Erfassung und Protokollierung von sinnlichen und apparativ wahrnehmbaren Aspekten menschlicher Handlungen und Reaktionen« (Gehrau 2017: 17). Damit werden vordergründig Fragen rund um intendierte Wirkungsstrategien thematisiert. Kappelhoffs medienästhetischer Ansatz versucht vielmehr, Grundmuster und Poetologien verschiedener Affizierungsformen aufzudecken und damit das unmittelbare Wahrnehmungserleben und das Filme-Sehen als Praxis und Poiesis zu verdeutlichen.⁶

Die folgende Analyse betrachtet deshalb die Inszenierung aus einer Zuschauer*innenperspektive als *Modulation einer medialen Wirklichkeitserfahrung von Interaktion*, die die Rolle von Produzent*innen und möglichen Intentionen außen vor lässt und vielmehr die audiovisuellen Bewegungsbilder in den Vordergrund stellt. Vor diesem Hintergrund wird nicht von der Repräsentation eines Gesprächs ausgegangen, sondern von einer audiovisuellen Inszenierung und Orchestrierung: Kamera und Montage setzen die Personen, deren Körperbewegungen und räumliche Situierung in ein Verhältnis, welches von den Zuschauer*innen als Interaktion, als Gespräch erfahren wird. Im Bewegungsbild verschmilzt die Multimodalität des Spre-

3 Zur ausführlichen Theoriegeschichte und Begriffskonzeption siehe Müller/Kappelhoff (2018: Kap. 9) sowie Greifenstein (2020: Kap. 2.3).

4 Siehe für eine Diskussion der Forschung bei Kappelhoff und Bakels (2011) und Greifenstein (2020).

5 Zur Unterscheidung und Ausrichtung der verschiedenen Ansätze, vgl. Gehrau (2008). Ebenfalls verweise ich auf die Zeitschriftenreihe »Rezeptionsforschung« (2004- heute) des Nomos Verlags.

6 Die theoretischen Ansätze und methodischen Vorgehensweisen wurden maßgeblich in der Kolleg-Forschungsgruppe »Cinepoetics. Poetologien audiovisueller Bilder« entwickelt. Sie sind in zahlreichen Publikationen in den von De Gruyter publizierten Schriftenreihen (Cinepoetics (2016-heute), Cinepoetics – English Edition (2018-heute), Cinepoetics Essay (2020-heute) sowie das Online-Journal mediaesthetics (2016-heute); alle zu finden unter <https://www.cinepoetics.fu-berlin.de/publications/index.html> (08.07.2024) dokumentiert.

chens in Interaktion mit der audiovisuellen Multimodalität⁷ zu einer sich zeitlich entfaltenden Bedeutungsgestalt – einem Erfahrungsraum. Diese Gestalt, dieser Erfahrungsraum *ist* die mediale Wirklichkeit, die sich im Moment des Sehens im Körper der Zuschauer*innen entfaltet.

3. Audiovisuelle Modulation einer Gesprächserfahrung

3.1 Daten und methodische Vorgehensweise

Im Folgenden analysiere ich einen Videobeitrag (5:28 Min), welcher im Rahmen der DBDK der Partei BÜNDNIS 90/DIE GRÜNEN 2020 ausgestrahlt wurde.⁸ Er gehört zu einer Reihe von Einspielern, die unter dem Titel »Neue Zeiten, neue Antworten« immer zwei Parteimitglieder, hier Claudia Roth und Georg Kurz, im Gespräch zeigen. Das Gespräch fand in der vorfilmischen Realität als Videokonferenz statt, wird im Einspieler jedoch in einer Split-Screen-Montage in editierter Form zusammengeführt, wodurch eine Art ›Gespräch‹ von AzA inszeniert wird. Die audiovisuelle Inszenierung tritt somit deutlich als modulierendes Element hervor.

Die Analyse geht deshalb von der Audiovisualität des Materials aus und versucht über die Bewegungsdynamik des sich entfaltenden Bewegungsbildes herauszuarbeiten, wie es zur Erfahrung eines ›Gesprächs‹ kommt, wie diese moduliert ist und an welchen Stellen der Erfahrungsraum durch Irritationsmomente hinterfragbar wird. Methodisch werden sowohl medienästhetische als auch gesprächs- bzw. interaktionsanalytische Ansätze zusammengeführt. Die entsprechenden Bezüge werden in den einzelnen Abschnitten erläutert. Die Analyse folgt den folgenden Schritten: Zunächst (3.2) wird das Video in Bezug auf seine audiovisuelle Inszenierung als ein vereinfachtes Einstellungsprotokoll (Hickethier 2012: 37f.) dargestellt. Weiter (3.3) wird die Dynamik des sich entfaltenden Bewegungsbildes beschrieben (vgl. Greifenstein 2020; Müller/Kappelhoff 2018; Kindler-Mathôt et al. im Erscheinen). Der daraus hervorgehende Rhythmus wird dann in 3.4-3.5 mit Erkenntnissen der Gesprächs- und Interaktionsanalyse in Verbindung gebracht (vgl. Auer 2020; Kendon

7 Die Begrifflichkeiten »Multimodalität des Sprechens« und »audiovisuelle Multimodalität« werden in Müller (im Erscheinen) und Kindler-Mathôt et al. (im Erscheinen) entwickelt: »Auf der einen Seite ist da die Multimodalität des Sprechens in der Interaktion von Angesicht zu Angesicht, die sich als dynamisches Zusammenspiel von Blick, Hand- und Körpergesten mit der gesprochenen Äußerung darstellt. Auf der anderen Seite ist da die audiovisuelle Multimodalität, d.h. die audiovisuelle Inszenierung des Sprachgebrauchs, die sich in der Orchestrierung von Kameraführung, Cadrage, Einstellungen, Montage und Ton realisiert.« (Kindler-Mathôt et al. im Erscheinen: Kap.2.3).

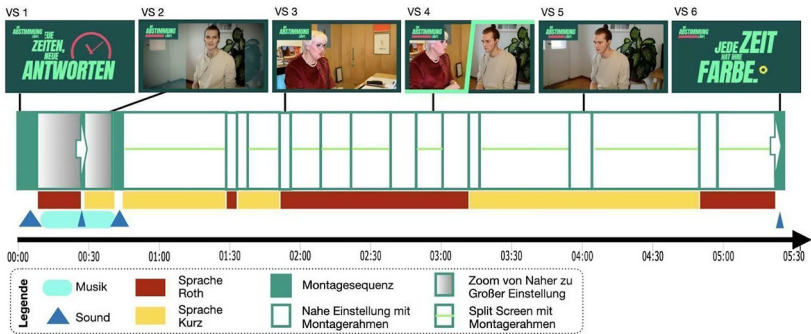
8 Abrufbar auf dem YouTube Kanal von BÜNDNIS 90/DIE GRÜNEN unter: <https://www.youtube.com/watch?v=WpofTzwIjT8> (27.02.2024).

1990; Müller/Bohle 2007; Sacks et al. 1974; Selting et al. 2009). Abschnitt 3.6 bringt die Ergebnisse zusammen und beschreibt damit den Erfahrungsraum ›Gespräch‹ als komplexen Aushandlungsprozess von wechselseitigen Positionierungen.

3.2 Einstellungsprotokoll eines ›Gesprächs‹

Abbildung 1 zeigt schematisch die audiovisuelle Inszenierung des Videobeitrags als ein vereinfachtes Einstellungsprotokoll im zeitlichen Verlauf. Dies erlaubt einen ersten Überblick über die zeitliche Entfaltung der Inszenierung. Ich orientiere mich hier an Hickethier (2012: 37f.), eine ähnliche Vorgehensweise ist ebenfalls bei Kinder-Mathôt et al. (im Erscheinen) geschildert.

Abb. 1: Einstellungsprotokoll des Videobeitrags (05:28 Min).



Oberhalb des Zeitstrahls ist die auditive Ebene annotiert. Neben Musik (türkis) und anderen Soundeffekten (blau) umfasst dies auch die verbalen Äußerungen (rot für Roth, gelb für Kurz). Darüber liegt die visuelle Inszenierung. Sie stellt eine Abfolge von Montagesequenzen (dunkelgrün) und nahen Kameraeinstellungen (weiß) dar, die von Schnitten getrennt werden (senkrechte Linien). Einige Einstellungen sind exemplarisch als Videostills (in der Abbildung als VS gekennzeichnet) abgebildet, die auf die Bildraumkomposition verweisen. Die gesamte Sequenz wird von einem dunkelgrünen Montagerahmen umfasst (vgl. VS 3 und 4), ebenfalls wird im linken oberen Bildrand durchgehend ein kleines Fenster mit der Inschrift »Die Abstimmung läuft« und einem Prozessbalken eingeblendet.

Der Beitrag folgt in seiner Inszenierung einem bestimmten Muster. Er beginnt mit einer ins Bild fliegenden Einblendung eines dunkelgrünen Hintergrunds, die den Titel der Reihe zeigt (VS 1). Daraufhin folgen zwei frontale nahe Einstellungen, die die einzelnen Sprecher*innen im Bild zeigen (VS 2). Der Bildraum ist teilweise transparent-dunkelgrün kaschiert (vgl. Hickethier 2012: 48), sodass die Personen

im Bild in einem hellen Quadrat hervorgehoben sind. Mit in die Kamera gerichteten Blick stellen Roth und Kurz sich vor, während sich die Einstellung in einem langsamen Zoom bis fast in eine Großaufnahme verkleinert. In Sekunde 40 endet die Vorstellung mit einer Abblende, die das Thema des Einspielers (»Selbstbestimmung«) zeigt. Nun beginnt das »Gespräch«, welches als eine Abfolge von nahen Einstellungen auf Roth (VS 3) und Kurz (VS 5) inszeniert wird. Dominant ist dabei die Split Screen Montage (hellgrüne Linie, Bsp. VS 4), in der das Bild mittig in zwei Felder (rechts-links) aufgeteilt ist und somit beide Personen im Bildraum zeigt. Das »Gespräch« ist durch insgesamt 17 Schnitte bei stabiler Kameraeinstellung und -perspektive unterteilt. Der Beitrag endet mit einer letzten Abblende mit dem Motto der DBDK (»Jede Zeit hat ihre Farbe«).

3.3 Bewegungsbild »Gespräch«

Das Einstellungsprotokoll zeigt, dass sich die Inszenierung des »Gesprächs« deutlich von der vorangehenden »Vorstellung« unterscheidet: Die Montage der dunkelgrünen Hintergründe, die damit einhergehenden Soundeffekte sowie der frontale Zoom auf die Personen, stehen im Gegensatz zur Split Screen Montage und den vermeintlich zueinander orientierten Körpern rechts und links im Bildraum. Das »Gespräch« entfaltet dabei eine andere Dynamik, ein eigenes Bewegungsbild, welches im weiteren Verlauf der Analyse im Vordergrund steht. Wie kann das Bewegungsbild beschrieben werden, ohne von der Ebene der narrativen Repräsentation, also dem Gespräch als solchem auszugehen? Ich beziehe mich hier auf die Herangehensweise von Müller und Kappelhoff (2018) sowie Greifenstein (2020) und beschreibe die distinkte Bewegungsdynamik und Qualität des Bewegungsbildes. Greifenstein hat hierfür, in Anlehnung an Daniel Sterns (2011) Formen der Vitalität, ein spezifisches Beschreibungsvokabular entwickelt, mit dem die spezifisch-rhythmischen Bewegungsqualitäten verdeutlicht werden. Greifenstein formuliert exemplarisch:

Methodisch-deskriptiv fokussiere ich entsprechend Bewegungsverben (oder ihre Substantivierungen) für die gesamte Verlaufsform (z.B. das Klopfen) und Adjektive zur Präzision der sinnlichen Eigenschaften dieser Bewegung (das sanfte Klopfen, das eilig hastige Klopfen, das harte-energische Klopfen). (Greifenstein 2020: 69)

Mit Blick auf das Bewegungsvokabular kann die Dynamik des Bewegungsbildes des oben betrachteten Gesprächs als ein züiges, geschmeidiges »Hin und Her« beschrieben werden, welches in seinem Rhythmus zunächst gleichmäßig fließend die Seiten auf einer horizontalen Ebene wechselt, dann aber ruckhafter wird und mehr und mehr einseitig dominiert. Das Bild eines Ping-Pong Spiels hilft, die spezifische Bewegung zu verstehen: Ein zügiger Ballwechsel lenkt die Aufmerksamkeit

alternierend von rechts nach links. Dabei wird der Ball jedoch nicht jedes Mal zurückgespielt, sondern es kommt zum neuen Aufschlag. Genauer: Während Kurz auf der rechten Seite den Ball annimmt und zurückspielt, führt Roth konsequent immer neue Aufschläge aus. Die Bewegung ist folglich zwar wechselseitig, aber dennoch einseitig dominierend.

Um diese deskriptive Ebene empirisch zu begründen, werfe ich in den folgenden beiden Abschnitten einen genauen Blick auf die Orchestrierung von audiovisueller Inszenierung und sprachlicher Multimodalität. Dabei werden das Zusammenspiel von Schnittmontage und Gesprächsstruktur sowie die Bildraumkomposition mit Blick auf die körperliche Koordination eines geteilten Interaktionsraums untersucht. Der Rückgriff auf gesprächsanalytische Forschung bietet dabei auch Erklärungsansätze für das sich entfaltende Dominanzgefühl und warum dieses aus Zuschauer*innenperspektive als Irritation in der Gesprächsdynamik sowie der Gesprächspraktik erfahren wird.

3.4 Rhythmus und Gesprächsstruktur

Das Zusammenspiel von filmischer Inszenierung und Gesprächsstruktur zeigt sich als starke Synchronisation: Jeder Sprecher*innenwechsel geht mit einem Schnitt einher (s. Abb.1), welcher immer den*die Sprecher*in in das Zentrum der Aufmerksamkeit rückt, da der Beginn eines neuen Turns nie im Split Screen, sondern immer in der Einzelansicht inszeniert ist. Das wahrgenommene Turn-Taking ist damit ein Produkt der Inszenierung. Das heißt, es ist kein Abbild der vorfilmischen Realität, sondern ein Teil der medialen Wirklichkeit und damit des Erfahrungsraums der Zuschauer*innen. Dennoch können die abwechselnden Turns vor dem Hintergrund der Turn-Taking Systematik von Sacks et al. (1974) betrachtet werden. Denn, wie im Falle einer dritten Person, die einem Gespräch von Aza zuhört, haben auch die Zuschauer*innen am Bildschirm ein Gefühl dafür, ob ein Gespräch ›rund‹ läuft oder Inkongruenzen auftreten.

Es wird zwischen zwei Möglichkeiten des Turn-Takings unterschieden (vgl. ebd.: 703): Erstens der *Fremdwahl* (ein*e Sprecher*in wählt den*die nächste*n Sprecher*in z.B. durch direkte Anrede oder eine Geste⁹) und zweitens der *Selbstwahl* (ein*e Sprecher*in wählt sich selbst als nächste*n). Kommt es zu keiner der Möglichkeiten, kann der*die vorherige Sprecher*in den Turn weiterführen (*current speaker continues*; Auer 2020: 165) oder das Gespräch unterbricht, bzw. endet. Die Übergänge zwischen Turns sind in der Regel fließend, ohne große Pausen oder simultanes Sprechen (vgl. Sacks et al. 1974: 700f.). Dies setzt voraus, dass sich die

9 Henne und Rehbock verweisen bereits 1979 auf die Relevanz der Multimodalität, insbesondere in Bezug auf Strategien des Turn-Takings. Zum multimodalen Turn-Taking, vgl. auch Schmitt (2005).

Sprecher*innen miteinander koordinieren und den nächstmöglichen Moment für einen Übergang, einen sogenannten *transition-relevance place* (TRP; ebd.: 703) zum einen antizipieren und zum anderen ihren nächsten Turn anschlussfähig vorbereiten: Die Interaktionsteilnehmer*innen »verarbeiten [...] die gerade entstehende TKE [Turn-Konstruktionseinheit] **während** ihrer Produktion und machen dabei bereits Vorhersagen darüber, wann sie zu einem Abschlusspunkt kommen könnte« (Auer 2020: 115; Herv. i. Orig.). Übergangsstellen gehen insbesondere mit syntaktischen und diskursiven Strukturen einher (vgl. ebd.: 117ff.) und sind meist prosodisch gekennzeichnet. Auch bei der Selbstwahl kündigen Sprecher*innen i.d.R. ihre Turn-Übernahme bereits an und machen sie somit antizipierbar, etwa durch die Veränderung der Körperhaltung in den Interaktionsraum hinein oder verstärkten Rezeptionssignalen (*pre-starters*) wie Nicken oder *hm_hms*.

Betrachten wir diese Aspekte mit der Turn-Taking Erfahrung des Videobeitrags, können zunächst vier Punkte festgehalten werden, auf welche im Folgenden eingegangen wird:

- 1) Insgesamt finden fünf Wechsel statt, die Turns sind dabei gleichmäßig in Anzahl und etwa proportional in Länge verteilt;
- 2) Vier der fünf Wechsel erfolgen durch Selbstwahl (ausgenommen zwischen Turn 2 und 3);
- 3) Trotz unmittelbarer Anschlüsse (*Latching*; Selting et al. 2009: 376) kommt es zu keinen Überlappungen;
- 4) Die meisten der Übergangsstellen sind prosodisch auffällig.

Zur genaueren Betrachtung werden Ausschnitte aus dem Gespräch als gesprächsanalytische Transkripte (nach GAT 2; Selting et al. 2009) aufbereitet.¹⁰

¹⁰ Die Orthografie wird zu Gunsten prosodischer Elemente außer Acht gelassen. So werden Pausen, Intonationsakzente und Phrasenenden mittels Kapitalen und Satzzeichen annotiert.

Abb. 2: Turn-Taking im Gespräch zwischen Roth und Kurz.

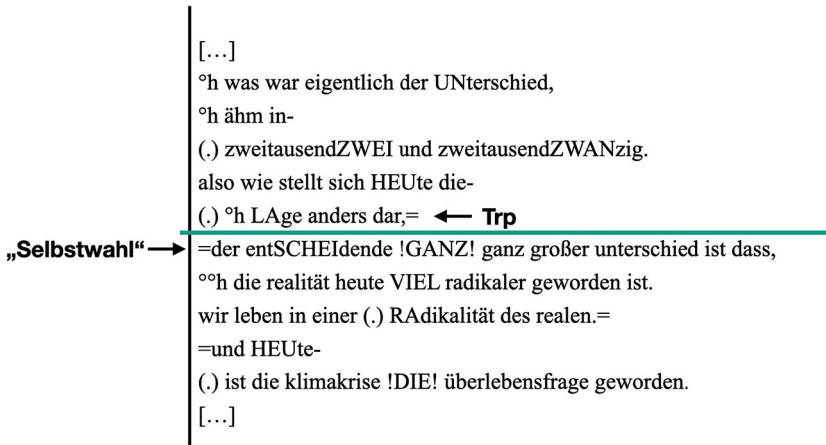
Turn Nr.	Sprecher*in	Zeile	Anfang und Ende des Turns als gesprächsanalytisches Transkript (GAT 2) (Auslassungen durch [...] gekennzeichnet)	Sprecherwechsel
1	GK:	001 002 003 [...] [...] 017 018 019 020 021 022	also ich bin ja so_n hoffnungslöser ÖKO gewesen; ich bin ja wegen so: so KLIma und umwelt sachen- (.) zu den GRÜnen gekommen- [...] [...] da macht man sich nich (.) FREI von- wenn man die part=wenn man die tür zum PARTEIbüro zu macht, das ist ja natürlich KLAR- aber die halt vor allem polItisch dafür kämpfen dass- (.) ALle menschen da- (.) ähm sich frei entfalten KÖNnen=	(Selbstwahl)
2	CR	023 024 025 026	=was BEDEUtet denn- der WERT selbstbestimmung; HEUte- für DICH,=	Selbstwahl
3	GK	027 028 [...] [...] 033 034 035 036	=ich KOMM ja aus nem- (.) nem kleinen (.) KAFF bei berchtesgaden, [...] [...] ähm (.) und für mich war das ne toTALE:- (.) äh beFREIung; irgendwie festzustellen dass das ja EIgentlich- auch ganz ANDers geht;	Fremdwahl (Frage mit Anrede)
4	CR	037 038 039 [...] [...] 072 073 074 075	für MICH war natürlich selbstbestimmung- äh ganz ganz ganz STARK- verKNÜPFT mit dem selbstbestimmungs!RECHT!. [...] [...] =und HEUte- (.) ist die klimakrise !DIE! überlebensfrage geworden. aber was bedeutet !DAS! für selbstbestimmung. was bedeutet das für selbstbestimmung für (.) MEInen LEbensstil.=	Selbstwahl
5	GK	076 077 078 079 [...] [...] 124 125 126 127 128 129 130	=was du GRAD- ähm geSAGT hast zum- (.) ähm selbstbestimmungsrecht der FRAU und so weiter eingehen.= =weil- [...] [...] = GLAUbe ich, !EI!gentlich ein schritt hin zu MEHR freiheit. (.) wenn man NICHT mehr geZWUNGen ist;= =zum beispiel warten bis man VOLLjährig ist;= Bis man FÜHRRerschein hat;= =TAUsende euro ausgibt für ein auto,= =und dann erst irgendwie moBIL zu sein?= [...] [...]	Selbstwahl
6	CR	131 132 133 [...] [...] 141 142 143 144 145 146 147 148	=selbstbestimmung verwirklicht FREIHEIT. (.) !SELBST! (.) bestimmung verwirklicht auch- die !WÜRDE! jedes menschen; [...] [...] wo es gruppen GIBT, die (.) die würde nicht JEDES menschen,= =sondern die WÜRde des DEUTschen menschen- oder des MÄNNlichen menschen- oder des CHRIStlichen menschen- oder des Heterosexuellen- oder des WEIßen, in den VORdergrund stellen.	Selbstwahl

In der Zuschauer*innenerfahrung wechseln sich Kurz und Roth in ihren Turns ab. Alle Wechsel (ausgenommen zwischen Turn 2 und 3) finden durch Selbstwahl statt. Die Fremdwahl entsteht aus einer Frage-Antwort Sequenz: Roth fragt, was Selbstbestimmung für Kurz heute bedeute. Syntaktisch und prosodisch mit leicht steigender Intonation zum Phrasenende (gekennzeichnet durch ein Komma), stellt der Turn eine abgeschlossene Einheit mit Handlungsaufforderung an Kurz dar. Entsprechend folgt ein Turn von Kurz mit einer inhaltlich passenden Antwort und womit eine Form von Intersubjektivität hergestellt wird: Aus Zuschauer*innenperspektive erscheint es, als ob Kurz den Turn als Fremdwahl verstanden und den TRP entsprechend antizipiert hat. Es kommt zu einem Latching, also einem unmittelbaren Anschluss der Antwort (gekennzeichnet durch Gleichheitszeichen).

Latching stellt innerhalb des ›Gesprächs‹ die Regel der Turn-Anschlüsse dar (Ausnahme zwischen Turn 3–4). Damit gehen für TRPs ungewöhnliche Tonhöhenbewegungen zum Ende der Intonationsphrasen einher, da sie gleichbleibende (gekennzeichnet als Bindestrich) oder steigende (Komma für leicht steigende, Fragezeichen für stark steigende Intonation) Tendenzen aufweisen, z. B. bei *KÖNnen* (022) oder *moBIL zu sein?* (130). Auer (2020: 153f.) bemerkt, dass solche Formen in Gesprächen der finalen Tonhöhenbewegung als »weiterweisend« gelten, also auf Unabgeschlossenheit des Turns verweisen. Die Inszenierung, genauer die Schnitte, die die Erfahrung von TRPs modulieren, führen also zu ersten feinen Irritationsmomenten, welche prägend für die Dynamik des Bewegungsbildes sind: Die schnellen kontinuierlichen Hin- und Herbewegungen zwischen den Turns, welche durch die damit einhergehenden Schnitte und die Wechsel von Split Screen zu Einzelansicht, zum einen akzentuieren, zum anderen auch eine fast rastlose Zügigkeit entfalten.

Mit Blick auf die TRPs tritt noch eine weitere Stelle hervor: In Turn 7 kommt es zu einem TRP, welcher durch erneute Selbstwahl von Roth, also einer Weiterführung, nicht zum Sprecher*innenwechsel führt. Das folgende Transkript verdeutlicht die Stelle.

Abb. 3: Ausschnitt aus Turn Nr. 4 (CR) mit ungenutztem TRP.



Wieso ist diese Stelle auffällig? Der TRP resultiert aus einer syntaktischen Fragekonstruktion mit entsprechend steigender Tonhöhenbewegung zum Phrasenende: [...] *also wie stellt sich HEUTE die- (.) °h LAge anders dar*, Er stellt damit, im Gegensatz zu den zuvor beschriebenen Äußerungen, eine abgeschlossene Turn-Konstruktion dar. Darauf beantwortet Roth die Frage selbst, indem sie direkten lexikalischen Bezug zum ersten Teil des Turns nimmt (*°h was war eigentlich der UNterschied, [...] =der entSCHEidende !GANZ! ganz großer unterschied ist dass, [...]*). Auch wenn das Selbstbeantworten von Fragen durchaus in Gesprächen vorkommt, ist an dieser Stelle die sprachliche Markierung einer Übergangsstelle durch die Phrasenintonation (insbesondere mit Blick auf die sonst auftretenden Unregelmäßigkeiten) recht markant. Ebenfalls gibt es keinen Hinweis darauf, dass es sich bei der Frage um eine rhetorische Frage handelt, da sie weder prosodisch noch metakommunikativ als solche markiert wird.

Die untypischen Formen der TRPs durch finale Tonhöhenbewegung oder Latching ohne Überlappung sind Produkte der Schnittmontage, die bspw. den Turn von Roth als Beantwortung der eigenen Frage gestaltet (in Transkript 2 als horizontale Linie). Sie stellen Irritationsmomente dar, die ausschließlich auf der Ebene der Zuschauer*innenerfahrung bestehen und kein Phänomen der vorfilmischen Realität sind. Sie sind Formen der von Holly (2015) angesprochenen Transkription, die zu einer Bedeutungsanreicherung führt. Im Falle der selbstbeantworteten Frage besteht dies in einer Irritation, die nicht die grundsätzliche Dynamik des Gesprächs unterbricht, aber dennoch als vielleicht sogar etwas unhöflich auffällt. Holly fügt in diesem Zusammenhang an, dass die Medialität damit dem*der Sprecher*in die alleinige Auktorialität entzieht, da die »performative Letztfassung« seiner Rede nicht mehr kontrolliert« (Holly 2015: 89) werden kann.

In Bezug auf den Rhythmus des Bewegungsbildes im Zusammenspiel mit der Gesprächsstruktur kann also festgehalten werden, dass die Erfahrung der Hin- und Her Dynamik eng mit der audiovisuellen Inszenierung des Turn-Takings verknüpft ist. In der Zuschauer*innenerfahrung wird, trotz der Schnitte und der anhaltenden Irritationsmomente durch die ungewöhnlich intonierten TRPs, ein Aushandlungsprozess wie in einem ›natürlichen‹ Gespräch von AzA deutlich. Die Videokonferenz wird – zumindest mit Blick auf das Turn-Taking – in der Erfahrung nicht relevant gesetzt, da keine hierfür typischen Phänomene wie lange Pausen oder Simultansprechen auftreten. In dieser Hinsicht ist das Turn-Taking kein Produkt der Repräsentation, sondern wird vielmehr im Rezipieren der Zuschauer*innen körperlich-affektiv erfahren und unmittelbar verstanden: nicht durch die ›Darstellung‹, sondern als eine spezifische Bewegungsdynamik, die die Wechselseitigkeit einer Gesprächserfahrung von AzA aufgreift, wenn auch in ungewöhnlicher Weise.

Ein weiterer Aspekt, der diese Erfahrung hervorhebt, zeigt sich in der Gestaltung des Bildraumes, insbesondere durch die bereits benannte Split Screen Montage. Die folgende Analyse zeigt, wie in der Montage die beiden Körper in einer für das Gespräch von AzA geläufigen Ausrichtung in der Zuschauer*innenerfahrung einen gemeinsamen Interaktionsraum zu etablieren scheinen.

3.5 Bildraumkomposition und Interaktionsraum

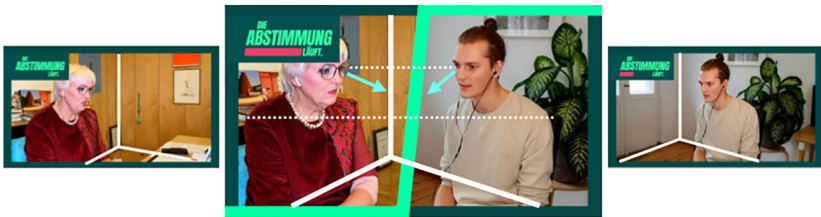
Die Aufteilung des Bildschirms in zwei unterschiedlich gestaltete Flächen sowie deren jeweilige Bildraumkomposition ist ein deutlich zu bestimmender Fall von medialer Wirklichkeit, denn die Perspektive auf zwei unterschiedliche Orte gleichzeitig ist in der vorfilmischen Realität schlicht unmöglich (s. Abb. 4). Der Split Screen bringt dabei zwei kontrastierende Aspekte zum Vorschein: Zum einen hebt die prägnante hellgrüne Linie quer durch die Bildmitte eine klare Trennung hervor, die ebenfalls mit Blick auf die einzelnen Flächen und deren Räumlichkeiten offensichtlich macht, dass sich die Personen an unterschiedlichen Orten befinden. Zum anderen schafft er dennoch eine Verbindung zwischen Personen und Räumen, die in der Zuschauer*innenerfahrung einen geteilten Interaktionsraum formen. Im Folgenden wird, diesmal mit Bezug auf interaktionsanalytische Forschung, die Konstitution dieses Raumes erläutert und damit ein weiteres Argument für die Erfahrung eines audiovisuell orchestrierten Gesprächs von AzA angeführt, da damit die zweite konstituierende mediale Bedingung inkludiert wird: Kopräsenz durch Anwesenheit.

In einem Gespräch von AzA orientieren sich die Teilnehmer*innen nicht nur sprachlich aneinander, sondern bilden in der Ausrichtung ihrer Körper einen geteilten Interaktionsraum. Müller und Bohle (2007) bezeichnen diese körperliche Koordination als das *Fundament fokussierter Interaktion*. Kendon (1990) beschreibt in seiner Arbeit zu *Spatial organization in social encounters* äußerst differenziert verschie-

dene wiederkehrende Konfigurationen: Für das Gespräch zwischen zwei Personen sind dies bspw. *vis-a-vis arrangement* (Personen stehen sich gegenüber), *L-arrangement* (Personen stehen in einem Winkel nebeneinander) oder *side-by-side arrangement* (Personen stehen nebeneinander) (vgl. ebd.: 213). Er prägt in diesem Zusammenhang den Begriff der *F-formation systems*. Diese Systeme treten immer dann auf, wenn zwei oder mehr Personen einen gemeinsamen Handlungsraum (*o-space*; ebd.: 211) zwischen sich etablieren und diesen kooperativ durch die Organisation ihres körperlichen Verhaltens aufrechterhalten. Eine F-Formation ist keine Bedingung für ein Gespräch, sie ist jedoch präferiert (vgl. Auer 2017: 10).

Wie ist der Interaktionsraum im Bildraum nun gestaltet? Mit Blick auf die zwei Bildflächen zeigen sich im Hintergrund zwei verschiedene büroartige Räume, die sich in ihrer Farbgebung abheben. Roth und Kurz sitzen jeweils aus den Außenseiten der Felder mit dem Rücken zu einer Wand an einem Tisch (bei Roth je nach Einstellung) sichtbar. Beide tragen kleine Kopfhörer, die ein Hinweis auf die Videotelefonie sind. Die Inszenierung hebt dies aber nicht weiter hervor, nur in wenigen Einstellungen ist das Endgerät von Roth im rechten Bildrand zu sehen. Die starke örtliche Trennung wird mithilfe einer Split Screen Montage überbrückt, sodass visuell ein Miteinander inszeniert wird (s. Abb. 4).

Abb. 4: Audiovisuelle Überwindung der räumlichen Trennung zu Gunsten eines geteilten Interaktionsraums.



Zunächst sind beide Personen so im Bildraum ausgerichtet, dass es im Zusammenschnitt im Split Screen erscheint, als säßen sie sich in einem offenen Winkel gegenüber: Beide nehmen jeweils etwa gleich viel Raum ein, Augen und Schultern sind auf gleicher Höhe im Bildraum gesetzt, beide sind in sitzender Position halbfrontal von der Kamera aufgenommen. In Verlängerung formen ihre Schultern deshalb eine L-Konfiguration, die in Richtung der Zuschauer*innen (also der Kameras) geöffnet ist. Verlängert werden die Achsen des Ls zusätzlich durch die Kameraperspektive in den jeweiligen Raum, die den Blick in eine Ecke freigeben. Im Split Screen sind beide Bildfelder so nebeneinander gesetzt, dass nicht nur die Körper auf gleicher Höhe ausgerichtet sind, sondern auch die Sockelleisten an den Wänden ineinander übergehen und so die beiden Räume visuell zu einem verbunden werden (in

der Abbildung als weiße durchgezogene Linie gekennzeichnet). Es entsteht so nicht nur ein L-förmiger Interaktionsraum zwischen den Körpern mit Innen und Außen, sondern auch ein gemeinsamer Raum mit zwei Wänden. Durch die Interaktion zwischen den beiden Teilbildern des Split Screens wird eine Erfahrung von Parallelität und Gleichzeitigkeit evoziert (vgl. Wulff 1991: 282), welche im *Mise-en-Scène* zum Eindruck von gemeinsamer Anwesenheit und Kopräsenz verbunden wird.

Der scheinbar geteilte Raum und die körperliche Ausrichtung zueinander sind folglich ein Produkt der Inszenierung. Die Modulation stellt einen wortwörtlich un/realen Interaktionsraum dar: unreal in der vorfilmischen Realität, real in der Erfahrung der Zuschauer*innen. Doch auch in der Konstruktion des Interaktionsraumes ergibt sich ein weiteres Irritationsmoment, welches im Blickverhalten besteht. Die aus der Abbildung 4 hervorgehende Blickrichtung auf Schulter- und Brustbereich (türkise Pfeile) würde in einem natürlichen Gespräch von AzA als unangebracht empfunden werden. Im Gegensatz zur oben beschriebenen Irritation beim Turn-Taking tritt der Blick jedoch nicht so stark hervor, da die L-Konfiguration und Ausrichtung der Körper, vor allem der gleichen Höhe der Augen und Schultern, dominanter ist. Es lässt sich zwar folglich eine intramediale Umdeutung nach Holly (2015) erahnen – der Blick auf Brusthöhe wird nur durch die spezifische mediale Konfiguration möglich – sie tritt jedoch nicht prominent hervor, sondern schmiegte sich eher in die modulierte Kopräsenz ein.

3.6. Multidimensionaler Erfahrungsraum ›Gespräch‹

Sowohl die Analyse des Rhythmus als auch der Bildraumkomposition aktivieren die Erfahrung eines Gesprächs von AzA, indem dessen mediale Bedingungen, Wechselseitigkeit und Kopräsenz durch Anwesenheit, audiovisuell konstruiert und/oder hervorgehoben werden und dadurch zu einem affektiven Zuschauer*innenerleben werden. Mit Blick auf das in 3.3 beschriebene Bewegungsbild bieten beide Ebenen der Analyse bereits empirische Anhaltspunkte, die sich vordergründig auf die medialen Bedingungen als eine Art ›Gerüst‹ des Gesprächs beziehen. In diesem abschließenden Analyseteil möchte ich zusätzlich noch auf den Inhalt, bzw. die pragmatische Ebene eingehen, in dem ich mich auf den Bedeutungsaushandlungsprozess innerhalb dieses Gerüsts beziehe. Damit wird neben den Rhythmen und Bewegungsrichtungen auch die Qualität und zunehmende Intensität des Bewegungsbildes adressiert und somit auch die spezifisch affektive Dimension des Erfahrungsraums. Abbildung 5 greift dazu das Bild des Ping-Pong Spiels auf und führt die Ebenen der Analyse zusammen.

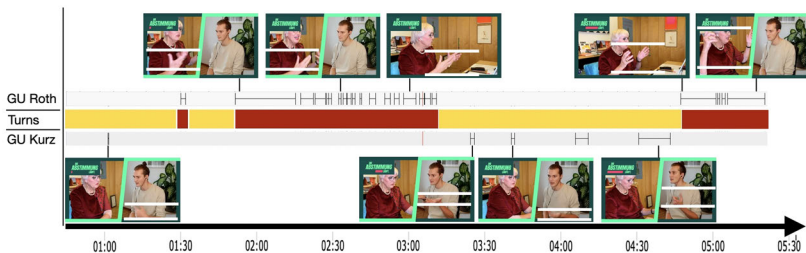
Abb. 5: Modulation einer einseitig dominierenden Hin und Her Bewegung.



Untereinander stehen jeweils Videostills der mit dem Turn-Taking einhergehenden Einstellungen. Pfeile in rot (Roth) und gelb (Kurz) visualisieren die jeweiligen Turns, also die Ballwechsel. Eine orange Verbindungslinie tritt dann auf, wenn ein Ball weitergespielt und nicht fallen gelassen wird (vgl. weiter unten: dynamische Anschlussfähigkeit). Ausschnitte des Transkripts verdeutlichen die Sprachhandlungen, die als Überschriften neben den nummerierten Turns aufgeführt sind. Dabei treten drei Aspekte zu Tage, die die Bewegungsqualität greifbar machen.

1.: Die *Fokussierung der Aufmerksamkeit ist wechselseitig* zwischen den jeweils sprechenden Personen. Während eines Turns ist mehr Bewegung auf entweder der rechten oder linken Seite des Bildraums erfahrbar. Diese ist audiovisuell, da sie zum einen durch sprachliche, hörbare »Bewegung«, zum anderen durch sichtbare Bewegung im Bild aktiviert wird. Dabei rücken Handgesten in den Fokus, die auf Grund der Kameraeinstellung und Perspektive gut sichtbar sind. Annotiert wurden die gestischen Einheiten (GU), die von Ruhepositionen der Hände abgegrenzt werden. Ohne auf eine detaillierte Gestenanalyse (zur Methodologie vgl. Müller 2024) einzugehen, wird deutlich, dass Roth mehr und tendenziell größere bzw. raumeinnehmendere Handgesten ausführt als Kurz. Zum Ende des Gesprächs nehmen sie außerdem an Intensität zu, sowohl in der Bewegungsqualität (ruckhafte, akzentuierte Schläge) als auch in der Größe (Armbewegungen durch den gesamten Bildraum, in der Abbildung markiert durch weiße Linien). Durch den Split Screen und auch die Einzeldarstellung der Personen auf der rechten oder der linken Seite im Bild wird die Aufmerksamkeit zusätzlich gelenkt. Der Blick wandert so je nach Sprecher*in von rechts nach links.

Abb. 6: Annotation der Gesteneinheiten (GU) im zeitlichen Verlauf des »Gesprächs« mit exemplarischen Stills.



2.: Das bereits beschriebene audiovisuell induzierte Turn-Taking erfolgt zwar zügig, dennoch sind die Übergänge *nicht durchgehend dynamisch anschlussfähig*. So nimmt nur Kurz zweimal den vorangehenden Turn dynamisch auf und führt ihn hin weiter, einmal als Erfüllung der Paar-Sequenz (s. Abb. 5, Turn 3) und einmal mittels eines metakommunikativen Kommentars zur vorangehenden Äußerung (s. Abb. 5, Turn 5). Solche expliziten Anknüpfungen finden bei Roth nicht statt. Sie bildet immer neue, unabhängige Einheiten, die, abgesehen von der Orientierung am Thema »Selbstbestimmung«, nicht in direktem Bezug zum vorherigen Turn stehen.

3.: Thematisch kann ein *Aushandlungsprozess* rund um das Thema »Selbstbestimmung« verfolgt werden. Die Turns stellen, vor allem ab Turn 3, Positionierungshandlungen dar. Bei Roth findet eine Steigung dieser Positionierung statt,

die in Turn 6 ihren Höhepunkt findet, welcher in prosodisch (und gestisch) beinah überbetonten, kurzen Aussagephrasen als eine Art Abschlussplädoyer formuliert wird (=selbstbestimmung verwirklicht FREIHEIT. (.) !SELBST! (.) bestimmung verwirklicht auch- die !WÜRDE! jedes menschen; [...]). Daraufhin endet das Gespräch sofort, da sich qua Montage ein grünes Bild einschleibt. Roth behält damit das letzte Wort. Im Vergleich dazu, und mit Blick auf die dynamische Anschlussfähigkeit der Turns, positioniert sich auch Kurz zunächst in Turn 3, schwächt dies jedoch in Bezug auf Roths darauf folgenden Turn (Nr. 4) deutlich ab (*das möchte ich vielleicht da noch vorne ANgestellt haben; das darf man nicht verWECHseln;=wenn ich mich da ein bisschen eingeschränkt fühle ist das was völlig anderes als wenn man dafür kämpfen muss dass zum beispiel paraGRAphen abgeschafft werden.*). Während Roth somit konsequent einen Standpunkt formuliert, von dem sie nicht abweicht, sondern diesen zum Ende hin noch deutlich prosodisch hervorhebt, ordnet sich Kurz im Verlauf diesem unter.

Das sich im Verlauf steigernde Dominanzgefühl der linken Bildseite wird folglich durch dreierlei Aspekte moduliert:

- 1) durch die wiederholte Selbstwahl ohne dynamischen Anschluss sowie durch die Selbstbeantwortung der eigenen Frage, die als Irritationsmoment markiert wurde;
- 2) durch die Intensivierung der Bewegungserfahrung (sprachlich und gestisch), die die Aufmerksamkeit fokussiert;
- 3) durch die Verstärkung der sprachlichen Positionierungen in Bezug auf das Thema des Gesprächs.

Alle drei Aspekte modulieren den Erfahrungsraum. Die Verschmelzung von sprachlicher und audiovisueller Multimodalität (vgl. Müller im Erscheinen; Kindler-Mathôt im Erscheinen) formt das Bewegungsbild, und damit die Zuschauer*innenerfahrung, eine unmittelbar affektive Erfahrung von Dominanz als gefühlter Ausdrucksqualität. Es entfaltet sich eine mediale Wirklichkeit, die nicht der vorfilmischen Realität entspricht, über die, ausgehend vom Material, nur wenige Aussagen getroffen werden können. Der einzige Anhaltspunkt hierfür ist die Videokonferenz, die aber nur am Rande Teil des Erfahrungsraums wird, sowie die Irritationsmomente, die diesen immer wieder erschüttern und so, zumindest für kurze Momente, die Medialität hinterfragbar machen.

