

ars digitalis

Christoph Borbach

Timo Kaerlein

Robert Stock

Sabine Wirth *Hrsg.*

Akustische Interfaces

Interdisziplinäre Perspektiven auf
Schnittstellen von Technologien, Sounds
und Menschen



Springer Vieweg

ars digitalis

Die Reihe ars digitalis wird herausgegeben von Prof. Dr. Dr. Peter Klimczak.

Sollen technische und kulturelle Dispositionen des Digitalen nicht aus dem Blickfeld der sie Erforschenden, Entwickelnden und Nutzenden geraten, verlangt dies einen Dialog zwischen den IT- und den Kulturwissenschaften. Ausgewählte Themen werden daher jeweils gleichberechtigt aus beiden Blickrichtungen diskutiert. Dieser interdisziplinäre Austausch soll einerseits die Kulturwissenschaften für technische Grundlagen, andererseits Entwickler derselben für kulturwissenschaftliche Perspektiven auf ihre Arbeit sensibilisieren und den Fokus auf gemeinsame Problemfelder schärfen sowie eine gemeinsame ‚Sprache‘ jenseits der Fachbereichsgrenzen fördern. Notwendig ist eine solche interdisziplinäre Auseinandersetzung nicht zuletzt deshalb, um den vielfältigen technischen Herausforderungen an Mensch, Kultur und Gesellschaft ebenso informiert wie reflektiert zu begegnen.

In dieser Reihe finden nicht nur Akteure aus Wissenschaft, Forschung und Studierende aktuelle Themen der Digitalisierung fundiert aufbereitet und begutachtet, auch interessierte Personen aus der Praxis werden durch die interdisziplinäre Herangehensweise angesprochen.

Peter Klimczak, Dr. phil. et Dr. rer. nat. habil., ist außerplanmäßiger Professor an der Brandenburgischen Technischen Universität und IT-Verfahrensverantwortlicher und IT-Infrastrukturverantwortlicher für das Berliner Schulwesen.

Christoph Borbach • Timo Kaerlein •
Robert Stock • Sabine Wirth
Hrsg.

Akustische Interfaces

Interdisziplinäre Perspektiven auf
Schnittstellen von Technologien,
Sounds und Menschen

Hrsg.

Christoph Borbach
SFB 1187 Medien der Kooperation
Universität Siegen
Siegen, Deutschland

Timo Kaerlein
Institut für Medienwissenschaft
Ruhr-Universität Bochum
Bochum, Deutschland

Robert Stock
Kultur-, Sozial- und Bildungswissenschaftliche
Fakultät, Institut für Kulturwissenschaft
Humboldt-Universität zu Berlin
Berlin, Deutschland

Sabine Wirth
Fachbereich Medienwissenschaft
Bauhaus-Universität Weimar
Weimar, Deutschland

ISSN 2662-5970

ISSN 2662-5989 (electronic)

ars digitalis

ISBN 978-3-658-47634-2

ISBN 978-3-658-47635-9 (eBook)

<https://doi.org/10.1007/978-3-658-47635-9>

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

© Der/die Herausgeber bzw. der/die Autor(en), exklusiv lizenziert an Springer Fachmedien Wiesbaden GmbH, ein Teil von Springer Nature 2025

Diese Publikation wurde gefördert durch die Deutsche Forschungsgemeinschaft (DFG)—Projektnummer 262513311—SFB 1187 Medien der Kooperation.

Das Werk einschließlich aller seiner Teile ist urheberrechtlich geschützt. Jede Verwertung, die nicht ausdrücklich vom Urheberrechtsgesetz zugelassen ist, bedarf der vorherigen Zustimmung des Verlags. Das gilt insbesondere für Vervielfältigungen, Bearbeitungen, Übersetzungen, Mikroverfilmungen und die Einspeicherung und Verarbeitung in elektronischen Systemen.

Die Wiedergabe von allgemein beschreibenden Bezeichnungen, Marken, Unternehmensnamen etc. in diesem Werk bedeutet nicht, dass diese frei durch jede Person benutzt werden dürfen. Die Berechtigung zur Benutzung unterliegt, auch ohne gesonderten Hinweis hierzu, den Regeln des Markenrechts. Die Rechte des/der jeweiligen Zeicheninhaber*in sind zu beachten.

Der Verlag, die Autor*innen und die Herausgeber*innen gehen davon aus, dass die Angaben und Informationen in diesem Werk zum Zeitpunkt der Veröffentlichung vollständig und korrekt sind. Weder der Verlag noch die Autor*innen oder die Herausgeber*innen übernehmen, ausdrücklich oder implizit, Gewähr für den Inhalt des Werkes, etwaige Fehler oder Äußerungen. Der Verlag bleibt im Hinblick auf geografische Zuordnungen und Gebietsbezeichnungen in veröffentlichten Karten und Institutionsadressen neutral.

Planung/Lektorat: Petra Steinmueller

Springer Vieweg ist ein Imprint der eingetragenen Gesellschaft Springer Fachmedien Wiesbaden GmbH und ist ein Teil von Springer Nature.

Die Anschrift der Gesellschaft ist: Abraham-Lincoln-Str. 46, 65189 Wiesbaden, Germany

Wenn Sie dieses Produkt entsorgen, geben Sie das Papier bitte zum Recycling.

Vorwort

Christoph Borbach, Timo Kaerlein, Robert Stock und Sabine Wirth

Die Computer der Gegenwart begegnen uns – wenn sie uns überhaupt noch begegnen – in den meisten Fällen nach wie vor in Form grafisch aufbereiteter Gebrauchsoberflächen: Sei es am klassischen PC, Smartphone oder Tablet, aber auch in den immersiven Bildumgebungen von Virtual-Reality-Anwendungen. Dies ist Resultat einer langen Entwicklung der Human-Computer Interaction (HCI), in der allerdings auch wiederholt alternative Interaktionsweisen mit analogen und digitalen Maschinen erprobt, aber anschließend meist wieder verworfen wurden. Die Geschichte der User Interfaces ist voller spekulativer Entwürfe, die regelmäßig weit über die Grenzen des jeweils technisch Realisierbaren hinauszielten, bevor sie sich zu nachhaltigen Dispositiven der Handhabung und Konventionen der Repräsentation verfestigten. Hans Dieter Hellige hat bereits 2008 in seinem noch immer lesenswerten „Überblick über die Langzeitentwicklung der MCI“ (Hellige, 2008, S. 10) (Mensch-Computer-Interaktion) diese als von gegenläufigen Entwicklungslogiken gekennzeichnet beschrieben: Auf kürzere Krisen- und Innovationsphasen, in denen eine große Zahl neuer technischer Lösungen parallel entwickelt und erprobt werden, folgen historisch jeweils längere „Stabilisierungs- und Reifephasen mit ausgeprägt konvergenter Entwicklungstendenz“ (Hellige, 2008, S. 20). So kann man verbindlich konstatieren, dass akustische Schnittstellen in der Geschichte der HCI gleich mehrmals erfinden worden sind, insbesondere die Möglichkeit der natürlichsprachlichen Interaktion mit einem Computersystem: ob von Vannevar Bush als Teil seiner Memex-Vision (Bush, 1945), als „phonetic typewriter“ im Kontext der Bemühungen um die Automatisierung von Büroarbeit in den 1950er-Jahren (Hellige, 2008, S. 30), als ein zentrales Element der seit den 1970er-Jahren wiederkehrenden Vorstellungen sog. Natural User Interfaces (NUI) (Caudill, 1992) sowie neuerdings als multimodale Aufrüstung konversationeller Künstlicher Intelligenzen (sog. Large Language Models, LLM), die den verbalen Dialog mit dem Computer (erneut) als „fait accompli“ inszenieren.

Obwohl also Visionen und praktische Umsetzungen von akustischen Interfaces die HCI-Geschichte im Grunde von Beginn an begleitet haben, spielen sie weder im Selbstverständnis der sich für Bedienschnittstellen interessierenden Informatik noch in den sich mit Technik reflexiv auseinandersetzenen Geistes-, Kultur- und Sozialwissenschaften

eine zentrale Rolle. Dies steht im starken Kontrast zu den auf direkte Manipulation von Bildschirmobjekten ausgerichteten GUIs (Graphical User Interfaces) seit den späten 1960er-Jahren, die bis heute als leitendes Paradigma der HCI gelten und entsprechend viel Aufmerksamkeit in der Forschung erfahren haben. Dieses Desiderat war Anlass für eine arbeitsgruppenübergreifende Kooperation innerhalb der Gesellschaft für Medienwissenschaft (GfM), aus der zunächst eine Tagung mit dem Titel „Akustische Interfaces“ am 4. und 5. Mai 2022 im Medientheater des Instituts für Musikwissenschaft und Medienwissenschaft der Humboldt-Universität zu Berlin hervorgegangen ist. Zusammengefunden hatten sich dort u. a. Beteiligte dreier Arbeitsgruppen der GfM: Interfaces, Auditive Kultur und Sound Studies sowie Medienwissenschaft und Dis/Ability Studies. Das gemeinsame Anliegen war ein Ausloten von Schnittstellen in den jeweiligen Forschungsinteressen der AGs. Die eingegangenen Beiträge widmeten sich gleichermaßen der realisierten Technizität akustischer Interfaces in Gegenwart und Geschichte wie ihren spekulativen Entwürfen, ihren zahlreichen Verbindungen mit den Dis/Abilities ihrer Entwickler:innen und Nutzer:innen, ihrer materiell-semiotischen Verfasstheit sowie ihrer Einbettung in weitgreifende Medienökologien.

Die vorliegende Publikation demonstriert die Produktivität des in Berlin begonnenen Austauschs, soll aber eher den Auftakt weiterer Forschungen markieren statt deren Abschluss. Daher versammelt dieses Buch auch Beiträge, die nicht Teil des Tagungsprogramms waren, sowie deutsche Übersetzungen englischsprachiger Beiträge, deren Erkenntnis- und Analysegehalt für Fragen akustischer Interfaces zentral sind: Mara Mills Beitrag „Als mobile Kommunikationstechnologien noch neu waren“, Liz Fabers Untersuchung „Klang (be)schreiben: Sprachsynthese im goldenen Zeitalter der Science-Fiction (1930–1959)“ sowie Coreen McGuires Aufsatz „Die Kategorisierung von Schwerhörigkeit durch Telefonie im Großbritannien der Zwischenkriegszeit“. Die Autor:innen der Beiträge dieses Bandes erzählen situierte Geschichten akustischer Interfaces; sie diskutieren ihren medientheoretischen Status und ihre medienpraktische Verfertigung. Zusammengekommen bilden die Beiträge des vorliegenden Bandes eine erste Kartierung und ein Tableau akustischer Interfaces als von medienwissenschaftlichen Perspektiven gerahmtes Angebot für einen interdisziplinären Dialog. Dieser berücksichtigt explizit auch die Perspektive technischer Disziplinen, die u. a. um den Begriff des Auditory Displays organisiert sind (Kramer, 1994).

Wir danken Shintaro Miyazaki und dem Team des Instituts für Musikwissenschaft und Medienwissenschaft der Humboldt-Universität zu Berlin für ihre Gastfreundschaft im Mai 2022, Wolfgang Ernst und Viktoria Tkaczyk für ihre Grußworte und allen Beteiligten der Tagung für ihre vielfältigen Perspektiven und anregenden Diskussionsbeiträge. Peter Klimczak hat uns nach der Tagung das freundliche Angebot unterbreitet, eine thematische Publikation in der Reihe *ars digitalis* im Springer-Verlag zu planen. Petra Steinmüller und Ulrike Butz haben das daraus resultierende Publikationsprojekt von Beginn an anlagenseitig begleitet – auch ihnen sei gedankt für die hervorragende Betreuung. Zur Realisierung der Publikation beigetragen haben weiterhin Gina Pirsig, Celine Keuer und Kevin Onland. Unser besonderer und ausdrücklicher Dank gilt Charlotte Bolwin, die mit ihrem gewissen-

haften und umsichtigen Engagement in Korrektorat und formaler Endredaktion die Finalisierung des Bands maßgeblich unterstützt hat.

Literatur

Bush, V. (1945, 9 10). As we may think. *Life Magazine*, 112–124.

Caudill, M. (1992). Kinder, gentler computing. Natural I/O technologies make your computer work for you, instead of the other way around. *BYTE*, 17(4), 135–150.