4. Diskussion

Welche Rolle spielt nun die Zuschauer*innenerfahrung medialer Wirklichkeiten vor dem Hintergrund zeitgenössischen Medienkonsums? Welche Implikationen gehen damit einher? Eingangs wurde bereits die zunehmende Digitalisierung

angesprochen, aus welcher immer neue mediale Formen hervorgehen. Seien es politische Reden im Bundestagsfernsehen, positive Affirmationen von Instagram Influencer*innen oder populistische YouTube-Shorts: Im Konsum werden diese alle zu Zuschauer*innenerfahrung neuer, medialer Wirklichkeiten. Eingebettet in spezifische Interfaces und thematische Kontexte gehen sie immer mit einer grundsätzlichen Affizierung der Zuschauer*innen einher. Hier schließt sich die Diskussion um die Wirkmacht von Medien an, insbesondere mit Blick auf die vermeintliche »Objektivität« (Luginbühl/Schneider 2020: 60) von Massenmedien. Aus medienästhetischer Perspektive zeigt sich, dass es die affektive Grundierung ist, die die affektiven Bewertungen und Haltungen der Zuschauer*innen gegenüber moralisch-politischen Sachverhalten strukturieren. Scherer (2024) bezeichnet dies auch als Formen zeitgenössischer Propaganda. Wenn vor diesem Hintergrund mediale Wirklichkeiten nicht mehr hinterfragbar sind und mit vorfilmischer Realität gleichgesetzt werden, wenn der Abstand verschwindet und wir als Zuschauer*innen der »Illusion« einer scheinbar objektiv beobachtenden Kamera aufsitzen, werden die Potenziale der medialen Wirklichkeiten offenbar. Die Modulation von politischer Bildung, Wissensgenerierung sowie Meinungen und Bewertungen, sind dann abhängig von der Medienkompetenz der Zuschauer*innen – nicht nur in Bezug auf die Überprüfung inhaltlicher Richtigkeit, sondern vor allem mit Blick auf die Fähigkeit, die Medialität zu hinterfragen und etwa Irritationsmomente aufzudecken. Sonst werden Meinungen zu Fakten und Deep Fakes zu Realität. Es bleibt also nicht bei einem einfachen Übertragungsprozess, sondern im Konsum werden neue soziale Ordnung und Realitäten, neue Praktiken im Sinne einer Poesis des Filme-Sehens hervorgebracht (vgl. Kappelhoff 2018). McLuhan und Fiore bringen dies mit Blick auf die historische Entwicklung im Zusammenspiel von Medien und Gesellschaft auf den Punkt:

Gesellschaften wurden schon immer stärker von der Natur der Medien geprägt, mit denen Menschen kommunizieren, als vom Inhalt der Kommunikation. [...] Ohne Kenntnis der Wirkungsweise von Medien ist es unmöglich, Prozesse des sozialen und kulturellen Wandels zu verstehen. (2016: 8)

Als Praktik des Medienkonsums stehen die audiovisuellen Bilder damit im direkten Zusammenhang mit dem (demokratischen) Willensbildungsprozess. Vor diesem Hintergrund plädiert der Beitrag für eine Stärkung der Medienkompetenz, die insbesondere die charakteristische Medienspezifik audiovisueller Bilder als Bewegungsbilder vor dem Hintergrund von Verstehensprozessen berücksichtigt.

5. Schluss

Der Artikel hat aus einer medienästhetischen Perspektive heraus die Modulation medialer Wirklichkeit in der Zuschauer*innenerfahrung audiovisueller Bilder untersucht. Dabei wurde betont, dass diese Wirklichkeit weder bloßes Abbild der vorfilmischen Realität ist, noch sich auf der Ebene eines repräsentierten Narrativs entfaltet, sondern vielmehr als Bewegungsbilder mit spezifischen Dynamiken aus dem Zusammenspiel von sprachlicher und audiovisueller Multimodalität hervorgeht. Im Moment des Sehens und Hörens werden sie zu affektiv grundierten Bedeutungsgestalten, die nicht kognitiv dekodiert, sondern unmittelbar körperlich erfahrbar werden. Dies wurde am Beispiel eines Videobeitrags der digitalen Bundesdelegiertenkonferenz der Partei BÜNDNIS 90/DIE GRÜNEN 2020 illustriert. Das hier erfahrene Gespräch ist nicht nur durch die Inszenierung geformt, sondern – wie in der Analyse des Turn-Takings gezeigt – geht erst aus ihr hervor.

Obwohl in der vorfilmischen Realität als Videokonferenz stattgefunden, entfaltet sich vielmehr die Dynamik eines Gesprächs von AzA, dessen mediale Bedingungen (Wechselseitigkeit und Kopräsenz durch Anwesenheit) durch die audiovisuelle Inszenierung eines geteilten Interaktionsraums aktiviert werden. Nur in Irritationsmomenten wird die Erfahrung gestört und somit die Medialität für die Zuschauer*innen hinterfragbar. Aus dem Zusammenspiel sprachlicher und audiovisueller Multimodalität geht eine komplexe Gestalt – ein Erfahrungsraum medialer Wirklichkeit – hervor, womit sowohl affektive als auch pragmatische Aushandlungsprozesse in der Zuschauer*innenerfahrung aktiviert und relevant gesetzt werden. Die Erfahrung dieser medialen Wirklichkeiten geht immer mit Prozessen des sozialen und kulturellen Wandels einher, womit die gesellschaftliche Relevanz dieser komplexen Gestalten im Kontext zeitgenössischen Medienkonsums deutlich zu Tage tritt.

Literatur

- Auer, Peter (2017): »Anfang und Ende fokussierter Interaktion: Eine Einführung« in: InLiSt 59, S. 1–56.
- Auer, Peter (2020): »Die Struktur von Redebeiträgen und die Organisation des Sprecherwechsels«, in: Karin Birkner/Peter Auer/Angelika Bauer/Helga Kotthoff (Hg.), Einführung in die Konversationsanalyse, Berlin/Boston: De Gruyter, S. 106–235.
- Bakels, Jan-Hendrik (2017): Audiovisuelle Rhythmen. Filmmusik, Bewegungskomposition und dynamische Affizierung des Zuschauers, Berlin/Boston: De Gruyter.

- Bühler, Karl (1933): *Ausdruckstheorie. Das System an der Geschichte aufgezeigt*, Jena: Gustav Fischer.
- Deleuze, Gilles (1989): *Das Bewegungs-Bild: Kino 1*, Frankfurt a. M.: Suhrkamp.
- Deleuze, Gilles (1990): *Das Zeit-Bild: Kino 2*, Frankfurt a. M.: Suhrkamp.
- Gehrau, Volker (2008): »Rezeptionsforschung«, in: Uwe Sander/Friederike Von Gross/Kai-Uwe Hugger (Hg.), *Handbuch Medienpädagogik*, Wiesbaden: VS Verlag für Sozialwissenschaften, S. 341–345.
- Gehrau, Volker (2017): *Die Beobachtung als Methode in der Kommunikations- und Medienwissenschaft*, 2. Aufl., Stuttgart: utb GmbH.
- Greifenstein, Sarah (2020): *Tempi der Bewegung – Modi des Gefühls. Expressivität, heitere Affekte und die Screwball Comedy*, Berlin/Boston: De Gruyter.
- Haddington, Pentti/Tuire Oittinen (2022): »Interactional Spaces in Stationary, So-bile, Video-Mediated and Virtual Encounters« in: Andreas H. Jucker/Heiko Hausendorf (Hg.), *Pragmatics of Space*, Berlin/Boston: De Gruyter Mouton, S. 317–361.
- Henne, Helmut/Helmut Rehbock (1979): *Einführung in die Gesprächsanalyse*, Berlin/New York: Walter de Gruyter.
- Hickethier, Knut (2012): *Film- und Fernsehanalyse*, 5., aktualisierte und erweiterte Aufl., Stuttgart: J.B. Metzler.
- Holly, Werner (2015): »Sprache, Bild, Text. Visualität und Intermedialität von Sprache«, in: Ludwig Eichinger (Hg.), *Sprachwissenschaft im Fokus. Positionsbestimmungen und Perspektiven*, Berlin/Boston: De Gruyter, S. 71–92.
- Horst, Dorothea (2018): *Meaning-Making and Political Campaign Advertising: A Cognitive-Linguistic and Film-Analytical Perspective on Audiovisual Figurativity*, Berlin/Boston: De Gruyter.
- Kappelhoff, Hermann (2004): *Matrix der Gefühle. Das Kino, das Melodram und das Theater der Empfindsamkeit*, Berlin: Vorwerk 8.
- Kappelhoff, Hermann (2008): »Zuschauergefühl. Die Inszenierung der Empfindung im dunklen Raum des Kinos«, in: Beate Söntgen/Geraldine Spiekermann (Hg.), *Tränen*, München: Wilhelm Funk, S. 195–206.
- Kappelhoff, Hermann (2018): *Kognition und Reflexion. Zur Theorie filmischen Denkens*, Berlin/Boston: De Gruyter.
- Kappelhoff, Hermann/Jan-Hendrik Bakels (2011): »Das Zuschauergefühl. Möglichkeiten qualitativer Medienanalyse«, in: *Zeitschrift für Medienwissenschaft* 5:2, S. 78–96.
- Kappelhoff, Hermann/Cornelia Müller (2011): »Embodied Meaning Construction. Multimodal Metaphor and Expressive Movement in Speech, Gesture, and Feature Film«, in: *Metaphor and the Social World* 1:2, S. 121–153.
- Kendon, Adam (1990): »Spatial Organization in Social Encounters. The F-Formation System«, in: Adam Kendon (Hg.), *Conducting Interaction. Patterns of Behavior in Focused Encounters*, Cambridge: Cambridge University Press, S. 209–238.

- Kindler-Mathôt, Clara/Cornelia Müller/Jana Junge/Katerina Papadopoulou (im Erscheinen): »Sprechen in audiovisuellen Medien: Systematik für eine medienästhetische Analyse multimodaler Kommunikation«, in: Rita Vallentin/Dorothea Horst (Hg.), *Sprache entgrenzen. Beiträge zu einer kulturwissenschaftlichen Linguistik*, Bielefeld: transcript.
- Koch, Gertrud (2003): »Nachstellungen – Film und historischer Moment«, in: Eva Hohenberger/Judith Keilbach (Hg.), *Die Gegenwart der Vergangenheit. Dokumentarfilm, Fernsehen und Geschichte, Vorwerk 8*, S. 216–229.
- Krämer, Sybille (1998): »Über die Kommunikation im Internet. Überlegungen zur telematischen Interaktion«, in: *Netzdiskurs – Das Internet und der Strukturwandel von Kommunikation und Öffentlichkeit 67/97*, S. 11–22.
- Lanwer, Jens Philipp (2019a): »Blended joint-attention in medial-vermittelter Interaktion«, in: Konstanze Marx/Axel Schmidt (Hg.), *Interaktion und Medien. Interaktionsanalytische Zugänge zu medienvermittelter Kommunikation*, Heidelberg: Winter, S. 99–124.
- Lanwer, Jens Philipp (2019b): »Erzählen im virtuellen Interaktionsraum«, in: *Networx 84:1.0*, S. 4–44.
- Luginbühl, Martin (2019): »Mediale Durchformung. Fernsehinteraktion und Fernsehmundlichkeit«, in: Konstanze Marx/Axel Schmidt (Hg.), *Interaktion und Medien. Interaktionsanalytische Zugänge zu medienvermittelter Kommunikation*, Heidelberg: Universitätsverlag Winter, S. 125–146.
- Luginbühl, Martin/Jan Georg Schneider (2020): »Medial Shaping from the Outset: On the Mediality of the Second Presidential Debate, 2016«, in: *Journal für Medienlinguistik/Journal for Media Linguistics Vol. 3:1*, S. 57–93.
- Marx, Konstanze/Axel Schmidt (Hg.) (2019): *Interaktion und Medien. Interaktionsanalytische Zugänge zu medienvermittelter Kommunikation*, Heidelberg: Winter.
- McLuhan, Marshall/Quentin Fiore (2016): *Das Medium ist die Massage. Ein Inventar medialer Effekte*, 4. Aufl., Stuttgart: Tropen.
- Meier, Christoph (2000): »Neue Medien – neue Kommunikationsformen? Strukturmerkmale von Videokonferenzen«, in: Werner Kallmeyer (Hg.), *Sprache und neue Medien*, Berlin/New York: Walter de Gruyter, S. 195–221.
- Müller, Cornelia/Ulrike Bohle (2007): »Das Fundament fokussierter Interaktion. Zur Vorbereitung und Herstellung von Interaktionsräumen durch körperliche Koordination«, in: Reinhold Schmitt (Hg.), *Koordination. Analysen zur multimodalen Interaktion*, Tübingen: Narr, S. 129–165.
- Müller, Cornelia/Hermann Kappelhoff (2018): *Cinematic Metaphor. Experience – Affectivity – Temporality*, Berlin/Boston: De Gruyter Mouton.
- Müller, Cornelia (2024): »A Toolbox of Methods for Gesture Analysis«, in: Alan Cienki (Hg.), *Handbook of Gesture Studies*, Cambridge: Cambridge University Press, S. 182–216.

- Müller, Cornelia (im Erscheinen): »Sprache-Sprechen, hören, sehen, fühlen. Eine medienästhetische Perspektive auf die Multimodalität des Sprechens in audiovisuellen Medien«, in: Rita Vallentin/Dorothea Horst (Hg.): Sprache entgrenzen. Beiträge zu einer kulturwissenschaftlichen Linguistik. Bielefeld: transcript.
- Plessner, Helmut (1982): »Die Deutung des mimischen Ausdrucks. Ein Beitrag zur Lehre vom Bewußtseins des anderen Ichs« [in Zusammenarbeit mit J. J. Buijten-dijk], in: Philosophischer Anzeiger 1, S. 72–126.
- Sacks, Harvey/Emanuel A. Schegloff/Gail Jefferson (1974): »A Simplest Systematics for the Organization of Turn-Taking for Conversation«, in: Language 50:4, S. 696–735.
- Scherer, Thomas J. J. (2024): Inszenierungen zeitgenössischer Propaganda. Kampagnenfilme im Dienste des Gemeinwohls, Berlin/Boston: De Gruyter.
- Scherer, Thomas/Sarah Greifenstein/Hermann Kappelhoff (2014): »Expressive Movements in Audiovisual Media. Modulating Affective Experience«, in: Cornelia Müller/Alan Cienki/Ellen Fricke/Silva H. Ladewig/David McNeill/Jana Bressemer (Hg.), Body – Language – Communication. Handbücher zur Sprach- und Kommunikationswissenschaft, Berlin/Boston: De Gruyter Mouton, S. 2081–2092.
- Schmitt, Reinhold (2005): »Zur multimodalen Struktur von *turn-taking*«, in: Gesprächsforschung – Online Zeitschrift zur verbalen Interaktion 6, S. 17–61.
- Selting, Margret/Peter Auer/Dagmar Barth-Weingarten/Jörg Bergmann/Karin Birkner/Elizabeth Couper-Kuhlen/Arnulf Deppermann/Peter Gilles/Susanne Günthner/Martin Hartung/Friederike Kern/Christine Mertzluft/Miriam Morek/Frank Oberzaucher/Jörg Peters/Uta Quasthoff/Wilfried Schütte/Anja Stukenbrock/Susanne Uhmann (2009): »Gesprächsanalytisches Transkriptionssystem 2 (GAT 2)«, in: Gesprächsforschung 10, S. 353–410.
- Sobchack, Vivian (1992): The address of the eye. A Phenomenology of Film Experience, Princeton: Princeton University Press.
- Stern, Daniel (2011): Formen der Vitalität. Die Erforschung dynamischen Erlebens in Psychotherapie, Entwicklungspsychologie und den Künsten, Frankfurt a. M.: Brandes & Apsel.
- Wulff, Hans J. (1991): »Split Screen: Erste Überlegungen zur semantischen Analyse des filmischen Mehrfachbildes«, in: Kodikas/Code 14:3-4, S. 281–290.
- Wundt, Wilhelm (1975): Völkerpsychologie. Eine Untersuchung der Entwicklungsgesetze von Sprache, Mythos und Sitte, 3. Aufl., Band 1: Die Sprache, Aalen: Scientia Verlag.
- Xia, Qiang (2023): »Turn-taking in Video-Mediated and Co-Present Dialogues. A Corpus-Based study of German«, in: Register Aspects of Language in Situation (REALIS) 2:4, S. 1–29.

Ludische Dividuationen

Eine antiholistische Betrachtung von CULTIST SIMULATOR und DEAD SPACE

Dominic Brakelmann, Arvid Kammler

Abstract Videospiele sind multimodal hochgradig hybride mediale Phänomene. Jenseits ihrer produktionsseitigen Teilungen konstituieren sie rezeptionsseitig Erfahrungsräume geteilter und verteilter Körperlichkeiten, deren performative Repräsentation in Form der Spieler-Spielraum-Spiel-Relation wir uns in diesem Text widmen wollen. Mittels des Konzeptes der körper-räumlichen Verteilung richten wir den Blick auf die mediale Vermittlungssituation des Videospieles und der in ihr involvierten Körper. In der Folge argumentieren wir, dass der Akt der Dividuation grundlegend für das reflexive Potenzial des Mediums hinsichtlich analog-digitaler Interaktionsräume und deren Akteure ist. Dieses medienontologische Spezifikum spiegelt sich auf den Ebenen von Spielmechanik, Erzählung, Affekt und Ästhetik wider. Anhand ausgewählter Fallbeispiele (u.a. CULTIST SIMULATOR und DEAD SPACE) spüren wir Formen dieser Teilungen nach, um zu einer dividuationalistischen Lesart des Videospieles und des Videospieles zu gelangen.

Keywords Dividuation; Verteilung; Teilung; Körper; Raum; Dividuum

1. Dividuationen der Spielkörper

Im Hinblick auf die Vielzahl an holistischen Erzählungen zu Videospiele – die oftmals aus den Marketingabteilungen von Publishern entstammen und in ihrer bild- und wortgewaltigen Ästhetik das vollkommene Eintauchen in neueste Spielerfahrungen und mithin das Einswerden mit übermächtigen, platzhalterartigen Spielfiguren versprechen¹ – wollen wir den Blick darauf werfen, wie die einzelnen am

1 Hierfür sind beliebige Werbeclips zu Spielen wie CYBERPUNK 2077 (Cyberpunk 2077: 2020) oder STARFIELD (IGN: 2020) eindrücklich, deren Machart sich in unzähligen anderen Vertretern ihrer Art wiederfinden lässt.

Prozess des Spielens beteiligten Entitäten produktiv zusammenfinden, um das Videospiele existent werden zu lassen. Dabei wollen wir die spielenden Subjekte, das Spiel und die daraus entstehenden ästhetischen Erfahrungsräume nicht als unteilbar oder individuell und vor dem Zeithorizont als stillstehend begreifen, sondern den Blick auf ihre Einzelheiten, Teilbarkeit und auch körperräumliche Verteilung lenken. Daraus ergibt sich die Möglichkeit, ihr prozessuales Zusammenwirken aus einem Blickwinkel zu betrachten, der die Positionalisierung der Spielenden, die Potenzialität des Spiels, sowie den sich daraus ergebenden Erfahrungsraum auf Basis einer kontinuierlichen Veränderung offenbart. Der Spielraum ist dabei im Grunde mehr als ein Erfahrungsraum, durch welchen den Spielenden Affekte, Erlebnisse und Repositionalisierungen widerfahren. Es ist vielmehr ein Möglichkeitsraum, durch den und mit dem die Spielenden handelnd tätig werden und der sich zwischen der virtuellen Spielwelt und dem tatsächlichen Raum entfaltet. Ein Interaktionsraum, der Reales und Unreales aktuell werden lässt.

Als Grundlage dieser Überlegungen dienen uns Michaela Otts (2018: 173) Ausführungen zur *Dividuation*, die sie über Whiteheads Zelltheorie, Gilbert Simondons Technikphilosophie und schließlich den daraus von Gilles Deleuze weiterentwickelten Begriff des Dividuellen entwirft. Dieser erlaubt es uns, unser grundlegendes Selbstverständnis des Individuums von einer anderen Seite her zu denken. Das Dividuelle von Deleuze bezieht sich Ott zufolge auf filmische und auch musikalische Kompositionen, die sich aufgrund der Veränderungen über ihren zeitlichen Ablauf hinweg nicht als individuelle ästhetische Erfahrung beschreiben lassen (vgl. ebd.: 169). Das Dividuelle dient als Ausgangspunkt für eine übergeordnete Reflexion, die die Frage der Subjektivierung vor dem Hintergrund alltäglicher Existenzweisen und Teilbarkeiten neu stellt. Angesichts der Vielzahl an digitalen Handlungs- und Erfahrungsweisen, die unser alltägliches Leben durchziehen, betrachtet Ott unser permanentes, teilhabendes Umherschweifen und Eingebundensein in die verschiedensten soziokulturellen Gegebenheiten als Ausgangspunkt ihrer Fragestellung:

Wie also, so die nicht mehr abweisbare Frage, können wir uns angesichts dessen noch als ungeteilte Akteure und autonome Entscheider unserer Teilhaben verstehen? Müssen wir diesem »Schicksal« nicht zuallererst sprachlich Rechnung tragen und zugeben, dass wir längst freiwillige und unfreiwillige Dividuationen, Vielfach-Teilhabende und -Zugeteilte, sind? (ebd.: 172f.)

Ausgehend vom vermeintlichen Individuum als Mittelpunkt, lenkt Ott den Blick auf die mikroskopische Ebene, auf der wir von Milliarden von Kleinstlebewesen bewohnt und durch die unsere körperlichen Funktionsweisen grundlegend mitbestimmt werden. Sie stellt so heraus, dass das menschliche Subjektwerden nicht unwesentlich bereits vor einer Wahrnehmungsschwelle stattfindet, über die hinaus wir uns dieser Vorgänge bspw. erst über eine im Nachhinein stattfindende Refle-

xion gewahrt werden können (vgl. ebd.: 171). Auf makroskopischer Ebene hingegen werden wir durch mediale Vermittlungsstrategien in eben jene bereits erwähnten weltumspannenden soziokulturellen Handlungs- und Erfahrungsmomente eingebunden (vgl. ebd.). Daraus wird deutlich, dass das menschliche Subjektwerden immer schon und fortwährend und auf allen Ebenen verschränkt ist.

Für diesen Beitrag ergeben sich daraus Fragestellungen, die sowohl das Videospiel als potenzialgeladenen Möglichkeitsraum als auch die Spielenden selbst betreffen, die diesen Raum erkunden und mitgestalten.

2. Videospiele als körperliche Erfahrungsräume

Die Annahme des grundlegenden Erfahrungsmodus des Videospieles als geteiltes Spielen, ermöglicht eine Erweiterung des Analysefeldes, innerhalb dessen sich Entitäten unterschiedlicher Realitäten und Medialitäten zu einer hybriden Assemblage² verbinden. Um den Fokus auf die subjektiven und körperlichen Aspekte des Spielens zu richten, die sowohl die Spielfigur als auch die Spielenden selbst betrifft, hilft es zunächst, das Verhältnis von Körper und Raum im Prozess des Spiels genauer zu betrachten. Hierbei sind Dichotomien, wie sie etwa immersionstheoretischen (Murray 1997; Ryan 2001; McMahan 2003; Thon 2008), narratologischen (Backe 2008; Newman 2004) oder dispositivtheoretischen (Liebe 2008; Mosel 2009) Zugängen zugrunde liegen, wenig hilfreich, weil sie der Komplexität der Spielsituation nur schwer gerecht werden. Ihnen ist die Konstruktion von Spiel und Spielenden als in sich geschlossene Entitäten eigen, die im Spielprozess auf je unterschiedliche Art in Kontakt treten. Dabei blenden diese Perspektiven aus, wie spielspezifische Körper, Räume und Zeitlichkeiten grundsätzlich aus dem Spielprozess emergieren.

Es lohnt sich daher, im ersten Schritt das mit dem Spiel interagierende Subjekt genauer zu betrachten, das im Moment des Spiels aus externer Sicht mit einer anderen – audiovisuell vermittelten – Körperlichkeit interagiert. Was geschieht in diesem Prozess mit dem spielenden Subjekt, das wir in unseren alltäglichen Lebensweisen oftmals als abgeschlossene, einzelne Entität erfassen und denken? Bojan Peric weist aus einer rezeptionsästhetischen Perspektive in seinen Ausführungen zum Immersionssubjekt darauf hin, »dass ein Subjekt als nichts in sich Abgeschlossenes, sondern vielmehr als etwas ständig im Wandel Seiendes gedacht werden muss« (Peric 2016). Er beschreibt weiter, dass das Subjekt durch ein simples Erleben geformt wird, sich verändert und damit aus einzelnen Versatzstücken seine

2 Assemblage rekurriert hier auf die Zusammensetzung des Videospieles aus Praktiken und technischen Objekten/Interfaces. Der Begriff ist den Arbeiten von Deleuze und Guattari (1992) entlehnt, die ihn zur Beschreibung von Handlungsmacht in soziotechnischen Gefügen entwickelten.

jeweils momentane Kohärenz bildet (vgl. ebd.). Peric macht dies zunächst am lesenden Subjekt deutlich, das sich durch »Input, den es aus literarischen [...] Geschichten erhält, konstituiert; diese bilden einen Teil der Gesamtheit der eigenen Geschichten, welche das Subjekt als solches ausmachen« (ebd.). Übertragen auf das Videospiele zieht er jedoch interessanterweise in Zweifel, ob für das dargestellte Subjekt, sprich die Spielfigur, dasselbe gilt, da es rein als Darstellung existiere und demzufolge kein Subjekt sein kann. Dieser Punkt ist deshalb wichtig, weil Peric im Versuch der cartesianischen Dichotomie zu entkommen, diese zunächst scheinbar lediglich hin zu einer Dichotomie von spielendem Subjekt und Spielfigur verschiebt. Im weiteren Verlauf etabliert er das Konzept des *Immersionssubjekts*, das die Grenzen zwischen den Spielenden und dem Medium unscharf werden lässt (ebd.), um schließlich über eine dreigeteilte Interaktionsmatrix zum Schluss zu kommen, »dass das [spielende] Subjekt eine Figur [...] nicht nur spielt, sondern tatsächlich ist« (ebd.).

Anhand von Perics Argumentationsverlauf wird deutlich, dass wir es beim Videospiele mit einem Medium zu tun haben, das gemeinsam mit dem spielenden Subjekt paradoxe Interaktionsräume aufspannen kann, in denen spielende und gespielte Körper ineinander übergehen, obwohl sich die Darstellungs- und durch sie vermittelten Erfahrungsweisen so grundlegend unterscheiden. Das auf den ersten Blick unter Umständen als unreal erscheinende virtuelle Bildschirmgeschehen wird nicht zuletzt im Sprechen über das Spielen evident – z.B. wenn beim Spielen jemand sagt: *Jetzt bin ich tot; Du hättest dort springen müssen*. Brendan Keogh beschreibt die Voraussetzung für »an engagement with virtual objects is still dependent on actually handling actual objects alongside actually viewing and listening to actual images and sounds« (Keogh 2018: 61). Beim Spielen geht es also um eine Gleichzeitigkeit von Virtuellem und Aktuellem. Das Reale und das Unreale verschränken sich im Spiel untrennbar miteinander und werden wechselseitig wirksam.

Darüber hinaus wird ebenfalls offensichtlich, dass wir, wenn wir verkürzende Dichotomien vermeiden wollen, das spielende Subjekt gar nicht an einem bestimmten Punkt erfassen können, sondern uns dem Spiel immer nur als Prozess annähern können. In dieser Hinsicht kann das Videospiele als Erfahrungsraum beschrieben werden, der Körperlichkeit auf je spielspezifische Art und Weise moduliert und somit Anteil am steten Werden des spielenden Subjekts besitzt. Wichtig ist hierbei ebenfalls, dass wir es mit einem wechselseitigen Verhältnis von Spielenden und Spiel zu tun haben, denn der Erfahrungsraum Videospiele wird auch von den Spielenden gestaltet und verändert, basierend auf der im Spiel angelegten handlungsorientierten Potenzialität.

Um die Potenzialität hinsichtlich der Dividuation und ihrer Prozesshaftigkeit genauer beschreiben zu können, bietet sich das Konzept der *Transduktion* von Gilbert Simondon (2011) an. Dieses beschreibt die Verkettung dynamischer Operationen, in denen technische Objekte und Menschen zueinander in Beziehung gebracht wer-

den und metastabile³ Relationen bilden, die durch reziproke ontogenetische Transformationsvektoren gekennzeichnet sind. Entscheidend ist, dass der transduktive Prozess ein vorhandenes Potenzial realisiert, wobei die Struktur im Potenzial nie von vornherein festgelegt ist, sondern nur als eine Möglichkeit von vielen besteht. Im Detail wird ein Milieu, das sich in einem metastabilen Zustand befindet (und demzufolge ein nicht realisiertes Energiepotenzial aufweist) durch eine Informationsspannung strukturiert und damit transformiert (vgl. Simondon 2011: 238). Das Ergebnis dieses Prozesses sind offene Strukturen, die weitere transduktive Prozesse durchlaufen können und sich so beständig verändern.

3. Erfahrungsraum CULTIST SIMULATOR

Mit diesem theoretischen Fundament möchten wir uns im Folgenden dem konkreten Erfahrungsraum des Videospieles CULTIST SIMULATOR (Weather Factory 2018) zuwenden und diesen eingehend beschreiben, um im Anschluss das ästhetische Phänomen des geteilten Spielens anhand dieses Beispiels zu diskutieren.

CULTIST SIMULATOR vereint Spielelemente ganz unterschiedlicher Spieltypen und lässt sich am ehesten mit einer Mischung aus Echtzeitstrategie, Adventure-Rollenspiel und Kartenspiel beschreiben. Die Präsentation der Spielelemente in ihrer Verteilung sowie die aus der genannten Mischform resultierende Komplexität sind mithin ein Grund, weshalb dieses Spiel ein geeignetes Beispiel für geteiltes Spielen und für körperräumliche Verteiltheit in Bezug auf die Beziehung zwischen Spielenden und Spiel ist. In seiner Form als Kartenspiel, präsentiert sich das Spiel bereits als individualistisch. Zumal das, was einer Avatarfigur am nächsten kommt, bereits ein aus einzelnen Elementen Zusammengesetztes ist – sowohl in seiner audiovisuellen Präsentation als auch in seinen spielmechanischen Relationen zu anderen Spielelementen.

Die Handlung spielt etwa zu Beginn des 20. Jahrhunderts in einer nicht näher bezeichneten europäischen Großstadt. Eine der zentralen Aktivitäten im Spiel besteht darin, auf Basis von mythischen Weltvorstellungen einen Kult zu gründen und potenzielle Mitglieder mithilfe des im Selbststudium erarbeiteten, esoterischen Kanons zum Beitritt zu überzeugen. Als Kultoberhaupt sind die Spielenden ständig auf der Suche nach verborgenem und verbotenem Wissen, das im weiteren Spielverlauf immer mehr die rational-wissenschaftlichen Sphäre verlässt. Insgesamt ähnelt

3 Metastabilität bezeichnet den Zustand eines Systems, in dem die darin enthaltenen Elemente durch einen Katalysator eine Zustandsveränderung erfahren können. Der Katalysator muss dabei nicht unbedingt von außen kommen, sondern kann aus dem System selbst hervorgehen. Auch sind grundlegende Eigenschaften, die der anschließend erneut metastabile Zustand des Systems aufweist, bereits in der vorherigen Version als Potenziale angelegt.

das Spiel in seiner Präsentation – die über viele einzelne, auf einem Tisch angeordnete Karten realisiert wird – einer narrativ weit ausgearbeiteten Version eines Tarot-Kartendecks. Wie beim esoterischen Tarot präsentiert *CULTIST SIMULATOR* stets nur abstrakte Ideen durch Karten wie *Reason*, *Passion*, *A Fleeting Reminiscence*, *Streets Strange by Moonlight* oder *Fascination*. Auch die näheren Beschreibungstexte der einzelnen Karten sind so gestaltet, dass sie weiterführenden Interpretationsraum anbieten: »Sometimes, the light of the moon is a key to other spaces. I've found a place where, for a night or two, the streets curve in unfamiliar ways. If I walk here, I might find insight, or I might be touched by madness« (*Cultist Simulator*, Weather Factory 2018). Der überwiegende Teil der Texte im Spiel sind in dieser vignettenhaften, poetischen Sprache geschrieben, die eine dem Setting angemessene Atmosphäre heraufbeschwört und die Diegese des Spiels bildhaft ausschmückt.

Abb. 1: Der Spieltisch in *CULTIST SIMULATOR*.



3.1 Verteilung und Spielfeld

Spannend ist die Frage, wo genau die Spielenden in *CULTIST SIMULATOR* als Oberhaupt des Kults sowohl in der Spielwelt als auch im Spielraum verortet sind. Grundlegend kann zwischen zwei unterschiedlichen Kategorien von Entitäten im Spiel unterschieden werden. Die Erste ist die Spielfigur, die die Geschicke des Kults lenkt, Wissensfragmente analysiert, Mitglieder rekrutiert und sowohl in der innerdiegetischen physisch realen Welt der Hauptstadt der Spielwelt existiert als auch in der diffusen, traumartigen Sphäre des Okkulten wandelt. Die zweite Kategorie umfasst alle anderen Figuren, mit denen die Spielenden im Spielverlauf interagieren kön-

nen: Gefolgsleute, Gegenspieler*innen⁴, Unterstützer*innen und beschworene Wesen. Jede dieser Figuren wird durch eine eigene Karte auf dem Spieltisch repräsentiert, die Spielfigur des Kultoberhaupts ist jedoch nicht in Form eines klassischen Avatars in der Darstellungsform eines menschlichen Körpers in der Spielwelt präsent, wie etwa in Städtebausimulationen. Demzufolge funktioniert die räumliche Vermittlung der Spielwelt in *CULTIST SIMULATOR* auch nicht über die Spielfigur in Form einer einzelnen Karte, sondern offensichtlich über die Verteilung auf viele verschiedene Karten.

Abb. 2: Mögliche Anordnung der Spielkarten.



Tritt man einen Schritt zurück und betrachtet die visuelle Präsentation des Spiels, so fällt in Abbildung 1 der Spieltisch auf, auf dem die Karten liegen. Die Anordnung des Tisches vor dem Hintergrund der körperräumlichen Verteiltheit der Spielfigur in Form unterschiedlicher Karten vermittelt eine bemerkenswerte Position. Das Blickregime spiegelt zumindest teilweise die realweltliche Situierung der Spielenden am Schreibtisch vor dem Monitor wider und verweist darüber hinaus auf vergleichbare spielräumliche Situationen, in denen ein Kartenspiel, bspw. Patience, an einem Tisch gespielt wird.

Eine Ebene weiter sind die Karten sowie die Aktionen auf dem Tisch liegend sichtbar. Wie bereits beschrieben, gibt es Karten verschiedener Kategorien, deren Anordnung auf dem Tisch durch das Spiel willkürlich vorgenommen wird und von

4 Das Spiel ist ein reines Singleplayer-Spiel, der Begriff Gegenspieler*in bezieht sich hier auf die Charaktere, die gegen den Kult ermitteln.

den Spielenden bei Bedarf nach eigenen Vorlieben verändert werden kann. Am unteren Bildschirmrand befindet sich als eine weitere Ebene das Interface mit einer Reihe von Anzeigen und Schaltflächen darüber. Hier kann unter anderem der Name der Spielfigur geändert, das Spiel pausiert oder beschleunigt und das Optionsmenü aufgerufen werden. In der Mitte des Interface sind die Ressourcen des Spiels aufgelistet, die später auch jeweils als Karten auf dem Tisch dargestellt werden: *Health*, *Passion*, *Reason* und *Funds*. Zwischen dieser und der Ebene der Karten befinden sich die Aktionsfenster, die detaillierte Erläuterungen zu laufenden Aktionen enthalten, sowie die Beschreibungstexte der einzelnen Karten anzeigen. Ein Teil der Komplexität des Spiels entsteht also nicht durch eine übermäßige Anzahl von Darstellungsebenen, sondern vielmehr durch die Vielzahl der unterschiedlichen Karten und die Kombinationsmöglichkeiten, die mit den verschiedenen Aktionen verbunden sind.

Hinsichtlich der Dividuation spielt der Tisch in *CULTIST SIMULATOR* als Grundlage aller Handlungen eine sehr interessante Rolle. Gehen wir davon aus, dass sich das spielende Subjekt ständig verändert und neu formiert, so spiegelt der Tisch mit den daraufliegenden Karten diesen Prozess wider. Der Tisch und das Tableau bilden hierbei den Wechsel zwischen der momentanen Verfestigung – dem metastabilen Zustand nach Simondon – im Tableau und der Möglichkeit der freien Neuordnung und der Bildung neuer Relationen der auf dem Tisch sich befindenden Dinge.⁵

Die Anordnungen der Karten auf dem Tisch in *CULTIST SIMULATOR* veranschaulichen einen ganz ähnlichen Vorgang. Auch hier sind Prozesse der Neuordnung und des Nebeneinanderstellens zu beobachten, die schließlich jeweils für einen gewissen Zeitraum in einen metastabilen Zustand übergehen, um dann wieder in Bewegung zu geraten und sich einer weiteren Neuordnung zu unterziehen. Der Tisch ist sowohl die Fläche dieser potenziellen Neuordnung als auch Tableau als Basis metastabiler Spielsituationen. Das Besondere an *CULTIST SIMULATOR* im Speziellen und an Videospiele im Allgemeinen ist dabei, dass Tisch und Tableau kongruent sind (Unterhuber 2018: 211). Was *CULTIST SIMULATOR* in Form der Spielkarten und des Tisches offenbart, sind genau diese Prozesse der Spielregeln, die gewissermaßen wortwörtlich in Bild und Ton gefasst werden. Während andere Spiele diese Vorgänge durch Inventarmenüs, Lebensenergiebalken und Ausrüs-

5 Tisch und Tableau referenzieren hierbei Georges Didi-Hubermans Arbeit zu Aby Warburgs kolossalem Projekt des Mnemosyne Atlas, mit dem Warburg versuchte, das Nachleben antiker Bildsprache in den Werken der nachfolgenden europäischen Kunst sichtbar zu machen (vgl. Didi-Huberman 2018). Hierbei bediente Warburg sich Fotografien von Kunstwerken ganz unterschiedlicher Art und Themen und setzte diese durch physisches Nebeneinanderplatzieren auf Tableaus miteinander in Beziehung. Durch diesen topologischen Prozess des In-Beziehungen-Setzens wurden Gemeinsamkeiten deutlich, die vorher teilweise verborgen geblieben waren. Der Tisch diente Warburg hierbei als Basis freier Assoziation, woraus sich schließlich metastabile Anordnungen in Form der Tableaus entwickelten.

tungsgegenstände verkleiden, interagieren die Spielenden in *CULTIST SIMULATOR* immer mit dem offen gelegten Spielgerüst an sich.

3.2 Dividuation: Körperräumliche Verteilung und Neuordnung

Die Körperlichkeit der Spielfigur hinsichtlich der Überlagerung der diegetischen Räumlichkeiten spielt eine interessante Rolle im Spielgeschehen. Wie oben in Abbildung 1 zu sehen ist, ist die Körperlichkeit der Spielfigur im Gegensatz zu der der Kultmitglieder auf verschiedene Karten verteilt und setzt sich aus *Health*, *Passion*, *Reason* und ferner *Decrepitude* zusammen. Während die ersten drei gleichzeitig drei Viertel der grundlegenden Ressourcen und Attributskarten darstellen, die immer wieder zur Erfüllung bestimmter Aktionen erforderlich sind, ist Letzteres ein umgewandeltes *Health*, das durch einen Fluch über die Spielfigur hereingebrochen ist. Im Gegensatz zur körperräumlichen Verteiltheit des Avatars in Städtebausimulationen, der dort vor allem als Ergebnis raum-zeitlicher Planungstätigkeiten in Form einer modellhaft ästhetisierten Stadt auftritt, ist die körperräumliche Verteiltheit des Avatars in *CULTIST SIMULATOR* die notwendige Grundlage nahezu aller Handlungsoperationen. Die Aufteilung des Avatars, der in vielen anderen Spielen als weitgehend intakter Körper dargestellt wird, ist jedoch nicht von Spielbeginn angegeben, sondern entwickelt sich erst langsam durch aufeinanderfolgende Aktionen des Arbeitens, Träumens und Studierens. Auch hier gibt es folglich eine Potenzialität einer immer weiter ausufernden Verteiltheit. Nun wäre es möglich zu argumentieren, dass die körperräumliche Verteiltheit des Körpers der Spielfigur in seiner Fragmentierung auf drei oder vier verschiedene Kartentypen gegeben sein mag, die Figur aber letztlich noch immer von allen anderen Spielelementen klar abgegrenzt werden kann. Erstens ergibt sich die rein visuelle Abgrenzung des Avatakörpers von anderen Spielelementen auf dem Spieltisch in Abbildung 2 möglicherweise aus der persönlichen Ordnungsliebe der Spielenden im Rahmen des Spiels und verschleiert sehr effektiv das dem Spielablauf zugrunde liegende Beziehungsgeflecht der Spielelemente untereinander: Nur ein Bruchteil der laufenden Aktionen im Spiel kommen ohne eine Beteiligung der körperlichen Attributskarten aus. Das bedeutet, dass die Spielfigur in ständiger Interaktion mit den anderen Spielelementen steht und grundsätzlich körperlich in die ablaufenden Prozesse eingebunden und mit ihnen verflochten ist. Der zweite Punkt betrifft das Paradoxon der Kongruenz von Spielfigur und Spielenden als Bedingung der Möglichkeit des Spiels an sich. Dies ist bspw. dann zu beobachten, wenn eine dedizierte, von ihrer Umgebung abgrenzbare Spielfigur als Vermittlungsinstanz zwischen Spielenden und Spiel identifiziert werden kann, und sei es nur in Form von Karten. Die Dividuation der Spielenden in Videospielen resultiert indes gerade nicht allein aus einer möglichen Verteiltheit des spielfigurlichen Körpers innerhalb der Spielwelt. Die Dividuation ergibt sich vielmehr aus dem gesamten Vermittlungsgefüge, in dessen konstituierende Prozesse

die Spielenden eingebunden sind und dessen wechselseitig wirkende Strukturen sie gemeinsam mit dem Spiel über einen bestimmten Zeitraum hinweg kollaborativ herausarbeiten.

So spielt im konkreten Spielverlauf von *CULTIST SIMULATOR* die Körperlichkeit der Spielfigur im Rahmen der Handhabung der Spielmechaniken (Bewegen von Karten auf einer Tischoberfläche) keine vordergründige Rolle (wohl aber im Rahmen der Diegese). Im Grunde bilden die drei Kartentypen *Health*, *Reason* und *Passion* schlicht einen Teil der zu verwaltenden Ressourcen, deren oben herausgestellte, tiefergehende Bedeutung rein auf diegetischer Ebene relevant ist. Die Verteilung der *Spielenden* ist in der gesamten Verteilung der *Spielkarten* auf dem Kartentisch zu verorten und somit ist auch die Vermittlungsinstanz nicht auf die Spielfigur des Kultoberhaupts beschränkt. Gerade durch die aus persönlichen Präferenzen resultierende Ordnungspraxis des Spieltischs entsteht eine spezifische Ästhetik, ähnlich dem Modellbauaspekt bei Städtebausimulationen. Auch wenn sich das eigentliche Spielziel um den Aufstieg des Individuums (was der Argumentation folgend genau genommen als Dividuum beschrieben werden müsste) zu einer höheren Existenzebene dreht, frei von körperlicher und damit auch sterblicher Gebundenheit, durchläuft diese Spielfigur immer wieder Prozesse der Auflösung und des Wiederzusammensetzens und wird so zu einem gleichwertigen Teil des gesamten Kartenensembles. *CULTIST SIMULATOR* lässt seine Spielfigur fortwährend zwischen Metastabilität und Auflösung oszillieren, worauf ebenfalls die unterschiedlichen Perspektiven der Begleittexte zu den Aktionen und Karten verweisen, die entweder in Form interner oder externer Fokalisierung (vgl. Genette 1980) auftreten und damit Raunigs (2015) Idee einer *dividuellen Autorschaft* folgen.