Hellige, H.-D. (2008). Krisen- und Innovationsphasen in der Mensch-Computer-Interaktion. In H.-D. Hellige, *Mensch-Computer-Interface. Zur Geschichte und Zukunft der Computerbedienung* (S. 11–92). transcript.

Kramer, G. (1994). *Auditory display. Sonification, audification, and auditory interfaces*. Addison-Wesley.

Siegen, Deutschland
Bochum, Deutschland
Berlin, Deutschland
Weimar, Deutschland
Im Dezember 2024

Christoph Borbach
Timo Kaerlein
Robert Stock
Sabine Wirth

Inhaltsverzeichnis

Akustische Interfaces. Eine Lagebestimmung 1
Christoph Borbach, Timo Kaerlein, Robert Stock und Sabine Wirth

Technik/Operationen

Akustische versus implizit sonische Schnittstellen. Eine innertechnische Sicht . . . 37
Wolfgang Ernst

Appunns Tonometer. Zu den Politiken eines zeitkritischen akustischen Interfaces im Phonogramm-Archiv Berlin 55
Christopher Klauke

Werkzeuge und Medienpraktiken. Intelligente persönliche Assistenten und das Paradigma objektorientierten Programmierens 75
Benedikt Merkle und Tim Hector

Körper/Sinne

Der Mensch als akustisches Interface: Über Prozesse der Einhörnung, Übertragung und Übersetzung bei der Liveaudiodeskription und im Blindenfußball 97
Judith Willkomm

Voice Interfacing: Zum Ermöglichungspotenzial digitalen Spielens mit der Stimme für Menschen mit Behinderungen 123
Markus Spöhrer

Körper, Stimmen, Prothesen. Eine Geschichte sprechender Interfaces als Assistenztechnologien 145
Christoph Borbach und Benjamin Lindquist

**Die Kategorisierung von Schwerhörigkeit durch Telefonie
im Großbritannien der Zwischenkriegszeit 167**

Coreen Anne McGuire

Umwelten/Räume

Mobiles Musikhören als Interface 191

Eva Schurig

Als mobile Kommunikationstechnologien noch neu waren 207

Mara Mills

**Schnittstellen-Hören. Auditory Display, Interfacedisplay und Sonic
Display. Eine medienpraktische Verknüpfung 225**

Sebastian Schwesinger

Materialitäten/Gestaltung

**Ton auf Band: Raum-Zeit-Manipulationen und Materialwiderstände
im BBC Empire Service 245**

Viktoria Tkaczyk und Christina Dörfling

**Sound be-greifen. Ein Versuch eines akustischen Interface
für die (Un-)Hörbarmachung digitaler Signale 269**

Jan Claas van Treeck

**Design Tinkering, akustische Interfaces und Crip Computing:
Den Turn zum Auditiven weiterdrehen 287**

Daniel Wessolek und Thomas Miebach

Visionen/Spekulationen

**Klang (be)schreiben: Sprachsynthese im goldenen Zeitalter
der Science-Fiction (1930–1959) 303**

Liz Faber

**A book that speaks of its own – Bücher als akustische Interfaces. Eine
spekulative Annäherung an Repräsentationen von Sound im Kontext
von Wissenskommunikation 321**

Margarethe Maierhofer-Lischka

**Der Maschinensemiotik-Ansatz für ein akustisches Mensch-Maschine-
Interface 329**

Peter beim Graben und Peter Klimczak

Über die Autor:innen

Benedikt Merkle verfasste gemeinsam mit Tim Hector den Beitrag „Werkzeuge und Medienpraktiken. Intelligente persönliche Assistenten und das Paradigma objekt-orientierten Programmierens“. Er ist wissenschaftlicher Mitarbeiter am Lehrstuhl für Virtual Humanities an der Ruhr-Universität Bochum und forscht zu Geschichte, Theorie und Ästhetik digitaler Medien.

Benjamin Lindquist, Dr. verfasste zusammen mit Christoph Borbach den Beitrag „Körper, Stimmen, Prothesen: Eine Geschichte sprechender Interfaces als Assistenztechnologien“. Er ist ein Mellon Postdoctoral Fellow mit Northwestern University's Science in Human Culture Program. Seine Forschungsgebiete sind KI, Emotionen und Sprachsynthese.

Christina Dörfling, Dr. phil. verfasste zusammen mit Viktoria Tkaczyk den Beitrag „Ton auf Band: Raum-Zeit-Manipulationen und Materialwiderstände im BBC Empire Service“. Sie promovierte mit einer Arbeit über die Schaltungsgeschichte des elektrischen Schwingkreises. Zuvor studierte sie Musik-, Medien und Geschichtswissenschaft an der Humboldt-Universität zu Berlin und war u. a. wissenschaftliche Mitarbeiterin an der Hochschule für Musik Weimar im BMBF-Projekt „Musikobjekte der populären Kultur“.

Christoph Borbach, Dr. phil. ist Mitverfasser der Beiträge „Akustische Interfaces. Eine Lagebestimmung“ sowie „Körper, Stimmen, Prothesen. Eine Geschichte sprechender Interfaces als Assistenztechnologien“. Er ist Postdoktorand am SFB 1187 „Medien der Kooperation“ an der Universität Siegen und forscht, lehrt und publiziert in den Bereichen Medientheorie und Technikgeschichte sowie Medienepistemologie und -praxeologie.

Christopher Klauke, M.A. verfasste den Beitrag „Appunns Tonometer. Zu den Politiken eines zeitkritischen akustischen Interface im Phonogramm-Archiv Berlin“. Er ist Doktorand am Max-Planck-Institut für Wissenschaftsgeschichte in Berlin und forscht in den Bereichen Geschichte und Politiken musikethnologischer Wissenstechniken sowie über die Genealogie von *Music Information Retrieval*.

Coreen McGuire, PhD verfasste den Beitrag „Die Kategorisierung von Schwerhörigkeit durch Telefonie im Großbritannien der Zwischenkriegszeit“. Sie lehrt am Department of History der Universität Durham und forscht zu Disability History, Medizingeschichte und Wissenschafts- und Technologiestudien.