In der Relation von Spielenden und Spielfigur, die sich erst im Prozess des Spielens entfaltet, kommen unterschiedliche Entitäten zusammen und schaffen durch wechselseitige Handlungsoperationen neue Räumlichkeiten, Körperlichkeiten und Subjekte. Eine dividualistische Lesart erlaubt einen Blick auf die Entstehungsprozesse der Körperlichkeiten dieser Entitäten. Im Vergleich dazu können Theorien, die medial vermittelte Zugänge bspw. als Prothesen oder körperliche Erweiterungen begreifen, dies nicht leisten, da sie oftmals auf einer Vorannahme eines abgeschlossenen Körperbildes basieren. Gail Weiss, hier zitiert von Keogh, argumentiert in der Weiterentwicklung von Maurice Merleau-Pontys Theorien

for an *essential instability* of bodies; rather than ›a cohesive and coherent phenomenon that operates in a fairly uniform way in our everyday existence‹ it is important to account for ›a multiplicity of body images, body images that are co-present in any given individual, and which are themselves constructed through a series of corporeal exchanges that take place both within and outside of specific bodies‹ (Weiss 1999: 1f., zitiert n. Keogh 2018: 48)

Keogh argumentiert, dass es, um verkörperte Erfahrungen beschreiben zu können, notwendig sei zu betrachten, »how embodiment emerges from phenomena in a dynamic coming-together of multiple corporealities« (Keogh 2018: 48). Diese Herangehensweise erlaubt es, das Spielen über reine Immersionserfahrungen⁶ hinaus als eine durch Wechselseitigkeit und Offenheit gekennzeichnete ästhetische Erfahrung zu beschreiben. Dieser offenen Spielerfahrung, von der wir vorher nicht genau wissen, was sie für uns als Spielende bereithält, wohnt eine Potenzialität inne, der wir nur gerecht werden, wenn wir alle am Prozess beteiligten Entitäten als eine offene Assemblage, jene metastabile Anordnung, begreifen. Für Keogh resultiert diese Herangehensweise in der Dezentrierung des menschlichen Körpers als Mittelpunkt medialer Vermittlung im Spiel und rückt dadurch die anderen, nicht-menschlichen Körper in den Fokus, die im Spiel ebenfalls Teil des Prozesses konstanter Neukonfiguration sind (vgl. ebd.: 50).

Im folgenden Teil möchten wir anhand eines weiteren Beispiels deutlich machen, inwieweit diese geteilte Erfahrung des Videospelns insbesondere medienreflexiv gefasst werden kann. *DEAD SPACE* geht hierbei über die bereits im Spielprinzip von *CULTIST SIMULATOR* angelegte Teilung der einzelnen Spielelemente hinaus und hält die Spielenden im gesamten Erfahrungsverlauf dazu an, Teilungen aktiv vorzunehmen. Aus dem Verhältnis der Spielfiguren von *DEAD SPACE* und *CULTIST SIMULATOR* und deren jeweiligen auf ihre Umwelten gerichteten Handlungsweisen ergibt sich hinsichtlich der dividualistischen Lesart des Spielens ein äußerst interessantes Spannungsfeld.

4. *Dead Space*: Die holistische Narration

Im Fall von *DEAD SPACE*⁷ handelt es sich im Gegensatz zu *CULTIST SIMULATOR* um ein avatarzentriertes Videospiel mit einer figurenzentrierten Erzählung (vgl. Beil 2010), das sowohl auf spielmechanischer und ästhetischer als auch diegetischer Ebene Themen der Einheit und Trennung behandelt. Hiermit zeigt das Spiel in für das Medium seltener Klarheit, dass es sich nicht um ein homogenes System in einer fiktionalen Welt handelt, sondern um einen Komplex konkurrierender Potenziale, die sich situativ überlagern, verdecken und ihre hierarchische Ordnung verändern können.

-
- 6 Die stets lediglich von einem Eintauchen abgeschlossener und mehr oder weniger unveränderlicher Körperlichkeiten in ein bereits fertig- und bereitgestelltes anderes Milieu sprechen.
- 7 Die folgenden Überlegungen nehmen das *DEAD SPACE*-Remake aus dem Jahr 2023 als Untersuchungsgrundlage, das in weiten Teilen inhaltlich dem ersten *DEAD SPACE* (2008) entspricht.

Die Ereignisse in *DEAD SPACE* werden aus der Perspektive von Isaac Clarke erlebt. Clarke befindet sich im Jahr 2508 auf einer Rettungsmission, die das Minenschiff *USG Ishimura* zum Ziel hat. Nach Ankunft auf dem Raumschiff wird die Besatzung mit einer Infektion vermeintlich außerirdischen Ursprungs konfrontiert. Weite Teile der Besatzung wurden von einer Art Parasit infiziert, der mithilfe der Ausbildung grotesker Extremitäten die leblosen menschlichen Körper in monströse und feindselige Lebensformen verwandelt. Ursprung dieser Kontamination ist ein Artefakt, das während der Minenarbeiten auf dem angrenzenden Planeten hervorgebracht wurde. Dieses Artefakt, der sogenannte Marker, steht im Zentrum der Rahmenhandlung von *DEAD SPACE*. Clarke hat zudem ein starkes persönliches Motiv: Seine Lebensgefährtin Nicole ist als Ärztin auf der *USG Ishimura* stationiert. Im Verlauf der Handlung von *DEAD SPACE* wird jedoch deutlich, dass diese zu Anfang des Spiels kommunizierte Motivation elliptischem Erzählen geschuldet ist. Wie sich im Spielverlauf herausstellt, hat Nicole nach dem Ausbruch Suizid begangen und ihr wiederholtes Auftreten ist auf Isaacs mentale Kontamination mit den Kräften des Markers zurückzuführen. Das teleologische Streben nach Wiedervereinigung auf persönlicher Ebene erfüllt sich im Spielverlauf folglich nicht.

Gleichzeitig mit dieser persönlichen Tragödie entfaltet sich auch die Geschichtlichkeit der Spielwelt. Dieser Marker, der die als Necromorph bezeichneten Mischwesen heraufbeschwört, ist in ein überpersönliches Narrativ eingebettet: So ist das Auftauchen und Verschwinden des originären Markers in der Welt von *DEAD SPACE* Initiator eines das Weltall umspannenden Glaubensphänomens, das unter dem Namen Unitology firmiert. Die Verheißung dieses Glaubens besteht in der Konvergenz von Lebenden und Toten in einem prophezeiten utopischen Kulminationsmoment. Dieses Konvergenzphantasma wird durch die tatsächliche Wirkung der Marker verneint, deren eigentlicher Zweck darin besteht, durch Kontamination und Transformation toter Lebewesen ausreichend biologisches Material anzusammeln, um eine letzte mondformige Entwicklungsstufe der Necromorph zu erreichen.

Im Rückgriff auf die Erzählung von *DEAD SPACE* wird deutlich, dass diese auf verschiedenen Ebenen, nominell einer persönlichen und einer universellen oder kosmischen, Wünsche und Begierden nach Vereinigung negiert.

Hans-Joachim Backe verweist darauf, dass die Handlung von *DEAD SPACE* meta-referenzielles Potenzial auf unterschiedlichen Ebenen besitzt und beklagt, dass die Forschungsliteratur zu *DEAD SPACE* diesen Aspekt bisher nicht dezidiert berücksichtigt hat: »[I]ts audiovisual presentation, its gameplay, and its cultural subtext – are seamlessly integrated with one another and central themes of the narrative« (Backe 2021: 206). Diese These, die in der Denktradition des Gesamtkunstwerks unausweichlich erscheint, weist jedoch Sollbruchstellen auf, wie im Folgenden gezeigt werden soll. So stellt Andrei Nae eine Interpretation der Ereignisse in *DEAD SPACE* als Kolonialismusparabel mit religionskritischer Note vor: »In *Dead Space*, the colonization of space is a functional project until religion interferes with the mission

and derails the Ishimura from its initial purpose« (Nae 2018: 165). Allerdings liefert er keine Verbindung mit der medialen Ontologie des Spiels.

Backe hebt die Kongruenz auch hervor und betont die Nähe des Videospiele als multimediale Assemblage zum historischen Konzept des Gesamtkunstwerks, indem er konstatiert:

DEAD SPACE [...] is much better characterized as a crystalline *Gesamtkunstwerk*. The palimpsest of many distinct communicative layers points the finger at the difficulty of discussing computer games in terms of »conventionally distinct media.« By complicating and defamiliarizing the mixture of media and messages, the game foregrounds the overly complex relationship between its elements that preclude a meaningful discussion in terms of media. (Backe 2021: 208f.)

Mit Bezug auf das übergeordnete Narrativ des Spiels sowie seine audiovisuell belastende und kakophonische Natur, erscheint es fraglich, DEAD SPACE als gelungene mediale Anordnung in der Tradition ganzheitlicher Kunst- und Medientheorien zu interpretieren. Viel eher scheint das Spiel selbst eine Lesart anzubieten, die seine Brüche, Widersprüche, Redundanzen, Dopplungen und Inkonsequenzen in eine metaphysische Ethik des Uneinheitlichen und Fragmentarischen überführt. So schreibt Backe: »The Marker's promise of ›being made whole again‹ is metaphorically shown to be unattainable and dangerous when forced. [...] The role of Nicole finally forms a nexus between narrative and mechanics« (Backe 2021: 211).

Abb. 3: Isaac und Nicole auf unterschiedlichen Seiten der Kluft.



Er identifiziert in Nicoles Sprache zudem

wording of doubtlessly intentional ambiguity, not only referring to the apparent situation – both of them being separated by an unbridgeable gap on the ship – but the metaphysical separation between the living and the dead. This thematic unsettling resonates in the gameplay, as Nicole can die in this situation, which results in a ›game over,‹ despite the fact that she is already dead at this point and her appearance is only a hallucination of Isaac's. (Backe 2021: 211f.)

Backe verweist somit mehrfach auf die eigentliche Unmöglichkeit der Vereinigung auf unterschiedlichen Ebenen, die den Kern der Erzählung von *DEAD SPACE* ausmachen (s. Abbildung 3).

4.1 Spielmechanische Trennungen

Neben dem narrativen Inhalt von *DEAD SPACE* ist die Spielmechanik eine dieser Ebenen. Das Spiel hat nicht nur im Bereich der Videospieforschung aufgrund seines sogenannten diegetischen Interfaces besondere Aufmerksamkeit erregt. Diegetisch bedeutet in diesem Zusammenhang die Integration spielrelevanter Informationen auf der Erzählebene. So werden u.a. gefundene Texte, Videoanrufe, Munitionsvorrat und Inventar als Hologramme vor die Avatarfigur projiziert:

Instead of pausing the game time and accessing an extradiegetic screen, in *Dead Space* the inventory is a projection of the playable character's suit [...]. This projection is, once again, part of the text-actual-world which emerges through gameplay. Activating the inventory does not draw players away from the storyworld of the video game, but rather keeps them immersed in the virtual storyworld of the game. In opposition to the hypermediacy that characterizes the inventory in traditional survival horror video games, *Dead Space* narrativizes the inventory so as to maintain the illusion of medium transparency. (Nae 2018: 162f.)

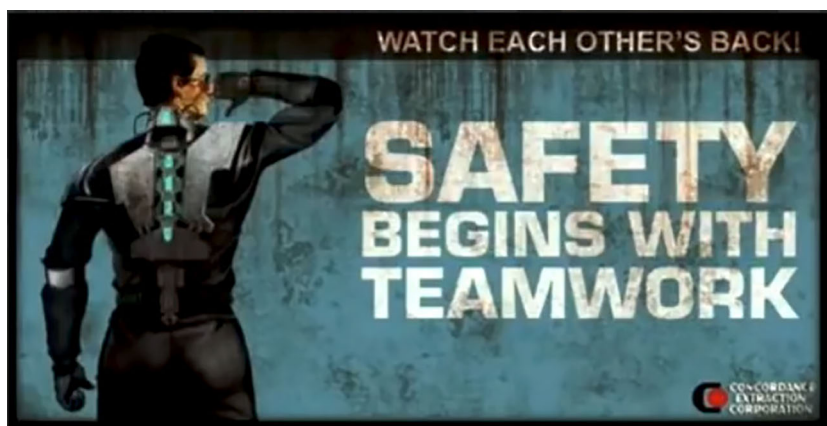
Nae betont hier die zeitliche Dimension potenziell spielunterbrechender Elemente wie Inventarbildschirme und deutet somit eine wichtige Funktion des diegetischen Interfaces jenseits narrativer Plausibilisierung an. Durch das ununterbrochene Fortlaufen der Spielhandlung im Hintergrund wird die Involvierung in die situativen Zusammenhänge der Spielwelt aufrechterhalten. Auch Benjamin Beil zufolge diene diese Interfacestrategie dazu, »den Tiefensog der diegetischen Spielwelt mit der spielerisch funktionaleren und flexibleren Bildgestaltung der Strukturbild-Elemente eines HUD zu verbinden« (Beil 2012: 110), vermöge aber »ihren spielfunktionalen Charakter kaum zu kaschieren« (ebd.). Worauf die letztgenannte, lakonische Bemerkung verweist, wird bei genauerer Betrachtung der Lebensenergieanzeige deutlich: Ein »leuchtende[r] Balken, der sich über den Rücken der Avatarfigur

erstreckt« behalte ebenfalls den Strukturbildstatus, auch zugunsten der Lesbarkeit, beeinträchtigt aber »die narrative Überformung und damit die vollständige Integration in die diegetische Spielwelt« (ebd. 118f.). Die Fragilität und Brüchigkeit dieser Interfacelogik bringt Backe mit einer folgenschweren Bemerkung auf den Punkt:

Dead Space's diegetic interface has been praised for its clarity and functionality [...]: eschewing extradiegetic overlays indicating the avatar's status, inventory, or goals, the game puts all this information into the gameworld as holograms projected by Isaac's protective space-suit. While this system allows for both a superficially more cinematic and more naturalistic impression, on closer inspection, it enforces a notion of epistemological opacity. Isaac's suit is a closed system which communicates with the player as an outside observer – a configuration stressed by the location of indicators on Isaac's back, where they can be seen by anyone but himself. (Backe 2021: 207)

Dieser Malus wird durch ein Sicherheitsplakat im siebten Level des Spiels ironisch aufgegriffen und reflektiert (Abbildung 4).

Abb. 4: Sicherheitsplakat, das die Interfacelogik ironisch reflektiert.



Die zugrundeliegende Positionslogik ist jedoch eine andere: Mit erhöhtem Grad an diegetischer Relevanz befinden sich die Elemente vor der Avatarfigur, sodass die Informationen, die für Isaac auf der Handlungsebene relevant sind, in seinem Blickfeld liegen. Dies liegt auch daran, dass die gewählte Perspektive konventionsbedingt ist. Als Third-Person-Shooter ist der Avatar in DEAD SPACE ständig im Bild präsent. Die Blickperspektiven von Spielenden und Avatar unterscheiden sich. Eine Plausibilisierung von Informationen über Lebensenergie, Munitionsvorrat und Videote-

lefonie über ein möglicherweise in einen Helm integriertes Head-up-Display, wie im Genre von First-Person-Shootern konventionalisiert, ist folglich nicht naheliegend.⁸ Als Folge dieser anderen Logik finden die Interfacelemente ihre Position im Raum jeweils in Abhängigkeit von der steuerbaren Avatarfigur. Im Extremfall liegen diese sogar an der Hinterseite des Avatarkörpers selbst, wie im Fall von Energieleiten in *DEAD SPACE*. Videobotschaften werden als zweidimensionale Hologramme vor der Spielfigur platziert, um auch in Ermangelung klassischer, das Spielgeschehen unterbrechender Filmsequenzen, an den medialen und narrativen Funktionen filmischer Bewegtbildmedien zu parasitieren. Dies produziert gleichzeitig Opazität und Störungen, in den Momenten, in denen sich beide Formen überlagern. Dasselbe gilt für Textelemente, die in Form von Tagebucheinträgen, Briefen und Ähnlichem ebenfalls vor die Spielfigur projiziert werden und dabei das zuvor Sichtbare semi-transparent befallen. Insbesondere die Filmsequenzen sind in diesem Zusammenhang durch eine Störästhetik markiert und verzerrt. Das Abgebildete kann zudem doppelt im Bild präsent sein, wenn die Quelle der Übertragung als Figur hinter einer Glasscheibe in der Spielwelt sichtbar ist, was das mediale Zeichenverhältnis von Abbild und Abgebildetem weiter verkompliziert (Abbildung 5).

Abb. 5: Holographische Dopplung der Nichtspielercharaktere.



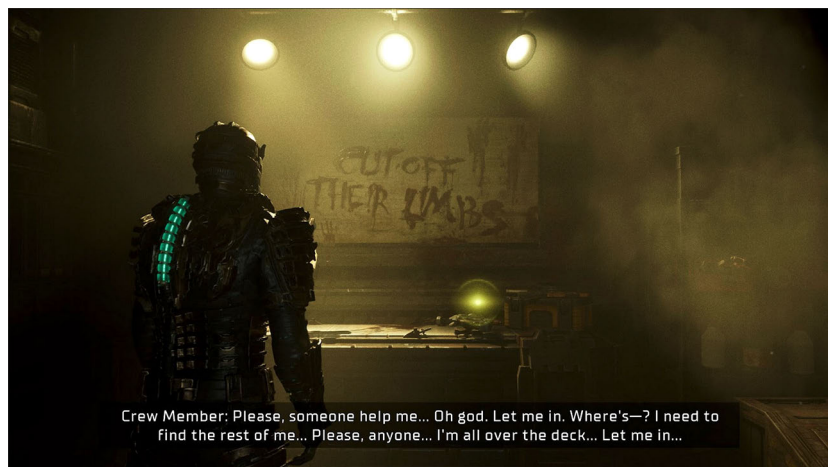
Dass das Ideal der Vereinigung von Interface und Spielwelt nicht konsequent funktionieren kann, merkt auch der führende UI Designer von *DEAD SPACE*, Dino

8 So wird beispielsweise regelmäßig mit Verweis auf kybernetische Implantate, Helmdisplays oder AR-Brillen erklärt, wieso in den simulierten subjektiven Kameraperspektiven vieler Videospiele Anzeigen über Lebensenergie, Munitionsvorrat und ähnliches sichtbar sind.

Ignacio, in einem Vortrag auf der Game Developers Conference im Jahr 2013 an (vgl. Ignacio 2013).

Die Tatsache, dass dieser Eindruck schon während der Produktion präsent war, lässt darauf schließen, dass es nicht konsequent darum ging, mithilfe dieser Interfacelogik ein homogenes und geschlossenes Ganzes zu konstruieren. Vielmehr reproduziert sich auf der Ebene der visuellen und akustischen Kommunikation die Unmöglichkeit von totaler Konvergenz, ganz im Sinne der Spielhandlung. Das Hauptmenü von *DEAD SPACE*, welches das Spiel pausiert und semi-transparent hinter dem Avatar positioniert ist, kann als deutlicher Beleg für diese Diagnose gelten. Das diegetische Interface von *DEAD SPACE* erweist sich also bei genauerer Betrachtung als durchaus uneinheitlich, getrennt, wenn nicht gar vereinzelt und hat damit auch Konsequenzen für die Avatarlogik. Diese kann weder in einem prothesen-theoretischen Verständnis als Erweiterung eines spielenden Subjekts aufgehen, noch körperphänomenologisch als Werkzeug zwischen Spielwelt und realer Welt (vgl. Klevjer 2006).⁹ Vielmehr kommt dem Avatar die Rolle eines Vermittlers mehrseitiger Einflüsse zu, der ständig unter dieser Masse auseinanderzufallen droht, was sich im Fail State des Spiels, also den ausführlichen Todesanimationen von Isaac Clarke, wiederfinden lässt. Diese Motive der Trennung und Vereinzelung spiegeln sich auch in der primären Spielmechanik von *DEAD SPACE* wider.

Abb. 6: Schriftlicher Verweis auf Spielmechanik und gleichzeitige Handlungsanweisung.



9 Vgl. auch folgendes Zitat in Michel Serres (1987:100f.) »Der Parasit«: »Vulgäre Banalitäten wie die Verlängerung der Körperorgane zu Hammer, Fernrohr oder Zange. Beinahe obszön – oder naiv: Wie wird diese Philosophie des Hackbeils, ich sage erst gar nicht einem Computer, sondern allein schon einer einfachen Maschine gerecht?«

Als Third-Person-Shooter liegt die privilegierte Kontrollfunktion von DEAD SPACE in der Steuerung seiner Avatarfigur. Das bewegliche Abbild Isaac Clarkes ist nicht nur zu jedem Zeitpunkt im Bild präsent, seine Bewegung selbst ist Grundbedingung und Kerninhalt der Spielhandlungen. Die Körperteile des menschlichen Avatars plausibilisieren die Aktionen, die durch Eingaben in den DEAD SPACE ausgeführt werden können, wie z.B. Laufen, Zielen, Schießen und Stampfen. Wird die Integrität des dargestellten Körpers und somit des Kontrollverhältnisses gestört oder ganz aufgelöst, endet damit auch die Steuerbarkeit des Avatars und führt zum absoluten *Fail State* des Spiels. Dies wird ebenfalls durch die explizite Natur der Protagonistentode vermittelt: »He [Isaac] is dismembered, then beheaded, then chewed, and then he explodes. His body leaks, splashes and splatters, goes to pieces, struggles and splits« (Carr 2013: 35). Jenseits dieser Dimension der Kontrollvermittlung wird der Avatakörper in DEAD SPACE nicht als kompetente Einheit gerahmt, sondern als Kompositum aus Schutzanzug und Mensch. »The suit is obviously armor. It is designed to keep things out [...]. It also keeps things in. Without his second skin, Isaac falls apart in very messy ways« (ebd.). Diese prothetische Haut (vgl. ebd.)¹⁰ ist wie das Waffenarsenal des Spiels an dafür vorgesehenen Bänken aufrüstbar und ist neben den wiederholten und vergebliehen Versuchen, das Minenschiff oder zumindest einzelne Teile davon zu reparieren und instand zu setzen, die einzige Repräsentation konstruktiver Kompetenz von Isaac Clarkes ausgewiesener Profession als Raumschiffingenieur.

Die zentrale Spielmechanik von DEAD SPACE ist jedoch eine andere. Bei ihr handelt es sich um *strategic dismemberment* (Schofield 2008; Andrews et al. 2024). Diese Mechanik bezieht sich auf das Kampfsystem des Spiels. Im Gegensatz zu der etablierten Konvention in Shootergenres, den Kopf feindlicher Entitäten als Schwachstelle und Zielpunkt zu priorisieren, kommuniziert DEAD SPACE unmissverständlich mit dem Auffinden eines laserbetriebenen Schneidwerkzeugs, der ersten Schusswaffe im Spiel, die Notwendigkeit, die grotesken Körperformen der Necromorph durch Abtrennen ihrer Gliedmaßen auszuschalten (Abbildung 6). Ergänzt wird diese primäre Kampfmechanik durch das Zerstampfen besiegter Necromorph, das Ressourcen wie Währung oder Munition freisetzt, während es die feindlichen Körper weiter zerteilt. Statt Maschinen zu konstruieren, obliegt es nun dem Ingenieur, biologisches Material zu trennen. Dieses *strategic dismemberment* der Necromorph beraubt sie ihrer Mobilität und trennt die Verbindung zu den spielinternen Berechnungsprozessen, die diese Nichtspielercharaktere steuern

10 Hierbei fällt auf, dass das von Carr beschriebene Verhältnis von Isaac und seinem Schutzanzug im Spiel auch begrifflich genau der Art von totaler Prothese entspricht, in dem das Verhältnis von Avatar und Spielenden nicht vollständig aufgeht. Ich danke Elizabeth Neumann für diesen Hinweis.

und kontrollieren, um im Gegenzug die Steuerbarkeit der Avatarfigur durch Aufrechterhaltung ihrer körperlichen Integrität zu gewährleisten. Backe hat dieses strategische Zergliedern als *différance*-Akt beschrieben:

A vivisection of the undead body that makes an impermeable whole manageable through segmentation. [...] [This] carries overtones of deconstructivism in it, with the violence committed against the individual part of the hivemind representing one step in the chain of *différance*, identifying one signifier that is [sic!] dissembles another one with every kill. (Backe 2021: 210)

Im Kontext der vorliegenden Arbeit ist festzuhalten, dass dieses Spielhandeln durchaus als Akt der Dekonstruktion verstanden werden kann, zumindest in einer dividualistischen Logik, welche die spielmechanischen Aktionen unter dem Gesichtspunkt der Aufspaltung, Trennung und Teilung unvereinbarer Einzelelemente versteht.

4.2 Abjekte Ästhetik

Diese dividualistische Logik hat jenseits ihrer spielmechanischen Dimension ebenfalls deutlich erkennbare Auswirkungen auf die Ästhetik von *DEAD SPACE*. Die Ästhetik des Spiels kann angesichts der folgenden Beobachtungen ebenfalls in mehrfacher Hinsicht als dividualistisch bezeichnet werden.

So konstatiert Backe für *DEAD SPACE* eine *Palimpsest-Ästhetik*, welche die Problematik des Videospieles als eher disruptive Assemblage unzähliger medialer Einflüsse denn als distinktes Medium reflektiert:

People converse with the avatar from behind splintered windows and in front of informational posters, and layers of graffiti and scrawls in different alphabets form actual palimpsests throughout the gameworld, resulting in a visual cacophony of unintelligible bits and pieces that echoes both the unreliability of the game's interface and its themes in the sheer texture of the environment. [...] Once this carefully orchestrated cacophony of communication has been noticed, it acts as an alienating impression that works towards evoking an aesthetic attitude in players. (Backe 2021: 208f.)

Dieser Verfremdungseffekt verdeutlicht die Ambivalenz der Ästhetik von *DEAD SPACE*. Die Masse der sich gegenseitig überlagernden, parasitär von Störungen befallenen und unzuverlässigen Zeichen unterschiedlicher medialer Natur (s. Abbildung 7), ist gleichermaßen Attraktionsmoment und Rätselkomponente des Spiels. Schließlich sind es diese Zeichen, die es auch im Kontext der Erzählung zu entschlüsseln gilt, obwohl sie bereits anfänglich als unzuverlässig markiert und ihre Interpretation und Bewertung problematisch bis unmöglich sind.

Abb. 7: Verzerzte Videobildübertragung als Störungsbefall.



Hinzu kommt, dass die Physiognomie und Ikonographie der Necromorph wiederholt mit dem Konzept des *Abjekts* in Verbindung gebracht worden sind:

The Necromorphs seem to embody in the most viscerally and graphically way conceivable, Kristeva's ›abject‹, that centers around the concepts of blood, body fluid, excrements and corpses. The Necromorphs embody Kristeva's ›abject‹ since, as the characters die a bodily death in terms of their loss of a normal, human bodily embodiment, they also in the process, give birth to something new, i.e. the hideous Necromorph. It seems like a most monstrous illustration of Kristeva's process of becoming the abject. (Patra 2021: 475f.)

Dieses Abjekte in der Ästhetik des Spiels ist ebenfalls als ambivalent zu bewerten, da sie einerseits durch die inszenierte Ontogenese der Necromorph und andererseits durch die Spielmechanik des strategischen Zergliederns um eine audiovisuell spektakuläre Note ergänzt wird.

Was beide Ansätze gemeinsam haben, ist ihre dividualistische Logik. Die abjekt Form der Necromorph stellt ein Spannungsverhältnis dar, das nur durch Zerteilung aufgelöst werden kann (s. Abbildung 8): »These undead yet agile monsters are made from twisted and recycled human flesh. Legs wobble on shoulders. Heads pop out of chests. Tentacles spring from spines. Freud's uncanny, Kristeva's abjection and Bakhtin's grotesque are all applicable« (Carr 2013: 34).

Dabei ist das Abjekte als repulsives Horrormotiv selbst mit Bestandteilen der Separation und Teilung ontologisch verbunden: »The abject would thus be the ›object‹ of primal repression. But what is primal repression? Let us call it the ability of the

speaking being, always already haunted by the Other, to divide, reject, repeat« (Kris-teva 1982: 12).

Teilen, Abstoßen, Wiederholen sind fundamentale Elemente sowohl der multi-medialen Zeichenkakophonie als auch der Konfrontation mit den Necromorph und ihrer transformativen Körperlichkeit. Beide Seiten der bisherigen ästhetischen Interpretation von *DEAD SPACE* können also als Signifikanten einer Ästhetik der un-einheitlichen Vielheit von Brüchen und Trennungen verstanden werden. Eine Ästhetik, die unter Voraussetzung der Unmöglichkeit einer harmonischen Verschmelzung und Ganzwerdung das Prinzip der Teilung als ontologische Existenzweise erfahrbar macht.

Abb. 8: *Abjekter Necromorph.*



5. Fazit

CULTIST SIMULATOR und *DEAD SPACE* ermöglichen einen ersten Einblick in die Bandbreite der dividualistischen Dimension von Videospiele. Während die Spielfigur in *DEAD SPACE* über den gesamten Spielverlauf damit beschäftigt ist, mehr oder weniger körperlich unversehrt zu bleiben, sich nicht selbst der Teilung zu unterziehen und stattdessen die teilenden Handlungsoperationen nach außen zu tragen, lässt sich anhand der Spielfigur in *CULTIST SIMULATOR* eine Inversion dieses Verhältnisses von Körper und Raum beobachten. Gerade deshalb ergänzen sich beide Spiele in ihrer je eigenen dividualistischen Lesart und bilden somit eine große Bandbreite möglicher Dividuationen des Spielprozesses ab. *CULTIST SIMULATOR*

offenbart das Prozessuale der Spielerfahrung, indem es seine Darstellungsweisen der bildlichen Tiefe beraubt und damit sein Regelwerk berührbar werden lässt. Teilungen und Verteilungen werden hier zu konzeptionellen Erfahrungen, die im Vergleich zu den körperlich intensiven Teilungshandlungen in *DEAD SPACE* weitaus weniger ekstatischen Schrecken bereithalten.

DEAD SPACE auf der anderen Seite zelebriert das viszerale Moment der spielerischen Dividuation und definiert das Dividuiere als eines der zentralen Handlungselemente. Beide Untersuchungsgegenstände weisen in spieltypischer Manier, raumzeitlich komprimiert, auf unsere bereits eingangs mit Ott und Raunig erwähnten Umstände des zeitgenössischen Seins und der Entledigung alter Existenzweisen hin und offenbaren darüber hinaus die Praxis der Dividuation in spielerischer Form. Hierdurch wird ebenfalls deutlich, dass an Stellen wie z.B. Interfaces, der Spiegelung der Spielsituation der Spielenden oder der Offenlegung von Spielstrukturen das Unreale und das Reale in Videospiele zusammengebracht werden. So verschwimmen im Spielprozess die Grenzen zwischen Virtuellem und Tatsächlichem und es eröffnen sich Interaktionsräume, durch die sowohl die Handlungen der Spielenden, als auch die des Spiels an sich gleichermaßen von Bedeutung sind.

6. Schlussbemerkungen

Die vorangegangenen Untersuchungen haben zwei Videospiele als un/reale Möglichkeits- und Interaktionsräume begriffen. Hierzu wurden konkrete Aspekte und Potenziale beider Vertreter unter den Begriff der Dividuation gestellt. Unter anderem wurde die dividualistische Dimension von Erzählung, Spielmechanik und Ästhetik des Videospiele anhand von *CULTIST SIMULATOR* und *DEAD SPACE* in den Blick genommen. Mithilfe einer detaillierten Einzelbetrachtung ausgewählter Elemente dieser Spiele unter Berücksichtigung des gegenwärtigen Forschungsstandes und theoretischer Überlegungen wurde eine spezifische Qualität des medien-spezifischen Interaktionsraumes Videospiele herauszuarbeiten versucht. Videospiele wurden als offene Assemblage auf ganz unterschiedlichen Ebenen begriffen. Zudem wurde der Frage nachgegangen, welche epistemischen Potenziale erschlossen werden können, wenn der holistischen Lesart, die das Medium als Gesamtkunstwerk oder Teil einer Konvergenzkultur verstehen (vgl. Backe 2021; Jenkins 2006), mit einer Perspektive der radikalen Teilung und Geteiltheit begegnet wird. Es zeigt sich, dass verschiedene Spiele unterschiedliche Aspekte dieses mehrdimensionalen Charakteristikums u.a. in Bezug auf Ästhetik und Medialität des Mediums Videospiele offenlegen.

Literatur

- Andrews, Nate/PlamzDooM/Jesuhibu (2024): Strategic Dismemberment. Abrufbar unter: <https://www.giantbomb.com/strategic-dismemberment/3015-418/> (Stand: 07.05.2024).
- Backe, Hans-Joachim (2008): Strukturen und Funktionen des Erzählens im Computerspiel. Eine typologische Einführung, Würzburg: Königshausen & Neumann.
- Backe, Hans-Joachim (2021): »Gesamtkunstwerk in Digital Games. The Palimpsestine Aesthetics of *Dead Space*«, in: Massimo Fusillo/Marina Grishakova (Hg.), *The Gesamtkunstwerk as a Synergy of the Arts*, Brüssel: Peter Lang, S. 197–215.
- Beil, Benjamin (2010): *First Person Perspectives. Point of View und figurenzentrierte Erzählformen im Film und im Computerspiel*, Münster/Hamburg/Berlin/London: LIT.
- Beil, Benjamin (2012): *Avatarbilder. Zur Bildlichkeit des zeitgenössischen Computerspiels*, Bielefeld: transcript.
- Carr, Diane (2013): »Bodies augmentation and disability in *Dead Space* and *Deus Ex: Human Revolutions*« [sic!], in: Konstantin Mitgutsch/Simon Huber/Michael Wagner/Jeffrey Wimmer/Herbert Rosenstingl (Hg.), *Context Matters! Exploring and Reframing Games in Context Proceedings of the 7th Vienna Games Conference FROG 2013*, Wien: new academic press, S. 31–41.
- Carr, Diane (2014): *Ability, Disability and Dead Space*. Abrufbar unter: <https://gamestudies.org/1402/articles/carr> (Stand: 31.01.2024).
- Cyberpunk 2077 (2020): *Cyberpunk 2077 — Seize the Day*. Abrufbar unter: <https://www.youtube.com/watch?v=U8qJc6znzZc> (Stand: 07.05.2024).
- Deleuze, Gilles/Félix Guattari (1992): *Kapitalismus und Schizophrenie: Tausend Plateaus*, Berlin: Merve Verlag.
- Didi-Huberman, Georges (2018): *Atlas, or the Anxious Gay Science*, Chicago/London: The University of Chicago Press.
- Genette, Gérard (1980): *Narrative Discourse. An Essay in Method*, Ithaca/New York: Cornell University Press.
- IGN (2023): *Starfield – Todd Howard Stage Presentation | gamescom 2023*. Abrufbar unter: <https://www.youtube.com/watch?v=LoVcl8hbcDk> (Stand: 07.05.2024).
- Ignacio, Dino (2013): *Crafting Destruction. The Evolution of the Dead Space User Interface*. Präsentation auf der Game Developers Conference 2013. Abrufbar unter: <https://www.youtube.com/watch?v=pXGWJRV1Zoc> (Stand: 31.01.2024).
- Jenkins, Henry (2006): *Convergence Culture. Where Old and New Media Collide*, New York/London: New York University Press.
- Keogh, Brendan (2018): *A Play of Bodies. How we Perceive Videogames*, Cambridge: MIT Press.
- Klevjer, Rune (2022): *What is the Avatar? Fiction and Embodiment in Avatar-Based Singleplayer Computer Games*, Bielefeld: transcript.

- Kristeva, Julia (1982): *Powers of Horror. An Essay on Abjection*, New York: Columbia University Press.
- Liebe, Michael (2008): Die Dispositive des Computerspiels, in: Jan Distelmeyer/Christine Hanke/Dieter Mersch (Hg.): *Game over!? Perspektiven des Computerspiels*, Bielefeld: transcript, S. 73–94.
- McMahan, Alison (2003): Immersion, Engagement, and Presence. A Method for Analyzing 3-D Video Games, in: Mark Wolf/Bernard Perron (Hg.): *The Video Game Theory Reader*, New York: Routledge, S. 67–86.
- Mosel, Michael (2009): Das Computerspiel-Dispositiv, in: Michael Mosel (Hg.): *Gefangen im Flow? Ästhetik und dispositive Strukturen von Computerspielen*, Boitzenburg: Verlag Werner Hülsbusch, S. 153–179.
- Murray, Janet H. (1997): *Hamlet on the Holodeck. The Future of Narrative in Cyberspace*, Cambridge: MIT Press.
- Nae, Andrei (2018): »Mission Objective. Carry the White Man's Burden to Outer Space – The Gamification of Colonization in *Dead Space*«, in: *Ekphrasis. Cinema, Cognition and Art 2* (2018), S. 157–167.
- Neitzel, Britta (2017): Involvierungsstrategien des Computerspiels, in: GamesCoop (Hg.): *Theorien des Computerspiels zur Einführung*, Hamburg: Junius Verlag, S. 75–103.
- Newman, James (2004): *Videogames*, New York: Routledge.
- Ott, Michaela (2018): »Dividuationen – des Films, des menschlichen Subjekts«, in: Lars Friedrich/Karin Harrasser/Céline Kaiser (Hg.), *Szenographien des Subjekts*, Wiesbaden: Springer Fachmedien Wiesbaden, S. 167–177.
- Patra, Indrajit (2021): »Horror Within and Without. Studying the Role of the Monstrous and the Grotesque in the Horror World-building of *Dead Space Universe*«, in: (TOJQI) *Turkish Online Journal of Qualitative Inquiry 12*, S. 454–486.
- Peric, Bojan (2016): Das dreifache Immersionssubjekt. Abrufbar unter: <https://paidia.de/das-dreifache-immersionssubjekt-subjektspaltung-und-konstruktion-im-computerspiel/> (Stand: 30.01.2024).
- Raunig, Gerald (2015): *Maschinischer Kapitalismus und molekulare Revolution*, Wien: Transversal Texts.
- Ryan, Marie-Laure (2001): *Narrative as Virtual Reality. Immersion and Interactivity in Literature and Electronic Media*, Baltimore, MD: Johns Hopkins.
- Schofield, Glenn (2008): *Dead Space. Strategic Dismemberment*. Abrufbar unter: <https://www.youtube.com/watch?v=cn4XfU1Vt5Y> (Stand: 07.05.2024).
- Serres, Michel (1987): *Der Parasit*, Frankfurt a.M.: Suhrkamp.
- Thon, Jan-Noël (2008): Immersion Revisited. On the Value of a Contested Concept, in: Amyriss Fernandez/Olli Leino/Hanna Wirman (Hg.), *Extending Experiences. Structure, Analysis and Design of Computer Game Player Experience*, Rovaniemi: Lapland University Press, S. 29–43.

Unterhuber, Tobias (2018): »Von Regenschirmen und Nähmaschinen. Das Computerspiel und sein Raum als Tableau des Wissens«, in: Martin Hennig/Hans Krah (Hg.), *Spielzeichen II. Raums Spiele/Spielräume*, Glückstadt: Verlag Werner Hülsbusch, S. 204–211.

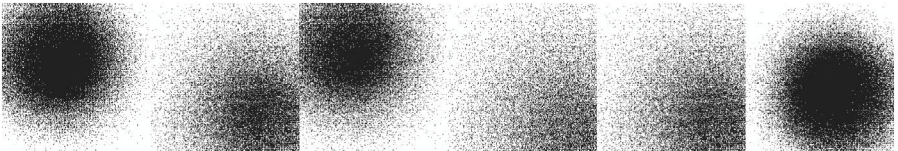
Filmverzeichnis

ALIEN (Ridley Scott 1979: 20th Century-Fox).
 EVENT HORIZON (Paul W. S. Anderson 1997: Paramount).
 THE THING (John Carpenter 1982: Universal).

Spielverzeichnis

CULTIST SIMULATOR (Weather Factory 2018: Humble Bundle, Linux/OS X/Windows/Android/iOS/Switch).
 CYBERPUNK 2077 (CD Projekt Red 2020: CD Projekt, Windows/PlayStation 4/Xbox One/Google Stadia//PlayStation 5/Xbox Series X/S).
 DEAD SPACE (EA Redwood Shores 2008: Electronic Arts, PlayStation 3/Xbox 360/Windows).
 DEAD SPACE (Motive Studios 2023: Electronic Arts, PlayStation 5/Xbox Series/Windows).
 STARFIELD (Bethesda Game Studios 2023: Bethesda Softworks, Windows, Xbox Series X/S).

Sektion 3: Partizipation





Excuse Us While We Improve Your View, Atlantis

Giacomo Marinsalta

Keywords *Politische Ästhetik; Visuelle Kunst; Performative Räume; Urbane Medien-Ökologie; Video-essay*

Atlantis where are you?

Confined to the sea, your islands lie on the horizon

They form the ruins of a world that travels from lands to countries
holding a passport for passengers only

Atlantis where will you go?

A single day of misfortune and you vanished away
leaving only images to fill the void where you once lay.

I have seen you elsewhere, Atlantis and maybe I have found you here

confined to the see your islands lays on the horizon
halfway built, they form the ruins of a passenger world
everlasting fiction of a backward progress

Atlantis where are you?

No longer here
Gone in but a single day
now the images flow and they reveal us your absence

Excuse us while improve your view, Atlantis but we can't see you anymore
time and distance dissolved

your grounds are sinking your water
bloomed and now recedes

all lives streaming down together

Atlantis where will you go?

Your life hangs like a thread in the deep currents' creed

A single day of misfortune and you vanished away

leaving only images to fill the void where you once lay.

Der experimentelle Kurzfilm ›Excuse Us While We Improve Your View, Atlantis‹ überträgt die mythische Vorstellung des verlorenen Inselreiches Atlantis auf die Ästhetik der internationalen Freihandelszonen und des Sightseeings rund um die Expo 2020 Dubai. Dabei wird dokumentarisches Filmmaterial einer Rechercheise zur Weltausstellung kontrastiert. Das Video-Essay führt die Zuschauer*innen von der Ankunft mit dem Flugzeug in Dubai durch die Innenstadt und zeigt die Reise zum Veranstaltungsort.

Der Kurzfilm deutet in seiner traumhaften und avantgardistischen Erzählung rund um den kommerziellen Charakter der Expo auf eine Verbindung mit dem Mythos von Atlantis hin, einem mächtigen Stadtstaat, der außerhalb des Atlantiks lag und ein Gegenbeispiel für die Politik der Polis von Athen darstellte. Die Selbstverherrlichung und die gut verpackte Vermarktung wirtschaftlicher Interessen für eine nachhaltigere Gesellschaft bei solchen internationalen Großereignissen schaffen Assoziationen zum Schicksal des utopischen Atlantis, das an seinem zivilisatorischen Höhepunkt als Idealstaatsmacht von Poseidon im Meer versenkt wurde, nachdem seine Bürger*innen zu gierig und machthungrig geworden waren. Denn: Die meisten Orte, an denen in der Vergangenheit Weltausstellungen stattfanden, verblissen in der Erinnerung zu grandiosen, temporären Ereignissen.