Daniel Wessolek, Dr. verfasste zusammen mit Thomas Miebach den Beitrag „Design Tinkering, akustische Interfaces und Crip Computing: Den Turn zum Auditiven weiterdrehen“. Als *Digital Creative* leitet er das Media Lab an der Akademie der Bildenden Künste Nürnberg und forscht und lehrt in den Bereichen Interaktionsdesign und Mensch-Maschine-Interaktion. Zentral sind für ihn Co-Design-Prozesse, Dis-/Ability sowie offene, inklusive und ethische Hard- und Softwareansätze.

Eva Schurig, Dr. verfasste den Beitrag „Mobiles Musikhören als Interface“. Aktuell ist sie Postdoktorandin an der Carl von Ossietzky Universität Oldenburg, zuvor arbeitete sie als Postdoktorandin an der Hochschule für Musik, Theater und Medien Hannover. Sie forscht im Bereich systematische Musikwissenschaft zu Themen wie Mobiles Musikhören, Wellbeing und musikalische Aktivitäten.

Jan Claas van Treeck, Prof. Dr. verfasste den Beitrag „Haptischer Sound. Ein Versuch eines akustischen Interface für die (Un-)Hörbarmachung der digitalen Signale“. Er ist Professor für digitale Transformation und Medien an der Media School der Hochschule Fresenius Campus Hamburg und forscht in den Bereichen Künstliche Intelligenz, Kybernetik und digitale Medienökonomie.

Judith Willkomm, Dr. verfasste den Beitrag „Der Mensch als akustisches Interface: Über Prozesse der Einhörung, Übertragung und Übersetzung bei der Liveaudiodeskription und im Blindenfußball“. Sie ist akademische Mitarbeiterin an der Universität Konstanz im Fachbereich Literatur-, Kunst- und Medienwissenschaften und forscht aktuell an der Schnittstelle zwischen Medienwissenschaft, Disability Studies und Sportwissenschaft, neben den Themen des Beitrags u. a. auch zu barrierefreiem Gaming.

Liz W. Faber, PhD verfasste den Beitrag „Klang (be)schreiben: Sprachsynthese im goldenen Zeitalter der Science-Fiction (1930–1959)“. Liz Faber ist Assistant Professor of English & Communication am Dean College, aktuelle Forschungsthemen sind Darstellungen von Computern und künstlicher Intelligenz in der Science-Fiction. Buchveröffentlichungen: *The Computer's Voice: From Star Trek to Siri* (University of Minnesota Press, 2020) und *Robot Suicide: Death, Identity, and Artificial Intelligence in Science Fiction* (Lexington Press, 2023).

Mara Mills, PhD verfasste den Beitrag „Als mobile Kommunikationstechnologien noch neu waren“. Sie ist Associate Professor für Media, Culture, and Communication an der New York University, wo sie das Center for Disability Studies mitbegründete und leitet. Sie ist Gründungsmitglied des Journals *Catalyst: Feminism, Theory, Technoscience* und Mitherausgeberin der Bände *Testing Hearing: The Making of Modern Aurality* (Oxford, 2020) und *Crip Authorship: Disability as Method* (NYU, 2023). Sie forscht und lehrt in den Bereichen Disability Studies, Science and Technology Studies und Sound Studies.

Margarethe Maierhofer-Lischka, PhD verfasste den Beitrag „A book that speaks of its own – Bücher als akustische Interfaces. Eine spekulative Annäherung an Repräsentationen von Sound im Kontext von Wissenskommunikation“. Sie ist als freischaffende Musikerin, Klangkünstlerin und Wissenschaftlerin tätig und forscht zu den Themen musikalisches Hören, musikalisch-mediale Klanginszenierungen, Klangkunst und zeitgenössische Musikpraktiken.

Markus Spöhrer, Dr. verfasste den Beitrag „Voice Interfacing: Zum Ermöglichungspotenzial digitalen Spielens mit der Stimme für Menschen mit Behinderungen“. Er ist akademischer Mitarbeiter am Internationalen Zentrum für Ethik in den Wissenschaften der Universität Tübingen. Seine Forschungsgebiete umfassen die Beziehungen zwischen Behinderung und digitalen Spielen, die Schnittstellen zwischen Game Studies und Surveillance Studies sowie Gegenwartsfilme.

Peter beim Graben, Dr. verfasste zusammen mit Peter Klimczak den Beitrag „Der Maschinensemiotik-Ansatz für ein akustisches Mensch-Maschine-Interface“. Er ist *associated PI* am Bernstein Center for Computational Neuroscience in Berlin. Seine Forschungsgebiete sind Neuroinformatik, Sprachtechnologie sowie mathematische Musiktheorie und -ästhetik.

Peter Klimczak, Prof. Dr. Dr. verfasste zusammen mit Peter beim Graben den Beitrag „Der Maschinensemiotik-Ansatz für ein akustisches Mensch-Maschine-Interface“. Er leitet kommissarisch das Referat „Digitale Lösungen und Infrastruktur“ in der Senatsverwaltung für Bildung, Jugend und Familie (= Kultusministerium des Landes Berlins) und ist außerplanmäßiger Professor an der Brandenburgischen Technischen Universität. Er forscht in den Bereichen Digitale Medien, Künstliche Intelligenz/Kognitive Systeme sowie Medien- und Kultursemiotik.

Robert Stock, Prof. Dr. phil. ist Mitverfasser des Beitrags „Akustische Interfaces. Eine Lagebestimmung“. Er ist Juniorprofessor für Kulturen des Wissens am Institut für Kulturwissenschaft an der Humboldt-Universität zu Berlin und forscht, lehrt und publiziert in den Bereichen Dis/Ability, digitale Medien und Zugänglichkeit, Environmental Dis/Humanities und Luso-Afrikanische Dekolonisierungsprozesse.

Sabine Wirth, Jun.-Prof. Dr. ist Mitverfasserin des Beitrags „Akustische Interfaces. Eine Lagebestimmung“. Sie ist Juniorprofessorin für Digitale Kulturen an der Fakultät Medien der Bauhaus-Universität Weimar. Ihre Forschungsschwerpunkte umfassen die Mediengeschichte und -theorie digitaler Kulturen, Handhabungsdispositive des Personal, Mobile und Ubiquitous Computing, Ansätze der Interface Studies, Digitale Bildkulturen sowie die Veralltäglichung von KI-Technologien.