Das Video verknüpft die Moral der mythologischen Legenden aus Platons Dialogen ›Timaios‹ und ›Kritias‹ mit der Bildsprache des Dubaier Luxus-Immobilienmarktes und Tourismus, wobei eines der größten Hotels des Emirats, genannt ›Atlantis‹, als Hauptfigur dient. Das begleitende Gedicht deutet auf ein Gefühl der Verdammnis hin, das sich in den Bildern zeigt, die über den Bildschirm fließen, und vermischt persönliche Reflexionen über das katastrophale Ende von Atlantis einschließlich der Zusammenbrüche künstlicher Inseln in Dubai oder die Entstehung riesige vertikaler Infrastrukturen, die aus der Ölwirtschaft des Emirats hervorgehen. Die zitierten Worte Platons »Ein einziger Tag des Unglücks bist du verschwunden« wecken einen Sinn für Vergangenheit oder düstere Zukunft, die sich zirkulär bedingen. Das idyllische Empire fungiert in dem Vergleich mit der Stadt Dubai als Symbol für neue soziale Ordnungen und Grenzen des modernsten kosmopolitischen Urbanismus: Zuschauende werden zu Tourist*innen und deren spektatorische Betrachtung wirkt wie bei einer Reiseführung.



GO
somewhere
exotic

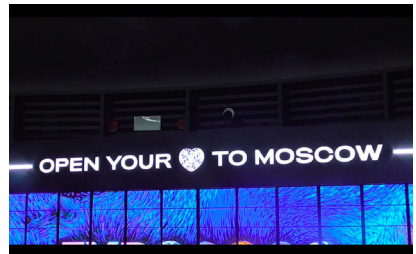
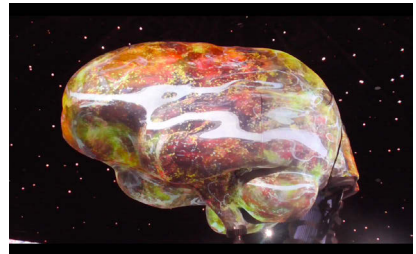
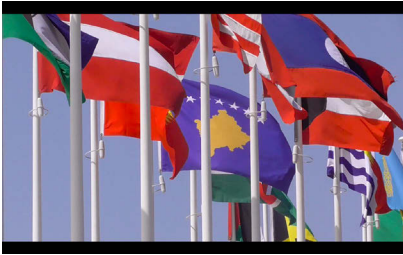


Experience the
EXCEPTIONAL



Thematic Pavilions
A World of Entertainment
Page 1 of 2





Das neue Atlantis besteht aus Unterwasser-Spaß, rekordbrechenden Wolkenkratzen oder künstlichen Landschaftsprojekten, wie dem sinkenden Archipel der ›World Islands‹. Das Reich der Menschen ist nicht von Wissenschaft vergrößert, wie Francis Bacons Utopie es in ›The New Atlantis‹ von 1623 skizziert, sondern aus Gründen der Bequemlichkeit zusammengeschrumpft als bedienungsfreundliches, aalglatte, glitch-freies Taucherlebnis auf 40 Hektar Land inmitten der Wüste. Die morbiden Fantasien einer reibungslosen Welt vor dem Schatten der Klimakrise werden in zahlreichen eklatanten Werbespots der Expo genussvoll ausgebreitet: Als Bühne für Unterhaltung und Konsumismus, hier immersive Erfahrungen und Rhetorik über Nachhaltigkeit verzaubern romantisch virtuell. Globale Herausforderungen an Mobilitätsinfrastrukturen und Produktionsautomatisierung propagieren eine verheißungsvolle Zukunft und werden mit leeren Phrasen über Empathie als Werkzeug für Erfolg oder Energie als universelle Kraft hinter der Menschheit vorgestellt.

Fortschritt bedeutet bei der Expo Dubai 2020 nicht Kommunikation, sondern eben nur Spaß zum Preis künstlicher Stille, für den sogenannten Frieden. Jede Form von Konflikt wird durch hohe Sicherheitszäune und das Verbot jeglicher Arten von Protest unterbunden. Reisepässe verwandeln sich in Marketingstrategien und modische Accessoires, die scheinbare Bewegungsfreiheit in ›the world in one place‹ suggerieren. Jeder, der für dieses Gadget bezahlt, kann einen Stempel von verschiedenen Länder-Palästen erhalten, alles Teil der Marke Expo. Zwar investiert die Wirtschaftsdiplomatie Millionen in großartige Pavillons als Alleinstellungsmerkmale, doch die reine nationalistische Darstellung der Welt ist nicht immer wirklichkeitsgetreu: So werden in den Pavillons von Syrien, Palästina und Russland klare Aussagen über deren politische Realität öffentlich nicht ausgestellt. Die Expo, die ursprünglich für das Jahr 2020 geplant war, wurde aufgrund der Covid-Pandemie auf das Jahr 2022 verschoben und fiel somit in die Zeit des russischen Angriffs auf die Ukraine. Dies scheint jedoch dem omnipräsenten Unterhaltungsversprechen in dieser politikfreien Zone keinen Abbruch zu tun.

Alle Versprechen von Fortschritt erscheinen viel mehr wie verwaschene retrofuturistische Bilder aus den 1980er Jahren. Einige Slogans der Expo erinnern sogar eher an die 1950er Jahre, wie im russischen Pavillon, wo die Installation eines großen menschlichen Gehirns neben dem Satz ›Nuclear for a better life‹ auftaucht. Der Pavillon der Vereinigten Staaten bleibt rhetorisch kohärent und präsentiert mögliche Lösungen, unseren Planeten mit der Mission von Elon Musks SpaceX zu verlassen. Riesige Wasserfontänen schießen in den Himmel, doch im Video laufen sie rückwärts: In dieser Version von Atlantis scheint alles zurück in Richtung Vergangenheit zu fließen. Die Expo selbst fungiert zwar als machtvolle Plattform für diplomatische Begegnungen, die es Nationen ermöglichen, öffentliche Unterstützung für ihre Handelsaktivitäten zu gewinnen, präsentiert aber gleichzeitig die aus dem

Kolonialismus resultierenden, technologischen und industriellen Fortschritte, ohne diesen Kontext offenzulegen.

Der Mensch wird in diesem Atlantis zur bloßen Kulisse für die Videos und Bilder anderer Leute und das Einzige, was in der heutigen Expo wichtig zu sein scheint, ist die Sicht auf die Dinge und nicht ihre eigentliche Gestalt. Wie soll in den örtlichen Interaktionsräumen Gemeinschaft entstehen, wenn das Publikum zu reinen Zuschauenden und die Welt in einen vermeintlichen Vergnügungspark transformiert wird?

Das Gemeinschaftsgefühl, das die Expo fördert, manifestiert sich nur durch die Teilnahme an immersiven Erlebnissen. Passant*innen entdecken nationalistische Pavillons, in denen die Interaktion auf kulturelle Lernformen durch digitale Erfahrungen wie Audio- und Videospaziergänge beschränkt ist, wodurch der Austausch zwischen den Teilnehmenden auf bloße Beobachtung und Staunen reduziert wird. Einerseits zieht die Veranstaltung Tourist*innen aus der ganzen Welt an, auf der anderen Seite verwandelt sich das Konzept von »the world in one place« in ein Schema von Spaziergängen, Warteschlangen und Flanieren. Eine Szene, in der Menschen eine Bank fotografieren, auf der Statuen von Menschen sitzen, die auf ihre Smartphones schauen, dient als euphemistisches Beispiel für die interaktiven Räume, die rund herum stattfinden. Die majestätische Aussicht auf Orte wie das Hotel Atlantis, die Worlds Islands und das Gelände der Expo versucht allegorisch das Bild eines alten Reiches wiederzugeben, wobei die innovativen Seiten jedoch mit ihrer Exklusivität kollidieren. Die Spannung zwischen Vergangenheit und Zukunft wird zudem durch das Sprühen von Fontänen verstärkt, die rückwärtslaufen. Wasser fasziniert die Zuschauer*innen und seine wundersamen Ströme können über das Bildschirm verfolgt werden- von kolossalen Wasserspielen bis hin zu künstlichen Wasserfällen wird die Aufmerksamkeit durch den technologischen Blick gefangen genommen, der sich auf das kleinste Detail konzentriert, als einzige Möglichkeit, sich vor dem fatalen Schicksal von Atlantis zu retten und seine Verjüngung inmitten der Wüstenlandschaft von Dubai und der Expo zu imaginieren.

Als zärtlicher sozialer Kommentar zum kapitalistischen Wachstum angelegt, hinterfragt »Excuse Us While We Improve Your View, Atlantis« die vereinfachenden Logiken des technologischen Glaubens als Grundlage für eine gemeinsame Zukunft. Allegorische Vergangenheit und innovative Visionen verschmelzen in einem Reisebericht, der die illusorische und exotisierte Vorstellungskraft eines glückseligen »Woanders« anzweifelt. Der Zusammenhang zur mythologischen Geschichte einer Ruine ermöglicht eine kritische Perspektive auf die Reise, durch eine (alb-)traumhafte Miniaturwelt, der ihr eigener Verfall bereits eingeschrieben zu sein scheint: Die Expo 2020 in Dubai wird als neoliberale Dystopie entlarvt, die nur virtuell, als ihr eigenes Werbevideo funktionieren kann. Atlantis bleibt dann vielleicht nur ein flüchtiges Bild, eine Vision des Fortschritts, verknüpft mit nationalistischen Träumereien, stets nach Perfektion strebend.

Abbildungen

1. Spalte:

Abb. 1: Werbespot des Hotels Atlantis, »Atlantis – The Palm Dubai«

Abb. 2: »Go somewhere exotic and even hypnotic« – Werbespot von Emirates

Abb. 3: Werbespot der Expo

Abb. 4: »Experience the exceptional« – Werbespot von Emaar

Abb. 5: »Dreams begin here« – Werbespot von Emaar

2. Spalte:

Abb. 6: Werbespot des Hotels Atlantis, »Atlantis – The Palm Dubai – A world away from your everyday«

Abb. 7: »Welcome to the world's greatest show« – Werbespot von Emirates

Abb. 8: »A world of entertainment« – Touchscreen-Rubrik über die Expo auf Flugzeugbildschirmen

Abb. 9: »Priceless« – Mastercards Werbespot für die Expo

Abb. 10: »The Heart of Europe-Dubai« – Detail aus der Google Earth-Ansicht der »World Islands«

3. Spalte:

Abb. 11: Fahnenreihe außerhalb des Expo-Eingangs

Abb. 12: Der »Explorer«-Zug

Abb. 13: »200 Hektar umzäuntes Gelände«

Abb. 14: Pavillon von Syrien

Abb. 15: Pavillon der Vereinigten Staaten mit dem SpaceX-Projekt

4. Spalte:

Abb. 16: Verbot politischer Schilder und Schriftzüge

Abb. 17: Reisepässe mit Expo-Branding

Abb. 18: Eine Installation des russischen Pavillons

Abb. 19: »Open your heart to Moscow« am Eingang des russischen Pavillons

Abb. 20: Pavillon des Staates Palästina

Mobile Crowdsensing

Interaktionsmodelle des mobilen Sensing und die infrastrukturelle Allgegenwart der ›mobile sensor networks‹

Vesna Schierbaum

Abstract *Beim Mobile Crowdsensing wird in großem Maße auf sensorische Daten zurückgegriffen, die Nutzer*innen mit ihren mobilen Devices produzieren. Auch wenn in Industriediskursen immer wieder auf die partizipatorischen Potenziale der Technik hingewiesen wird, ist das Mobile Crowdsensing stärker im Kontext des Pervasive Computing anzusiedeln. Ihm liegt in der Regel ein Interaktionsmodell zugrunde, in welchem Datenverarbeitungsprozesse unterhalb der Wahrnehmungsschwelle der Nutzer*innen ablaufen, um deren Interaktion mit den mobilen Endgeräten auf interfacegestützte Service-Funktionen zu beschränken. Der Beitrag versteht sich als eine kritische medienwissenschaftliche Hinführung zur Technik des Mobile Crowdsensing, das auf der Prämisse einer Allgegenwart mobiler sensorischer Infrastruktur aufbaut. Tatsächlich handelt es sich dabei um das Zusammenkommen verschiedener Technologien, Infrastrukturen, Medien und Wissensbestände, was die Frage nach einer Verflechtungsgeschichte des mobilen Sensing aufwirft.*

Keywords *mobile sensor networks; Partizipation; Kontrolle; Mobilität; Pervasive Computing*

1. Einleitung

Der Begriff ›Mobile Crowdsensing‹ taucht erstmals 2011 in einer Publikation von drei IBM-Mitarbeitenden auf – neben einer Reihe von weiteren Artikeln, die sich mit der Implementierung und den Einsatzgebieten des *Internet of Things* auseinandersetzen. Ganti et al. betonen gleich zu Beginn des Aufsatzes die Relevanz des Mobile Crowdsensing für ›the evolution of the Internet of Things‹ als emergierende Technik und Innovation:

An emerging category of devices at the edge of the Internet are consumer-centric mobile sensing and computing devices, such as smartphones, music players, and in-vehicle sensors. These devices will fuel the evolution of the Internet of

Things as they feed sensor data to the Internet at a societal scale. In this article, we examine a category of applications that we term mobile crowdsensing, where individuals with sensing and computing devices collectively share data and extract information to measure and map phenomena of common interest. (Ganti et al. 2011: 32)

Ganti et al. verweisen auf die rasche Ausbreitung mobiler und mit Sensoren ausgestatteter Computer, die vor allem von Endverbraucher*innen verwendet werden und aufgrund ihrer Internet-, Rechen- und Kommunikationsfähigkeit mehr noch als stationäre Geräte Lieferanten für wichtige sensorische Daten über die physische Umwelt darstellen. Ihr Mehrwert liege darin, Relationen zwischen einer Vielzahl physisch voneinander unabhängig agierender Individuen und den von ihnen mit den mobilen Devices produzierten sensorischen Daten zu erfassen, die somit die Grundlage für sich anschließende Wissensprozesse über Umweltphänomene größeren Maßstabes liefern.

Entlang rezenter Industriediskurse zum Mobile Crowdsensing soll in den folgenden Ausführungen eine medienwissenschaftliche Hinführung zur Technik vollzogen werden. Dabei werden zunächst die in Publikationen evozierten Verheißungen auf Partizipation und Demokratisierung einer kritischen Betrachtung unterzogen. Denn im starken Kontrast zu diesen Versprechen stehen die ebenso prominenten Aussagen der Entwickler*innen darüber, dass sich die massenhaft verfügbare Hardware dazu eigne, unbemerkt auf sensorische Daten zuzugreifen. Nutzer*innen sollen gerade nicht bewusst an der Datenerhebung und -verarbeitung teilnehmen, sondern ihre Interaktionen mit den Geräten und ihren Umgebungen sollen selbst Gegenstand skalierbarer Datenerfassungen werden. Deshalb werden anschließend die betreffenden Kontrolldiskurse in den Blick genommen, in denen die ›Crowd‹ als infrastrukturell gegebenes mobiles Sensornetzwerk entworfen wird, das durch die Implementierung von Modellen zum Zwecke der Datenerhebung instrumentalisiert werden soll. Von diesen Beobachtungen ausgehend wird argumentiert, dass die Technik des Mobile Crowdsensing vor dem Hintergrund einer Verflechtungsgeschichte verschiedener Technologien, ihrer Infrastrukturen, Wissensbestände und Praktiken Betrachtung finden muss. Zum Schluss werden mögliche Entwicklungstrajektorien und Einsatzpunkte dieser Verflechtungsgeschichte aufgezeigt.

2. ›Aktive‹ und ›opportunistische‹ Partizipation

Ganti et al. (2011) schließen mit ihrem Begriffsvorschlag an bereits existierende Diskussionen um die epistemischen Potenziale sowie die gesellschaftliche Relevanz mobiler Sensoren an. Ein zentrales Argument stellt seither die durch mobile Sen-

soren ermöglichte Partizipation von Nutzer*innen an den Wissensprozessen dar. Solche Bottom-Up-Modelle der sensorischen Datenerhebung firmieren unter dem Konzept des Participatory Sensing und haben in Schlagwörtern wie dem *Citizen Sensing* oder *Urban Sensing* heute auch in der medienwissenschaftlichen Forschung Resonanz gefunden (Burke et al. 2006; Boulos et al. 2011). So hat etwa Jennifer Gabrys 2013 das an der University of Cambridge assoziierte Projekt *Citizen Sense*¹ ins Leben gerufen, in dem untersucht wird, inwiefern die Interaktion mit Sensoren *in situ* die Beziehung von Bürger*innen zu ihren Umwelten und darüber hinaus entsprechende Verantwortlichkeiten beeinflussen kann (vgl. Gabrys 2016: 305). Die Verwendung von Sensoren durch Forscher*innen und Citizen Scientists versteht Gabrys als Antwort lokaler Milieus auf das Problem einer sich verändernden Umwelt (vgl. ebd.: 99). Die sich der Datenerhebung anschließende, kollaborative Auswertung und Kommunikation von Messergebnissen wird bei Gabrys zur Problembewältigungs- und Ermächtigungsstrategie für Bürger*innen, die sich für den Erhalt ihrer Lebensräume einsetzen. Gabrys kritisiert andererseits solche Smart City-Projekte, innerhalb derer Citizenship auf ein Bündel operationalisierter, datafizierter Interaktionen von Bürger*innen mit vernetzten Sensortechnologien und smarten Umgebungen reduziert wird, die dann vor allem Unternehmen und Verwaltungen ökonomisch und politisch nutzbares Wissen liefern (vgl. ebd.). Daran anschließend haben Daniela van Geenen und Timo Kaerlein (2023) erst kürzlich argumentiert, dass Praktiken des Urban Sensing auf Grundlage der »soziomateriellen Involviertheit« von Bürger*innen in die technischen und datenbezogenen Wissensprozesse bewertet werden müssen. Ein kritisches Potenzial hätten demzufolge insbesondere solche Initiativen, bei denen »sensing citizens« über die Situation der Datenproduktion hinaus mit interaktiven Praktiken des Bastelns und Experimentierens mit Sensoren als »offene« Objekte in Berührung kämen und an der Datenauswertung als kollektives »sense-making« beteiligt seien (vgl. ebd.: 93, 95, 103; vgl. Mondada 2021).

Derartige Projekte firmieren unter dem Begriff des Mobile Crowdsensing seit Kurzem auch in Deutschland. Beispielhaft kann hier ein 2020 am Bodensee durchgeführtes, vom baden-württembergischen Umweltministerium gefördertes Projekt mit dem Namen »Crowdsensing für Bodensee Online«² genannt werden, das die

1 Siehe <https://citizensense.net/>.

2 Bei Bodensee Online handelt es sich um ein Informationssystem für die »Beschreibung, Quantifizierung und Prognose der Auswirkungen anthropogener Eingriffe« (Landesanstalt für Umwelt Baden-Württemberg 2008, 3) in das Gewässer des Bodensees, die durch die dreidimensionale Berechnung von Wassertemperaturen und Strömungsverhältnissen abgeschätzt werden können. Grundlage des Informationssystems sind unterschiedliche Modelle des Wasserkörpers sowie echtzeitlich erhobene Daten, die durch ein stationäres Sensornetzwerk am Rande des Bodensees erhoben werden. Siehe <https://www.lubw.baden-wuerttemberg.de/wasser/bodenseeonline>.

Aggregation von Wassertemperaturdaten durch ansässige Bootsfahrer*innen zum Ziel hatte. In der Vorbereitung wurden die Teilnehmer*innen zu Workshops eingeladen, in denen man ihnen den Umgang mit den Sensoren und den eigens für das Projekt entwickelten Dashboards erklärte. Neben dem Zweck, wichtige Daten für die Klimaforschung zu sammeln, formulierten die Teilnehmer*innen im Rahmen des Projekts insbesondere auch ein technisches Interesse an der Sensortechnologie selbst. Solche Beispiele deuten auf eine hohe Involviertheit und Interaktivität insbesondere in der Vorbereitungsphase derartiger Messvorhaben hin. Der Projektleitung zufolge sollten die Teilnehmer*innen nach der Anbringung der Sensoren an den Booten jedoch weitgehend ungestört ihren Freizeitaktivitäten mit den Booten nachgehen können, als seien die Sensoren überhaupt nicht da.

Dass das Mobile Crowdsensing nicht einfach als synonyme Begriff für partizipative Sensing-Modelle angesehen werden darf, wird deutlich, wenn Partizipation als ein möglicher ›Modus‹ der Datensammlung in den Hintergrund tritt. Ganti et al. etwa klassifizieren das Crowdsensing anhand der ›aktiven‹ oder ›bewussten‹ Mitwirkung von Teilnehmer*innen auf der einen Seite sowie der ›opportunistischen‹ oder passiven Beteiligung auf der anderen Seite (vgl. 2011: 32). Sebastian Vehlken (vgl. 2016: 59f.) schreibt hierzu, dass die Involvierung von Teilnehmer*innen im Sinne einer bewussten Partizipation in der Regel mit einer Aktivierung menschlicher Sinneswahrnehmung einhergehe, wenn Nutzer*innen etwa gezielt Fotos aufnehmen, diese mit zusätzlicher Information versehen und auf digitalen Plattformen teilen. Eine aktive Partizipation wirkt sich also auf die Aktionspfade der Nutzer*innen in der Interaktion mit den mobilen Devices selbst aus. Sie partizipieren über die Situation der Datensammlung hinaus durch Datenpraktiken des *sharing*, *tagging* und *mapping* sowie der kollaborativen Erarbeitung und Repräsentation von Ergebnissen. Derlei Praktiken weisen stets einen starken Bezug auf die Sensordaten selbst auf, die als solche zu sogenannten »matters of concern« werden (vgl. Latour 2005).

Mit der zweiten Kategorie eines ›opportunistischen‹ oder passiven Sensing hebt sich das Mobile Crowdsensing bei Ganti et al. (2011) jedoch vom Participatory Sensing ab. Als ›passive Partizipation‹ werden Fälle bezeichnet, bei denen auf digitalen Plattformen geteilte Daten ohne das Wissen der Nutzer*innen aggregiert oder schlichtweg permanent Positionsdaten erhoben werden, wie dies bei heutigen Smartphones und anderen Wearables der Fall ist. Folglich betrifft das Mobile Crowdsensing explizit auch diejenigen permanenten und ubiquitären Datenerhebungsprozesse, bei denen Nutzer*innen mehr oder weniger unbewusst am Prozess der Datenerhebung teilnehmen, während sie die Services ihrer mobilen Geräte nutzen. In der Regel erhalten sie dabei keine Einsicht in die weitere Verwendung der Daten. Opportunistische Partizipation meint hier also nicht zwingend eine aktive ›Involvierung‹ in den Wissensprozess, sie entspricht vielmehr einem minimalistischen Verständnis von Ko-Operation, bei der die *in situ*-Datenproduktion in automatisierten Prozessen der Datenverarbeitung und -auswertung angeeignet

wird, ohne dass dabei von einem Konsens zwischen den beteiligten Akteur*innen ausgegangen werden kann (vgl. Borbach/Thielmann 2019; Schüttpelz/Gießmann 2015). Eine kritische Auseinandersetzung mit der »soziomaterielle[n] und organisatorische[n] Ausgestaltung« (van Geenen/Kaerlein 2023: 89) dieser Prozesse erscheint deshalb zwar umso dringlicher, aufgrund ihrer Opazität aber auch schwieriger.

Die Gegenwartsdiagnose einer Ubiquität der sensorischen Datenerhebung hat der Medientheoretiker Marc B. N. Hansen im Rahmen seines reichlich rezipierten Aufsatzes zur »Ubiquitous Sensibility« einmal so formuliert: »[...] our phones and the sensors that now inhabit our lived space passively register massive amounts of behavioral and environmental data without any active involvement, decision to initiate, or even awareness on our part« (2012: 53). Laut Hansen hat die technisch-sensorische Durchdrungenheit von Umwelt eklatante wahrnehmungs- und medientheoretische Konsequenzen. Gegenwärtige »Sensormedien« (Thielmann 2019; Scholz 2021) versetzen uns in eine »condition of transparent ubiquity« (Hansen 2012: 53), welche nicht mehr unsere Sinne als solche medial vermitteln, sondern »sensitivity itself« (ebd.: 54). Mit der Beschreibung der ubiquitären algorithmischen Prozesse sensorischer Erfassung spielt Hansen auf das *Ubiquitous Computing* an, ein in den frühen 1990er Jahren philosophisch unterfüttertes Konzept des Computerwissenschaftlers Mark Weiser, der am Xerox PARC in Palo Alto mit seinem Entwickler*innenteam die heutigen smarten Geräte, insbesondere auch das iPhone mit seinen zahlreichen integrierten Sensoren, entscheidend vorbereitete (vgl. Sprenger 2015). Weiser hatte hierfür die Idee einer ubiquitären Computerisierung von Umgebungen entworfen, wobei die Rechenkraft als »calm technology« (vgl. Weiser/Brown 1995) die menschliche Wahrnehmung unterlaufen sollte, um ihren einstigen Zweck des dem Menschen Dienenden endlich zu erfüllen. Weisers idealistische Vision der Mensch-Maschine-Interaktion wurde Mitte der 1990er Jahre von der Industrie aufgegriffen und im Feld des sogenannten Pervasive Computing weiterentwickelt. Am IBM fokussierte man sich dabei aber stärker auf die Mobilität der Computer sowie auf den exzessiven Einsatz und die Integration von Sensortechnologie (vgl. Ark/Selker 1999; Dourish/Bell 2011: 15). Es ist also nicht verwunderlich, dass Industriediskurse zur Technik des Mobile Crowdsensing aus eben jenem Dunstkreis von Entwickler*innen am IBM hervorgegangen sind. Gewissermaßen soll das Mobile Crowdsensing dann ein Systemmodell für die epistemische Nutzbarmachung jener von Hansen beschriebenen gegenwärtigen Medienkonstellation liefern, bei der große Massen von Nutzer*innen in der sensorisch-medial vermittelten Interaktion mit ihrer Umgebung permanent sensorische Daten produzieren. Das Mobile Crowdsensing verweist hier auf ein Spannungsverhältnis zwischen ubiquitärer Datenerhebung und den zuvor geschilderten Partizipationsversprechen. Es steht in der Tradition des Ubiquitous und Pervasive Computing, bei denen Datenverarbeitungsprozesse unterhalb der Wahrnehmungsschwelle der

Nutzer*innen ablaufen, um deren Interaktion mit den mobilen Endgeräten auf interfacegestützte Service-Funktionen zu beschränken und diese zu erleichtern.

3. Mobile Sensor Networks

Die Verbreitung von mobilen, sensorgestützten Endgeräten hat seit der Lancierung des iPhones als vielleicht größte Erfolgsgeschichte dieser Entwicklung nicht nur Imaginationen einer Demokratisierung von digitaler und datengetriebener Umwelt- und Klimaforschung befeuert, sondern auch von einer durch die Verbreitung ermöglichte Skalierbarkeit der Datenerhebung insgesamt (vgl. Lane et al. 2010). Sensing-Projekte sollen dadurch flexibel auf unterschiedliche Informationsbedarfe und politische oder gesellschaftliche Belange antworten (vgl. ebd.: 141ff.). Die »sensorische Crowd« erscheint vor diesem Hintergrund als ein Makro-Modell sensorischer Datenerhebung auf Ebene einer »societal scale« (Ganti et al. 2011: 32). Mit ihr soll die räumliche Distribution von Sensoren möglich werden, sodass Messungen feinmaschiger und Modelle optimiert werden können (vgl. Gabrys 2016: 32). Als mobil agierende, räumlich distribuierte Vielheit von Nutzer*innen-Technik-Netzwerken erfüllt sie also allem voran eine epistemische Funktion: Die sensorische Crowd erlaubt die Verteilung und Verdichtung sensorischer Messungen und damit die Skalierung der sensorischen Datenerhebung. Ganz im Sinne der sensorisch-technischen Landnahme antwortet sie auf ein zentrales – und einst militärisch motiviertes – Anliegen der drahtlosen Sensornetzwerke, nämlich auf die Notwendigkeit der stetigen Erweiterung stationärer sensorischer Messungen. Dort, wo Forscher*innen in Hinblick auf ein zu erfassendes Gebiet mit dem Abdeckungsproblem, d.h. mit dem »problem of deployment and coverage« (Dantu et al. 2005: 408) konfrontiert sind, soll mit der Technik des Mobile Crowdsensing Abhilfe geleistet werden. Teilnehmer*innen sollen Umweltparameter eines geographischen Gebietes über einen bestimmten Zeitraum hinweg feinmaschiger erfassen, als dies mit kosten- und wartungsintensiven Sensorstationen möglich ist. Mit der Crowdsensing-Studie am Bodensee wollte man beispielsweise der unzureichenden sensorischen Erfassung des Wasserkörpers entgegenwirken. Bootsbesitzer*innen sollten eigens für das Projekt entwickelte Sensoren an ihren Booten befestigen, um damit Wassertemperaturdaten auch aus den tieferen Bereichen des Sees zu messen. Dadurch erhoffte man sich einen größeren Detailreichtum der Daten, die perspektivisch zur Validierung und Anpassung vorhandener Strömungsmodelle des Wasserkörpers verwendet werden sollten. Die Crowd avanciert so zu einer »productive force« (Vehlken 2016: 62), mittels derer die Präzisierung von Modellen des Umweltmonitoring, aber auch allgemeiner der Mobilität und umgebungsspezifischer Interaktion vorangetrieben werden kann.

Dass die eingesetzten Sensoren und Interfaces im Kontext der Klimaforschung wie im Bodensee-Beispiel erst entwickelt werden müssen, steht einer häufig wiederholten Prämisse des Mobile Crowdsensing entgegen. Entwickler*innen stützen sich auf die Gewissheit, dass der pragmatische Rückgriff auf Sensordaten durch die bereits im Umlauf befindliche und vernetzte Hardware möglich wird. Dabei sollen Nutzer*innen im Idealfall »ungestört« anderen Aktivitäten nachgehen, die dann wiederum selbst zum Gegenstand sensorischer Erfassung werden.

The massive use of smart phones, and the adoption of on-board sensors in modern vehicles, enable the use of non-specialized crowds as mobile sensor networks. Ideally, people will go to their daily activities by walking or driving and the sensors attached to them will collect data without interrupting their activities. (Chakeri et al. 2022: 278)

In zentralen Publikationen zur Technik des Mobile Crowdsensing werden Nutzer*innen immer wieder als *mobile sensor networks* diskursiviert. Aufgrund ihrer Vernetzung erscheinen die »non-specialized crowds« als *gegebene* infrastrukturelle Grundlage für die mobile Datenerhebung, sodass sich nicht mehr nur von einer Rhetorik der Gegebenheit der Daten selbst (vgl. Rosenberg 2013; Thielmann 2014: 355), sondern auch von einer Rhetorik der Gegebenheit von Infrastruktur sprechen ließe, die so jedoch transparent zu werden droht (vgl. Star/Ruhleder 1996: 113). Entgegen einem klassischen Infrastrukturverständnis umfassen diese Infrastrukturen dann nicht lediglich die in einem »infrastrukturellen Rückraum« (Schabacher 2022: 8) operierenden digitalen Netze und Server, sondern auch die Endgeräte und ihre Nutzer*innen selbst. Sie bleiben weitgehend anonym und werden zu Komponenten einer Infrastruktur des Mobile Crowdsensing, insofern sie aus Perspektive von Unternehmen und Institutionen als »durable object-constellations« (Seberger/Bowker 2021: 171ff.) erachtet werden, die durch geeignete Incentives in Form von Apps adressiert und mobilisiert werden können.

Mit der sensorischen Crowd als *mobile sensor network* ist eine soziotechnische Kategorie rezenter Datenerhebungsprozesse aufgerufen, die sich weniger leicht in ihrer Sequenzialität als Reihe situierter, verkörperter und interaktiver Praktiken *on the ground* untersuchen lassen, eben weil sie stärker auf Top-Down-Operationen der Datenakquise verweisen. Die Crowd erscheint hier als Aggregat, mittels dessen Daten über Umweltphänomene großen Maßstabs durch Dritte akquiriert werden können. Das Mobile Crowdsensing kann deshalb als eine Technik der »sensor society« par excellence angesehen werden, mit der Marc Andrejevic und Mark Burdon den Wandel von einer zielgerichteten Informationsgenerierung hin zu den »always-on, ubiquitous, opportunistic ever-expanding forms of data capture« (2015: 19) beschrieben haben. Insbesondere dem Versprechen einer medieninhärenten Interaktivität digitaler mobiler Gadgets erteilen die Autoren eine Absage und stellen statt-

dessen eine »passive-ication of interactivity« (ebd.: 20) fest, bei welcher der Umfang der beispielsweise mit einem Smartphone generierten sensorischen Daten bei weitem die durch die Nutzer*innen aktiv kommunizierten Informationen übersteigt:

If, once upon a time, the mobilization of the promise of interactivity was characterized by the enthusiastic portrayal of heightened forms of active participation on the part of users, the automated collection of data »passive-izes« this interactivity. Our cars, phones, laptops, Global Positioning System (GPS) devices, and so on allow for the comprehensive capture of the data trails users leave as they go about the course of their daily lives. In the business world, this device-driven data—combined with new techniques for putting it to use—have been enthusiastically greeted as a valuable economic resource. (Andrejevic/Burdon 2015: 20)

Dass mobile Sensoren, insbesondere die der heutigen Smartphones, aber auch diejenigen der immer autonomer werdenden Autos, ständig sensorische Daten über Interaktionen von Nutzer*innen mit ihrer Umgebung generieren, evoziert entsprechende Kritik am datenbasierten Überwachungskapitalismus bzw. an einem Capture-Kapitalismus, der im Wesentlichen auf kommerziell motivierte Praktiken des Tracking, der Intervention und Steuerung aufbaut (vgl. Zuboff 2020; Agre 1994; Friedrich 2022). Laut Andrejevic und Burdon überschreitet die Sensorgesellschaft aber die Rekonfiguration von in Diskursen zu digitalen Überwachungstechnologien ohnehin schon aufgerufenen Kategorien von Privatheit, Überwachung und »sense-making« (2015: 20f.): »There are structural asymmetries built into the very notion of a sensor society insofar as the forms of actionable information it generates are shaped and controlled by those who have access to the sensing and analytical infrastructure« (ebd.: 21). Sie schlagen deshalb vor, die medienwissenschaftliche Aufmerksamkeit stärker auf die »costly infrastructures« (ebd.) zu richten, welche die Datensammlung, -speicherung und -verarbeitung überhaupt erst ermöglichen und es den jeweiligen Unternehmen erlauben, auf sie zurückzugreifen – also eben jene »kostspieligen Infrastrukturen«, die in der Rhetorik des Crowdsensing entweder affirmativ als im Überfluss verfügbare Infrastruktur gerahmt oder aber weitgehend diskursiv verworfen werden.

4. Interventionen des Crowdmanagements

Neben den gegenwärtigen Medieninfrastrukturen des »Signal Traffic« (vgl. Parks/Starosielski 2015) – im spezifischen die seit Mitte der 2000er Jahre entstandenen »Infrastrukturen des Cloud Computing« (Sprenger/Engemann 2015: 18) – müssen angesichts der permanenten, unter der Wahrnehmungsgrenze ablaufenden, automatischen und umfassenden sensorischen Datenerhebung dann insbesondere die

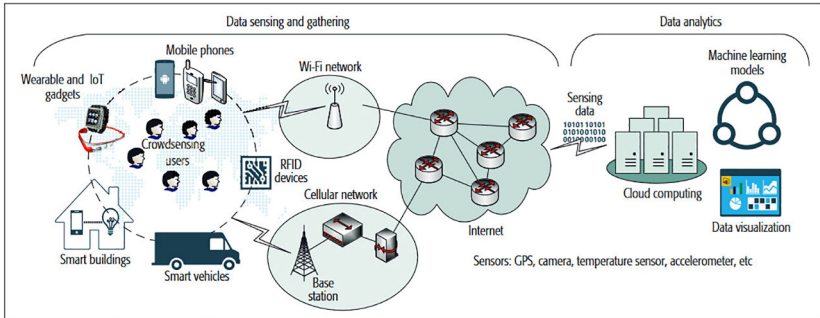
für die Zirkulation der Daten notwendigen algorithmischen Prozesse und Operationen in den Blick genommen werden, die den Datenverkehr datenlogistisch organisieren. Mithilfe von »Big Sensor Data«-Algorithmen sollen etwa die Daten von Sensoren wie Mikrofonen, Kameras, Beschleunigungssensoren von Autos, Gyroskopsensoren in Smartphones oder von GPS-Signalen innerhalb eines *networked sensor system* erhoben, gefiltert, integriert, ausgewertet, interpretiert und repräsentiert werden (vgl. Ang/Seng 2016). In der Industrieliteratur wird Mobile Crowdsensing als System aus dynamischen On-/Offline-Konstellationen von bewegten Körpern, Objekten, stationären Kommunikationsnetzwerken und Software entworfen, das auf softwarebasierten Verfahren zur Organisation und Prozessierung von Daten und Datenflüssen basiert. Eindrücklich wird dies insbesondere durch die Publikationen begleitenden graphischen Systemdarstellungen, die eine vermeintlich geschlossene Datenökologie suggerieren (Abbildung 1). Damit evoziert die Technik durchaus Vorstellungen einer vollständig autonomisierten Organisation von Datenflüssen, die dann auch reziprok – im Sinne des Closed Loop Sensing (vgl. Friedrich 2022: 645) und auf Basis entsprechender Machine Learning Algorithmen – zur Steuerung mobiler sensorischer Messungen und allgemeiner zum Management von datafizierter Mobilität reichen soll, ohne dass die Intervention menschlicher Akteure vonnöten ist. Solche Entwürfe werfen deshalb, wie Kathrin Friedrich bemerkt, nicht zuletzt »Fragen nach der verbleibenden menschlichen Handlungsmächtigkeit und veränderten Interaktionsbedingungen von Lebewesen« (ebd.) allgemein auf. Als zentraler Treiber für derartige Systemvorstellungen fungiert dabei stets die Annahme einer umfassenden Programmierbarkeit, das heißt Implementierbarkeit formalisierten Wissens, wie Jennifer Gabrys sie im Kontext aktueller digitaler Sensortechniken untersucht hat:

»Programmability«, as I employ the term, has a somewhat wider use than just software or code. Instead, this expanded engagement with programmability considers how code is not a discursive structure or rule that acts on things, but rather is an embodied and embedded set of operations that are articulated across devices, environments, practices, and imaginations. Programmability then exceeds software (and even computation) to encompass the formation of events, spaces, and things. (Gabrys 2016: 11)

Gabrys versteht die Programmierbarkeit jedoch gerade nicht als schlichte Annäherung performativer Prozesse an die Logik algorithmischer Operationen, sondern begreift sie vielmehr in ihrer Kontingenz als »fraught process, [...] continually unfolding the complex if at times troubling attempts to make environments into spaces of observation, distributed experience, and even automated management« (ebd.: 281). Vor diesem Hintergrund stellt sich beim Mobile Crowdsensing die Frage nach den zugrundeliegenden System Designs sowie entsprechenden

Modellierungs- und Formalisierungspraktiken, die sich in Diagrammen, Modellen und Graphen äußern und aus medienwissenschaftlicher Perspektive spannende Gegenstände darstellen.

Abbildung 1: System model of mobile crowdsensing (Screenshot). In: Alsheikh, Mohammad Abu et al. (2017): »The Accuracy-Privacy Trade-off of Mobile Crowdsensing«, in: *IEEE Communications Magazine*, 55(6), S. 132–139, S. 133.



Auf der Prämisse einer vermeintlich im Überfluss verfügbaren, jederzeit »anzapfbaren« Infrastruktur sensorisch ausgestatteter User-Netzwerke aufbauend, fokussiert die Industrieliteratur zum Mobile Crowdsensing also verstärkt die auf Modellen basierenden Verfahren des Crowdmanagements sowie die stete Optimierung und Anpassung von Apps unter Verhandlung etwaiger privatschutzrechtlicher Bedenken (vgl. Roitman et al. 2011; Vehlken 2016). Sie eröffnet deshalb nicht zuletzt einen interessanten Einblick in die zugrundeliegenden Modellierungen von Nutzer*innen und die daran anknüpfenden Überlegungen zu umgebungs- und medien-spezifischen Affordanzen. Dass dies mitunter bizarre Züge annimmt, lässt sich an einem Aufsatz zum *Vehicular Crowdsensing* veranschaulichen, in welchem die Autoren nach möglichen »incentive mechanisms« suchen, »that encourage [...] vehicles to deviate from their pre-planned trajectories in order to visit and collect data from the isolated places« (Chakeri et al. 2022: 278). Es wird auf das höhere »involvement« der Nutzer*innen verwiesen, die von ihrem eigentlichen Wegeziel abweichen müssen, damit das Messvorhaben gelingen kann. Ganz in der Tradition des Ubiquitous und Pervasive Computing schwingt hier die negative Bewertung einer hohen Nutzer*innen-Involvierung mit, während Partizipation als ein möglicher Anreiz nicht genannt wird. Im Gegenteil wird für die Anreizschaffung die Annahme zugrunde gelegt, dass sich AVs (Autonomic Vehicles) »in a selfish and rational way« verhalten, »i.e., they always look to maximize their own utilities« (ebd.: 280). Als ein möglicher Mechanismus wird die Gamifizierung der Datenproduktion vorgeschlagen, bei der die Nutzer*innen mit »rewards« (ebd.: 278) ausgezeichnet werden, wenn

sie den vorgeschlagenen Trajektorien folgen und die gewünschten Daten erheben. Die mobilen Sensorknoten erscheinen hier als User-Vehicle-Hybride, deren Verhaltenszuschreibung als ›egoistische‹ Agenten auf eine gängige Prämisse des Mobilitätsmanagements zurückgeht. Sie basiert auf der modellierenden Schwarmforschung bzw. genauer auf agentenbasierten Simulationsmodellen, die medienhistorisch eingehend durch Sebastian Vehlken betrachtet wurden. Diesem zufolge stellen die in den 1970er Jahren naturwissenschaftlich beforschten tierischen ›Selfish Herds‹ relevante Modelle für die geometrisch-mathematische »Entstehung globaler Muster als Ergebnis individueller, lokaler Verhaltensprozesse« (2012: 289) dar. Der integrierte Einsatz von Sensoren innerhalb sogenannter »sensory integration systems« (vgl. Schilt/Norris 1997) hingegen habe dazu geführt, so Vehlken weiter, dass das ›Soziale‹ selbst »als Funktion verschiedener Informationsinputs« (2012: 290) bestimmt sei. So erscheint die algorithmische Implementierung geeigneter Incentives beim Mobile Crowdsensing als Korrekturoperation gegenüber dem Verhalten mobiler Agenten, das hier nicht mehr eindeutig dem Menschen oder der Maschine zugeordnet werden kann. Dabei wird aber nicht nur recht banalen, vor allem aus der Verhaltensbiologie entnommenen Verhaltensmodellen gefolgt, sondern offenbar auch weitgehend opak bleibenden Prämissen aus der Spieltheorie.

Wegen dieser die Kontrolle und das Mobilitätsmanagement betreffenden Dimensionen hat Sebastian Vehlken das Crowdsensing allgemeiner in den Kontext einer Genealogie der Masse und des ›crowd behavior‹ gestellt. Dabei kontrastiert er moderne, aus der Massenpsychologie des späten 19. und frühen 20. Jahrhunderts hervorgegangene Beschreibungen der Masse als affektives und emotional aufgeladenes Kollektiv gegen die oben bereits angesprochene, sich durch Entwicklungen im Bereich des Agent-Based-Modelling (ABM) und der Simulation wandelnde wissenschaftliche Behandlung der Crowd als »mere quantifiable ›connectives« (Vehlken 2016: 52), die insbesondere für das Verfahren des Crowd Capturing relevant werden. Als »zweite Transformation« dieser Genealogie nennt Vehlken den Übergang zum sogenannten Multimodal Crowdsensing (vgl. Roitman et al. 2011).³ Hierbei geht er insbesondere auf die medial ermöglichte Selbstreflexivität der Crowd ein, der beispielsweise mittels interaktiver Karten in den betreffenden Crowdsensing-Applikationen und Dashboards bestimmte Informationen in annähernder Echtzeit zurückgespielt werden können, »leading to autonomous decision making« (Vehlken 2016: 61). Mehr noch aber dienen die gesammelten Daten, die dann mithilfe integrierter Modelle verarbeitet werden, der Visualisierung, Auswertung und Vorhersage dynamischer Umweltgeschehen, »to support actions, assessments and decision-making by experts« (ebd.: 60). Im Crowdsensing, welches schließlich ›alle Arten‹ von durch

3 Beim *Multimodal Crowdsensing* liegt der Schwerpunkt begrifflich stärker auf den multimodalen Verfahren der Datenverarbeitung, die beim Mobile Crowdsensing aber ebenso zentral sind.

Menschen und mobile Geräte erhobene Sensordaten umfasse, zeige sich laut Vehlken der Versuch einer Autonomisierung allumfassender Organisation von Umweltgeschehen (vgl. ebd. 61).

5. Mobilität, Crowd, Sensing: Trajektorien einer Vorgeschichte

Das *Mobile Crowdsensing* verweist vor diesem Hintergrund auf eine politische Formel gegenwärtiger Datenprozesse, welche vom MIT-Professor Alex Pentland einmal so zusammengefasst wurde: »[...] sense the situation, combine observations with dynamic demand and reaction models, and use the resulting predictions to control the system« (2012: 31). Die massenhaft verfügbaren »behaviour-sensing-data« sollen die Grundlage für die zukünftige Autonomisierung gesellschaftlicher Kontrollsysteme bilden. Derlei umfassende Zukunftsvisionen von einem durch mobil erhobene Sensordaten gefütterten – letztlich global operierenden – Kontroll- und Feedbacksystem fußen dabei auf einer gegenwärtigen »Konstellation bestehender mobiler Medientechniken« (Buschauer 2014: 406), welche die ständig wiederkehrenden Verweise auf das weitreichende Vorhandensein nötiger technischer Hardware zu legitimieren scheinen. Dabei umfasst das Mobile Crowdsensing – wie meine Ausführungen gezeigt haben – eine ganze Reihe von Technologien, Infrastrukturen, Praktiken und Diskurse, die für konkrete Durchführungen in Anschlag gebracht und im Sinne einer prototypischen Datenzirkulationsarchitektur und -logistik unter teils hohem Aufwand aufeinander abgestimmt werden müssen. Um eine schlichte Übernahme der Rhetorik einer Gegebenheit von Infrastruktur in der medien- und kulturwissenschaftlichen Auseinandersetzung mit dem Mobile Crowdsensing vorzubeugen, erscheint eine Kontextualisierung und historische Einordnung im Sinne einer *Vorgeschichte* der Technik umso relevanter. Eine solche Verflechtungsgeschichte zeigt, dass das komplexe Zusammenspiel der hier eingesetzten Technologien, Infrastrukturen und Wissensbestände niemals unproblematisch ist.

Von Pentlands Artikel ausgehend hat Regine Buschauer die Visionen eines sensordatenbasierten Kontrollsystems allgemeiner in eine diskurs- und medien-geschichtliche Betrachtung des Entwicklungszusammenhangs von Mobilitätsinfrastrukturen und Telekommunikation gestellt. Sie setzt bei der frühen Mobilfunktechnologie an, die technisch gesehen im Wesentlichen durch das Location Sensing bei gleichzeitiger Bewegung der Nutzer*innen als »Kernfunktionalität« (ebd.: 419) definiert ist. Am frühen Mobilfunk lässt sich demnach das technische *Apriori* des Sensing verdeutlichen, das notwendig ist, um drahtlose Kommunikation überhaupt zu gewährleisten. Interessant ist, dass die Adressierung der mobilen Einheiten für die Mobiltelefonie noch bis in die 1990er Jahre hinein eine technologische Herausforderung darstellt, insbesondere wenn zu viele Nutzer*innen innerhalb einer Funkzelle das System potenziell überlasten. Die Crowd kommt in

der Technologie des Mobilfunks als Störung der wechselseitigen Kommunikation zwischen zwei mobilen Einheiten zum Tragen. Der Problematik des »spectral overcrowding« begegnet man mit dem exzessiven Ausbau und einer Ubiquitisierung von stationärer Infrastruktur sowie mit der Ausstattung der mobilen Einheiten mit kleinen Rechnern, die eine eigene Adresse aufweisen (vgl. Cooper/Nettleton 1983). Zugleich wird eine komplexe datenlogistische »infrastructure« (Curien 1998) für das Mobilitätsmanagement entwickelt. Im Verlauf der 1980er Jahre beginnt man außerdem mit der Suche nach geeigneten Mobilitätsmodellen, die zum »Tracking [of] Mobile User in Wireless Communications Networks« (Bar-Noy/Kessler 1993) dienen sollen und bereits hier auf eine Virtualisierung des Diskurses zur mobilen Telekommunikation hindeuten (vgl. Gießmann 2012: 331). Bis heute werden in der Mobilfunktechnik Simulationen zur Optimierung infrastruktureller Performance eingesetzt, um Datentransfer überhaupt zu gewährleisten.

Parallel zur Genese der mobilen Netzwerke in der Telekommunikation und der Entstehung des Internets beschäftigt man sich ab den 1980er Jahren im Feld der *Wireless Sensor Networks* mit Fragen der Mobilmachung von Sensoren. Im Kontext der durch die US-amerikanische Defense Advanced Research Projects Agency (DARPA) entwickelten Large Scale Sensor Networks zum Zwecke der militärischen Gebietsüberwachung werden die Möglichkeiten einer hybriden Strukturierung der Sensornetzwerke erforscht, wobei mobile Sensorknoten den statischen Sensornetzwerken hinzugefügt und mittels geeigneter Verfahren des »collaborative signal processing« (Kumar/Shepherd 2001) integriert werden sollen. Man erhofft sich davon insgesamt eine Flexibilisierung der militärischen Operation sowie eine gewisse Redundanz des Erfassungsgebiets (vgl. Dostal 2001). Für die drahtlosen Sensornetzwerke sind Forschungen im Feld des Pervasive Computing zentral, die insbesondere die Miniaturisierung und Energieeffizienz und damit einen breiten Einsatz von Sensoren für kommerzielle, militärische und wissenschaftliche Zwecke vorantreiben. Im Pervasive Computing befasst man sich mit der Integration von mobilen Sensoren in stationäre *wireless sensor networks*, die als Verbindungsinstanz zwischen physischer Umgebung und Kommunikationsnetzen fungieren sollen (vgl. Estrin et al. 2002). Auch in der Umweltforschung zu distribuierten Sensornetzwerken werden explizit die epistemischen Potenziale mobiler Sensoren befragt, wobei zwischen sogenannten *Mobile Adhoc Networks (MANETs)* und solchen einer *Controlled Mobility* unterschieden wird. Die Mobilisierung von Sensornetzwerken bezieht sich dabei zunächst nicht etwa auf die Involvierung von Nutzer*innen, sondern reicht vielmehr in das Feld der Robotik hinein. *Controlled Mobility* meint also »the ability of a network to move intentionally, and without human assistance« (Dantu et al. 2005: 404). Beispielhaft kann hier das 2002 lancierte Forschungsprojekt Robomote der University of California genannt werden, bei welchem die Eignung kleiner und mit Sensoren ausgestatteter fahrender Roboter für den Einsatz innerhalb statischer Sensornetzwerke erprobt wurde (vgl. Sibley et al. 2002). Dazu gehören ebenso Überlegungen zu geeigneten Model-

len und Algorithmen einer ›Sensing Coverage Optimization‹, die sich zum Beispiel am Bewegungsverhalten von Ameisen orientieren (vgl. Lu et al. 2021).

Zeitgleich werden mit den MANETs im Ausgang des 20. Jahrhunderts auch die bereits existierenden Mengen von Handybesitzer*innen referiert, die jedoch keinesfalls als steuerbare mobile Netzwerke erscheinen, sondern im Gegenteil eher als sich unablässig und unkontrolliert bewegend und damit schwer zu steuernde Crowd. Erst die Verschiebung hin zu einem stärker systemadaptiven und ortsspezifischen Verständnis von Sensoren als pervasive und vernetzte Medientechnologie, die unter anderem auch mit einer massenhaften Implementierung von GPS-Technologie in Fahrzeuge einhergeht, bringt neue Formen der Lokalisierung hervor, welche die Umgebungen der Geräte stets mit thematisieren und insofern laut Florian Sprenger (2018: 264) »Verfahren der Positionierung« darstellen. Hier werden die Nutzer*innen-Mengen zu einem vielversprechenden Gegenstand für etwaige Sensing-Unternehmungen, weil sie als in konkreten Umgebungen positionierbare, adressierbare und steuerbare Sensorknoten erscheinen.

Wenn das Mobile Crowdsensing also im Kontext einer längeren Vorgeschichte des mobilen Sensing Betrachtung finden soll, so kann es nicht sinnvoll auf den technischen Einsatz oder Rückgriff auf Smartphones oder andere *hand-held gadgets* als Einzelmedien reduziert werden. Es erweist sich vielmehr als heterogene Technik, die auf die Verflechtungsgeschichte einer ganzen Reihe von Technologien, Infrastrukturen und Wissensbeständen aufbaut. Wie bereits angedeutet, müssen diese – entgegen der Rhetorik der Allgegenwart von Hardware – bis heute je nach Anwendungsfall erst entwickelt und getestet werden. Bei der Machbarkeitsstudie am Bodensee galt dies nicht nur für die neu entworfenen Sensoren, sondern auch für ein entsprechendes LoRaWan-Netzwerk für die mobile Übertragung der Daten, mit der außerdem ein hoher Energieverbrauch verbunden ist. Sensing-Unternehmungen sind damit stets auch selbst Projekte der Infrastrukturierung, die nicht nur den Ausbau von IoT-Netzwerken, sondern auch die Gewährleistung von Funk- und Energieversorgung betrifft. Insofern, so möchte ich hier argumentieren, muss das Mobile Crowdsensing stärker im Kontext der komplexen Genese mobiler Medien allgemein und spezifischer einer Mediengeschichte des mobilen Sensing betrachtet werden, die sich laut Tristan Thielmann immer schon als »komplementäre Systeme und Erweiterungen zu stationären elektrischen/elektronischen Medien und Kommunikationsnetzwerken« (2014: 351) entwickelt haben.⁴ Der Fokus auf die Gestaltung und

4 Tristan Thielmann (2014: 352) erklärt sich daraus die gegenwärtig anhaltende »Konvergenzbewegung von Medienkulturwissenschaft und Science and Technology Studies«, bei der sich erstere der medialen Perspektive auf die Medien-Werdung alltäglicher Gegenstände verpflichtet, während man sich in der STS einer infrastrukturellen Sichtweise bediene, die untersucht, wie das Internet durch Mobilien Erweiterungen im Sinne von »Netzwerkstrukturen, Produktions- und Vertriebsprozesse« ausbildet.

Nutzung von und den in situ-Interaktionen *mit* heutigen mobilen Medien kann so durch eine Betrachtung ihrer jeweiligen Einbettung in mediale Infrastrukturen und digitale Systeme der Datenverarbeitung geweitet werden.

6. Fazit

In dem Beitrag habe ich Diskurse zum Mobile Crowdsensing untersucht und darin Dimensionen aufgezeigt, welche die Technik für eine medienwissenschaftliche Betrachtung interessant machen. Aufgrund des häufigen Verweises auf bereits vorhandene mobile Sensing-Infrastruktur und den von Nutzer*innen »unbemerkten« Zugriff auf ebendiese, der als »opportunistische« Partizipation deklariert wird, reihen sich Industriediskurse zum Mobile Crowdsensing in das Feld des Pervasive Computing ein. Das Mobile Crowdsensing zielt so auf die sensorische Erfassung von Interaktionen mit der Umwelt selbst ab, die aufgrund der Verbreitung mobiler Sensoren skalierbar und durch Modelle verallgemeinerbar werden. Damit erscheint das Mobile Crowdsensing als Systemmodell und Datenlogistik komplexer algorithmischer Operationen, welche die Datenflüsse zwischen mobilen Endgeräten und fixierten digitalen Infrastrukturen in immer smarter werdenden Umgebungen organisieren. Folglich rücken die wissensbasierten Verfahren der Datenverarbeitung in den Fokus, also Modelle und Graphen, die nicht nur zur autonomen Entscheidungsfindung beitragen sollen, sondern auch weitere mobile Datenerhebungen mit mobilen Sensoren koordinieren und damit die Steuerung und Kontrolle als zentrale mediale Dimensionen des Mobile Crowdsensing ausweisen (vgl. Vehlken 2016). Das hierfür notwendige massenhafte Vorhandensein mobiler Devices wird dabei als materielle Grundlage für eine pragmatische Aggregationsmethode entworfen, die als kosteneffizient und ressourcenschonend erscheint.

Demgegenüber habe ich schlaglichtartig versucht, Trajektorien und mögliche Einsatzpunkte eines Entwicklungszusammenhangs zwischen der Genese mobiler Kommunikationssysteme, statischer Sensornetzwerke und Crowddiskursen aufzuzeigen. Das diese Bereiche verbindende Motiv kann als ein Wechselspiel zwischen stationärer Kommunikationsinfrastruktur und mobilen Medien bezeichnet werden, das durch eine Datenlogistik des ko-operativen Sensing bewerkstelligt und abgestimmt werden muss. Ausstehend ist dabei die historische Einordnung der Entwicklung mobiler Apps und digitaler Plattformen in diesem Kontext geblieben, die jedoch im historischen Zusammenhang von Mobilität, Sensing und Crowd und allgemeiner in der *Vorgeschichte* des Mobile Crowdsensing als gegenwärtige Medientechnik eine zentrale Rolle einnehmen, insbesondere weil sie den in-situ-Praktiken der Nutzer*innen, also ihre medienspezifischen Interaktionen, neue Ausprägungen verleihen.

Gabriele Balbi und Paolo Magaudda (2018) haben in ihrer *History of Digital Media* auf die Ideologie eines historischen »presentism« (ebd.: 11) hingewiesen, der insbesondere die wissenschaftliche Auseinandersetzung mit digitalen Medien prägt. Sie stützen sich dabei auf Graham Murdoch und Michael Pickering (2009), die mit Blick auf Arbeiten zu den heutigen Mobilfunksystemen von einem »cultural presentism« (ebd.: 8) sprechen. Ein solcher »kultureller Presentismus« zeuge – so argumentieren die Autoren – von einer durch gegenwärtige Mobilitäts- und Telekommunikationspraktiken veränderten Wahrnehmung von Raum- und Zeithorizonten. Gerade die exzessive Aufzeichnung und Speicherung von Ereignissen habe eine kulturelle Dissoziation von der eigenen technologischen Geschichte zur Folge, während »our sense of long-term continuities diminishes and we become increasingly present-focused« (ebd.).

New media seem to change so quickly and are considered so radically »revolutionary« (thus provoking a drastic break with the past) that historical analysis has often been difficult to apply and, ultimately, useless for a vision concentrated on the present and future (not by chance named »presentism«). (Balbi/Magaudda 2018: 11)

Wenngleich das Mobile Crowdsensing in den betreffenden Entwickler*innentexten durchaus als Innovation diskursiviert wird, verweist es rhetorisch dennoch immer wieder auf eine etablierte Medienkonstellation, die nicht nur in ihrer räumlich-infrastrukturellen Ausprägung als Totalität erscheint (»societal scale«), sondern auch in ihrer Zeitlichkeit und Historizität. Auch kritische Gegenwartsdiagnosen fallen dieser Rhetorik bis zu einem gewissen Maße anheim, wenn sie beispielsweise die »Ära« der »Sensor Society« als Ausgangspunkt ihrer Kritik nehmen. Die Ubiquität der mobilen, pervasiven und vernetzten Sensoren evoziert nicht die ständige Rede einer *newness* der Medientechnologien, sondern im Gegenteil die Rede von ihrer – im Sinne einer anhaltenden Gegenwart verstandenen – *givenness* als Kulminationspunkt einer kontinuierlichen, auf bestimmten Zukunftsvisionen ausgerichteten Technologieentwicklung sowie darüber hinaus als Ausgangspunkt weiterer Innovationen. Derartige Rhetoriken vernachlässigen erstens die lokale und situative Materialität von Sensing-Vorhaben und allgemein der Datafizierungsprozesse (vgl. Burkhardt et al. 2022). Die Gegebenheitsrhetorik geht darüber hinaus zusätzlich mit einer Naturalisierung der konkreten Wissenstechnologien einher, insbesondere der Modellierung und Kontrolle sogenannter *mobile sensor networks*, die – wie dies ein kurzer Exkurs in die Geschichte der mobilen Adressierung und der drahtlosen Sensornetze gezeigt hat – eng mit der Genese der Mobilfunknetze und der Sensornetze verstrickt sind, ohne dabei immer schon auf die Involvierung von menschlichen Nutzer*innen abzielen. Die Prämisse einer *givenness* der gegenwärtigen Medienkonstellation wird dann aber auch zum Ausgangspunkt einer »discursive produc-

tion of future-making« (Dourish/Mainwaring 2012), insofern sie das Zukunftsdenken innerhalb dieser Diskurse und Schauplätze weiter kolonialisiert. Eine Medien- und Wissensgeschichte des Mobile Crowdsensing hingegen kann die heterogenen historischen Verstrickungen von Infrastrukturentwicklung, Medien- und Wissenstechnologien und ihren Diskursschauplätzen ab der zweiten Hälfte des 20. Jahrhunderts insbesondere im Hinblick auf die gegenseitige Bezüglichkeit der drei Kernaspekte – Mobilität, Sensing, Crowd – in ihren Kontinuitäten und Diskontinuitäten sichtbar machen (vgl. Schröter 2023).

Literatur

- Agre, Philip E. (1994): »Surveillance and Capture: Two Models of Privacy«, in: *The Information Society*, 10(2), S. 101–127.
- Andrejevic, Mark/Mark Burdon (2015): »Defining the Sensor Society«, in: *Television & New Media*, 16(1), S. 19–36.
- Ang, Li-Minn/Kah P. Seng (2016): »Big Sensor Data Applications in Urban Environments«, in: *Big Data Research*, 4, S. 1–12.
- Ark, Wendy S./Ted Selker (1999): »A Look at Human Interaction with Pervasive Computers«, in: *IBM Systems Journal*, 38(4), S. 1–5.
- Balbi, Gabriele/Paolo Magaudda (2018): *A History of Digital Media. An Intermedia and Global Perspective*, New York/London: Routledge.
- Bar-Noy, Amotz/Ilan Kessler (1993): »Tracking Mobile Users in Wireless Communications Networks«, in: *IEEE Transactions on Information Theory*, 39(6), S. 1877–1886.
- Borbach, Christoph/Tristan Thielmann (2019): »Über das Denken in Ko-Operationenketten. Arbeiten am Luftlagebild«, in: Sebastian Gießmann/Tobias Röhl/Ronja Trischler (Hg.), *Materialität der Kooperation*, Wiesbaden: Springer, S. 115–167.
- Boulos, Maged N Kamel/Bernd Resch/David N Crowley/John G Breslin/Gunho Sohn/Russ Burtner/William A Pike/Eduardo Jezierski/Kuo-Yu Slayer Chuang (2011): »Crowdsourcing, Citizen Sensing and Sensor Web Technologies for Public and Environmental Health Surveillance and Crisis Management: Trends, OGC Standards and Application Examples«, in: *International Journal of Health Geographics*, 10(67), S. 1–29.
- Burke, Jeffrey A./Deborah Estrin/Mark Hansen/Andrew Parker/Nithya Ramanathan/Sasank Reddy/Mani B Srivastava (2006): »Participatory Sensing«, in: *Proceedings of the Fourth International Conference on Embedded Networked Sensor Systems*: Oct. 31 – Nov. 3, 2006, Boulder, Colorado, USA.
- Burkhardt, Marcus/Daniela van Geenen/Carolin Gerlitz/Sam Hind/Timo Kaerlein/Danny Lämmerhirt/Axel Volmar (Hrsg.) (2022): *Interrogating Datafication. Towards a Praxeology of Data*, Bielefeld: transcript Verlag.

- Buschauer, Regine (2014): »(Very) Nervous Systems. Big Mobile Data«, in: Ramón Reichert (Hg.), *Big Data: Analysen zum digitalen Wandel von Wissen, Macht und Ökonomie* (Digitale Gesellschaft 3), Bielefeld: transcript Verlag, S. 405–436.
- Chakeri, Alireza/Xin Wang/Luis G. Jaimes (2022): »A Vehicular Crowdsensing Market for AVs«, in: *IEEE Open Journal of Intelligent Transportation Systems*, 3, S. 278–287.
- Cooper, George R./Ray W. Nettleton (1983): »Communications: Cellular Mobile Technology: The Great Multiplier: It Can Greatly Expand Mobile Communications, but Some Technical and Regulatory Issues Remain Unsolved«, in: *IEEE Spectrum*, 20(6), S. 30–37.
- Curien, Nicolas (1998): »L'économie des réseaux«, in: *Panorama des Télécommunications* (11).
- Dantu, Karthik/Mohammad Rahimi/Hardik Shah/Sandeep Babel/Amit Dhariwal (2005): »Robomote: Enabling Mobility in Sensor Networks«, in: *IPSN 2005. Fourth International Symposium on Information Processing in Sensor Networks*, S. 404–409.
- Dostal, M. Brad C.A.J. (2001): *Enhancing Situational Understanding through the Employment of Unmanned Aerial Vehicles*. Abrufbar unter https://www.globalsecurity.org/military/library/report/call/call_01-18_ch6.htm, (Stand: 29.08.2023)
- Dourish, Paul/Genevieve Bell (2011): *Divining a Digital Future. Mess and Mythology in Ubiquitous Computing*, Cambridge, Mass: The MIT Press.
- Dourish, Paul/Scott D. Mainwaring (2012): »UbiComp's Colonial Impuls«, in: Anind K. Dey (Hg.), *Proceedings of the 2012 ACM Conference on Ubiquitous Computing*, New York, NY, S. 133–142.
- Estrin, Deborah/David Culler/Kris Pister/Gaurav Sukhatme (2002): »Connecting the Physical World with Pervasive Networks«, in: *IEEE Pervasive Computing*, 1(1), S. 59–69.
- Friedrich, Kathrin (2022): »Tracking«, in: Heiko Christians/Matthias Bickenbach/Nikolaus Wegmann (Hg.), *Historisches Wörterbuch des Mediengebrauchs: Band 3*, Göttingen: Böhlau, S. 631–649.
- Gabrys, Jennifer (2014): »Programming Environments: Environmentalism and Citizen Sensing in the Smart City«, in: *Environment and Planning D: Society and Space*, 32(1), S. 30–48.
- Gabrys, Jennifer (2016): *Program Earth. Environmental Sensing Technology and the Making of a Computational Planet* (Electronic Mediations 49), Minneapolis/London: University of Minnesota Press.
- Ganti, Raghu/Fan Ye/Hui Lei (2011): »Mobile Crowdsensing: Current State and Future Challenges«, in: *IEEE Communications Magazine*, 49(11), 32–39.
- Gießmann, Sebastian (2012): *Die Verbundenheit der Dinge*, Dissertation (Kaleidogramme Band 114), Berlin: Humboldt-Universität zu Berlin.

- Hansen, Mark (2012): »Ubiquitous Sensibility«, in: Jeremy Packer/Stephen B. Crofts Wiley (Hg.) *Communication Matters: Materialist Approaches to Media, Mobility and Networks*, Abingdon, Oxon/New York, N.Y.: Routledge.
- Kumar, Srikanta/David Shepherd (2001): »SensIT: Sensor Information Technology for the Warfighter«, in: *Proceedings of the 4th International Conference on Information Fusion*.
- Lane, Nicholas et al. (2010): »A Survey of Mobile Phone Sensing«, in: *IEEE Communications Magazine*, 48(9), S. 140–150.
- Latour, Bruno (2005): »From Realpolitik to Dingpolitik or How to Make Things Public«, in: Bruno Latour/Peter Weibel (Hg.), *Making Things Public: Atmospheres of Democracy*, Cambridge, Mass.: MIT Press, S. 4–31.
- Lu, Chao/Xunbo Li/Wenjie Yu/Zhi Zeng/Mingming Yan/Xiang Li (2021): »Sensor Network Sensing Coverage Optimization with Improved Artificial Bee Colony Algorithm Using Teaching Strategy«, in: *Computing*, 103(7), S. 1439–1460.
- Mondada, Lorenza (2021): *Sensing in Social Interaction. The Taste for Cheese in Gourmet Shops*, Cambridge: Cambridge University Press.
- Murdock, Graham/Michael Pickering (2009): »The Birth of Distance: Communications and Changing Conceptions of Elsewhere«, in: Michael Bailey (Hg.), *Narrating Media History*, London: Routledge.
- Parks, Lisa/Nicole Starosielski (Hrsg.) (2015): *Signal Traffic. Critical Studies of Media Infrastructures*, Champaign: University of Illinois Press.
- Pentland, Alex (2012): »Society's Nervous System: Building Effective Government, Energy, and Public Health Systems«, in: *Computer*, 45(1), S. 31–38.
- Roitman, Haggai/Iván Cantador/Miriam Fernández (2011): »Workshop on Multimodal Crowd Sensing (CrowdSens 2012)«, in: *Proceedings of the 21st ACM International Conference on Information and Knowledge Management (CIKM '12)*, S. 2770–2771.
- Rosenberg, Daniel (2013): »Data before the Fact«, in: Lisa Gitelman (Hg.), »Raw data« *Is an Oxymoron*, Cambridge, Massachusetts: The MIT Press, S. 15–39.
- Schabacher, Gabriele (2022): *Infrastruktur-Arbeit. Kulturtechniken und Zeitlichkeit der Erhaltung*, Berlin: Kulturverlag Kadmos.
- Schilt, Carl R./Kenneth S. Norris (1997): »Perspectives on Sensory Integration Systems: Problems, Opportunities, and Predictions«, in: Julia K. Parrish/William M. Hamner (Hg.), *Animal Groups in Three Dimensions*, Cambridge: Cambridge University Press, S. 225–244.
- Scholz, Sebastian (2021): »Sensing the ›Contemporary Condition‹: The Chronopolitics of Sensor-Media«, in: *Krisis | Journal for Contemporary Philosophy*, 41(1), S. 135–156.
- Schröter, Jens (2023): »Medienarchäologie der digitalen Medien«, in: Sven Stollfuß/Laura Niebling/Felix Raczkowski (Hg.), *Handbuch Digitale Medien und Methoden*, Wiesbaden: Springer, S. 1–15.

- Schüttpelz, Erhard/Sebastian Gießmann (2015): »Medien der Kooperation. Überlegungen zum Forschungsstand«, in : Navigationen – Zeitschrift für Medien- und Kulturwissenschaften, 15(1), S. 7–55.
- Seberger, John S./Geoffrey C. Bowker (2021): »Humanistic Infrastructure Studies: Hyper-Functionality and the Experience of the Absurd«, in: Information, Communication & Society, 24(12), S. 1712–1727.
- Sprenger, Florian (2018): »Zehn Elemente einer Medientheorie mobiler Adressierung«, in: Oliver Ruf (Hg.), Smartphone-Ästhetik, Bielefeld: transcript Verlag, S. 243–268.
- Sprenger, Florian/Christoph Engemann (2015): »Im Netz der Dinge. Zur Einleitung«, in: Florian Sprenger/Christoph Engemann (Hg.), Internet der Dinge: Über smarte Objekte, intelligente Umgebungen und die technische Durchdringung der Welt, Bielefeld: transcript Verlag, S. 7–58.
- Star, Susan L./Karen Ruhleder (1996): »Steps Toward an Ecology of Infrastructure: Design and Access for Large Information Spaces«, in: Information System Research, 7(1), S. 111–134.
- Thielmann, Tristan (2014): »Mobile Medien«, in: Jens Schröter (Hg.), Handbuch Medienwissenschaft, Stuttgart: Metzler, S. 350–359.
- Thielmann, Tristan (2019): Sensormedien. Eine medien- und praxistheoretische Annäherung, Siegen: Universität Siegen.
- van Geenen, Daniela/Kaerlein, Timo (2023): »Modalitäten von Kritik in Praktiken des Urban Sensing«, in: Andreas Beinsteiner/Nina Grünberger/Theo Hug/Suzanne Kapelari (Hg.), Ökologische Krisen und Ökologien der Kritik, Innsbruck: Innsbruck University Press, S. 89–108.
- Vehlken, Sebastian (2012): Zootechnologien. Eine Mediengeschichte der Schwarmforschung, Zürich: diaphanes.
- Vehlken, Sebastian (2016): »Multimodal Crowd Sensing«, in: Mathias Denecke/Anne Ganzert/Isabell Otto/Robert Stock (Hg.), ReClaiming Participation: Technology, Mediation, Collectivity, Bielefeld: transcript Verlag, S. 51–66.
- Weiser, Mark/Brown, John S. (1995): Designing Calm Technology. Abrufbar unter: <https://people.csail.mit.edu/rudolph/Teaching/weiser.pdf> (Stand: 23.07.2024).
- Zuboff, Shoshana (2020): The Age of Surveillance Capitalism. The Fight for a Human Future at the New Frontier of Power, New York: Public Affairs.

Digital Fashion im Online-Shop

Der un reale Interaktionsraum als sozialer Handlungsraum

Helga Behrmann

Abstract Dieser Beitrag erörtert Online-Shops als un realen Interaktionsraum, in welchem Digital Fashion zum Verkauf angeboten wird. Hier treten Anbietende und Kund*innen in einen interessensgeleiteten medialen Austausch. Kauf und Nutzung der virtuellen Kleidungsstücke finden in einem medialen Dispositiv statt. Visuelle Kommunikation entfaltet sich in der Interaktion mit einem distanzierten Dritten. Während Goffmann eine Face-to-Face-Interaktion als einen Zustand der gegenseitigen Offensichtlichkeit beschreibt, zielt Online-Kommunikation gerade darauf ab, alltägliche Routinen mit Nicht-Anwesenden zu teilen. Vorstrukturierte Navigationspfade, digitale Körperpraktiken und die Komposition von Sehnsuchtsorten ermöglichen den Nutzer*innen, virtuelle Kleider in ihrem Sinn zu formen. Digital Fashion unterliegt subjektiven Qualitätskriterien, welche über Erlebnisqualität, Neuheit, Status und Partizipation online ausgehandelt werden.

Keywords Digital Fashion; Virtual Dress; Kleiderhandeln; Online-Shop; Sozialer Raum; Synthetische Situation

1. Digital Fashion als Gegenstand mediatisierter Interaktion

Im Mai 2019 verkauft das niederländische Unternehmen *The Fabricant* das Kleid mit Namen *Iridescence* für 9500 Dollars. Ein immaterielles¹ Kleidungsstück, ein Datensatz, der auf einem Screen sichtbar wird, aber nie auf dem Körper getragen werden kann. In nur wenigen Jahren – beflügelt durch die Pandemie – etabliert sich Digital Fashion zu einer neuen Kleiderpraxis, die auf eine komplett digitale Infrastruktur angewiesen ist. Während der Anwendung entsteht medienbedingt ein temporärer Interaktionsraum, in welchem real-anwesende, nicht-anwesende und un reale Akteur*innen über virtuelle Kleider miteinander kommunizieren. Dieser Beitrag

1 Immaterialität bezieht sich hier auf die fehlende textile Materialität, in einer medienmaterialistischen Perspektive beruhen auch mediale Prozesse auf materialen Ressourcen und Stoffströmen für Energie und Infrastruktur (Peters 2015: 38).

untersucht die Frage, wie die medialen Bedingungen der Digital Fashion mit ihrer medienspezifischen Konstitution des Interaktionsraums, ihren Handlungsoptionen und ihrem Zusammenspiel der ungleichen Akteur*innen soziale Ordnung hervorbringen. Virtuelle Kleidungsstücke werden unter anderem im Online-Shop *DRESSX*² zum Verkauf angeboten, dieser steht beispielhaft für die Untersuchung. Methodisch werden anhand von Screen-Kontext-Analysen die Text-Bild-Bezüge im Online-Shop herausgearbeitet, um im Vergleich zu textilen Kleidungspraktiken die relevanten medienspezifischen Besonderheiten darzustellen.

Ein virtuelles Kleid lässt sich als Kleiderform definieren, welche digitale un reale Körper bedeckt. Physikalische Eigenschaften von Textilien wie Glanz, Gewicht oder Elastizität werden simuliert. Das virtuelle Kleid spricht mehrere Sinne an und lässt sich über mehrere Bild-, Text- und Symbolebenen erschließen, es bildet ein multimodales und multicodales Angebot (vgl. Lobinger 2012: 72). Digital Fashion oszilliert ebenso wie traditionelle Mode zwischen Kleid, Bild, Neuheit, Investition, Medium, Konsum und Kommunikation. Während sich traditionelle Kleiderpraktiken in eine Auseinandersetzung mit Modebildern und textilen Kleidungsstücken auffächern, entfällt bei Digital Fashion die physische Erfahrung der leiblich-körperlichen Aneignung. Zugleich bedingen Mediengeräte den Zugang. Folglich findet eine neue, medial distanzierte Auseinandersetzung mit Kleidung und Körperlichkeit statt (vgl. Crewe 2013: 17): die digitale körperliche Aneignung.

Digital Fashion unterliegt einer Zweckbestimmung. Bildautor*innen kreieren die digitalen Kleiderformen und Programmierer*innen binden die Bilder in Online-Plattformen ein, die den Zugang möglich machen. Im Zusammenwirken entsteht gemäß den Anforderungen von Unternehmen ein Medienprodukt, in der Definition von Gläser – ein Produkt zur Befriedigung von Informationsbedürfnissen³, welches über technische Medien verbreitet wird (vgl. 2008: 79). Das Medienprodukt ist zugleich ein Verbundprodukt: Im gleichen Herstellungsvorgang entstehen – untrennbar – zwei unterschiedliche Produkte⁴, weil sie auf zwei unterschiedlichen Märkten agieren. Auf dem Rezipient*innenmarkt und auf dem Werbemarkt (vgl. ebd.: 153) überlappen sich mehrere Funktionen⁵. Auf der einen Seite stehen die Hersteller*in-

-
- 2 Quelle *DRESSX* (o.J.): Startseite. Abrufbar unter: <https://dressx.com/> (Stand: 11.01.2024).
 - 3 Neben Nachrichten, Bildung, Kultur und Unterhaltung umfasst der Begriff »Information« ausdrücklich auch die Botschaften der Werbung und Marktkommunikation (vgl. Gläser 2008: 79).
 - 4 Zum Vergleich: Auch eine Modezeitschrift ist ein Medienprodukt und damit ein Verbundprodukt. Die gleiche Zeitschrift wird an zwei Zielgruppen adressiert. Mit redaktionellem Inhalt werden Modeinteressierte angesprochen, welche den Kaufpreis für ein einzelnes Exemplar bezahlen; mit Anzeigenflächen werden Unternehmen angesprochen, welche ihre Werbung in der gesamten Auflage platzieren und dafür bezahlen.
 - 5 Gläser differenziert in Informations-, soziale, politische und ökonomische Funktionen (vgl. 2008: 340).

nen und Plattformbetreiber*innen als Anbieter*innen, auf der anderen Seite befinden sich zwei Gruppen von Adressat*innen: die Kund*innen, welche die Dienstleistung kaufen, und die Werbekund*innen, also Unternehmen, die für den Zugang zu Zielgruppen bezahlen, um personalisierte Werbung zu platzieren.

Digital Fashion ist dann auffindbar, wenn Nutzer*innen eigeninitiativ handeln. Erst das Aufsuchen und Auswählen der Angebote und das Hochladen persönlicher Fotos machen eine individualisierte Dienstleistung möglich. Nach Gläser beeinflussen Konsumierende das Resultat des Produktes und sind selbst für das Ergebnis verantwortlich (vgl. ebd.: 157). Damit wird deutlich, dass die Nutzung von Digital Fashion nicht vollumfänglich durch die Gestaltung und die Produktplatzierung vorgegeben werden kann, sondern erst durch individuellen Gebrauch, kulturelle Akzeptanz und andere soziale Faktoren (vgl. Ziemann 2011: 17f.) geformt wird. In der Perspektive der Nutzer*innen erscheint die soziale Funktion im Vordergrund. Rekreation durch Unterhaltung und Eskapismus überlappen sich mit dem Bedürfnis der Nutzer*innen, sich in einer medialen Öffentlichkeit privat zu zeigen. Wie Schroer herausarbeitet, braucht das Individuum die Bestätigung durch andere (vgl. 2010: 282).⁶ Weil aber Selbstbildnisse mit textilen Kleidern an den eigenen leiblichen Körper und an reale geografische Orte gebunden sind, welche nicht immer dem Ideal entsprechen, sollen virtuelle Kleider als »Postdigital Fashion« (Musto 2022: 137) nun diese Lücke schließen: Personalisierte Modebilder mit einer Distanz zum eigenen Körper an Orten der freien Wahl werden nun auf spielerische Weise möglich.

2. Die digitale Konstitution des unrealen Interaktionsraums

Die Dreiecksbeziehung aus virtuellem Kleid, Akteur*innen und Ort formt den unrealen Interaktionsraum in einem technisch-medialen Dispositiv. Nach Misoch werden damit virtuelle Räume generiert, welche »eine Rahmung für die an sich ortlosen Kommunikationsprozesse bieten« (2006: 59). Nutzer*innen befinden sich an mehreren Orten zugleich: an einem »real-geografischen Ort« (ebd.) vor dem Medengerät und an »virtuellen ›Orten‹ im Cyberspace, die sich ›hinter‹ dem Bildschirm eröffnen« (ebd.). Misoch bezeichnet dies als eine »Pluralisierung der kommunikativen Räume« (ebd.). Hier wird ein Handeln mit dem virtuellen Kleid möglich. Böhme spricht in diesem Zusammenhang von einem »Spiel- und Handlungsraum«

6 Schroer bezieht sich auf Goffmans Analysen zur Selbstdarstellung im Alltag (vgl. Goffman/Dahrendorf 2017: 231), welche das theatralische Moment der alltäglichen Interaktion darstellen. Auch König verweist auf den Wechsel vom Narzissmus zum Exhibitionismus und stellt heraus, wie im Schauen und Sich-Beschauen-Lassen auch soziale Verhältnisse offengelegt werden (vgl. 1985: 119ff.).

(2004: 134); für ihn handelt es sich »um eine Verschränkung von Darstellungsräumen und dem Raum leiblicher Anwesenheit« (ebd.: 139). Während der Raum der leiblichen Anwesenheit durch den eigenen Körper spürbar und damit »in reinem Selbstbezug« (ebd.: 133) gegeben ist, ermöglicht der Blick auf den Bildschirm die Wahrnehmung weiterer – unrealer – Räume. Weil immer nur Ausschnitte von Raumsituationen dargestellt werden, spricht Mersch von einer »Fraktalisierung von Räumen« (2011: 59). Die Kombination der Fragmente bezeichnet er als »Multiple Räume« (ebd.), welche entsprechend neue Handlungsformen hervorrufen, weil Nutzer*innen auf die gleichzeitige Anwesenheit in mehreren Räumen reagieren (vgl. ebd.). Löw beschreibt Raum als ein prozesshaftes Phänomen (vgl. 2023: 224ff.): Raum entsteht durch das Handeln von Menschen. Mit dem Begriff »Spacing« (ebd.: 158ff.) verdeutlicht sie einerseits die Platzierung von sozialen Gütern, Menschen und Orten in einer sichtbaren, flüchtigen oder dauerhaften Anordnung und andererseits den Prozess des Bauens selbst. Löw nimmt Bezug auf die alltäglichen Routinen der Raumerzeugung, wobei die Konstitution des Raums in den Hintergrund tritt und institutionell unsichtbar gemacht wird. Insbesondere stellen Plattformbetreiber*innen⁷ über eine technische Infrastruktur kurzfristige soziale Raumsituationen her, wobei die agierenden Institutionen wie etwa Suchmaschinen, Modeunternehmen, Mediendienste, Energieanbieter und Programmhersteller nicht als handelnde Personen wahrgenommen werden können, sie verbergen sich hinter der Oberfläche der Programmstruktur.

Erfahrungen aus traditionellen Kleiderkäufen im stationären oder Online-Handel werden als Leitmotive für die Programmstruktur thematisiert. Nach Gottschalk verschmelzen die beiden zuvor nach offline und online getrennten Erfahrungsbereiche an einem Terminal⁸ (vgl. 2018: 1): Menschliche Erfahrung und mediale Interaktion beeinflussen einander. Er weist darauf hin, dass Nutzer*innen nur mit ihren leiblichen Empfindungen navigieren können (vgl. ebd.). Alltagsroutinen wie die Suche und Auswahl von neuen Kleidungsstücken, der Kaufvorgang und das Sich-Kleiden werden zunehmend an die Kommunikationsstrukturen des Terminals angepasst und schließlich auch in der Virtualität ausgehandelt. Die Handlungen sind jedoch auf Hören, Sehen und Tasten beschränkt. Gottschalk verweist kritisch auf die Auswirkungen, denn »[...] our interactions with these technologies confuse the relation between mind and the body and sever the integration between the senses« (2021:

7 Diese neue Form der digitalisierten Märkte – als Plattformkapitalismus (vgl. Srnicek 2017), als Überwachungskapitalismus (vgl. Zuboff 2019) und als digitaler Kapitalismus (vgl. Staab 2019) – zeichnet sich dadurch aus, dass die Märkte nicht für alle zugänglich sind, sie von wenigen Unternehmen dominiert werden und Preise sich nicht zwingend aus Angebot und Nachfrage bilden.

8 Gottschalk bezeichnet Terminals als digitale Technologien, die Nutzer*innen aktivieren, auf eine Online-Umgebung zuzugreifen, Informationen zu finden und mit anderen zu interagieren (vgl. 2021: 101).

101f.). Weiter führt er aus, dass die Vielfalt und Anzahl der Funktionen, die zukünftig online erfolgen werden, exponentiell anwachsen werden, während die ausdifferenzierten manuellen Gesten auf sich wiederholende Bewegungen wie das Bedienen von Bildschirmen und Plastiktastaturen reduziert werden (vgl. ebd.). Für das digitale Modehandeln bedeutet dies, dass alle leiblich-körperlichen Bewegungsformen, welche sich aus der textilen Materialität bedingen – das Dehnen und Überstreifen von Kleidungsstücken, das Schließen von Knöpfen, das Sich-drehen im Faltenrock, das Hochschieben von Ärmeln und Hosenbeinen – dann in eine auf einem Bildschirm darstellbare Aktion überführt werden müssen. Die Interaktion, die sich auf physische Erfahrungen mit Kleidungsstücken gründet, splittet sich nun in haptische Erfahrungen mit dem Mediengerät und eine visuelle Auseinandersetzung mit den Bildern.