Sebastian Schwesinger, M.A. verfasste den Beitrag „Schnittstellen-Hören. Auditory Display, Interfacedisplay und Sonic Display. Eine medienpraktische Verknüpfung“. Er ist wissenschaftlicher Mitarbeiter im Bereich Musik und Medien der Abteilungen Medien- und Kommunikationswissenschaft sowie Musikwissenschaft an der Martin-Luther-Universität Halle-Wittenberg und forscht zur Mediengeschichte akustischer Simulationen, zur Geschichte und Theorie klanglichen Denkens sowie zu piratischen Medien- und Musikpraktiken und dem Verhältnis von Ökonomie und Spiel.

Thomas Miebach verfasste zusammen mit Daniel Wessolek den Beitrag „Design Tinkering, akustische Interfaces und Crip Computing: Den Turn zum Auditiven weiterdrehen“. Er ist UX-Consultant bei DB Systel GmbH und forscht in den Bereichen UX und Interfacedesign, Inklusiver Kommunikation und Barrierefreiheit.

Tim Hector verfasste zusammen mit Benedikt Merkle den Beitrag „Werkzeuge und Medienpraktiken. Intelligente persönliche Assistenten und das Paradigma objekt-orientierten Programmierens“. Er ist wissenschaftlicher Mitarbeiter am SFB 1187 „Medien der Kooperation“ an der Universität Siegen und forscht in den Bereichen Medienlinguistik, Gesprächsanalyse und linguistische Praxeologie.

Timo Kaerlein, Dr. phil. ist Mitverfasser des Beitrags „Akustische Interfaces. Eine Lagebestimmung“. Er ist Akademischer Rat am Institut für Medienwissenschaft der Ruhr-Universität Bochum. Er forscht und lehrt zur Geschichte, Theorie und Ästhetik von Interfaces, zu digitalen Nahkörpertechnologien und Sensormedien.

Viktoria Tkaczyk, Prof. Dr. verfasste zusammen mit Christina Dörfling den Beitrag „Ton auf Band: Raum-Zeit-Manipulationen und Materialwiderstände im BBC Empire Service“. Sie ist Professorin für „Medien und Wissen“ im Fachgebiet Medienwissenschaft der Humboldt-Universität zu. Sie lehrt und forscht zu Technologien und Wissenstechniken der Frühen Neuzeit und Moderne. Ihre Publikationen befassen sich mit Medien in wissenschaftlichen Experimenten und Testverfahren, der epistemologischen Bedeutung akustischer Medien und mit Medien der Natur- und Geisteswissenschaften.

Wolfgang Ernst, Prof. Dr. verfasste den Beitrag „Akustische versus implizit sonische Schnittstellen. Eine innertechnische Sicht“. Er war bis Herbst 2024 Inhaber des Lehrstuhls „Medientheorien“ an der Humboldt-Universität zu Berlin und forscht in den Bereichen radikale Medienarchäologie und Medienepistemologie. 2024 erhielt er den für Forschungen zum Thema „Acoustic Studies“ ausgeschriebenen Meyer-Struckmann-Preis der Philosophischen Fakultät der Heinrich-Heine-Universität Düsseldorf.



Akustische Interfaces. Eine Lagebestimmung

Christoph Borbach, Timo Kaerlein, Robert Stock
und Sabine Wirth

Zusammenfassung

Der Beitrag kartiert, welche medien- und kulturwissenschaftlichen Fragen rezente und emergente, aber auch historische Technologien und Praktiken akustischer Schnittstellen für die Interface- und Dis/Ability-Forschung evozieren und was sie aus Perspektive der Sound Studies charakterisiert. Als Einleitung zum ersten deutschsprachigen Band, der akustische Interfaces als eigenständiges Sujet und medienkulturwissenschaftliches Forschungsfeld identifiziert und problematisiert, wird aufgezeigt, welche medienökologischen Dimensionen auditorische Schnittstellen haben, als Teil welcher Daten- und Prozesslogistiken sie fungieren, welchen Designparadigmen und welchen historischen und epistemischen Trajektorien sie folgen. Nicht zuletzt werfen die aktuelle Ubiquität miniaturisierter Smart Speaker und die verbale Komponente diverser Large Language Models eine Reihe von Fragen für unsere mediale Lage auf. Was sind

C. Borbach (✉)

SFB 1187 Medien der Kooperation, Universität Siegen, Siegen, Deutschland

E-Mail: christoph.borbach@uni-siegen.de

T. Kaerlein

Institut für Medienwissenschaft, Ruhr-Universität Bochum, Bochum, Deutschland

E-Mail: timo.kaerlein@rub.de

R. Stock

Kultur-, Sozial- und Bildungswissenschaftliche Fakultät, Institut für Kulturwissenschaft,
Humboldt-Universität zu Berlin, Berlin, Deutschland

E-Mail: robert.stock@hu-berlin.de

S. Wirth

Fachbereich Medienwissenschaft, Bauhaus-Universität Weimar, Weimar, Deutschland

E-Mail: sabine.wirth@uni-weimar.de

© Der/die Autor(en), exklusiv lizenziert an Springer Fachmedien Wiesbaden
GmbH, ein Teil von Springer Nature 2025

C. Borbach et al. (Hrsg.), *Akustische Interfaces*, ars digitalis,
https://doi.org/10.1007/978-3-658-47635-9_1

akustische Interfaces und als Teil welcher Geschichten werden sie analysier- und rekonstruierbar? Welche Politiken, Ökologien und Umweltlichkeiten materialisieren sich in akustischen Interfaces? Was verbinden sie miteinander, was trennen sie voneinander? Welche Relationen ermöglichen akustische Interfaces und wie formatieren und limitieren sie diese? Wie werden akustische Schnittstellen ausgestaltet und welche ästhetischen und epistemologischen Paradigmen sind ihnen eingeschrieben? Wie adressieren, modellieren und ko-konstituieren akustische Interfaces menschliche Sinne und Körper? Und, nicht zuletzt, wie lassen sich die Praktiken und Technologien akustischer Interfaces überhaupt erforschen?

Schlüsselwörter

Akustische Interfaces · Interface Studies · Dis/Ability Studies · Sound Studies · Medienökologie · Human-Computer-Interaction · Digitalität

1 Im Dialog mit Maschinen

Am 25. September 2023 kündigt das US-amerikanische Softwareunternehmen Open AI auf seinem Entwickler:innenblog ein neues Interface zur Interaktion mit der Konversations-KI ChatGPT an. Der Titel des Eintrags lautet „ChatGPT can now see, hear, and speak“ (OpenAI, 2023) und wird, so das Versprechen, „a new, more intuitive type of interface“ darstellen. Mit dessen Hilfe soll fortan die bisher textbasierte Interaktion mit dem auf dem Large Language Model (LLM) GPT-4 basierenden Chatbot auf Basis gesprochener Sprache möglich sein. Nicht nur der freie Dialog über beliebige Themen wird mit dem neuen Voice Interface als Nutzungsszenario avisiert, sondern auch der verbale Austausch zu einem vorliegenden Bild, das beispielsweise ein zu behebendes handwerkliches Problem und das zur Verfügung stehende Werkzeug zeigt. Dabei, so die Entwickler:innen, sei die Gestaltung des Interface u. a. informiert von einer Auseinandersetzung mit den Nutzungserfahrungen mit der frei zugänglichen App *Be My Eyes*, die blinde und sehbehinderte Menschen durch die situationsbezogene Vernetzung mit sehenden Menschen bei der Identifizierung von Gegenständen und Navigation des Alltagslebens unterstützt (Bendel, 2024; Avila et al., 2016). Am 13. Mai 2024 wird schließlich mit GPT-4o („omni“) die nächste Generation des Chatbots veröffentlicht, die in der Lage sein soll, „natürlich“ anmutende Konversationen über Mediengrenzen hinweg zu führen. Das neue Produktfeature eines Voice Machine Interface (VMI) kann als gegenwärtiger Höhepunkt einer Konvergenz verschiedener medientechnischer Entwicklungen im Bereich der Voice Analytics, der Sprachsynthese, des Natural Language Processing (NLP) und der Large Language Models gelten. Dieses baut auf bereits etablierten assistiven Technologien wie Smart Speakern und Personal Digital Assistants auf, überschreitet diese aber zugleich in Richtung eines multimodalen Open-Domain-Dialogs. In Aussicht gestellt wird eine nunmehr freie, mediengestützte Konversation über jedes beliebige Thema. Mittlerweile haben

alle wesentlichen Entwickler:innen von LLMs Voice Interaction in ihre Modelle integriert. Damit eröffnet sich ein Möglichkeitsraum, der eine grundlegende Umstrukturierung der Human-Computer Interaction (HCI) denkbar macht: von der traditionell auf visuellen Gebrauchs- und Zeigeflächen basierenden Interaktion in Richtung eines akustischen Interface, das nicht nur von Open AI diskursiv als „natürlichere“ und „nutzer:innengerechtere“ Gestaltung des Mensch-Maschine-Verhältnisses apostrophiert wird.¹

Dabei handelt es sich bei der Einführung und Weiterentwicklung akustischer Mensch-Maschine-Schnittstellen möglicherweise um eine mehr als inkrementelle Verschiebung im Feld digitaler Medienpraktiken. Wenn sich der Zugang zum Komplex der Digitalität „zunehmend weniger distanziert via den Hand-Auge-Kreislauf (Bild, Schrift), sondern intuitiver via den Gehör-Stimme-Kreislauf, also über Sprechen, Affektivität und Adaptivität an die Performativität des Digitalen“ (Gramelsberger, 2023, S. 183) realisiert, steht mit dem Wandel des Interfaceparadigmas das Verhältnis von Anwender:in und digitaler Technologie grundlegend zur Debatte. Für Gramelsberger hat eine körpertechnische Ausrichtung digitaler Praktiken auf den „Gehör-Stimme-Kreislauf“ gar medienanthropologische Implikationen, verwirklicht sich doch mit ihr die „Interaktion mit der environmentalen, digitalen Wirklichkeit“ (Gramelsberger, 2023, S. 183) auf Kosten der (mit Plessner gesprochen) „exzentrischen Positionalität“ (Plessner, 1928) als basaler anthropologischer Qualität des In-der-Welt-Seins. Dies sei darin begründet, dass an die Stelle eines primär kognitiven Umgangs mit symbolverarbeitenden Maschinen in der digitalen Gegenwart zunehmend eine performative und environmentale Infrastruktur proaktiver Algorithmen trete, die weniger Subjekte adressiert, als vielmehr Affekte moduliert (vgl. dazu bereits Rouvroy, 2013). Akustische Interfaces haben an dieser Entwicklung Anteil, ohne sie alleinig zu determinieren.

Aus medienwissenschaftlicher Perspektive erscheint die sich am Eingangsbeispiel andeutende Konjunktur stimmbasierter Interaktion mit Computern, die auf eine bis in die Anfangszeit der HCI reichende Geschichte zurückverfolgt werden kann, als Anlass und Einladung zu einer systematischen Auseinandersetzung mit akustischen Interfaces und ihren gesellschaftspolitischen Implikationen im Kontext historischer Settings, gegenwärtiger Praktiken und digitaler Kulturen insgesamt. Der vorliegende Sammelband bündelt erstmals im deutschsprachigen Raum eine Reihe interdisziplinärer Beiträge, die sich aus vielfältigen Perspektiven mit der Kultur- und Technikgeschichte akustischer Interfaces, mit gegenwärtigen und emergenten Medienpraktiken, ihren Affordanzen und ihrer Plattformisierung auseinandersetzen und dabei ästhetische mit technischen sowie historische mit informatischen Analyseansätzen kombinieren. Es ist ab-

¹ Dass die Entwicklung hin zu sprechbasierter Interaktion mit KI-Systemen eine neue Stufe der Intimisierung und Anthropomorphisierung von Medientechnik darstellen kann, zeigt die Debatte über die Ähnlichkeit einer der möglichen Stimmen von ChatGPT mit der Stimme der Schauspielerin Scarlett Johansson (Tiku et al., 2024). Johansson hatte im Film *Her* (2013) einer fiktionalen Künstlichen Intelligenz ihre Stimme geliehen, die weitere Nutzung ihrer Stimme für Sprachassistenzen allerdings untersagt.

zusehen, dass die aktuelle Konjunktur akustischer Interfaces auch dahin gehend kritisch werden wird, als dass durch KI in Verbindung mit Technologien der Sprachsynthese bereits nonkonsensuelle Dialoge mit prominenten Stimmakteur:innen realisiert werden können. Die kritische Kehrseite dessen sind Deepfakes. Vokale Autorschaft wird damit künftig weiter problematisch, auch in Bezug auf existierende humane Stimmkörper. Und diesbezüglich werden sprechende Interfaces – nebst ihrer ethischen, juristischen, politischen, ökonomischen usw. Dimensionen – wohl zunehmend auch ein Forschungsfeld der Critical AI Studies bilden.