3. Mediale Bedingungen als Handlungsebenen zur unrealen Interaktion

Digital Fashion lädt zur ästhetischen Rezeption ein. Vermeintlich kostenlos und zielloos lassen sich leiblich anwesende Nutzer*innen treiben. Angelehnt an die Figur des Flaneurs⁹ oder der Passant*in lässt sich digitales Kleiderhandeln mit Flanieren im Internet¹⁰ vergleichen. Weil Nutzer*innen mit hoher Flanierneigung eine besondere Vorliebe für unterhaltende und kommunikative Angebote zeigen (vgl. Doedens 2010: 213), werden auch hier die Hauptmotive für die Nutzung vermutet. Bieger erläutert über die Ästhetik der Immersion die »Verwischung der Grenze zwischen Bildraum und Realraum« (2011: 75). Für sie sind es »Räume, in denen sich die Wirklichkeit der Welt und die Wirklichkeit des Bildes in der unmittelbaren Wirklichkeit körperlichen Erlebens konsolidieren« (ebd.).

Die vorgegebenen Handlungsoptionen im Online-Shop *DRESSX* sind über die entsprechenden Werkzeuge bereits aus anderen Online-Anwendungen bekannt. Sie reduzieren die Komplexität und ermöglichen vorstrukturierte Kommunikation: *Steuerungspfeile* sind über die Richtungen auch gleichzeitig Raumkoordinaten, mit

9 Die literarische Figur des Flaneurs und seine weibliche Entsprechung der Passante, wie sie bei Edgar Allan Poe, Charles Baudelaire, Walter Benjamin und anderen Autor*innen auftreten, sind von einem vorübergehenden Interesse an Dingen und Situationen, einer kurzen Verweildauer an einem Ort, selbstbestimmten Innehalten und Weiterziehen sowie dem Genießen des Augenblicks gekennzeichnet (vgl. Prose 2010). Der Flaneur, der die beginnende Konsumgesellschaft spiegelt, wurde in der Postmoderne zum Cyberflaneur (vgl. Hartmann 2004) beziehungsweise zum Phoneur (vgl. Luke 2005), der sich mit seinem Mediengerät durch die Infrastrukturen der modernen Lebenswelten treiben lässt.

10 Doedens unterscheidet prozessorientiertes und inhaltsorientiertes Surfen als zwei grundlegend verschiedene Surfstrategien. Nutzer*innen gehen jedoch nie ausschließlich nach nur einer der beiden Strategien vor (vgl. 2010: 219).

denen neue Teilräume eröffnet werden. Die *Lupe* dient dazu, sich dem Produkt zu nähern, das *Plus* zum Upload des Fotos ist eine Anweisung, aktiv zu werden und weitere Schaltflächen wie *Buy it now* ermöglichen den Gang zu Kasse. Kontaktfelder wie die *Sprechblase* eröffnen die Möglichkeit, Fragen an nichtanwesende Personen der Kund*innenberatung zu stellen. Über die *Sharing-Symbole* beschreitet man die virtuelle Bühne der Sozialen Medien. Aufgrund der technischen Bildbeschaffenheit können die Modelle auf dem Bildschirm gedreht, in Vorwärts- und Rückwärts-Bewegungen, im Wind, von innen oder auch in Vergrößerungen bis hin zu einzelnen Faserbewegungen gezeigt werden.

Weil der Datensatz nicht ohne Weiteres mit dem leiblichen Körper kompatibel ist, bedarf es medialer Operationen, um die Beziehung zwischen digitalen Kleidern und leiblichen Körpern herzustellen. Brachem und Stübbe unterscheiden grundsätzlich drei Möglichkeiten, virtuelle Kleidungsstücke zu tragen: Als *assembled* beschreiben sie das Hinzufügen virtueller Kleidung als eine zusätzliche Bildebene zu bereits aufgenommenen Fotos oder Videos des eigenen Körpers zu einem späteren Zeitpunkt; *simultaneous* ist das zeitgleiche Überblenden mit virtueller Kleidung an physischen, bekleideten oder unbekleideten Körpern, betrachtet durch Augmented-Reality-Brillen oder Social-Media-Filter; und als *omni-virtual* bezeichnen sie das Zeigen virtueller Kleidung an virtuellen Körpern, welche als anonyme oder individualisierte Avatare in Erscheinung treten (vgl. 2022: 1). Als Bildhintergründe können geografische oder fiktive Orte jeder Wahl eingesetzt werden. Bei *DRESSX* wird die *assembled*-Beziehung zwischen digital und körperlich in der medialen Interaktion erfahrbar: die Auswahl des virtuellen Kleides im Online-Shop, die Herstellung eines passenden Fotos mit dem leiblichen Körper, das Hochladen des Bildes, das Betrachten und Kommentieren der Bilder und das Weiterschicken in die Sozialen Netzwerke. Die Ergebnisse werden auf der Bildschirmoberfläche als Collage mit mehreren Ebenen sichtbar.

Als weitere Möglichkeit für virtuelles Kleiderhandeln bietet *DRESSX* ebenso wie andere Online-Shops an, Modelle als NFTs¹¹ zu erwerben. Damit werden die unendlich kopierbaren Daten zu originalen Sammlerstücken (vgl. Särämäkari 2021: 23f.) und dienen wie eine Aktie als Investitionsprodukt. Insbesondere Luxus-Labels greifen diese Variante von Digital Fashion auf, um Nutzer*innen den Zugang zu definierten Teilbereichen des Metaverse zu öffnen (vgl. Joy et al. 2022: 340).

Im Online-Shop werden persönliche Wünsche geweckt. Die Anordnung der Bedienfelder, die begleitenden Texte, das Aufrufen und Kombinieren von Raumsitua-

11 Designer*innen arbeiten an einer Online-Wertschöpfung: »They nicely echoed the footsteps of the blockchain technology, especially non-fungible tokens (NFTs), approaching towards the design and development of digital-only garments as tradable assets, and provided an excellent food for thought for educators and industry leaders to figure out the construct of tomorrow's fashion designers« (Sayem 2022: 139).

tionen und Produktvarianten in animierten Bildern regen Nutzer*innen an, immer weiter zu klicken. Über die getroffene Auswahl werden persönliche Vorlieben offengelegt, die im »Netzwerk der Begierde« (Kozinets 2022: 11) den Erlebniskontext bilden. Produktempfehlungen basieren auf Suchmaschinen, die algorithmisch gestützt möglichst passgenau auf die Bedürfnisse der Suchenden abgestimmt werden. »Die Arbeitsweise von Empfehlungssystemen macht es seitens ihrer Designer (sic!) erforderlich, eine große Bandbreite an Datenspuren mit Bedeutsamkeit zu versehen« erläutert Lehner (2017: 71). Das bedeutet, dass Aktionen aus den zuvor generierten Nutzer*innenprofilen über algorithmische Zuschreibungen als Konsumbedürfnisse interpretiert werden, um dann Waren als entsprechende Problemlösung präsentieren zu können. Ein unbestimmtes Interesse oder auch Neugierde werden in ein Begehren überführt, welches einen Produktkauf nach sich ziehen soll. In dieser Übersetzung liegt nach Lehner die Absicht, Nutzer*innen so einzustimmen, dass sie das gleiche Ziel wie der technische Aktant haben (vgl. ebd. 74f.), denn der »[...] Umschlagpunkt oder Hebel dieser gemeinsamen Zielsetzung ist die Übersetzung von Bedürfnissen in Empfehlungen und, wichtiger noch für die Rentabilität des Empfehlungssystems, von Empfehlungen in Bedürfnisse« (ebd.). Die bildbegleitende Argumentation in Textform hilft Kund*innen, ihren Kauf vor sich oder anderen zu rechtfertigen. Wenn Kund*innen sich vom Angebot angesprochen fühlen, liegt das also nicht zwingend an einem eigenen konkreten, vorher ausdifferenzierten Bedürfnis, sondern vor allem am Design der Produktempfehlung, der Ansprache der Zielgruppe und an den emotionalen Aufforderungen zu Aktionen. Wie Kheladi et al. in ihrer Studie zeigen, sind Freude, soziale Interaktivität und Neugier die wichtigsten Prädiktoren für den Kauf virtueller Kleidung, denn Nutzer*innen suchen im Rahmen des digitalen Modekonsumtrends vor allem nach hedonistischen, sozialen und technologischen Befriedigungen (vgl. 2023: 1397).

Traditionelle Modestädte verschmelzen mit neuen Moderäumen (vgl. Crewe 2018). Im Online-Shop DRESSX überführen die mediale Warenpräsentation und insbesondere die begleitenden Textaussagen die Konsumsphäre aus den geografisch-realen Erlebnisorten wie Warenhäusern und Einzelhandel (König 2009: 92ff.), welche von leiblich-körperlicher Fortbewegung, Face-to-Face-Situationen und haptischem Erspüren der Kleidungsstücke geprägt sind, in einen unrealen Erlebnisort mit immersiven Erlebnissen. Eine emotionale und ästhetische Auseinandersetzung mit den Bildern kann hier die eigentlichen Körperhandlungen ersetzen, die traditionell bei einem textilen Kleidungsstück ausschlaggebend oder bestätigend für eine Kaufentscheidung waren: »Das reale Anprobieren von Bekleidung und der haptische Kontakt zwischen Körper und Textil entfällt zugunsten einer Wahrnehmung von Ware und Körper als digitalem, animierbarem Bild« (Gaugele 2005: 445). Dabei gibt es offenbar sehr genaue Kenntnisse darüber, wie diese Bilder auszusehen haben, denn die Visualisierung der Kleider erfolgt gemäß »Standardisierung und sozialer Normierung von Körpern« (Döring/Draude 2012: 71) und als »Repräsentation

tion heteronormativer Geschlechterbilder« (Gaugele 2014: 125). Die soziokulturell geprägten Körperideale lösen sich auch in der »post-industriellen Fashion« (Atkinson 2017) nicht auf. In den Bildern reproduziert, bilden sie Normenkomplexe, die von den Nutzer*innen nicht hinterfragt werden. In der Anprobe-App¹² werden Vor-Ansichten als virtuelle Anprobe gezeigt, bevor man sich zum Kauf der Datei entschließt. In der App erfährt man, welche Posen sich für welche Kleider eignen und wie der Körper beleuchtet werden soll, um das Ergebnis, eine Fotomontage mit dem virtuellen Kleid, optimal zu gestalten. Der Wunsch, sich in neuen Kleidern möglichst ansprechend zu präsentieren, lässt sich im Bild schneller und leichter erfüllen, denn die »Anpassung an das Schönheitsideal, den Schock zur Differenz kann man sich ersparen, durch ein Sample vordefinierter Avatarbausteine kann man ein ›kreatives‹ virtuelles Alter-Ego gestalten« (Gaugele 2005: 445). Die Vorab-Präsentation hat eine persuasive Wirkung: Entgegen einem traditionellen Online-Kauf, der auf zukünftige Ereignisse ausgerichtet ist, wenn das Konto belastet wird und ein Paket ausgeliefert wird, dessen Inhalt für Überraschungen sorgen kann, wird bei Digital Fashion der volle Leistungsumfang bereits vor dem Kauf dargestellt. Die Kaufhandlung endet nicht mit einem Versprechen auf ein textiles Produkt, sondern findet erst nachträglich statt, wenn man mit der Bildinszenierung zufrieden ist und dann das Bild zur weiteren persönlichen Verwendung freischalten lassen möchte.

4. Interaktion mit leiblich anwesenden, abwesenden und unrealen Interaktionspartner*innen

Das Konzept der Digital Fashion im Online-Shop führt die historische Entwicklung fort, Waren ohne Verkaufspersonal zu präsentieren. Stihler erläutert im Zusammenhang mit der Entstehung der Kaufhäuser, wie durch Anonymität Kaufanreize verstärkt wurden¹³: »Aktiver und verbaler Austausch zwischen Verkäufer und Käufer wurde ersetzt durch die passive stumme Antwort des Konsumenten gegenüber Gegenständen« (1998: 82). Sie wertet diesen Prozess als Beispiel, »wie der ›Prozeß der Zivilisation‹ (sic!) Aggressionen und Gefühle gegenüber Mitmenschen zähmt und stattdessen auf Gegenstände gerichtete Wünsche und Gefühle verstärkt« (ebd.). Während für die Suchenden die Herausforderung darin besteht, das beste Angebot oder die am besten geeignete Hilfestellung für die Bewertung und Verarbeitung

12 Der Online-Shop *DRESSX* bietet die kostenlose Anprobe-App *DRESSX.AI* an, die eine Vor-schau der digitalen Kleidungsstücke auf hochgeladenen Fotos der Nutzer*innen generiert (Quelle: *DRESSX.AI* (o.J.): Startseite. Abrufbar unter: <https://dressx.ai/> (Stand: 11.01.2024)).

13 Jeder hatte Zugang durch freien Eintritt ohne Kaufverpflichtung, die Ware wurde zurückgenommen oder umgetauscht. Dank der vorproduzierten Größen war ein Wiederverkauf möglich (vgl. Stihler 1998: 82).

der Angebote zu erhalten, gilt es für Anbieter*innen, in der Vielzahl der Angebote aufzufallen. Ziemann erläutert, wie mediale Angebote, aber auch kommunikative Kontakte, soziale Verpflichtungen und individuelle Bedürfnisse Aufmerksamkeit erfordern und damit in Konkurrenzverhältnisse treten (vgl. Ziemann 2011: 114). Entscheidungen, die auf der ursprünglichen Face-to-Face-Situation basieren, sind von einer Unausweichlichkeit geprägt, denn sobald »ein Individuum [...] in die unmittelbare Gegenwart eines anderen gerät, tritt eine Grundbedingung des gesellschaftlichen Lebens außerordentlich deutlich hervor: seine folgenschwere Offensichtlichkeit« (Goffman 2001: 58). Goffman argumentiert, dass der Verlauf vom persönlichen Agieren und Reagieren der Akteur*innen abhängt und der Ausgang grundsätzlich offen ist. Mediale Räume wie Home-Office, Soziale Netzwerke, Chatrooms oder Online-Käufe unterscheiden sich grundlegend von der Face-to-Face-Situation, da hier alltägliches Handeln mit Nicht-Anwesenden und unrealen Interaktionspartner*innen häufig zeitversetzt stattfindet. Weil auf den Screens jedoch nur Bilder und Texte zur Verfügung stehen, lässt sich Imdahls Begriff der Simultanstruktur aus der Bildanalyse hier nutzbar machen (vgl. Imdahl 1996: 23). Das Layout des Screens ermöglicht einen Überblick, alle Informationen können quasi gleichzeitig erfasst werden. Die gegenseitige Offensichtlichkeit aus der Face-to-Face-Situation wird hier zu einer anonymen Komposition der Angebote.

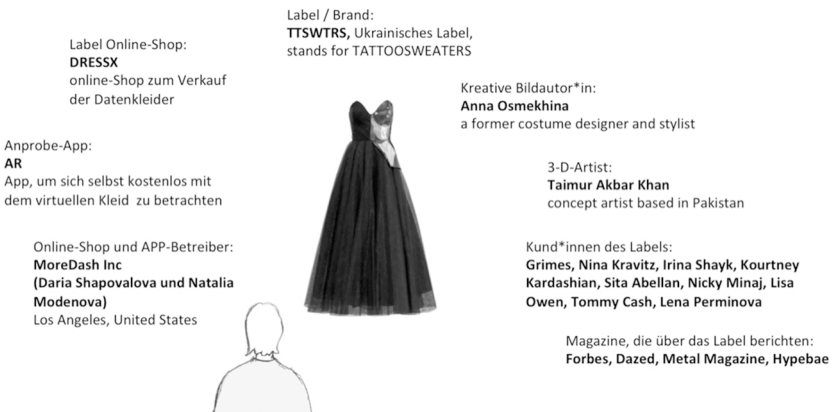
Am Beispiel des Online-Shops *DRESSX* wird durch das Design schnell klar, dass Rezipierende auf ihre sozialen Bedürfnisse angesprochen werden. Für die Kontaktaufnahme mit dem unrealen Interaktionspartner *DRESSX* steht das prominente Kontakt-Feld *CONTACT US* oder die unpersönliche E-Mail *hello@dressx.com* zur Verfügung. Deren Nutzung birgt jedoch Hemmschwellen, denn man muss den eigenen Namen und E-Mail-Adresse angeben sowie ein konkretes Anliegen formulieren. Selbst unter den Stellenangeboten finden sich keine realen Kontaktpersonen mit Firmenname, Adresse oder Sitz. Die höchst vertraulichen Inhalte einer Bewerbung sind an eine unpersönliche E-Mail-Adresse zu schicken. Dabei ist unklar, ob man mit menschlichen oder nicht-menschlichen Interaktionspartner*innen in den Austausch tritt, oder ab welchem Schritt ein Wechsel von einem zum anderen vollzogen wird. Selbst wenn man als Nutzer*in davon ausgehen darf, dass hinter den vorstrukturierten Kommunikationsangeboten letztendlich andere Menschen stehen, die ihrerseits Interesse an dem Austausch haben, so lassen sich diese an keiner Stelle als Personen wahrnehmen. Stattdessen bietet *DRESSX* umfangreiche Informationen über Kollaborationen, Labels und Partner sowie FAQ und Erklärungen zur Funktionsweise der Digital Fashion an.

Weil die Struktur der Online-Anwendungen weder eine Interaktion zwischen Menschen zulässt, noch einen in ein Medium übertragenen Prozess des Sich-Kleidens darstellen kann, spricht Knorr Cetina von einer synthetischen Situation (vgl. 2009: 65): eine Umgebung, welche durch Leistungsüberblick, Kommunikationselemente und Steuerungswerkzeuge angereichert und zeitweilig oder dauerhaft her-

gestellt wird und welche die Reaktionspräsenz des anderen ohne dessen physische Präsenz ermöglicht. So kann der Online-Kauf auch ohne Verkaufspersonal, Anprobe, den Gang zur Kasse und das Zücken der Geldbörse oder Kreditkarte stattfinden.

Der un reale Raum lädt zur medialen Interaktion mit nicht-anwesenden und un-realen Interaktionspartner*innen ein: Die familiäre Ansprache in den begleitenden Texten nutzt häufig Personalpronomen wie *wir, uns, unser*. Was als Personen wahrgenommen werden soll, entpuppt sich als unpersönliches Unternehmen, dessen Sitz und Gründer*innen nicht auf der Plattform zu finden sind. Erst weiterführende Recherchen fügen ein Netzwerk aus Namen, Personen, Berufsbezeichnungen und Institutionen mit weltweiten geografischen Verortungen über das virtuelle Kleid zusammen. Dies dient dazu, das vermeintlich immaterielle Produkt in irgendeiner Weise begreifbar zu machen, und wenn man es schon nicht auf dem Körper tragen kann, so kann man doch darüber reden, es im Bild zeigen und den sozialen Stellenwert aushandeln.

*Abbildung 1: Um die leiblich anwesende Nutzer*in gruppieren sich nicht anwesende Betreiber*innen der Verkaufs-Plattform und der Anprobe-App als Unternehmen, hinter dem zwei Gründer*innen und viele weitere namenlose Mitarbeiter*innen stehen. Weiterhin werden Label und Brand genannt, der Name der Designerin und des 3-D-Künstlers, die alle an der Herstellung des virtuellen Kleides beteiligt sind. Das Name-Dropping setzt sich fort, indem Kund*innen des Labels und Magazine, die bereits darüber berichtet haben, ebenfalls aufgezählt werden (eigene Darstellung basierend auf dressxx.com).*



Die raumzeitliche Einheit der Face-to-Face-Situation ist zwar unterbrochen, die Wirkzone des jeweils Anderen besteht jedoch in dieser Begegnung fort. Die abwesenden Anbieter*innen bestimmen über die Kommunikationstools als un-reale Kommunikationspartner*innen den narrativen Rahmen, in dem diese Handlungen

stattfinden. Vermeintlich menschliche Interaktion wird zunehmend von neuen Akteur*innen beeinflusst: »Mit der gesteigerten Informationalität synthetischer Situationen kommen auch Algorithmen mit in die Interaktionssituationen hinein« (Reichmann 2018: 103). Skopische Medien¹⁴ bilden einerseits den Raum, in welchem Interessierte auf das virtuelle Kleid treffen, andererseits greifen sie aber auch aktiv in die Interaktion ein. Die aus dem traditionellen Mode-Marketing bekannte Dreiteilung in Präsentation der Modefotos mit der Rezeption der Betrachter*innen, Kauf und Aneignung der textilen Kleider wird beim virtuellen Kleid überwunden (vgl. Behrmann 2024: 120). Reichmann betont dabei die Reflexivität der skopischen Medien: »Das, was die Teilnehmer in einer synthetischen Situation tun, spiegelt sich in dem, was das skopische Medium abbildet« (2018: 94). Zusätzlich zu den menschlichen Akteur*innen, die als Anbieter*innen, Zwischenhändler*innen oder Nutzer*innen über das virtuelle Kleid interagieren, übernehmen jetzt Algorithmen die Aufgabe, die Kommunikation zu strukturieren und auftretende Probleme zu lösen.¹⁵ »Tatsächlich ermöglichen die immensen personenbezogenen Datenmengen und immer intelligenter werdende Algorithmen [...] ganz neue Möglichkeiten für die Manipulation von Meinungen und Entscheidungen« (Reisch 2020: 87). Das Design der digitalen Schnittstellen verwendet Nudges¹⁶, um bestimmte Verhaltensweisen zu begünstigen oder zu verhindern. Über die illustrative Darstellung der Kommunikationsinhalte scheint eine freie Wahl möglich. Illusion und Immersion »sind in ihrer Wirkungsausrichtung von Grund auf manipulativ; beide setzen auf eine affektive Bewegung des Betrachters; beide sind auf dessen affirmative Haltung angewiesen« (Bieger 2011: 84). Das Interesse der Nutzer*innen wird in ein grundsätzliches Kaufinteresse überführt, indem ästhetische Erfahrungen die Bewegung in eine vorgegebene Richtung anstoßen.

Während die unrealen Akteur*innen für die Dauer der gesamten unrealen Interaktion gleich anonym bleiben, ändert sich das für die Nutzer*innen. Klick für Klick werden Interessen, Wünsche und Sehnsüchte bis zu den Kontodaten offenbart. Alle Interaktionspartner*innen werden von Interessen geleitet, die Begegnung ist jedoch keineswegs auf Augenhöhe. Im unrealen Raum, von einer institutionellen

14 Skopische Medien visualisieren auch nicht-physische Orte: »Sie nehmen also einerseits Informationen aus der Interaktion auf. Gleichzeitig vermengen (synthetisieren) sie diese Informationen mit Daten, die potenziell global gespeichert sein können, erstellen daraus etwas Neues und machen dieses neu gestaltete Daten- und Informationshybrid in der synthetischen Situation präsent und situational erfahrbar« (Reichmann 2018: 94).

15 »Wenn Algorithmen Probleme lösen, dann manipulieren sie Daten, das heißt sie sortieren, strukturieren, ordnen Daten oder wandeln diese um. Algorithmen und Datenbestände sind also eng miteinander verknüpft und ohne das jeweils andere nutzlos« (Reichmann 2018: 97f.).

16 Mit Nudges sind hier Steuerungselemente gemeint, die durch ihre Platzierung, ihre Voreinstellungen und ihre Abfolge auf Handlungen von Nutzer*innen einwirken (vgl. Reisch 2020: 88).

Unsichtbarkeit geprägt, findet die Aushandlung auf der Vorder- und der Hinterbühne statt.¹⁷ Kund*innen, welche die von ihnen bezahlten Datenkleider nutzen, nehmen dabei kaum oder gar nicht wahr, wie die Wirtschaftsunternehmen, welche die Plattformen betreiben, ihre Daten erheben und zum eigenen Vorteil benutzen. Digital Fashion hat das Potenzial, Kund*innen möglichst lange aktiv im Online-Shop zu halten und unterstützt Plattformbetreiber*innen dabei, Zugänge zu Zielgruppen zu öffnen und darüber hinaus Daten zu sammeln. Die Quantifizierung der Kund*innen (vgl. Zuboff 2019) fördert so eine neue Entwicklung, die Interaktion zwischen Nutzer*innendaten und Algorithmen, zwei unrealen Akteuren. Das Zusammenspiel von menschlichen anwesenden und abwesenden sowie nicht-menschlichen Akteur*innen, die jedoch in der synthetischen Situation nicht explizit voneinander unterschieden werden, bestimmt nun die Rezeption von Digital Fashion.

Allgegenwärtig ist das Angebot an Produkten, welches nahelegt, die vormaligen sozialen Bedürfnisse durch ein Konsumerlebnis zu kompensieren. Algorithmen als neue Akteur*innen schlagen immer neue Produkte vor, die das Ergebnis der Datenverarbeitung zeigen. Lehner konstatiert: »Das einzige Bedürfnis, das der menschliche Akteur als Nutzer (sic!) von digitalen Plattformen haben muss, ist das Bedürfnis nach einem Bedürfnis« (Lehner 2017: 78). Er erläutert, wie über den technischen Umweg aus den diffusen sozialen Bedürfnissen konkrete Konsumbedürfnisse generiert werden und gleichzeitig den Nutzer*innen die Anstrengung abgenommen wird, eigenständig dieses Bedürfnis zu konkretisieren; letzten Endes handelt es sich »um eine Befreiung davon, die Investition tätigen zu müssen, Bedürfnisse zu kultivieren« (ebd.). Nutzer*innen sind in dieser Situation weniger als Konsument*innen im Sinne von Käufer*innen zu betrachten, die sich mit einem klar ausformulierten Bedarf auf die Suche machen. Vielmehr lässt sich ein subjektiver Standpunkt¹⁸ der Rezipierenden ausmachen, nach dem jeweils das Rezeptionserlebnis ausschlaggebend für ein weiteres Verharren auf der Angebotsseite ist. Die algorithmische Kultur zielt darauf ab, Nutzer*innen ihre Agency abzunehmen und sie dauerhaft über ihre Einbindung in die Serviceleistung (vgl. Kozinets 2022: 27) am Ball zu halten.

17 Echtzeit-Datensammlungen werden verkauft, ohne dass Preise oder auch Strategien zur Vermeidung der Abschöpfung der Daten veröffentlicht werden. Die Online-Handlungen müssen lediglich »erst von einem physisch-kontinuierlichen Prozess in die Form diskreter Datenspur gebracht werden. Jeder Klick wird zum Beispiel als Spur in ein diskretes Datum übersetzt, mit welchem dem Nutzer Motive, Interessen und Neigungen unterstellt werden« (Lehner 2017: 71).

18 Hellmann unterscheidet hier die Entscheidungs rationalität als Außenansicht der Konsumierenden und die Erlebnisrationalität als Innenansicht der Konsumierenden (vgl. 2018: 75).

5. Der Selbstbezug als Form der sozialen Ordnung

Anstelle der traditionellen textilen Kleider werden Bilder zum eigentlichen Konsumgegenstand. Das Design ist auf Zielgruppen ausgerichtet, die schnell, intuitiv und plakativ mit Bildern umgehen und die Eigeninitiative, Mühe und Zeit in ihre Selbstinszenierung investieren: »Nicht die ›Sender‹ sind es, welche eine ›Botschaft‹ auf einen machtlosen ›Empfänger‹ einströmen lassen. Es sind die Nutzer (sic!) selber, die sich ihr Opus auf den Bildschirm bringen« (Rusch et al. 2007: 103). Und so werden die virtuellen Kleider durch Auswahl, Kombination und kreative Nutzung zum Werk der Kund*innen, auch wenn die Vorlagen der virtuellen Kleider von Bildautor*innen kreiert wurden. Weil der praktische Nutzen, den eigenen Körper zu bekleiden, beim virtuellen Kleid nicht gegeben ist, unterliegt das Design umso mehr der Herausforderung, dieses Produkt für die mediale Nutzung zu optimieren. »Design hat also nicht nur mit der Befriedigung praktischer Bedürfnisse zu tun, sondern immer auch mit der Gestaltung von Bedeutungen, Begehrlichkeiten und sinnlichen Erfahrungen«, konstatiert Prinz (2019: 290). Das virtuelle Kleid beeinflusst durch seine Gestaltung, ob und wie sich Akteur*innen zu ihrer Umwelt in Beziehung setzen (vgl. ebd.: 297).

Als Datensatz lässt das virtuelle Kleid zu, dass es am Bildschirm bearbeitet und individualisiert wird. Folglich liegt auch die Nutzung im unrealen Raum. Ein Raum, in welchem die Nutzer*in die einzig leibliche Anwesende ist. Diese Person wird über das Design des Online-Shops *DRESSX* nicht nur auf ihr Bedürfnis nach Kontakt und Austausch, sondern auch nach Kompetenz und Status angesprochen. Mediale Aufmerksamkeit wird in Aussicht gestellt, weil das Ergebnis des Sich-virtuell-Kleidens in den eigenen Social-Media-Kanälen verbreitet werden kann. Eine Verlinkung zu den gängigen Social-Media-Plattformen ist direkt im Online-Shop vorgesehen, der Aufwand ist entsprechend gering und erscheint als das Ziel, auf das man hinarbeiten soll. So sind die Aufforderungen zu verstehen, mit der sich die nicht-anwesenden Betreiber*innen des Online-Shops an potenzielle Kund*innen wenden:

- 1) »Create eye-catching content for social media. Make sure your social posts don't stay unnoticed. Create new fashion content instantly and post metalooks straight to your socials!«¹⁹
- 2) »Create your look of the day«²⁰
- 3) »Dressing in up to 45 s«²¹

19 Quelle: *DRESSX.AI* (o.J.): Startseite, <https://dressx.ai/> (Stand 23.07.2024).

20 Ebd.

21 Ebd.

- 4) »Dress your profile picture. Enhance your social profile avatar using digital outfits«²²

Nutzer*innen werden mit Headlines wie (1) und (2) dazu angehalten, sich in kürzester Zeit (3) in neuen virtuellen Outfits ihrer persönlichen medialen Öffentlichkeit (4) zu präsentieren. Digital Fashion unterstützt damit einen dauerhaften Selbstbezug. Indem man den eigenen Körper nach den Vorgaben der App fotografiert und mit virtuellen Kleidern medial inszeniert, entsteht ein neuer digitaler, unrealer, Körper, der sich vom leiblichen Körper distanzieren lässt. Das erwünschte Selbstbild lässt sich dann anderen Abwesenden – Bekannten und Unbekannten – präsentieren, um sich deren Akzeptanz zu versichern: »Since the need for social recognition is so central to well-being, hypermodern individuals turn to online interactions in the hopes of gratifying it« (Gottschalk 2021: 104). Bieger versteht die jeweilige Gestalt des Körpers »als Versuch der Aktualisierung eines imaginären Selbstbildes« (2008: 56). Dieser Zustand kann jedoch immer nur für einen kurzen Zeitraum stabilisiert werden: »Die auf diesem Weg erreichte Identität ist in einem ständigen Fluss; sie wird in einem Zuge erreicht und schon wieder überschritten« (ebd.: 65). Sobald sich durch ausbleibende Kommentare Unsicherheiten einstellen oder die Bilder von anderen kritisch kommentiert werden, muss Abhilfe mit neuen Inszenierungen geschaffen werden. Nutzer*innen unterliegen einem »vehementen Anerkennungs- und Individualisierungsimperativ« (ebd.), dem nur mit einem ständigen Erneuerungsprozess Rechnung getragen werden kann. Ein auf diese Weise erzeugtes Selbstbild muss immer ein performativer Prozess bleiben. Bieger konstatiert: »[...] die in ihrem Kraftfeld produzierte Identität ist stets nur eine Momentaufnahme, die aufgrund ihrer konstitutiven Unabgeschlossenheit nicht einfach zu einem fixen Selbstbild erstarren kann« (ebd.). Und so lässt sich der immerwährende Antrieb zu einer neuen Inszenierung als »Darstellungsnotwendigkeit« (Misoch 2006: 115) verstehen.

Sich selbst sichtbar machen heißt, sich in Kleidungsstücken zu inszenieren, die von den Betrachtenden in genau der Weise gelesen werden sollen, die man selbst vorgibt. Die eigenen Erwartungen an sein Bild, die als doppelte Kontingenz (vgl. Luhmann 1984: 166) immer in Abhängigkeit der vermuteten Erwartung der Anderen an einen selbst stehen, schränken nicht nur die Möglichkeiten der Inszenierung ein, sondern sie konstituieren die soziale Ordnung. Die Erwartungen sind von temporärer Natur, weil sie auch immer die Gefahr der Enttäuschung bergen. So ist man für sich selbst sein eigenes Publikum, indem man sein Selbstbildnis kritisch prüft, verändert und immer wieder optimiert. Die obligatorische Selbst-Produktion und ihre ständige Aktualisierung lassen sich dann besonders gut im unrealen Raum praktizieren, wenn auch hier die soziale Anerkennung angestrebt wird. Dies

22 Ebd.

wiederum kann nur über die Sichtbarkeit von Eigenschaften erfolgen, die im Programmcode der sozialen Netzwerke vorgesehen sind. Die so unter Druck stehenden Nutzer*innen können nicht einfach davon ablassen, ihre Online-Profile mit immer neuen Bildern von sich und ihren Aktivitäten zu aktualisieren, wenn sie die Online-Bestätigung als Zeichen von persönlichem Erfolg werten und diesen als eine ökonomische Größe wahrnehmen: »[...] to be seen by others is increasingly regarded as a requirement for personal and social advance. We are successful to the degree that our self-expression is consumed by others« (Roberts 2014: 136). Aufmerksamkeit ist ebenso wie Interesse und Feedback kein homogenes Gut, welches wie Geld einer eindeutigen Messbarkeit unterliegt. Um den Wert zu bestimmen, wurden über Umwege andere, technisch programmierte Messeinheiten entwickelt, was darauf hinweist, dass diese Vergleichbarkeit auf direktem Weg nicht zu erreichen war (vgl. Lehner 2017: 60). Durch die Anzahl der Likes, Dislikes, der Freunde, der Bewertungen und Empfehlungen, der Forenbeiträge oder Kommentare wird ein zahlenbasierter und damit vermeintlich objektiver Vergleich überhaupt erst möglich. Und so greifen Nutzer*innen auf Digital Fashion zurück, verbunden mit der Erwartung, positive Bestätigung zu erhalten. Reckwitz spricht von einem »doppelten Optionalitätshabitus« (2006: 102), der die Rezipierenden als expressive Subjekte kennzeichnet: »[...] es nimmt sich selbst als Subjekt von Wahlakten wahr und damit zugleich die anderen [...] als Objekte von Wahlakten. Gleichzeitig perzipiert es sich selbst als Objekt von Wahlakten anderer und formt sich entsprechend« (ebd.). Besonders deutlich wird dieser Prozess in den Empfehlungen, welche die Plattformbetreiber*innen an Designer*innen der virtuellen Kleider geben. So ist unter anderem zu lesen: »We remember that clients are not going to wear it [...] to the plastic computer game, but they will wear it to ›go out‹ to social media – they will post it on Instagram first of all. That’s why it should naturally match the visual language of the platform«²³. Zwar muss das virtuelle Kleid immer noch über ästhetische Merkmale verfügen, die es neu, aufregend und ansprechend wirken lassen, es soll aber zusätzlich Begehrlichkeiten wecken, die sich trotz der Ähnlichkeit zum textilen Kleid auf eine ausschließliche Online-Nutzung beziehen. Die Feedbacks der abwesenden Anderen finden im unrealen Interaktionsraum statt und sind damit auch wieder an die vorgegebene Programmstruktur angepasst, die eher auf Vernetzung und Aktionen ausgerichtet ist als auf kritische Kommentare. So lässt sich festhalten, dass die soziale Aushandlung der Selbstinszenierung im unrealen Interaktionsraum in erster Linie auf einer Aushandlung mit sich selbst beruht.

23 DRESSX und DRESSX.AI werden von MoreDash Inc bzw. Daria Shapovalova und Natalia Modenova betrieben. Um selbst kreativ zu werden, gibt es technische und gestalterische Richtlinien: DRESSX Check List for 3D designers requirements (Quelle: <https://docs.google.com/document/d/1mKyVJgesxt3fMzbjYUZ8ggiQ6av6-FoNto4pj7UUTKk/edit>; Stand: 29.11.2022).

6. Schlussfolgerung

Der Online-Shop ist ein interessengeleiteter unrealer Interaktionsraum, wo leiblich anwesende Nutzer*innen, abwesende anonyme Anbieter*innen und weitere unrea- le Akteure wie Programmtools und Algorithmen aufeinandertreffen, um über Bild- inszenierungen mit virtuellen Kleidern Kommunikation zu betreiben. Reale Bezie- hungen werden simuliert. Nutzer*innen streben ihre mediale Inszenierung über ein imaginäres, andauernd veränderbares Selbstbildnis an und formen damit die Dienstleistung, die sie dann bezahlen.

Das Angebot spiegelt die sozialen Bedürfnisse der Nutzer*innen nach Aufmerk- samkeit, Idealisierung und Kommunikation. Weil bei Digital Fashion während des digitalen Bekleidens – die Auswahl des Kleidungsstückes, das Posen, die Anprobe, der Kauf und die Präsentation – die Nutzer*innen die einzigen leiblich Anwesen- den sind, findet die Aushandlung letztendlich mit sich und an sich selbst statt. Erst die Ergebnisse werden dann einem nicht anwesenden Publikum – wiederum über Programmtools und Algorithmen – präsentiert. Selbst wenn die Programmstruktur bestimmte Handlungen begünstigt und die Agency der Nutzer*innen durch die An- bietenden gesteuert werden kann, so ist dies nicht als ein Ausgeliefert-Sein zu be- trachten. Der unrea- le Interaktionsraum wird zum sozialen Handlungsraum einer neuen kulturellen Praxis, die darin besteht, jeweils mit imaginären Selbstbildnis- sen eine kurzlebige und unbeständige Ordnung herzustellen, die nicht hinterfragt wird, und die ihre ständige Erneuerung in immer neuen Bildern erfährt.

Literatur

- Atkinson, Douglas (2017): »Post-Industrial Fashion and the Digital Body«, in: Susan Broadhurst/Sara Price (Hg.), *Digital Bodies. Creativity and Technology in the Arts and Humanities*, London: Palgrave Macmillan, S. 147–160.
- Behrmann, Helga (2024): *Digital Fashion. Das virtuelle Kleid radikalisiert die Mode*, Potsdam: Universität Potsdam.
- Bieger, Laura (2008): »Schöne Körper, hungriges Selbst – Über die moderne Wunschökonomie der Anerkennung«, in: Annette Geiger (Hg.), *Der schöne Kör- per. Mode und Kosmetik in Kunst und Gesellschaft*, Köln: Böhlau, S. 53–68.
- Bieger, Laura (2011): »Ästhetik der Immersion: Wenn Räume wollen. Immersives Er- leben als Raumerleben«, in: Gertrud Lehnert (Hg.), *Raum und Gefühl. Der Spati- al Turn und die neue Emotionsforschung*, Bielefeld: transcript-Verlag, S. 75–95.
- Böhme, Gernot (2004): »Raum leiblicher Anwesenheit – Raum als Medium von Dar- stellung«, in: Sybille Krämer (Hg.), *Performativität und Medialität*, München: Fink Verlag, S. 129–140.

- Brachem, Judith/Lucas Stübbe (2022): »Ways of Wearing«, in: Chinouk Filique de Miranda (Hg.), *Fashion Studies Journal (Fashion & Digital Engagement Summer Issue 2022)*. Abrufbar unter: <https://www.fashionstudiesjournal.org/digital-engagement-a/2022/8/15/ways-of-wearing> (Stand: 13.02.2023).
- Crewe, Louise (2013): »When Virtual and Material Worlds Collide: Democratic Fashion in the Digital Age«, in: *Environment and Planning A* 45 (4), S. 769–780.
- Crewe, Louise (2017): *The Geographies of Fashion. Consumption, Space and Value*, London: Bloomsbury Publishing Plc.
- Doedens, Sebastian (2010): *Flanieren im Internet. Surfstile und Surfstrategien junger Internetnutzer*, Baden-Baden: Nomos.
- Döring, Daniela/Claude Draude (2012): »Körper nach Zahlen. Vom Maßnehmen und der Simulation von Menschlichkeit«, in: HU Berlin (Hg.), *Zentrum für transdisziplinäre Geschlechterstudien*, Berlin: HU Berlin, S. 61–87.
- Gaugele, Elke (2005): »Changing Rooms oder: Was geschah in der Umkleidekabine?«, in: Beate Binder/Silke Götsch/Wolfgang Kaschuba/Konrad Vanja (Hg.), *Ort. Arbeit. Körper. Ethnografie Europäischer Modernen*, Berlin: SMPK – Museum Europäischer Kulturen, S. 439–446.
- Gaugele, Elke (2014): »Grenzkontrollen: »Fashion & Surveillance««, in: Birgit Haehnel (Hg.), *Anziehen! Transkulturelle Moden: Transcultural Fashion = Dressed up!*, Bielefeld: transcript-Verlag, S. 124–128.
- Gläser, Martin (2008): *Medienmanagement*, München: Vahlen (Vahle's Handbücher der Wirtschafts- und Sozialwissenschaften). Abrufbar unter http://bvbr.bib-bvb.de:8991/F?func=service&doc_library=BVB01&doc_number=010566477&line_number=0002&func_code=DB_RECORDS&service_type=MEDIA (Stand 15.03.2023).
- Goffman, Erving (2001): »Interaktion und Geschlecht«, in: Erving Goffman/Hubert Knoblauch/Helga Kotthoff (Hg.), *Interaktion und Geschlecht*, Frankfurt a.M.: Campus-Verlag, S. 50–104.
- Goffman, Erving/Ralf Dahrendorf (2017): *Wir alle spielen Theater. Die Selbstdarstellung im Alltag*, München/Berlin, Zürich: Piper.
- Gottschalk, Simon (2018): *The Terminal Self. Everyday Life in Hypermodern Times*, Abingdon/Oxon/New York, NY: Routledge.
- Gottschalk, Simon (2021): »CLICK, VALIDATE, AND REPLY. Three Paradoxes of the Terminal Self«, in: Dirk vom Lehn/Natalia Ruiz-Junco/Wil Gibson (Hg.), *The Routledge International Handbook of Interactionism*, London/New York: Routledge Taylor & Francis Group, S. 99–111.
- Hartmann, Maren (2004): *Technologies and Utopias. The Cyberflaneur and the Experience of »Being Online«*. Teilw. zugl.: London, University of Westminster, Dissertation, 2004, München: Fischer. Abrufbar unter: http://bvbr.bib-bvb.de:8991/F?func=service&doc_library=BVB01&doc_number=014737400&lin

- e_number=0001&func_code=DB_RECORDS&service_type=MEDIA (Stand: 08.03.2023).
- Hellmann, Kai-Uwe (2018): »Ein Casting der besonderen Art. Zum Menschenbild der Verbraucherwissenschaften«, in: Sebastian Nessel/Nina Tröger/Christian Friedrich/Renate Hübner (Hg.), *Multiperspektivische Verbraucherforschung. Ansätze und Perspektiven*, Wiesbaden: Springer VS, S. 53–81.
- Imdahl, Max (1996): *Giotto Arenafresken. Ikonographie, Ikonologie, Ikonik*, München: Fink Verlag.
- Joy, Annamma/Ying Zhu/Camilo Peña/Myriam Brouard (2022): »Digital Future of Luxury Brands: Metaverse, Digital Fashion, and Non-Fungible Tokens«, in: John Wiley & Sons, Ltd.: *Strategic Change* 31, S. 337–343.
- Khelladi, Insaf/Catherine Lejealle/Saeedeh Rezaee Vessal/Sylvaine Castellano/Domenico Graziano (2023): »Why Do People Buy Virtual Clothes?«, in: *Journal of Consumer Behaviour*.
- Knorr Cetina, Karin (2009): »The Synthetic Situation: Interactionism for a Global World«, in: *Society for the Study of Symbolic Interaction*, 32 (1), S. 61–87.
- König, Gudrun M. (2009): *Konsumkultur – inszenierte Warenwelt um 1900*, Wien: Böhlau.
- König, René (1985): *Menschheit auf dem Laufsteg. Die Mode im Zivilisationsprozeß*, München: Hanser.
- Kozinets, Robert V. (2022): »Algorithmic Branding Through Platform Assemblages: Core Conceptions and Research Directions for a New Era of Marketing and Service Management«, in: *Journal of Service Management* 33 (3), S. 437–452.
- Larosse, Michaela (2021): *Building A Digital Fashion House – A How To Guide*. Abrufbar unter: <https://www.thefabricant.com/blog/2021/7/28/building-a-digital-fashion-house-a-how-to-guide> (Stand: 18.03.2022).
- Lehner, Nikolaus (2017): »Empfehlungssysteme: Begehrlichkeiten auf Umwegen«, in: Holger Brohm/Sebastian Gießmann/Gabriele Schabacher/Sandra Schramke (Hg.), *Workarounds. Praktiken des Umwegs*, Hamburg: Philo Fine Arts, S. 59–80.
- Lobinger, Katharina (2012): *Visuelle Kommunikationsforschung. Medienbilder als Herausforderung für die Kommunikations- und Medienwissenschaft*, Wiesbaden: Springer VS.
- Löw, Martina (2023): *Raumsoziologie, Jubiläumsausgabe*, Frankfurt a.M.: Suhrkamp.
- Luhmann, Niklas (1984): *Soziale Systeme. Grundriss einer allgemeinen Theorie*, Frankfurt a.M.: Suhrkamp.
- Luke, Robert (2005): »The Phoneur: Mobile Commerce and the Digital Pedagogies of the Wireless Web«, in: Peter Pericles Trifonas (Hg.), *Communities of Difference*, New York: Palgrave Macmillan US, S. 185–204.