Zwar gibt es bereits seit Anfang der 1990er-Jahre umfangreiche Forschung der International Community for Auditory Display zu Technologien auditorischer Repräsentation (Kramer, 1994). Diese ist allerdings einerseits ingenieurs- und computerwissenschaftlich geprägt und andererseits, wie es der Begriff bereits besagt, fokussiert sie auf die klangliche Darstellung von Daten.² Auf diese Weise wird der Begriff des Auditory Display weitgehend deckungsgleich mit jenem der Sonifikation verwendet. Sonifikation wiederum bezeichnet vornehmlich Techniken der Darstellung nichtklanglicher Daten durch Sounds, um aus zumeist quantitativen Daten qualitative Aussagen auf Basis des Akustischen abzuleiten, wie beispielsweise in der seismologischen Forschung die Detektion von langzeitlichen mechanischen Vibrationen im Kontext der Erdbebenerkennung (de Campo, 2014, S. 217). Dahingegen widmet sich der vorliegende Band der medienkulturwissenschaftlichen Erforschung und Reflexion akustischer Interfaces und eröffnet einen disziplinenübergreifenden Dialog zur Analyse der Schnittstellen zwischen Technologien, Sounds und Menschen. Die Publikation geht zurück auf eine Kooperation dreier Arbeitsgruppen innerhalb der Gesellschaft für Medienwissenschaft (GfM): Interfaces, Auditive Kultur und Sound Studies, Medienwissenschaft und Dis/Ability Studies.³ Es sind die von den Forschungsinteressen dieser drei AGs motivierten übergreifenden Fragestellungen, die auch die Anlage des Bandes animieren und in der vorliegenden Einleitung näher bestimmt werden.

2 Akustische Interfaces

Wurde der Begriff „Interface“ lange Zeit von anwendungs- und designorientierten Ansätzen der HCI dominiert und war insbesondere auf Fragen der *Usability*, das heißt auf Anwendungsszenarien ausgerichtet (Laurel, 1990; Bonsiepe, 1996; Johnson, 1997), gewannen Fragen zum Interface gleichwohl in medientheoretischen Debatten zur Medialität

²Vergleiche hierzu den Beitrag von Sebastian Schwesinger im vorliegenden Band.

³Erste inhaltliche Überlegungen zum vorliegenden Band wurden im Rahmen der Tagung „Akustische Interfaces“ entwickelt, die vom 4. bis 6. Mai 2022 am Institut für Musikwissenschaft und Medienwissenschaft der Humboldt-Universität zu Berlin stattgefunden hat. Neben den Herausgeber:innen dieses Bandes waren Shintaro Miyazaki und Konstantin Haensch an Konzeption und Organisation der Veranstaltung beteiligt.

des Computers der frühen 1990er-Jahre Bedeutung (Flusser, 1993; Kittler, 1993; Halbach, 1994). Mit neuer Aufmerksamkeit wird der Interfacebegriff vor allem international in der Tradition der Critical Humanities befragt (Galloway, 2012; Hookway, 2014; Andersen & Pold, 2018). Aktuelle Beiträge insistieren auf der Notwendigkeit einer medienwissenschaftlich informierten Interfacekritik (Andersen und Bro Pold, 2011; Dieter, 2024), die in der Lage ist, den „Interface-Komplex“ (Distelmeyer, 2021) gegenwärtiger und historischer digitaler Kulturen in verschiedenen Dimensionen analytisch zu erschließen. Verstanden als Critical Interface Studies (Wirth, 2023) setzen sich medien- und kulturwissenschaftliche Ansätze über Fragen der Anwendbarkeit hinaus mit den Kulturen und Politiken von Interfaces auseinander (Hadler, 2018). Interfaces werden so als zentrale kulturelle Vermittlungsinstanzen begriffen, denen spezifische Userentwürfe genauso eingeschrieben sind wie auch Gebrauchsnormen und Anwendungsszenarien. Als „form of relation“ (Hookway, 2014, S. 4) vermitteln Interfaces nicht nur zwischen Menschen und Computern, sondern sind integraler Bestandteil der globalen digitalen Infrastruktur und regeln Beziehungen zwischen Software, Hardware, User und der Welt außerhalb von Computern – u. a. als User Interfaces, Application Programming Interfaces (APIs), Kabel- bzw. Funkverbindungen und Sensoren. Branden Hookway kommentiert diese Situation folgendermaßen:

„The interface is the zone of relation that comes into being between human beings and machines, devices, processes, networks, and even organizations. It exists in the drawing together of elements into a relation and in eliciting from these elements the properties, behaviors, and actions that constitute a state of augmentation, from which control is made possible.“ (Hookway, 2014, S. 39)

In Hookways Verständnis sind Interfaces als intraagierende Instanzen zu verstehen, deren Prozessualität erst die beteiligten Entitäten in ihren Eigenschaften näher bestimmbar werden lässt, während es sie zugleich einem Prozess der Augmentation und Kontrolle unterzieht. Mit einer solchen Dynamisierung des Interfacebegriffs, die sich auch an Galloways Fokussierung auf das Interface als Effekt (2012) oder an Druckers Verständnis von Interfaces als „space of affordances and possibilities“ (2013) nachvollziehen lässt, wenden sich die Critical Interface Studies gegen eine Fixierung auf den Objektcharakter von Interfaces und betonen ihre situative Dynamik und Prozesshaftigkeit. Die Plausibilität und Tragfähigkeit dieser kulturwissenschaftlichen Perspektivierung von Interfaces müsste sich also in der Auseinandersetzung mit akustischen Interfaces bewähren, deren Interaktionslogiken sich wesentlich schwieriger „dingfest“ machen lassen, sodass gerade die Ephemeralität auditiver Phänomene neue analytische Zugriffe erfordert.