- Mersch, Dieter (2011): »Fraktale Räume und multiple Aktionen. Überlegungen zur Orientierung in komplexen medialen Umgebungen«, in: Gertrud Lehner (Hg.), Raum und Gefühl. Der Spatial Turn und die neue Emotionsforschung, Bielefeld: transcript-Verlag, S. 49–62.
- Misocho, Sabina (2006): Online-Kommunikation, Konstanz: UVK.
- Musto, Michela/Chiara Scarpitti (2022): »Postdigital Fashion. Emerging Paradigms For Hybrid Bodies«, in: Convergences Journal 29, S. 131–139.
- Peters, John Durham (2015): The Marvelous Clouds. Toward a Philosophy of Elemental Media, Chicago/London: The University of Chicago Press.
- Prinz, Sophia (2019): »Design aus kulturosoziologischer Perspektive«, in: Stephan Moebius/Frithjof Nungesser/Katharina Scherke, Handbuch Kulturosoziologie. Theorien – Methoden – Felder, Wiesbaden: Springer VS, Bd. 2, S. 289–303.
- Prose, Stefanie (Hg.) (2010): Flaneure – Begegnungen auf dem Trottoir, Lizenzausgabe, Frankfurt a.M.: Edition Büchergilde.
- Reckwitz, Andreas (2006): »Die historische Transformation der Medien und die Geschichte des Subjekts«, in: Andreas Ziemann (Hg.), Medien der Gesellschaft – Gesellschaft der Medien, Konstanz: UVK-Verlags-Gesellschaft, S. 89–107.
- Reichmann, Werner (2018): »Interaktion in mediatisierten Welten. Von Face-to-Face-Kommunikation zur Interaktion in ›synthetischen‹ Situationen«, in: Andreas Kalina/Friedrich Krotz/Matthias Rath/Caroline Roth-Ebner (Hg.), Mediatisierte Gesellschaften. Medienkommunikation und Sozialwelten im Wandel, unter Mitarbeit von Andreas Kalina, Baden-Baden: Nomos Verlagsgesellschaft mbH & Co. KG, S. 87–106.
- Reisch, Lucia A. (2020): »Nudging hell und dunkel: Regeln für digitales Nudging«, in: Wirtschaftsdienst 100 (2), S. 87–91.
- Roberts, Paul (2014): The Impulse Society. What's Wrong with Getting What We Want?, London: Bloomsbury Publishing Plc.
- Rusch, Gebhard/Helmut Schanze/Gregor Schwering (2007): Theorien der Neuen Medien. Kino, Radio, Fernsehen, Computer, Paderborn: Fink Verlag.
- Sayem, Abu Sadat Muhammad (2022): »Digital Fashion Innovations for the Real World and Metaverse«, in: International Journal of Fashion Design, Technology and Education 15 (2), S. 139–141.
- Särmäkari, Natalia (2023): »Digital 3D Fashion Designers: Cases of Atacac and The Fabricant«, in: Fashion Theory, 27 (1), S. 85–114.
- Schroer, Markus (2010): »Individualisierung als Zumutung. Von der Notwendigkeit zur Selbstinszenierung in der visuellen Kultur«, in: Peter A. Berger/Ronald Hitzler (Hg.), Individualisierungen. Ein Vierteljahrhundert »jenseits von Stand und Klasse«, Wiesbaden: Verlag für Sozialwissenschaft, S. 275–289.
- Srnicek, Nick (2018): Plattform-Kapitalismus, Hamburg: Hamburger Edition.
- Staab, Philipp (2019): Digitaler Kapitalismus. Markt und Herrschaft in der Ökonomie der Unknappheit, Berlin: Suhrkamp Verlag.

- Stihler, Ariane (1998): Die Entstehung des modernen Konsums. Darstellung und Erklärungsansätze, Berlin: Duncker & Humblot GmbH.
- Ziemann, Andreas (2011): Medienkultur und Gesellschaftsstruktur. Soziologische Analysen. Zugelassen: Weimar, Bauhaus-Universität, Habilitations-Schrift, 2009, Wiesbaden: Verlag für Sozialwissenschaft.
- Zuboff, Shoshana (2018): Das Zeitalter des Überwachungskapitalismus, Frankfurt a.M., New York: Campus Verlag.

Appropriating the Ape

Sprachliche Praktiken der Gemeinschaftsbildung im Subreddit R/WALLSTREETBETS

Friederike Fischer

Abstract *Der vorliegende Beitrag untersucht sprachliche Praktiken der Gemeinschaftsbildung auf Reddit. Im Fokus steht eine soziale Bewegung, die Anfang 2021 im Forum R/WALLSTREETBETS entstand und mit einer konzertierten Aktion weltweit erhebliche Schwankungen an den Börsenmärkten auslöste. Anhand dieses Fallbeispiels werden sprachliche Praktiken untersucht, mit denen die Mitglieder der Bewegung ihre Zugehörigkeit ausdrücken und so den Zusammenhalt der Gruppe stärken. Als theoretischer Rahmen dient einerseits das Verständnis des virtuellen Interaktionsraums und andererseits das Konzept der Community of Practice (CoP), die beide herangezogen werden, um R/WALLSTREETBETS zu konzeptualisieren. Anhand korpusbasierter Untersuchungen wird analysiert, welche Rolle in der CoP die Verwendung von Beleidigungen spielt, die zentraler Bestandteil des gruppenspezifischen Sprachgebrauchs sind und in einem Akt der Aneignung als Ausdruck der Gruppenzugehörigkeit neu kontextualisiert werden.*

Keywords *r/wallstreetbets; Reddit; Community of Practice; linguistic appropriation*

1. Einleitung

#metoo, Black Lives Matter oder #ichbinhanna – diese und andere Protestbewegungen haben in den letzten Jahren immer wieder unter Beweis gestellt, welches Potential soziale Netzwerke bergen, (Protest)Bewegungen hervorzubringen (vgl. Fielitz et al. 2020; Fielitz/Staemmler 2020; Kneuer et al. 2015). Zwar entwickeln sich diese Gruppierungen häufig online, und doch sind ihre Auswirkungen mitunter auch offline spürbar. Oftmals fungiert die Plattform, der eine solche Bewegung entspringt, als Kanal, über den beispielsweise Straßenproteste oder andere Aktionen geplant und koordiniert werden (vgl. Dang-Anh 2019; Dang-Anh et al. 2022). Allerdings können sich Bewegungen auch exklusiv online entwickeln und, ohne jemals physisch

»auf die Straße« gegangen zu sein, spürbare Auswirkungen auf die Welt außerhalb der sozialen Medien haben (vgl. Baringhorst 2019).

Der vorliegende Beitrag widmet sich der Untersuchung sprachlicher Merkmale einer solchen Bewegung, die sich Anfang 2021 im Onlineforum R/WALLSTREETBETS entwickelte. Mit einer konzertierten Aktion sorgte diese weltweit für erhebliche Schwankungen an den größten Aktienmärkten, darunter auch die New Yorker Börse. Die Mitglieder bewerkstelligten dies ausschließlich mittels koordinierter Aktienkäufe, die sie im Onlineforum R/WALLSTREETBETS planten. Ohne also jemals offline miteinander in Kontakt getreten zu sein oder gar Protesthandlungen ausgeübt zu haben, entwickelte die Online-Bewegung nicht nur die Reichweite und den Einfluss, die notwendig sind, um auch außerhalb der sozialen Medien einen bleibenden Eindruck zu hinterlassen, sondern stellte auch einen starken Zusammenhalt unter Beweis, der sich in den koordinierten Aktienkäufen äußerte. Im vorliegenden Beitrag untersuche ich die Verwendung von Beleidigungen und Schmähungen als gruppenspezifische sprachliche Praktik, mittels derer die Mitglieder den Zusammenhalt der Gruppe konstituieren und ihre Zugehörigkeit signalisieren. Anhand dieses konkreten Beispiels führe ich aus, wie Mitglieder von R/WALLSTREETBETS im virtuellen Interaktionsraum des Subreddits, in dem sie agieren, sich als Community of Practice (im Folgenden CoP) identifizieren und von Außenstehenden abgrenzen.

Um dies zu bewerkstelligen, gliedert sich der Beitrag in sechs Abschnitte. Zunächst wird das soziale Netzwerk Reddit vorgestellt. Im Mittelpunkt stehen dabei die sogenannten Subreddits, untergeordnete Foren, die für die linguistische Analyse als virtuelle Interaktionsräume und als CoP aufgefasst werden können. Beide Konzepte sind zielführend, wobei hier v. a. auf das Konzept der CoP rekurriert wird, da der Fokus auf dem gemeinschaftsstiftenden Potential sprachlicher Praktiken und dem gruppenspezifischen Repertoire liegen soll. In Kapitel drei werde ich zunächst das Fallbeispiel erläutern, um dann näher auf die Besonderheiten des Korpus einzugehen, das die Datengrundlage für die Analyse bildet. Dem schließt sich in Kapitel vier eine korpusbasierte Analyse der Verwendung von Beleidigungen an, die eine charakteristische sprachliche Praktik auf R/WALLSTREETBETS ist, um Zugehörigkeit auszudrücken und so zur Konstitution einer Gemeinschaft beiträgt. Unter Praktiken verstehe ich »sozial determinierte Handlungsmuster« (Dang-Anh 2019: 67; vgl. Holly u. a. 1984), die »situative Vollzugsmomente einer übersituativen Praxis« darstellen (Dang-Anh 2019: 67). Demzufolge handelt es sich bei kommunikativen Praktiken um »Vollzugsmomente sozialer Praxis, in denen Praxis sozial wirksam verhandelt wird« (Dang-Anh 2019: 72). Anhand von Keyword- und Kollokationsanalysen auf der Grundlage eines randomisierten Sample-Korpus wird aufgezeigt, welchen Beleidigungen auf R/WALLSTREETBETS eine gemeinschaftsstiftende Funktion zukommt und welche weiterhin semantisch als Beleidigungen und Schmähungen gelesen werden müssen. In Anlehnung an die sprachliche Praktik der *linguistic appropriation* wird anhand der Befunde ein möglicher Erklärungsansatz elaboriert und

diskutiert. Abschließend folgt ein Fazit, in dem die Erkenntnisse und Grenzen der Untersuchung sowie mögliche Anschlussfragen herausgearbeitet werden.

2. Zwischen Interaktionsraum und Community of Practice: Das soziale Netzwerk Reddit aus medienlinguistischer Perspektive

Der vorliegende Beitrag nimmt sprachliche Praktiken in den Blick, die mit der Entwicklung einer sozialen Bewegung auf Reddit einhergehen. Um dieses soziale Netzwerk zu verstehen und für die Analyse nutzbar zu machen, sind zwei einander komplementär ergänzende Konzepte hilfreich: der (virtuelle) Interaktionsraum und die CoP. In Ergänzung zueinander ermöglichen sie die ganzheitliche Betrachtung von Subreddits und machen diese für die nähere Untersuchung nutzbar. Im Folgenden werde ich beide Perspektiven anhand des aktuellen Forschungsstandes erläutern, wobei der Fokus auf dem Konzept der CoP liegt, da dieses in der anvisierten Untersuchung vordergründig ist.

Reddit ist ein soziales Netzwerk, in dem Nutzer*innen sich in einer Vielzahl von untergeordneten Foren, sogenannten Subreddits, austauschen und vernetzen können. Inhaltlich sind Subreddits keine Grenzen gesetzt: Sie reichen von ganzen Themenbereichen bis hin zu Nischeninteressen. Grundsätzlich sind Subreddits virtuelle Interaktionsräume innerhalb des übergeordneten Interaktionsraums Reddit. Dieses Verständnis führt Dürscheids Ausführungen weiter, denen zufolge die Gesamtheit aller sozialen Netzwerke ein Interaktionsraum ist, der in Form einzelner sozialer Netzwerke wiederum in kleinere Interaktionsräume unterteilt ist (vgl. Dürscheid 2015). In ihrer infrastrukturellen Funktion ähneln sich Subreddits im Wesentlichen: Sie stellen die notwendige Infrastruktur zur Verfügung, die es den Mitgliedern ermöglicht, sich untereinander auszutauschen. Es handelt sich dabei freilich nicht um physische, sondern um metaphorische Räume, in denen Nutzer*innen online interagieren und mittels der Affordanzen, die Reddit ihnen zur Verfügung stellt, die Interaktion und mit ihr den Interaktionsraum gestalten (vgl. Storrer 2001; Hutchby 2001). Dabei entstehen als Produkt medial-vermittelter Interaktion und der jeweiligen Medien neue soziale Realitäten:

Medien [verändern] auf diese Weise nicht bloß Interaktion, sondern schaffen zudem neue soziale Realitäten wie virtuelle Räume, Öffentlichkeiten oder Populärkultur, die ohne mediale Vermittlung kaum denkbar sind. Allen drei sind je eigene Interaktionspraktiken, gesellschaftliche Funktionen und Dynamiken eigen, die im Zusammenspiel materialer Möglichkeiten, sozialer Gebrauchsweisen und gesellschaftlicher Kontexte entstehen (Marx/Schmidt 2019b: 12)

Zu diesen neuen sozialen Realitäten zählen zweifelsohne virtuelle Interaktionsräume, welche wiederum neue Möglichkeiten der Interaktion eröffnen. Die Räume als solche gehen mit einer Terminologie einher, die physisch greifbare und klar umrissene Orte beschreibt: Nutzer*innen diskutieren in Foren, Gruppen oder Chaträumen oder gelangen in Echokammern (vgl. Stegbauer 2018). Diese Metaphorik ist jedoch irreführend und konfligiert mit dem Umstand, dass es sich eben nicht um physische Orte handelt, sondern Nutzer*innen in digitalen Räumen interagieren, deren Bezeichnung metaphorisch ist. Diese haben zwar in ihrer Oberflächendarstellung das gleiche Erscheinungsbild und verfügen über die gleichen Affordanzen, dennoch verfügen Nutzer*innen nicht über eine geteilte intersubjektive Wahrnehmung. Dennoch oder gerade deswegen trägt die Verwendung lokaldeiktischer Ausdrücke zur Etablierung und Aufrechterhaltung eines common ground bei, der trotz erschwelter Bedingungen die Interaktion der Nutzer*innen ermöglicht (vgl. Beißwenger 2013). Indem Nutzer*innen entscheiden, wie sie die Affordanzen des jeweiligen Mediums nutzen, etablieren sie kommunikative Praktiken in den unterschiedlichen Interaktionsräumen. Wird eine solche kommunikative Praktik konventionalisiert, gestaltet diese den Interaktionsraum maßgeblich mit.

Ein solches Verständnis von Subreddits als Interaktionsräume reduziert diese im Wesentlichen auf ihre infrastrukturellen Gegebenheiten. Soziale und interpersonale Aspekte werden hingegen nur marginal berücksichtigt, obwohl sie für ein umfassendes Bild von Interaktionsräumen unabdingbar sind (vgl. Hausendorf/Schmitt 2018). Reddit selbst propagiert demgegenüber wiederum ein Selbstverständnis, das dem Fokus auf diejenigen Merkmale, die Subreddits als Interaktionsräume charakterisieren, grundlegend entgegengesetzt:

Reddit ist eine große Community, die aus Tausenden von kleineren Communities besteht. Diese kleineren Unter-Communities werden auch als »Subreddits« bezeichnet und werden von Redditoren wie dir selbst erstellt und moderiert. [...] Auf Reddit gibt es eine Community für so ziemlich alles, was du dir vorstellen kannst und verschiedene Arten, die Communities zu finden, die sich mit Dingen beschäftigen, für die du dich interessierst. (Reddit, Inc. 2023b)

Die Bezeichnung der Plattform selbst und der ihr untergeordneten Foren als *Communities* legt nahe, dass gerade interpersonellen und nutzer*innenzentrierten Aspekten eine wesentliche Bedeutung beigemessen wird. Zwar wird der Community-Begriff Leuckert und Leuckert (vgl. 2020) zufolge im Kontext sozialer Medien inflationär verwendet und verliert dadurch seine eigentliche Bedeutung, doch ihre Untersuchung regt ein Umdenken an: Weg von Subreddits als reine Interaktionsräume hin zur Konzeptualisierung als CoP (vgl. Leuckert/Leuckert 2020).

Insbesondere in der soziolinguistischen Forschung erfreut sich das Konzept der CoP seit Anfang der 1990er Jahre großer Beliebtheit (vgl. Lave/Wenger 1991; Eckert/

McConnell-Ginet 1992; Leuckert 2020; Leuckert/Leuckert 2020). Eckert und McConnell-Ginet zufolge handelt es sich bei CoPs um ein »aggregate of people who come together around mutual engagement in an endeavor« (1992: 464). Als Alleinstellungsmerkmal nennen sie einerseits die Mitgliedschaft einer CoP (i. S. v. der Gesamtheit der Mitglieder) und andererseits die konventionalisierten Praktiken der Mitglieder: »As a social construct, a [CoP] is different from the traditional community, primarily because it is defined simultaneously by its membership and by the practice in which that membership engages« (Eckert/McConnell-Ginet 1992: 464). Diese Praktiken, und mit ihnen auch sprachliche und kommunikative Praktiken, sind demzufolge das Resultat der Interaktion innerhalb einer CoP. Zugleich sind diese in der CoP etablierten und konventionalisierten Praktiken auch ein Alleinstellungsmerkmal, anhand dessen sich eine CoP von Außenstehenden abgrenzen kann. So dient Meyerhoff und Strycharz zufolge das sprachliche Handeln der Mitglieder u. a. als Ausdruck der Zugehörigkeit zu einer CoP und der Abgrenzung von anderen. Demzufolge labeln die Mitglieder selbst in einem andauernden Prozess eine CoP als solche (vgl. Meyerhoff/Strycharz 2013).

Das individuelle Zugehörigkeitsgefühl der Mitglieder ist also eine wesentliche Voraussetzung, um eine Gruppe als CoP bezeichnen zu können. Im Fall von Online-CoPs ist diese Zugehörigkeit i. d. R. anhand des Engagements ihrer Mitglieder nachvollziehbar, das sich primär im sprachlichen Handeln äußert. Indem die Mitglieder sich aktiv am Diskurs in der CoP beteiligen und ihre Zugehörigkeit explizit zum Ausdruck bringen, signalisieren sie ihr Zugehörigkeitsgefühl. In ihrer korpusgestützten, komparativen Untersuchung von Subreddits setzen sich Leuckert und Leuckert kritisch mit der Konzeptualisierung von Subreddits als CoPs auseinander. Ausgehend von Meyerhoff und Strycharz' Verständnis überprüfen sie, inwiefern Subreddits die konstitutiven Eigenschaften von CoPs erfüllen. Diese Eigenschaften umfassen Meyerhoff und Strycharz (2013) zufolge das *mutual engagement*, ein *jointly negotiated enterprise* sowie ein *shared repertoire* der Mitglieder einer CoP:

- *mutual engagement*, also die Teilhabe und das Engagement der Mitglieder
- *jointly negotiated enterprise*, also ein gemeinsames Vorhaben, das die Mitglieder der CoP zusammen verfolgen
- *shared repertoire*, also eine Reihe geteilter (sprachlicher und nicht-sprachlicher) Praktiken, die das Resultat des Austauschs innerhalb der CoP sind

Leuckert und Leuckert (2020) bringen mit Blick auf ihre Daten verschiedene Einwände gegen das Verständnis von Subreddits als CoPs vor, darunter auch das Fehlen sprachlicher Praktiken, welche die Zugehörigkeit zur CoP signalisieren. In der Tat erfüllen die von Leuckert und Leuckert untersuchten Subreddits die o. g. Kriterien einer CoP nur teilweise, wodurch das CoP-Framework in ihrem Fall nicht lückenlos greift.

Die Beobachtungen aus Kapitel 2 legen nahe, dass Subreddits weder als reine Interaktionsräume noch als reine CoPs konzeptualisiert werden können. Beide Konzepte erfassen Subreddits nicht gänzlich, und nur in Ergänzung zueinander ermöglichen sie ein umfassendes Verständnis. Diese Verschränkung von Konzepten legt nahe, dass Subreddits in Anlehnung an Marx und Schmitt als neue soziale Realitäten verstanden werden können, die wiederum eigene »Interaktionspraktiken, gesellschaftliche Funktionen und Dynamiken« (Marx/Schmidt 2019b: 12) hervorbringen können. Ebendies geschah auf r/WALLSTREETBETS, als Anfang 2021 eine soziale Bewegung entstand, die wiederum als CoP konzeptualisiert werden kann. Wie ich in Kapitel 3 erläutern werde, unterscheidet sich r/WALLSTREETBETS in dieser Hinsicht maßgeblich von den Subreddits, die Leuckert und Leuckert (2020) in den Blick genommen haben, sodass gerade dieses Forum alle Kriterien einer CoP erfüllt. Um dies zu verdeutlichen, werde ich im Folgenden näher auf den Subreddit eingehen und ausarbeiten, welche Ereignisse dazu führten, dass sich dort eine soziale Bewegung und mit ihr eine CoP entwickelte, die auch über Reddit hinaus Einfluss auf die Offline-Welt nahm.

3. \$GME to the moon! Der versuchte Short Squeeze auf r/WALLSTREETBETS

Der Subreddit r/WALLSTREETBETS wurde 2012 erstellt und zählt spätestens seit Anfang 2021 zu den mitgliederstärksten Foren auf Reddit (vgl. Reddit, Inc. 2023a; Subreddit Stats 2024). Die Mitglieder tauschen sich dort über Aktienhandel und aktuelle Entwicklungen an den internationalen Börsenmärkten aus. Das Forum ist auch darüber hinaus auf Reddit weithin als Anlaufstelle für Fragen rund um Aktienhandel bekannt (vgl. Biermann/Hegemann 2021). Anfang 2021 erlangte es schlagartig Berühmtheit, als die Mitglieder in einer konzertierten Aktion im großen Stil Aktien des US-amerikanischen Videospielehändlers Gamestop kauften. Das Wertpapier war zuvor von Hedgefonds leerverkauft worden, die auf den Wertverlust der Aktie spekuliert hatten. Leerverkäufe sind in Fachkreisen höchst umstritten, da sie eingesetzt werden können, um ohnehin finanzschwache Unternehmen gezielt in die Insolvenz zu treiben. Über r/WALLSTREETBETS erfuhren die Kleinanleger*innen von den fragwürdigen Methoden der Hedgefonds und begannen in die Gamestop-Aktie zu investieren, angetrieben durch den Wunsch, das Nostalgie-Unternehmen zu retten und die unlauteren Methoden der Hedgefonds zu untergraben sowie sich dabei selbst zu bereichern. Diese Praxis wird als *Short Squeeze* bezeichnet und zielt auf die Umverteilung von Vermögen ab: Die mächtigen Hedgefonds sollen finanzielle Verluste erleiden, während die vermeintlich Kleinen profitieren.

Der Erfolg eines Short Squeezes hängt maßgeblich von der Anzahl der Beteiligten ab: Je mehr Personen bei der Aktion mitmachen, desto erfolversprechender

sind die Aussichten. Im Fall von R/WALLSTREETBETS schien der Versuch vielversprechend. Ende Januar 2021 vervielfachte sich die Mitgliederzahl des Subreddits, wobei der größte Mitgliederzuwachs innerhalb weniger Tage geschah (von rund 1,8 Millionen auf über 9 Millionen Mitglieder). Zur selben Zeit nahm auch der Traffic in Form von Kommentaren zu: Ende Januar erreichte die Anzahl der veröffentlichten Kommentare mit über 394 000 Kommentaren pro Tag ihren Höhepunkt, nachdem sie zuvor bei ca. 30 000 bis 70 000 Kommentaren pro Tag lag (vgl. Subreddit Stats 2024). Die Beteiligung am versuchten Short Squeeze war in der Tat so groß, dass die Mitglieder durch ihr Vorgehen erhebliche Schwankungen an den weltweiten Börsenmärkten auslösten (vgl. Biermann/Hegemann 2021) und mehrere Hedgefonds in den finanziellen Ruin trieben (vgl. Robertson 2021).

Im Verlauf der Aktion fungierte der Interaktionsraum R/WALLSTREETBETS als zentraler Treffpunkt, der den Beteiligten die Infrastruktur bot, sich über physische Grenzen hinweg miteinander zu vernetzen und ihr Vorgehen zu koordinieren. Das Forum wurde im Verlauf der Aktion immer mehr zu einer Arena, in der die Mitglieder lautstark ihren Zusammenhalt einforderten und sich mittels eines gruppenspezifischen Jargons als eingeschworene Gemeinschaft inszenierten. Hier zeigt sich deutlich, dass R/WALLSTREETBETS nicht nur als Interaktionsraum konzeptualisiert werden kann, sondern das CoP-Framework mitgedacht werden muss, um den Subreddit vollumfänglich zu verstehen. Aus der analytischen Perspektive des Interaktionsraums bietet R/WALLSTREETBETS den Mitgliedern die notwendigen Affordanzen, um sich untereinander auszutauschen. Ergänzend dazu kann der Subreddit als CoP konzeptualisiert werden, deren Mitglieder mithilfe des Short Squeezes versuchten, ein gemeinsames Vorhaben in die Tat umzusetzen. Der versuchte Short Squeeze ist in diesem Fall also das *jointly negotiated enterprise* der CoP. Die rege Interaktion im Forum und die Organisation zeitlich koordinierter Aktionen verdeutlicht das *mutual engagement* der Mitglieder. Das *shared repertoire* der CoP zeichnet sich durch einen Jargon aus, der auf lexikalischer Ebene u.a. maßgeblich durch die Verwendung von Beleidigungen geprägt ist. Wie ich im Folgenden aufzeigen werde, handelt es sich im Fall von R/WALLSTREETBETS dabei um eine charakteristische sprachliche und kommunikative Praktik, anhand derer Mitglieder ihre Zugehörigkeit zur CoP signalisieren und den Beleidigungen durch Aneignung ein gemeinschaftsstiftendes Potential verleihen.

4. Das R/WALLSTREETBETS-Korpus

Um die Verwendung von Beleidigungen im Subreddit zu untersuchen, wurden sämtliche Beiträge, Kommentare und Antworten, die zwischen dem 10. Januar und 9. Februar 2021 auf R/WALLSTREETBETS veröffentlicht wurden, im Sommer 2021 gescraped und in einem Korpus zusammengetragen, das etwa 220 Millionen

Tokens umfasst.¹ Inkludiert ist somit im Wesentlichen der Zeitraum, in dem die Mitglieder des Forums versuchten, einen Short Squeeze auszulösen. Die Daten wurden über die Webseite pushshift.io erhoben, die über die API von Reddit die Daten aus diversen Subreddits sammelt und zur Verfügung stellt (Reddit, Inc. 2024). Das Part-of-Speech-Tagging wurde mit dem *TreeTagger* durchgeführt und berücksichtigt somit Wortform, Wortart und Lemma (vgl. Schmid 1997). Zusätzlich zur Annotation der positionellen Attribute mit dem *TreeTagger* wurden eine Reihe struktureller Attribute annotiert. Diese Attribute wurden als Metadaten beim Scraping mithilfe von Pushshift übernommen und bei der Aufbereitung für die Korpuserstellung als xml-Tags hinterlegt, sodass sie im Korpus als durchsuchbare Metadaten zur Verfügung stehen. Tabelle 1 zeigt eine Übersicht über die Attribute, die bei der Erstellung des Korpus berücksichtigt wurden:

Tabelle 1: Übersicht der positionellen sowie strukturellen Attribute des *wallstreetbets*-Korpus

Positionelle Attribute	Strukturelle Attribute		
Wortform (<i>word</i>)	Beitragstext (<i>text</i>)		
Wortart (<i>pos/part of speech</i>)	Autor:in (<i>text_author</i>)		
Lemma (<i>lemma</i>)	Datum und Uhrzeit der Veröffentlichung (<i>text_datetime</i>)		
	URL des Beitrags (<i>text_permalink</i>)		
	Reddit-ID (<i>text_reddit_id</i>)		
	Korpus-ID (<i>text_submission_id</i>)		
	Bewertung, zusammengesetzt aus Up- und Downvotes der Mitglieder des Subreddits (<i>text_score</i>)		
	Art des Posts:		
	Beitrag (<i>submission</i>)	Kommentar (<i>comment</i>)	
		Art des Kommentars	
		Kommentar auf höchster Ebene (<i>initial</i>)	Antwort auf übergeordneten Kommentar (<i>comment</i>)
	ID des übergeordneten Beitrags oder Kommentars (<i>text_parent_id</i>)		

1 Mein Dank gilt hier Jan Langenhorst, der mit seinem computerlinguistischen Fachwissen die Erhebung der Daten ermöglicht und mich bei der Erstellung des Korpus unterstützt hat.

Während die grau hinterlegten strukturellen Attribute sich auf grundlegende Metainformationen beziehen, bilden die weiß hinterlegten strukturellen Attribute die interaktionalen Eigenschaften ab, um ein umfassendes Verständnis für den Diskurs auf r/WALLSTREETBETS zu entwickeln. Sie konkretisieren, um welche Art von Beitrag es sich handelt, sodass mithilfe dieser Attribute Diskursstränge rekonstruiert werden können. Die Daten sind mithin so aufbereitet, dass die interaktionalen Eigenschaften repräsentiert sind und der qualitativen Analyse zugänglich bleiben.

Der vorliegende Beitrag verfolgt einen explorativen Ansatz, in dem mittels der Analyse von Keywords und Word Sketches v.a. erste quantitative Befunde einen grundlegenden Eindruck vermitteln sollen, wie Beleidigungen zur Konstitution und Aufrechterhaltung einer CoP beitragen. Wie ich in Kapitel sechs erläutern werde, erweist sich insbesondere das Konzept der *linguistic appropriation* als anschlussfähig, um den Gebrauch von Beleidigungen als gemeinschaftsstiftende sprachliche Praktik besser zu verstehen.

5. Herangehensweise

Für die quantitativen Untersuchungen wurde Sketch Engine verwendet, eine online verfügbare Software, in der sowohl bereits vorhandene als auch eigene Korpora für statistische Analysen genutzt werden können (vgl. Kilgarrieff et al. 2004; 2014). Über Sketch Engine können neben der Erstellung von Frequenzlisten auch Keywords, Word Sketches, Kollokationen und n-Gramme analysiert werden. Die Größe des Wallstreetbets-Korpus erwies sich bei der Untersuchung von Keywords und n-Grammen als Hürde, da statistische Berechnungen am Gesamtkorpus sehr träge sind und auf Sketch Engine eine Obergrenze von 41 Millionen Tokens für den Upload eigener Korpora gilt. Daher wurde für die Untersuchung ein randomisiertes Sample-Korpus (*wsb-sample*) erstellt, welches etwa eine Million Tokens umfasst. Über CQP (Corpus Query Processor) wurden per Zufallsprinzip Beiträge ausgewählt und anhand ihrer individuellen IDs aus dem Datensatz kopiert. Aus diesen Beiträgen wurde wiederum das Korpus *wsb-sample* erstellt.

Anhand einer Keyword-Analyse wurden Lemmata aus dem *wsb-sample*-Korpus identifiziert, die im Vergleich zu anderen Korpora gehäuft auftreten, um so Rückschlüsse auf den gruppenspezifischen Sprachgebrauch auf r/WALLSTREETBETS zu ziehen. Dabei interessieren mich keine Lemmata, die typisch für Aktienhandelsdiskurse oder die Interaktion in Subreddits im Allgemeinen sind. Von Interesse sind jene Lemmata, die typisch für den spezifischen Subreddit r/WALLSTREETBETS sind. Aus diesem Grund wurde als Referenz ein Korpus erstellt, das Daten aus r/STOCKS umfasst, einem weiteren Subreddit, dessen Mitglieder sich wie auf r/WALLSTREETBETS auch mit aktuellen Entwicklungen an den weltweiten Börsenmärkten befassen und dort untereinander austauschen (vgl. Reddit, Inc. 2024; Reddit, Inc. 2023a).

Das stocks-sample-Korpus umfasst etwa 28 Millionen Tokens und wurde auf dieselbe Weise erstellt wie das wsb-sample-Korpus. Die Datengrundlage bildet hier ein randomisierter Auszug der Beiträge und Kommentare aus R/STOCKS, der ebenfalls mithilfe von Pushshift erhoben und in einem Korpus zusammengetragen wurde. Bei der Erstellung des stocks-sample-Korpus wurde darauf geachtet, dass der randomisierte Auszug der Daten dasselbe Verhältnis zur Größe des Gesamtkorpus hat wie das wsb-sample-Korpus. Als Referenz ist das stocks-sample-Korpus insofern geeignet, als es dem wsb-sample-Korpus in der thematischen Ausrichtung und dem Zweck ähnelt: Sowohl R/WALLSTREETBETS als auch R/STOCKS sind Subreddits, in denen sich Reddit-Nutzer*innen über Aktienhandel und aktuelle Entwicklungen an den internationalen Börsenmärkten informieren und austauschen. Lemmata, die typisch für den Diskurs im gesamten sozialen Netzwerk Reddit sind und solche, die typisch für Aktienhandelsdiskurse sind, sind also in beiden Korpora frequent und verfälschen die Ergebnisse der Keywordanalyse nicht.² Der Vergleich des stocks-sample- und des wsb-sample-Korpus ermöglicht demzufolge eine möglichst detaillierte und zielführende Untersuchung der Keywords ohne größere Verfälschungen.

Beide Korpora wurden in Sketch Engine hochgeladen, um dort statistische Untersuchungen durchzuführen. Neben der Keyword-Analyse wurden auch Word Sketches berücksichtigt, sodass nicht nur Aussagen zu isolierten Lemmata, sondern auch zu Mehrwortverbindungen getroffen werden können. Die Ergebnisse der Analyse werden im Folgenden vorgestellt.

5.1 Keywords

Tabelle 2 zeigt die fünfzig Keywords mit dem höchsten Keyness-Wert des wsb-sample-Korpus, verglichen mit dem stocks-sample-Korpus. Sie wurden über Sketch Engine nach ihrer relativen Frequenz je Million Tokens aufgelistet. Insgesamt lassen sich die Keywords grob in vier Kategorien unterteilen, die der Übersichtlichkeit halber in der Tabelle farblich voneinander unterschieden werden: Börsenspezifische Termini (gelb), reddit- bzw. netzwerkspezifische Termini (blau), Ausdrücke, die charakteristisch für den Sprachgebrauch auf R/WALLSTREETBETS sind (grün) sowie pejorative Ausdrücke und Personenbezeichnungen (rot). Die erste und zahlenmäßig am stärksten vertretene Kategorie beinhaltet Begriffe, die im Aktienhandel geläufig und somit börsenspezifisch sind. Dabei handelt es sich um Kürzel, anhand derer ein Wertpapier auf dem Börsenmarkt gelistet und identifiziert wird (u.a. *gme*, *pltr*,

2 Wie andere soziale Netzwerke auch, gibt es auf Reddit spezifische Termini, die v.a. die strukturellen Gegebenheiten und Affordanzen der Plattform bezeichnen (z.B. *flair*, *karma* oder *hot*). Ihre Bedeutungen sind für die Ausführungen im vorliegenden Beitrag nicht relevant, weswegen ich hier davon absehe, sie im Detail zu erklären.

nok, *amc*) oder Begriffe, die Vorgänge und Praktiken im Aktienhandel bezeichnen. Wenige Keywords umfassen Termini, die auf Reddit und/oder in anderen sozialen Netzwerken geläufig sind und sich auf bestimmte Eigenschaften bzw. Funktionalitäten des Netzwerks beziehen (z.B. *flair*, *flaired*, *mod*, *upload*, *notify*, *gif*). Neben diesen spezifischen Termini finden sich weitere Begriffe, die charakteristisch für den Sprachgebrauch auf R/WALLSTREETBETS sind und dort auch im Glossar des Forums aufgelistet sind (z.B. *dfv*, *stonk*, *diamond*, *tendie*, *banana*, *porn*). In dieser Liste besonders augenscheinlich sind allerdings eine Reihe von pejorativen Personenbezeichnungen, die in diversen Kontexten v.a. als Beleidigungen verwendet werden (z.B. *ape*, *retard*, *autist*, *fucker*, *boomer*, *dick*, *pussy*).

Tabelle 2: Fünfzig Keywords mit dem höchsten Keyness-Wert in *wsb-sample* (Referenz: *stocks-sample*), absteigend nach ihrem Keyness-Wert sortiert.

	Lemma	Relative Frequenz wsb-sample	Relative Frequenz stocks-sample
1	gme	4.345,43	6,78
2	melvin	485,38	0,00
3	pltr	385,09	0,00
4	ape	489,97	0,00
5	flair	957,74	2,24
6	flaired	300,11	0,00
7	dfv	226,61	0,00
8	nok	2.304,39	13,31
9	sndl	152,35	0,00
10	stonk	320,78	1,10
11	amc	1.613,84	11,15
12	boyfriend	202,11	0,00
13	misuse	300,11	1,35
14	apha	124,02	0,00
15	retard	1.136,12	9,12
16	covid	106,42	0,00
17	bb	1.208,85	11,43
18	autist	190,63	1,03
19	citadel	144,69	0,00

	Lemma	Relative Frequenz wsb-sample	Relative Frequenz stocks-sample
20	diamond	591,03	5,82
21	biden	94,17	0,00
22	chamath	87,28	0,00
23	gang	194,46	1,85
24	tendie	278,67	3,87
25	bla	129,38	1,35
26	notify	304,70	4,54
27	icln	91,10	0,00
28	cohen	127,85	1,74
29	yolo	1.403,31	30,24
30	webull	163,83	2,77
31	ladder	172,26	3,19
32	gamma	115,60	2,20
33	gif	82,68	1,53
34	gamestop	458,58	13,28
35	gay	104,88	2,66
36	squeeze	1.039,66	38,23
37	upload	127,09	3,83
38	retarded	402,69	14,84
39	fuckin	161,54	5,40
40	porn	114,07	3,59
41	fucker	112,54	4,72
42	moon	1.002,91	40,79
43	boomer	230,44	9,23
44	banana	83,45	2,98
45	mod	522,13	24,17
46	biden	94,17	0,00
47	eod	111,01	4,97
48	dick	477,72	24,71
49	pussy	104,88	4,90
50	halt	297,81	17,78

5.2 Word Sketches

Aus der Liste der fünfzig häufigsten Keywords im *wsb-sample*-Korpus wurden Word Sketches für eine weiterführende Untersuchung der Beleidigungen erstellt. Kilgarriff et al. (2010: o.S.) zufolge handelt es sich bei einem Word Sketch um eine »automatic corpus-derived summary of a word's grammatical and collocational behaviour«. Dabei wird im Korpus nach den Lemmata gesucht, die häufig vor bzw. nach den Beleidigungen stehen. Wie ich in der Diskussion der Ergebnisse ausführen werde, ist die gezielte Suche nach Kollokationen eine aufschlussreiche Möglichkeit, mehr über den kontextuellen Gebrauch der Beleidigungen zu erfahren. Tabelle 3 zeigt eine Auflistung der Lemmata, die am häufigsten in Verbindung mit den Beleidigungen *ape*, *retard*, *autist*, *fucker*, *boomer*, *dick* und *pussy* auftreten. Neben der absoluten Frequenz sind in der Tabelle auch die relativen Häufigkeiten aufgelistet, die anzeigen, welchen Prozentsatz aller Okkurrenzen des jeweiligen Lemmas die Kollokation ausmacht.

Tabelle 3: Word Sketch der Lemmata *ape*, *retard*, *autist*, *fucker*, *boomer*, *dick* und *pussy* in *wsb-sample*

Keyword	Grammatische Relation	Kollokation	Frequenz (absolut)	Frequenz (relativ)
<i>ape</i>	modifiers of x	<i>fellow</i>	19	8,84 %
		<i>retarded</i>	12	5,58 %
		<i>strong</i>	7	3,26 %
		<i>dumb</i>	5	2,33 %
		<i>literal</i>	3	1,4 %
<i>retard</i>		<i>full</i>	22	9,09 %
		<i>here</i>	40	16,53 %
		<i>am</i>	6	2,48 %
	<i>enough</i>	7	2,89 %	
	<i>many</i>	3	1,24 %	
<i>autist</i>		<i>fellow</i>	27	23,28 %
		<i>true</i>	6	5,17 %
		<i>rookie</i>	2	1,72 %
		<i>random</i>	2	1,72 %
		<i>Patagucci</i>	1	0,86 %
<i>fucker</i>		<i>mother</i>	19	27,93 %
		<i>beautiful</i>	2	2,94 %
		<i>Romanian</i>	1	1,47 %
		<i>chuckhold</i>	1	1,47 %
		<i>lemon</i>	1	1,47 %
<i>boomer</i>		<i>Ok</i>	4	11,08 %
		<i>Okay</i>	3	8,31 %
		<i>REEEEEEEEE</i>	1	2,77 %
		<i>brain-dead</i>	1	2,77 %
		<i>cenile</i>	1	2,77 %
<i>dick</i>		<i>sucking</i>	2	2,99 %
		<i>fat</i>	2	2,99 %
		<i>own</i>	4	5,97 %
		<i>retarded</i>	2	2,99 %
		<i>diamond</i>	3	4,48 %
<i>pussy</i>		<i>fuckoff</i>	1	2,56 %
		<i>submissive</i>	1	2,56 %
		<i>Newb</i>	1	2,56 %
		<i>capitTOL</i>	1	2,56 %
		<i>dusty</i>	1	2,56 %

Hier fällt auf, dass die Lemmata *ape*, *retard* und *autist* u. a. in Verbindung mit anderen Lemmata stehen, die eher positiv konnotiert sind (*fellow*, *strong*, *true*) oder, im Falle von *retard*, zum Ausdruck bringen, dass Nutzer*innen diese auf sich beziehen (*am*, *here*). Darin unterscheiden sie sich von den anderen Beleidigungen, die bis auf *fucker*, das zweimal in Verbindung mit *beautiful* auftaucht, nicht mit eindeutig positiven Lemmata stehen. Auch ist die Häufigkeit des Lemmas *fellow* auffällig, das in Verbindung mit *ape* mit 19 Okkurrenzen und in Verbindung mit *autist* mit 27 Okkurrenzen mit Abstand am häufigsten auftritt. Demgegenüber stehen die Lexeme *fucker*, *boomer*, *pussy* und *dick*. Während *ape*, *retard* und *autist* primär durch positiv bewertete Lexeme modifiziert werden, sind es bei *pussy* v. a. negativ bewertete Lexeme wie *fuckoff*, *submissive* oder *Newb*. Ähnliches ist auch bei dem Lemma *dick* zu beobachten, während *boomer* primär mit *Ok* bzw. *Okay* steht, wobei es sich bei *Ok boomer* um einen in sozialen Netzwerken weit verbreiteten feststehenden Ausdruck handelt (vgl. Zeng/Abidin 2021; Diedrichsen 2023). Diese Beobachtungen deuten darauf hin, dass eine Reihe ausgewählter Beleidigungen auf R/WALLSTREETBETS nicht mehr als Schmähungen und Beleidigungen verwendet werden, sondern die Mitglieder diese vielmehr als Selbstbezeichnung nutzen. Inwiefern diese Beleidigungen auch zur Konstitution der CoP beitragen, wird im Folgenden Kapitel näher erläutert.

6. »I'm proud to be a fellow autistic«: Das gemeinschaftsstiftende Potential von Beleidigungen

Am 6. Februar 2021 postete ein*e Nutzer*in im Subreddit R/ANSWERS einen Beitrag betitelt mit der Frage »Why does r/wallstreetbets use terms like ›retard‹ and ›autist‹ so much?« (R/ANSWERS, 6. Februar 2021, 01:41:54 Uhr)³. Unter den insgesamt 79 Kommentaren und noch mehr untergeordneten Replies finden sich diverse Erklärungsansätze, darunter einige, die auf das Verhalten und die Eigenwahrnehmung der Mitglieder anspielen: »It goes hand-in-hand with the risky behaviors, DGAF attitudes and self deprecation that the sub was popularized by.« (R/ANSWERS, 6. Februar 2021, 04:27:21 Uhr). In der Tat deuten die in Abschnitt 5 vorgestellten Ergebnisse darauf hin, dass Beleidigungen im wsb-sample-Korpus wiederkehrend und häufig genutzt werden. Die Annahme liegt also nahe, dass es sich dabei um einen wesentlichen Bestandteil des Repertoires der CoP handelt. Im Folgenden werde ich versuchen, anhand des soziolinguistischen Konzepts der *linguistic appropriation* zu erklären, warum Beleidigungen auf R/WALLSTREETBETS eine wichtige Rolle spielen und wie sie zur Konstitution der CoP im Subreddit beitragen. Der Ansatz der *linguistic appropriation* beschreibt die Verwendung hochgradig beleidigender Termini

3 Um die Privatsphäre der Nutzer*innen zu wahren, sehe ich davon ab, ihre Usernamen zu nennen und die URLs zu den zitierten Beiträgen anzugeben.

durch betroffene stigmatisierte Minderheiten und wurde v.a. beforscht an PoC im angloamerikanischen Raum, die das n-Wort als Selbstbezeichnung verwenden (vgl. Anderson 2018; Galinsky et al. 2003). Indem Betroffene Schmähungen als Selbstbezeichnung verwenden, durchlaufen die Begriffe einen Prozess der semantischen Umdeutung: Sie verlieren ihre ursprünglich herabsetzende Wirkung und werden stattdessen als Ausdruck der (stolzen) Zugehörigkeit zu einer Minderheit verwendet.

Wie in Kapitel 2 erläutert, entstehen in CoPs eine Reihe von Praktiken, darunter auch sprachliche und kommunikative Praktiken, die unmittelbar aus dem gemeinsamen Vorhaben der CoP resultieren und in diesem Zuge eines ihrer wesentlichen Merkmale werden (vgl. Eckert & McConnell-Ginet 1992). Wie ich anhand der Analyse von Keywords und Kollokationen aufzeigen konnte, ist im Fall von R/WALLSTREETBETS eine der augenscheinlichsten sprachlichen Praktiken die Verwendung von Beleidigungen. Insbesondere die Termini *ape*, *retard* und *autist* sind im Subreddit geläufig und auffällig hinsichtlich ihrer Verwendung in Kollokationen. Dabei handelt es sich um Lexeme, die in alltäglichen Gesprächssituationen seitens der Rezipient*innen als verletzend gedeutet würden (sowohl in der Face to face als auch in der medial vermittelten Kommunikation). Vor allem *autist* und *retard* sind Schmähungen, die marginalisierte und von Diskriminierung betroffene Personengruppen herabsetzen und daher als hochgradig problematisch einzustufen sind. Allen drei Begriffen ist gemein, dass sie auf die vermeintlich eingeschränkten kognitiven Fähigkeiten einer Person anspielen.

Zwar handelt es sich bei *ape*, *retard* und *autist* um Beleidigungen, die Untersuchung der Kollokationen im wsb-sample-Korpus relativiert diese Beobachtung jedoch teilweise. Die Befunde legen nahe, dass *ape* und *autist* vorwiegend durch *fellow* modifiziert werden, insbesondere *autist* steht mit Abstand am häufigsten in Verbindung mit dem attributiven Adjektiv. Semantisch könnten die Begriffe gegensätzlicher nicht sein: Während die Beleidigungen i. d. R. eine ausgrenzende, diskriminierende Wirkung entfalten, ist *fellow* mit dem deutschen Präfix *Mit-* vergleichbar, das »in Bildungen mit Substantiven eine Person als jemanden [kennzeichnet], der bzw. die zusammen mit einer anderen Person gesehen wird« (Dudenredaktion 2024). Als solches indiziert *fellow* hier unmissverständlich die Zugehörigkeit einer Person zur Gruppe, und dementsprechend handelt es sich bei *autist* und *ape* um Selbstbezeichnungen, welche die Mitglieder der CoP sich selbst verleihen. Auch stehen *ape* und *autist* wiederholt in Verbindung mit *strong* und *true*, also positiv konnotierten Adjektiven, die dementsprechend eine positive Haltung zum modifizierten Subjekt zum Ausdruck bringen. Ähnlich verhält es sich mit *retard*, das in Verbindung mit *here* und *am* steht. Dies deutet darauf hin, dass in diesen Instanzen *retard* als Selbstbezeichnung verwendet wird und ebenfalls nicht als Beleidigung in der eigentlichen Bedeutung gelten kann, da es nicht als beleidigende Äußerung gegen eine andere Person verwendet wird.

Die Häufigkeit der Kollokation *fellow + ape/autist* führt zur Annahme, dass die Beleidigungen hier hinsichtlich ihres Gebrauchs, sowohl aus semantischer als auch aus pragmatischer Sicht, andere Funktionen erfüllen. Diese Schimpfwörter durchlaufen also einen Prozess der semantischen Umdeutung. Ähnlich wie in einem Prozess der *linguistic appropriation* eignen sich die Mitglieder des Subreddits Beleidigungen an und verwenden diese, um Zugehörigkeit zur CoP zu signalisieren. *Linguistic appropriation*, die vorrangig bei PoC oder der LGBTQIA+-Community untersucht wurde (vgl. Anderson 2018; Galinsky u.a. 2003) hat allerdings i.d.R. einen politisch motivierten Hintergrund, der auf die strukturelle Diskriminierung der Betroffenen zurückzuführen ist. Bei den untersuchten Communities handelt es sich z.B. um Bevölkerungsgruppen, die seit Jahrhunderten in hohem Maße von (mitunter institutionalisiertem) Rassismus und Diskriminierung betroffen sind. Dies ist nicht der Fall bei R/WALLSTREETBETS. Die CoP wird zwar nicht stigmatisiert, dennoch stößt sie einen Prozess an, der Ähnlichkeiten zur *linguistic appropriation* hat: Indem die Mitglieder Beleidigungen als Selbstbezeichnung verwenden, werden diese semantisch umgedeutet und verlieren im Diskurs auf R/WALLSTREETBETS ihre diffamierende Funktion.

Die in R/WALLSTREETBETS beobachtete Verwendung von Beleidigungen weist starke Ähnlichkeiten zu Technau (vgl. 2018b) Befunden auf, denen zufolge Schimpfwörter unterschiedliche Funktionen zukommen, je nachdem wie und in welchen Kontexten sie verwendet werden. Werden Schimpfwörter referentiell und pejorativ gegen Betroffene verwendet, handelt es sich also um einen diffamierenden Akt, anders kann es sich jedoch bei der Verwendung von Beleidigungen in Gruppen verhalten (vgl. Technau 2018a; Technau 2018b; Croom 2013). Diese Beobachtung beruht auf Spotorno und Bianchi, welche die gruppeninterne Verwendung von Schimpfwörtern danach unterscheiden, ob es sich dabei um einen bewusst politisch motivierten Gebrauch handelt oder nicht:

In-groups use their own slurs for non-derogatory purposes in at least two kinds of contexts: A. Friendship contexts – where the non-derogatory use has no conscious political or cultural intent [...]; B. Appropriation contexts – where civil rights groups reclaim the use of the slur as a tool of deliberate political and social struggle (Spotorno/Bianchi 2015: 244)

Demzufolge ist nicht jeder Akt, in dem Mitglieder einer Gruppe eine Beleidigung übernehmen, als Form der *linguistic appropriation* zu verstehen. Vielmehr können Beleidigungen in Gruppen als Ausdruck der Zuneigung dienen: »[there,] the terms are applied to non-members of the target group (non-referential uses) or to people for whom the speaker has an affection (non-pejorative uses)« (Technau 2018b: 39). Im Fall von R/WALLSTREETBETS ist der Gebrauch von Schimpfwörtern also als Ausdruck der Zuneigung zu verstehen, anhand dessen Gruppenzugehörigkeit zur CoP signa-

liert werden kann. Indem die Mitglieder ihre Zuneigung mithilfe diverser Beleidigungen signalisieren, versuchen sie ihre Zugehörigkeit zur CoP auszudrücken und so das gemeinsame Vorhaben zu unterstützen: den versuchten Short Squeeze, für den der möglichst große Zusammenhalt einer möglichst großen Gruppe von Beteiligten notwendig ist. Demgegenüber stehen andere Beleidigungen, die nicht der Adressierung anderer Mitglieder dienen, sondern weiterhin ihre beleidigende Wirkung entfalten und so der Abgrenzung und Distanzierung von Nicht-Mitgliedern dienen. Es gilt: Die Guten sind *apes*, *retards* und *autists*, die Schlechten hingegen *fuckers*, *boomers*, *dicks* und *pussies*.

Auf R/WALLSTREETBETS sind dementsprechend primär Schimpfwörter positiv behaftet, die auf die vermeintlich eingeschränkten kognitiven Fähigkeiten der Mitglieder anspielen (u.a. *ape*, *retard* und *autist*). Erklärungen hierfür sind auf umgangssprachliche Termini aus der Finanzwelt sowie popkulturelle Bezugnahmen zurückzuführen: In ihrem Jargon spielen die Mitglieder auf Terminologie an, die unter professionellen Aktienhändler*innen geläufig ist. Dazu zählt u.a. der Ausdruck *dumb money*, der als diffamierende Bezeichnung für Kleinanleger*innen verwendet wird (vgl. Oxford English Dictionary 2023a)⁴. Des Weiteren ist die Verwendung von *ape* eine Anspielung auf den Film PLANET OF THE APES (dt. PLANET DER AFFEN) aus dem Jahr 2001, der von einem Planeten handelt, auf dem Menschenaffen die Herrschaft über die Menschen innehaben (vgl. IMDb.com, Inc. 2024). Diese popkulturelle Anspielung greift also nicht nur den Terminus des Menschenaffen (engl. *ape*) auf, sondern ebenfalls die Grundidee des Films: Die vermeintlich unterlegenen Menschenaffen erlangen die Macht und setzen sich gegen die Menschheit durch.

7. Grenzen und Möglichkeiten der Studie

Ziel des vorliegenden Beitrags war eine explorative Untersuchung sprachlicher Praktiken, die zur Konstitution einer CoP auf R/WALLSTREETBETS beitragen. Die Ergebnisse der korpusbasierten Untersuchung von Beleidigungen bieten einen ersten Einblick in den Sprachgebrauch der Mitglieder. Die Befunde deuten darauf hin, dass Beleidigungen als Instrument eingesetzt werden, um sich als CoP zu konstituieren und die eigene Zugehörigkeit zu signalisieren. Zunächst handelt es sich dabei jedoch nur um Eindrücke, die auf einer explorativen Herangehensweise

4 Das OED online listet unter *dumb money* drei Belege. Einer davon bezieht sich auf den folgenden Artikel aus der New York Times vom 28. Januar 2021, in dem der versuchte Short Squeeze von R/WALLSTREETBETS thematisiert wird: »Dumb Money: Is on GameStop, and It's Beating Wall Street at Its Own Game« (Phillips/Lorenz 2021). Im Gegensatz dazu steht *smart money* für erfahrene und professionelle Aktienhändler*innen und Hedgefondsmanager*innen (vgl. Oxford English Dictionary 2023b).

beruhen. Eine wesentliche Schwäche der durchgeführten Untersuchungen liegt somit in der Generalisierbarkeit der Ergebnisse, die unmittelbar mit der Größe der Datengrundlage zusammenhängt. Zwar handelt es sich beim wsb-sample-Korpus um einen randomisierten Ausschnitt aus dem Gesamtkorpus, allerdings umfasst das Sample mit rund 1 Million Tokens nur einen Bruchteil der insgesamt über 220 Millionen Tokens. Ob sich diese Befunde auf das Gesamtkorpus übertragen lassen, müssen größer angelegte Untersuchungen zeigen.

Als erster Gradmesser sind die Ergebnisse der Analyse anschlussfähig für weiterführende Untersuchungen. Es gilt zunächst die vorläufigen Ergebnisse aus dem wsb-sample-Korpus anhand des Gesamtkorpus zu überprüfen, um abzugleichen, inwiefern die Befunde darauf übertragbar sind. Zudem erlaubt die Annotation des Korpus tiefergehende qualitative Untersuchungen, welche die Verwendung von Schimpfwörtern aus diskurslinguistischer Perspektive ermöglicht. Eine solche qualitativ ausgerichtete diskurslinguistische Herangehensweise erlaubt wiederum die genauere Auseinandersetzung mit dieser sprachlichen Praktik ebenso wie eine trennscharfe Definition und Differenzierung der Verwendung von Beleidigungen auf r/WALLSTREETBETS. In einem weiteren Schritt können komparative Untersuchungen mit anderen sozialen Bewegungen aufschlussreich sein, um sowohl netzwerkintern als auch -übergreifend sprachliche Praktiken der Gemeinschaftsbildung in diversen sozialen Bewegungen zu vergleichen.

8. Fazit

In seiner Reichweite und den daraus resultierenden Implikationen, die es auf die Offline-Welt ausübte, ist das Fallbeispiel r/WALLSTREETBETS hochaktuell. Es zeigt eindrücklich eine mögliche Ausprägung eines Phänomens, das mit #metoo, Black Lives Matter oder #ichbinhanna bereits seit einiger Zeit nicht nur innerhalb der sozialen Medien, sondern auch darüber hinaus Wellen in der Gesellschaft schlägt. Eine Gelingensbedingung für den Erfolg einer solchen sozialen Bewegung ist die Herstellung von Gemeinschaft und Zusammenhalt. Der vorliegende Beitrag hat zunächst gezeigt, wie die soziale Bewegung, die in r/WALLSTREETBETS entstand, konzeptualisiert werden kann. Hierbei erweisen sich zwei Konzepte als zielführend: Einerseits das Verständnis von r/WALLSTREETBETS als virtueller Interaktionsraum und andererseits die Perspektive auf r/WALLSTREETBETS als CoP. Um den Subreddit vollumfänglich zu verstehen, reicht es nicht aus, eines der beiden Frameworks isoliert zu betrachten, vielmehr müssen beide Konzepte berücksichtigt werden. Beide Perspektiven sind als komplementär zu erachten, um den Subreddit für die sprachwissenschaftliche Analyse nutzbar zu machen. Diese dichotome Betrachtungsweise eröffnet eine neue Perspektive auf den Subreddit, die in Anlehnung an Marx und Schmidt als »neue soziale [Realität]« (Marx/Schmidt 2019: 12) zu verstehen ist. Mit-

hilfe der darin entstehenden, individuellen Interaktionspraktiken brachte R/WALLSTREETBETS eine soziale Bewegung hervor, die über eine Reihe gesellschaftlicher Funktionen und Dynamiken verfügt (vgl. Marx/Schmidt 2019b: 12). Diese Dynamiken entwickelten eine große Durchschlagskraft und Reichweite, die wiederum dazu führten, dass die soziale Bewegung die Grenzen des sozialen Netzwerks überwand und mit ihrer konzertierten Aktion erhebliche Schwankungen an den weltweiten Börsenmärkten auslöste.

Basierend auf diesem dichotomen Verständnis wurde eine Untersuchung zur Verwendung von Beleidigungen als eine mögliche sprachliche Praktik der Gemeinschaftsbildung durchgeführt. Die Ergebnisse der korpusbasierten Analysen vermitteln einen Eindruck, wie Mitglieder mithilfe von Beleidigungen ihre Zugehörigkeit ausdrücken und so mit ihrem sprachlichen Handeln zur Konstitution und Aufrechterhaltung dieser CoP beitragen können. Zwar handelt es sich hier um einen ausschnitthaften Einblick in eine isolierte Bewegung, doch die herausgearbeiteten Charakteristika reflektieren die Individualität und die Besonderheiten dieser Bewegung. Zudem konnte die Analyse zeigen, wie eben diese Individualität und die Besonderheiten im Fall von R/WALLSTREETBETS sowohl zur Konstitution einer CoP als auch zum Zusammenhalt der Bewegung beitragen. Insofern gibt der vorliegende Beitrag Anlass dazu, auch weitere soziale Bewegungen aus linguistischer Perspektive in den Blick zu nehmen, um darüber hinaus zu untersuchen, welche Rolle der Sprachgebrauch und sprachliche Praktiken in sozialen Bewegungen spielen.

Literatur

- Anderson, Luvel (2018): *Calling, Addressing, and Appropriation*, Oxford: Oxford University Press.
- Baringhorst, Sigrid (2019): »Auswirkungen der Digitalisierung auf soziale Bewegungen: Annahmen, Befunde und Desiderata der Forschung«, in: Jeanette Hoffmann/Norbert Kersting/Claudia Ritzi/Wolf J. Schünemann (Hg.), *Politik in der digitalen Gesellschaft*, Bielefeld: transcript, S. 151–170.
- Beißwenger, Michael (2013): »Space in Computer-Mediated Communication. Corpus-Based Investigations on the Use of Local Deictics in Chats«, in: Peter Auer/Martin Hilpert/Anja Stukenbrock/Benedikt Szmeicsanyi (Hg.), *Space in Language and Linguistics. Geographical, Interactional, and Cognitive Perspectives*, Berlin: De Gruyter, S. 494–528.
- Biermann, Kai/Lisa Hegemann (2021): Wenn die Yolo-Kidz an der Börse zocken. Abrufbar unter: https://www.zeit.de/digital/internet/2021-01/reddit-gamestop-p-aktien-wallstreetbets-mitglieder-wachstum-aktienmarkt?utm_referrer=htt ps%3A%2F%2Fde.wikipedia.org%2F (Stand: 04.04.2024).

- Croom, Adam M. (2013): »How to do Things with Slurs: Studies in the Way of Derogatory Words«, in: *Language & Communication* 33(3), S. 177–204.
- Dang-Anh, Mark (2019): *Protest twittern. Eine medienlinguistische Untersuchung von Straßenprotesten*, Bielefeld: transcript.
- Dang-Anh, Mark/Dorothee Meer/Eva L. Wyss (Hg.) (2022): *Protest, Protestieren, Protestkommunikation*, Berlin/Boston: De Gruyter.
- Diedrichsen, Elke (2023): »Internet Memes – Funktionen und Motivationen«, in: Simon Meier-Vieracker/Lars Bülow/Konstanze Marx/Robert Mroczynski (Hg.), *Digitale Pragmatik*, Berlin/Heidelberg: Springer, S. 153–183.
- Dudenredaktion (2024): Mit-. Abrufbar unter: https://www.duden.de/rechtschreibung/Mit_ (Stand: 04.04.2024).
- Dürscheid, Christa (2015): »Interaktionsräume ohne Grenzen? Texte in den neuen Medien«, in: Martine Dalmas/Marina Foschi Albert/Marianne Hepp/Eva Neuland (Hg.), *Texte im Spannungsfeld von medialen Spielräumen und Normorientierung: Pisaner Fachtagung 2014 zu interkulturellen Perspektiven der internationalen Germanistik*, München: Iudicium, S. 74–88.
- Eckert, Penelope/Sally McConnel-Ginet (1992): »Think Practically and Look Locally: Language and Gender as Community-Based Practice«, in: *Annual Review of Anthropology* 21(1), S. 461–488.
- Fielitz, Maik/Leslie Gauditz/Daniel Staemmler/Verena Stern (2020): »Digitaler Aktivismus: Hybride Repertoires zwischen Mobilisierung, Organisation und Vermittlung«, in: *Forschungsjournal Soziale Bewegungen* 33(2), S. 397–400.
- Fielitz, Maik/Daniel Staemmler (2020): »Hashtags, Tweets, Protest? Varianten des digitalen Aktivismus«, in: *Forschungsjournal Soziale Bewegungen* 33(2), S. 425–441.
- Galinsky, Adam D./Kurt Hugenberg/Carla Groom/Galen V. Bodenhausen (2003): »The Reappropriation of Stigmatizing Labels: Implications for Social Identity«, in: Jeffrey T. Polzer (Hg.), *Identity Issues in Groups: Research on Managing Groups and Teams*, Bingley: Emerald Publishing Limited, S. 221–256.
- Hausendorf, Heiko/Reinhold Schmitt (2018): »Sprachliche Interaktion im Raum«, in: Arnulf Deppermann/Silke Reineke (Hg.), *Sprache im kommunikativen, interaktiven und kulturellen Kontext*, Berlin/Boston: De Gruyter, S. 87–118.
- Holly, Werner/Peter Kühn/Ulrich Püschel (1984): »Für einen »sinnvollen« Handlungsbegriff in der linguistischen Pragmatik«, in: *Zeitschrift für Germanistische Linguistik* 12, S. 275–312.
- Hutchby, Ian (2001): »Technologies, Texts and Affordances«, in: *Sociology* 35(2), S. 441–456.
- IMDb.com, Inc. (2024): *Planet der Affen*. Abrufbar unter: <https://www.imdb.com/title/tt0133152/> (Stand: 04.04.2024).
- Kilgarriff, Adam/Vojtěch Kovář/Simon Krek/Irena Srdanović/Carole Tiberius (2010): »A Quantitative Evaluation of Word Sketches«, in: Anne Dykstra/Tanneke Scho-

- onheim (Hg.), Proceedings of the 14th EURALEX International Congress (Leeuwarden, 6–10 July 2010), Leeuwarden: Fryske Akademy, S. 372–379.
- Kilgarriff, Adam/Pavel Rychlý/Pavel Smrž/David Tugwell (2004): »The Sketch Engine«, in: Geoffrey Williams/Sandra Vessier (Hg.), Proceedings of the 11th EURALEX International Congress (Lorient, 6–10 July 2004), Lorient Cedex: Université de Bretagne sud, S. 105–116.
- Kilgarriff, Adam/Vít Baisa/Jan Bušta/Miloš Jakubíček/Vojtěch Kovář/Jan Michelfeit/Pavel Rychlý/Vít Suchomel (2014): »The Sketch Engine: Ten Years on«, in: *Lexicography* 1, S. 7–36.
- Kneuer, Marianne/Saskia Richter/Melanie Rudolph (2015): *Soziale Medien in Protestbewegungen: neue Wege für Diskurs, Organisation und Empörung?*, Frankfurt a.M./New York: Campus Verlag.
- Lave, Jean/Etienne Wenger (1991): *Situated Learning: Legitimate Peripheral Participation*, Cambridge: Cambridge University Press.
- Leuckert, Sven (2020): »Rethinking Community in Linguistics: Language and Community in the Digital Age«, in: Bettina Jansen (Hg.), *Rethinking Community through Transdisciplinary Research*, Cham: Springer International Publishing, S. 111–125.
- Leuckert, Sven/Martin Leuckert (2020): »Towards a Digital Sociolinguistics: Communities of Practice on Reddit«, in: Sofia Rüdiger/Daria Dayter (Hg.), *Corpus Approaches to Social Media*, Amsterdam: John Benjamins, S. 15–40.
- Marx, Konstanze/Axel Schmidt (Hg.) (2019a): *Interaktion und Medien: Interaktionsanalytische Zugänge zu medienvermittelter Kommunikation*, Heidelberg: Winter.
- Marx, Konstanze/Axel Schmidt (2019b): »Interaktionsanalytische Zugänge zu medienvermittelter Kommunikation. Zur Einleitung in diesem Band«, in: Konstanze Marx/Axel Schmidt (Hg.), *Interaktion und Medien: Interaktionsanalytische Zugänge zu medienvermittelter Kommunikation*, Heidelberg: Winter, S. 1–31.
- Meyerhoff, Miriam/Anna Strycharz (2013): »Communities of Practice«, in: J. K. Chambers/Natalie Schilling (Hg.). *The Handbook of Language Variation and Change*, Chichester/Malden: Wiley, S. 428–447.
- Oxford English Dictionary (2023a): dumb money, n. Abrufbar unter: <https://doi.org/10.1093/OED/5683876762> (Stand: 04.04.2024).
- Oxford English Dictionary (2023b). smart money, n. Abrufbar unter: <https://doi.org/10.1093/OED/1627998512> (Stand: 04.04.2024).
- Phillips, Matt/Taylor Lorenz (2021): »Dumb Money« is on GameStop, and it's Beating Wall Street at its Own Game. Abrufbar unter: <https://www.nytimes.com/2021/01/27/business/gamestop-wall-street-bets.html> (Stand: 04.04.2024).
- Reddit, Inc. (2024): Pushshift-API pushshift.io. Abrufbar unter: <https://pushshift.io/signup> (Stand: 04.04.2024).
- Reddit, Inc. (2023a): r/wallstreetbets – wallstreetbets. Abrufbar unter: <https://www.reddit.com/r/wallstreetbets/> (Stand: 04.04.2024).

- Reddit, Inc. (2023b): Was sind Communities oder »Subreddits«? Abrufbar unter: <https://support.reddithelp.com/hc/de/articles/204533569-Was-sind-Communities-oder-Subreddits-> (Stand: 04.04.2024).
- Robertson, Harry (2021): Short-Sellers are Nursing Estimated Losses of \$19 Billion in 2021 After Betting on GameStop's Share Price to Fall. Abrufbar unter: <https://markets.businessinsider.com/news/stocks/short-sellers-sitting-on-19-billion-of-losses-on-gamestop-data-shows-2021-1-1030020684> (Stand: 04.04.2024).
- Schmid, Helmut (1997): »Probabilistic Part-of-Speech Tagging Using Decision Trees«, in: Daniel B. Jones/Harold Somers (Hg.), *New Methods in Language Processing*, London: Routledge, S. 154–164.
- Spotorno, Nicola/Claudia Bianchi (2015): »A Plea for an Experimental Approach on Slurs«, in: *Language Sciences* 52, S. 241–250.
- Stegbauer, Christian (2018): *Shitstorms. Der Zusammenprall digitaler Kulturen*, Wiesbaden: Springer.
- Storrer, Angelika (2001): »Sprachliche Besonderheiten getippter Gespräche: Sprecherwechsel und sprachliches Zeigen in der Chat-Kommunikation«, in: Michael Beißwenger (Hg.), *Chat-Kommunikation. Sprache, Interaktion, Sozialität & Identität in synchroner computervermittelter Kommunikation. Perspektiven auf ein interdisziplinäres Forschungsfeld*, Stuttgart: ibidem, S. 3–24.
- Subreddit Stats (2024): [r/wallstreetbets stats](https://subredditstats.com/r/wallstreetbets). Abrufbar unter: <https://subredditstats.com/r/wallstreetbets> (Stand: 04.04.2024).
- Technau, Björn (2018a): *Beleidigungswörter: Die Semantik und Pragmatik pejorativer Personenbezeichnungen*, Berlin/Boston: De Gruyter.
- Technau, Björn (2018b): »Going Beyond Hate Speech: The Pragmatics of Ethnic Slur Terms«, in: *Lodz Papers in Pragmatics* 14(1), S. 25–43.
- Zeng, Jing/Crystal Abidin (2021): »#OkBoomer, Time to Meet the Zoomers«: Studying the Memefication of Intergenerational Politics on TikTok«, in: *Information, Communication & Society* 24(16), S. 2459–2481.

Die Autor*innen

Dr. Helga Behrmann promovierte interdisziplinär zu Digital Fashion und virtuellen Kleidern an der Universität Potsdam. Sie forscht zu Ästhetik, Herstellungsprozessen und soziokulturellen Bedingungen der medialisierten Mode und zu Sprach-Bildbezügen auf den Screens.

https://publishup.uni-potsdam.de/opus4-ubp/frontdoor/deliver/index/docId/62026/file/behrmann_diss.pdf

* h.behrmann@netzwerk-mode-textil.de

Dominic Brakelmann, graduierte im Master Medienwissenschaft an der Bauhaus-Universität Weimar. 2017 gründete er mit Arvid Kammler den spielekulturwissenschaftlichen Internetblog ilinx.space. Zuvor studierte er Germanistische Literatur- und Kulturwissenschaft, Kunstgeschichte und Musikwissenschaft an der Technischen Universität Dresden. Er ist zurzeit Wissenschaftlicher Mitarbeiter am Lehrstuhl für Medientheorie an der Zeppelin Universität Friedrichshafen und promoviert zu medienreflexiven Angsträumen in Literatur, Film und Videospiel. Aktuelle und thematisch relevante Veröffentlichungen:

Brakelmann, Dominik (2022): »Der Avatar als phantastische Anordnung des Computerspiels«, in: Paidia. Zeitschrift für Computerspielforschung. Abrufbar unter: <https://paidia.de/der-avatar-als-phantastische-anordnung-des-computerspiels/> (Stand: 12.10.2024).

Brakelmann, Dominik (2023): »A player chooses, a gamer obeys... Poetologische Diskursivierungen spielerischer Freiheit in Deus Ex, Deus Ex: Human Revolution, System Shock 2 und Bioshock«, in: Paidia. Zeitschrift für Computerspielforschung. Abrufbar unter: <https://paidia.de/a-player-chooses-a-gamer-obeys-poetologische-diskursivierungen-spielerischer-freiheit-in-deus-ex-deus-ex-human-revolution-system-shock-2-und-bioshock/> (12.10.2024).

Brakelmann, Dominik (2024): »Black Box Bunker. Bunkerarchäologie in Fallout 3«, in: Arno Görgen/Rudolf Thomas Inderst (Hg.): Old World Blues. Fallout und das Spiel mit der Postapokalypse. Marburg: BÜCHNER 2024, S. 219–244.

* brakelmann1@aol.com

Friederike Fischer, geb. 1995, ist Doktorandin und wissenschaftliche Mitarbeiterin am Institut für Germanistik und Medienkulturen der Technischen Universität Dresden. In ihrer Forschung widmet sie sich sozialen Bewegungen in Onlinekontexten und untersucht, wie Nutzer*innen sich mittels sprachlicher Praktiken online solidarisieren. Ihre Forschungsschwerpunkte liegen im Bereich der Medien-, Sozio- und Diskurslinguistik. Aktuelle und thematisch relevante Veröffentlichungen:

Fischer, Friederike/Meier-Vieracker, Simon/Niendorf, Lisa (i.E.): TikTok – Memefication und Performance. Interdisziplinäre Perspektiven. Heidelberg: J. B. Metzler.
 Truan, Naomi/Fischer, Friederike (2023): "Die digitale Hashtag-Kampagne rund um #CoronaEltern und #CoronaElternRechnenAb: Twitter-Positionierungspraktiken in der Pandemie", in: Journal für Medienlinguistik (jfml) 5(1), S. 1–33. <https://doi.org/10.21248/jfml.2023.48>

* friederike.fischer@tu-dresden.de

Tim Hector ist Postdoc im SFB 1187 »Medien der Kooperation« an der Universität Siegen. Er forscht in den Bereichen Medienlinguistik, Gesprächsanalyse, Mediendomestizierung und Mensch-Maschine-Interaktion. Aktuelle und thematisch relevante Veröffentlichungen:

Habscheid, Stephan/Hector, Tim/Hoffmann, Dagmar/Waldecker, David (Hrsg.) (2024): *Voice Assistants in Private Homes. Media, Data and Language in Interaction and Discourse*. Bielefeld: transcript.

Hector, Tim (im Ersch.): *Smart Speaker im Dialog. Sprachliche Praktiken mit Voice User Interfaces*. Berlin/Boston: de Gruyter (zugl. Diss., Phil. Fak. Uni Siegen).

* tim.hector@uni-siegen.de

Dr. Timo Kaerlein ist Akademischer Rat am Institut für Medienwissenschaft der Ruhr-Universität Bochum. Er forscht zur Geschichte, Theorie und Ästhetik von Interfaces, zu digitalen Nahkörpertechnologien und Sensormedien. Aktuelle und thematisch relevante Veröffentlichungen:

Kaerlein, Timo (2018): »Social Bots und die Formalisierung von Sozialität auf Online-Plattformen«, in: Theo Hug/Günther Pallaver (Hg.): *Talk with the Bots. Gesprächsroboter und Social Bots im Diskurs*, Innsbruck: Innsbruck University Press, S. 17–28.

Kaerlein, Timo (2024): »Medientheoretische Aspekte von Mobilität und Wearables«, in: Christoph Ernst, Katerina Krtilova, Jens Schröter und Andreas Sudmann (Hg.): *Handbuch Medientheorien im 21. Jahrhundert*. Wiesbaden: Springer VS.

* timo.kaerlein@rub.de

Arvid Kammler arbeitet als freier Medienwissenschaftler und Mediengestalter und forscht aktuell zur Überlagerung und wechselseitigen Konstitution von Räumlichkeiten in digitalen Spielen sowie zur Rolle des Körpers in temporären Anordnungen des Spiels. Weiter ist er Mitherausgeber des Online-Magazins *ilinx.space*. Aktuelle

und thematisch relevante Veröffentlichungen:

Kammler, Arvid (2021): »Here be dragons« Spielerische Entwürfe be- und entgrenzter Zukünfte«, in: Zeitschrift für Medienwissenschaft, Jg. 13, Nr. 2, S. 45–55. DOI: <http://dx.doi.org/10.25969/mediarep/16788>.

Kammler, Arvid (2022): »Playing with Light. On the Materialities of Video Game Spaces.«, in: Spiel|Formen, Jg. 2, S. 52–64. DOI: <http://dx.doi.org/10.25969/mediarep/19003>.

* mail@arvidkammler.com

Clara Kindler-Mathôt, geb. 1993, ist Doktorandin und wissenschaftliche Mitarbeiterin an der Europa-Universität Viadrina in Frankfurt (Oder). Die Kulturwissenschaftlerin und Linguistin forscht empirisch zu komplexen Prozessen der multimodalen Bedeutungsentfaltung in medienspezifischen, vordergründig politischen Kontexten. Ihr Fokus liegt dabei auf dem Zusammenspiel audiovisueller und sprachlicher Multimodalität. Aktuelle und thematisch relevante Veröffentlichungen:

Kindler-Mathôt, Clara (2024): »*eins zwei drei vor wusch*. Dynamische Entfaltung multimodaler Bedeutung einer onomatopoetischen Interjektion«, in: Kabatnik, Susanne/Lars Bülow/Marie-Luis Merten/Robert Mroczynski (Hg.): Pragmatik multimodal, S. 167–196.

Kindler-Mathôt, Clara/Cornelia Müller/Jana Junge/Katerina Papadopoulou (im Ersch.): »Sprechen in audiovisuellen Medien: Systematik für eine medienästhetische Analyse multimodaler Kommunikation«, in: Vallentin, Rita/Dorothea Horst (Hg.): Sprache entgrenzen. Beiträge zu einer kulturwissenschaftlichen Linguistik, Bielefeld: transcript.

* kindler@europa-uni.de

Lynn Klemmer ist eine Multimedia-Künstlerin. Sie schloss ihr Studium am National College of Art and Design in Dublin mit einem BA in Fine Art & Visual Culture (2017) und an der Universität Potsdam mit einem MA in European Media Studies (2023) ab. Sie arbeitet an der Schnittstelle von analogen und digitalen Medien und versucht, den Platz des Menschen in einer sich schnell entwickelnden algorithmischen, urbanen und natürlichen Landschaft neu zu überdenken. Sie ist auch Mitbegründerin von Mnemazine, einem luxemburgischen interdisziplinären Forschungs- und Kunstkollektivs, das sich als Plattform für Experimente in den Bereichen Philosophie, Soziologie und zeitgenössische Kunstpraxis manifestiert. Als Co-Direktorin von Six Minutes Past Nine veranstaltet sie derzeit ein neues virtuelles Forschungs- und Atelierprogramm.

* lynn_klemmer@hotmail.com

Didem Leblebici, geb. 1996, ist Doktorandin und akademische Mitarbeiterin an der Europa-Universität Viadrina Frankfurt (Oder). Ihre Forschung konzentriert sich auf die soziolinguistischen Implikationen zeitgenössischer KI-Sprachtechnologien auf Sprachideologien und Mehrsprachigkeit. Aktuelle und thematisch relevante Veröffentlichungen:

Leblebici, Didem (2024): »You are Apple, why are you speaking to me in Turkish?«. The Role of English in Voice Assistant Interactions«, in: *Multilingua* 23(4). <https://doi.org/10.1515/multi-2023-0072>.

* leblebici@europa-uni.de

Giacomo Marinsalta, geb. 1993, ist Grafiker und Medienkünstler. Er hat seinen Master in Europäischen Medienwissenschaften an der Universität Potsdam absolviert, wo er als wissenschaftliche Hilfskraft im Forschungsprojekt »Weather Reports: Wind as Media, Model, Experience« tätig war. Seine Forschung vereint visuelle Kunst und Medienökologie und befasst sich mit politischer Ästhetik und urbaner Praxis.

* giacomo@marinsalta.it

Katerina Papadopoulou ist Doktorandin und akademische Mitarbeiterin an der Europa-Universität Viadrina in Frankfurt (Oder). Mit einem transdisziplinären Hintergrund in Kultur-, Sprach- und Medienwissenschaften widmet sie sich primär der Untersuchung multimodaler Bedeutungsentfaltung in face-to-face sowie audiovisuellen Kontexten. Besonderer Fokus liegt dabei auf der dynamischen Entwicklung der Rede-Geste-Interaktion, multimodaler Metaphern und dem komplexen Zusammenspiel von affektivem Wahrnehmen und Verstehen. Aktuelle und thematisch relevante Veröffentlichungen:

Kindler-Mathôt, Clara/Cornelia Müller/Jana Junge/Katerina Papadopoulou (im Ersch.): »Sprechen in audiovisuellen Medien: Systematik für eine medienästhetische Analyse multimodaler Kommunikation«, in: Vallentin, Rita/Dorothea Horst (Hg.): *Sprache entgrenzen. Beiträge zu einer kulturwissenschaftlichen Linguistik*, Bielefeld: transcript.

* papadopoulou@europa-uni.de

Till Rückwart, geb. 1994, ist wissenschaftlicher Mitarbeiter am Seminar für Kultur- und Medienmanagement an der Freien Universität Berlin und dem metaLAB (at) Berlin. In seiner Forschung verbindet er künstlerische Praxis mit Fragestellungen zu visueller Kultur und Interface Studies und fokussiert softwarebasierte Störungen als Ausgangspunkt für medienreflexive Analysen.

* t.rueckwart@fu-berlin.de

Vesna Schierbaum, geb. 1992, ist wissenschaftliche Mitarbeiterin an der Professur für Digitale Medienkultur in Bonn und arbeitet dort im Teilprek "Precision Farming: Ko-operative Praktiken des Virtual Fencing". Sie forscht zur Medien- und Verflechtungsgeschichte des mobilen Sensing, zur Modellierung von Mobility und Crowdverhalten sowie zu Imaginationen sensorischer Gebietsabdeckung.

* vesna.schierbaum@uni-bonn.de

Paul Schütze, geb. 1994, ist wissenschaftlicher Mitarbeiter in der Forschungsgruppe Ethics and Critical Theories of AI an der Universität Osnabrück. Seine Forschungsschwerpunkte liegen in der kritischen Sozialphilosophie, sowie der Technik- und Klimaphilosophie. Aktuell forscht er dazu, wie Künstliche Intelligenz strukturell mit der Klimakrise verbunden ist. Aktuelle und thematisch relevante Veröffentlichungen:

Schütze, Paul (2024): »The Impacts of AI Futurism: An Unfiltered Look at AI's True Effects on the Climate Crisis», in: Ethics Inf Technol 26 (23). <https://doi.org/10.1007/s10676-024-09758-6>

Schütze, Paul (2024): »The Problem of Sustainable AI: A Critical Assessment of an Emerging Phenomenon», in: Weizenbaum Journal of the Digital Society, 4(1). <https://doi.org/10.34669/WI.WJDS/4.1.4>

* paul.schuetze@uni-osnabrueck.de

Clara Kindler-Mathôt, geb. 1993, ist Doktorandin und wissenschaftliche Mitarbeiterin an der Europa-Universität Viadrina in Frankfurt (Oder). Die Kulturwissenschaftlerin und Linguistin forscht empirisch zu komplexen Prozessen der multimodalen Bedeutungsentfaltung in medienspezifischen, vordergründig politischen Kontexten. Ihr Fokus liegt dabei auf dem Zusammenspiel audiovisueller und sprachlicher Multimodalität.

Didem Leblebici, geb. 1996, ist Doktorandin und akademische Mitarbeiterin an der Europa-Universität Viadrina Frankfurt (Oder). Ihre Forschung konzentriert sich auf die soziolinguistischen Implikationen zeitgenössischer KI-Sprachtechnologien auf Sprachideologien und Mehrsprachigkeit.

Giacomo Marinsalta, geb. 1993, ist Grafiker und Medienkünstler. Er hat seinen Master in Europäischen Medienwissenschaften an der Universität Potsdam absolviert, wo er als wissenschaftliche Hilfskraft im Forschungsprojekt »Weather Reports: Wind as Media, Model, Experience« tätig war. Seine Forschung vereint visuelle Kunst und Medienökologie und befasst sich mit politischer Ästhetik und urbaner Praxis.

Till Rückwart, geb. 1994, ist wissenschaftlicher Mitarbeiter am Seminar für Kultur- und Medienmanagement an der Freien Universität Berlin und dem metaLAB (at) Berlin. In seiner Forschung verbindet er künstlerische Praxis mit Fragestellungen zu visueller Kultur und Interface Studies und fokussiert softwarebasierte Störungen als Ausgangspunkt für medienreflexive Analysen.

Anna Zaglyadnova, geb. 1990, ist wissenschaftliche Mitarbeiterin im Kooperationsstudiengang Europäische Medienwissenschaft der Universität Potsdam und der Fachhochschule Potsdam. Die Medienwissenschaftlerin forscht zu Mensch-Maschine-Konfigurationen in hybriden Umgebungen, Verhältnis von Bildlichkeit und Materialität in erweiterten Realitäten und zur medialen Konfiguration des Politischen.