Demgegenüber stehen Bemühungen innerhalb der HCI, die sich seit der Frühzeit des Computings in den 1940er- und 1950er-Jahren, verstärkt aber seit dem Aufkommen grafischer Benutzeroberflächen (Graphical User Interfaces, GUIs) Ende der 1960er-Jahre vornehmlich als Geschichte von Visualisierungstechniken sowie von Gebrauchs- und Zeigeflächen (Wirth, 2014) darstellen lässt und von den historischen Akteur:innen selbst so ver-

standen wurde.⁴ Auch in den Critical Interface Studies lässt sich eine Schwerpunktsetzung auf visuelle Phänomene feststellen, die sich beispielsweise als Interesse an den diagrammatischen Operationen (Hadler et al., 2023) oder Ästhetiken (Andersen und Bro Pold, 2011) von Interfaces artikuliert. Damit stehen Fragen nach der räumlichen Anordnung semiotischer Elemente und den epistemischen Eigenschaften von visuell zugänglichen Zeichenrelationen im Fokus. Matthew Kirschenbaum attestierte den New Media Studies im Allgemeinen einen „prevailing bias ... toward display technologies“ (Kirschenbaum, 2007, S. 31). Erkki Huhtamos begrifflicher Vorschlag einer „Screenology“ für ein disziplinäres Feld, welches sich mit Screens als „information surfaces“ beschäftigen solle (2006, S. 32), kann dafür als ebenso programmatisch gelten wie Lev Manovichs prospektiver Attest, unsere Gesellschaft sei eine „society of the screen“ (Manovich, 2001, S. 99). Mittlerweile hat sich dieser Diskurslage geschuldet ein kultursemiotisch, bildwissenschaftlich und diagrammatisch differenziertes Beschreibungsvokabular für die Analyse visueller HCI-Metaphern und ästhetischer Gestaltungskonventionen herausgebildet; filmwissenschaftliche Konzepte werden auf Interfaces übertragen (so z. B. Distelmeyer (2017, S. 81), der von einer „Interface-Mise-en-Scène“ spricht); und selbst die Beschäftigung mit APIs und anderen infrastrukturellen Tiefenaspekten des Interfacekomplexes bedient sich primär Netzwerkdiagrammen und Datenvisualisierungen.

Der originäre naturwissenschaftliche Interfacebegriff, der sich in der zweiten Hälfte des 19. Jhs. als Fachbegriff für Phänomene der Zirkulation, Verteilung und Übertragung etablierte (Schaefer, 2011, S. 164), hat – so muss festgehalten werden – keine auditive Dimension. Als deutscher Begriff Phasengrenze bezeichnet er die Grenzfläche zwischen miteinander nicht kompatiblen chemischen Substanzen. Für den technologischen Begriff der Schnittstelle ist dies nicht allein etymologisch zu verstehen, sondern epistemologisch fruchtbar: Das Interface ist nicht einem von zwei benachbarten Stoffen oder Sphären zuzuordnen, sondern situiert sich als diaphane *relationale*, statt als manifest materiale Grenzfläche, im Dazwischen und figuriert damit zum eigentlichen Medium im Sinne der altgriechischen Terminologie des aristotelischen „to metaxy“ (Hagen, 2008). Als dieses Dazwischen verkörpert der Topos der Phasengrenze ein mediales Prinzip, wie es für Kanäle allgemein gelten darf, die sich schematisch und räumlich zwischen Sender:in und Empfänger:in situieren. Das Interface ist in diesem Sinne nicht allein die verbindende, sondern auch die abgrenzende Instanz: das Dazwischen zweier inkompatibler (chemischer) Sphären. Stellt also eine Membran (eines Lautsprechers, eines Mikrofons, eines Kopfhörers) bereits ein akustisches Interface dar, da hier mechanischer Druck (Klangwellen) in elektrische Spannungsdifferenzen übersetzt wird? Eine solche materialistische und technik-deterministische Engschnürung akustischer Interfaces scheint nicht zielführend, um dem vielseitigen Gegenstands- und Phänomenbereich gerecht zu werden. Die sonische Qualität der in diesem Band betrachteten akustischen Interfaces impliziert, dass sie nur angemessen als *räumlich* und *umweltlich* situierte Phänomene adressiert werden können. Während optische oder haptische Interfaces beispielsweise im tendenziell luftleeren Weltraum mehr

⁴Vergleiche den Beitrag von Tim Hector und Benedikt Merkle im vorliegenden Band.

oder weniger problemlos funktionieren, sind akustische Interfaces an elementare Medien (Peters, 2015) gebunden und konzeptuell environmental zu denken.

Angesichts der medialen Spezifik auditiver Phänomene scheint es angebracht, anstelle einer objektivierenden Beschreibung akustischer Interfaces das Programm einer medienwissenschaftlichen „Analytik des Interfacing“ (Lipp, 2017) zu verfolgen. Diese interessiert sich 1.) für die *relationierenden Effekte* akustischer Interfaces: Welche menschlichen und nichtmenschlichen Entitäten werden auf welche Weise mittels akustischer Interfaces miteinander in Beziehung gebracht? Welche dieser Relationen sind sichtbar, welche bleiben opak? Welche Körper und Materialitäten setzen akustische Interfaces voraus und ins Werk? Aus Perspektive der Critical Interface Studies lässt sich darüber hinaus 2.) nach den veränderten Paradigmen der *Steuerung und Kontrolle* von und mittels akustischer Interfaces fragen: Worin unterscheiden sich primär akustische Interfaces von primär visuellen und/oder haptischen? In welchem Verhältnis steht die Konjunktur akustischer Interfaces zur parallelen Entwicklung einer Autonomisierung nichtmenschlicher Akteure wie Chatbots und Roboter, deren Interfaces nicht unbedingt als vermittelnde Instanzen in Erscheinung treten, sondern an der Verfertigung eines synthetischen Gegenübers beteiligt sind? Was entzieht sich mit dem Rückzug des Visuellen, welche Möglichkeitsräume eröffnen sich – nicht zuletzt mit Blick auf nichtnormative Körper bzw. „differently-abled bodies“⁵? Insofern gerade User Interfaces als Bestandteil routinierter Medienpraktiken tief in das Alltagsleben ihrer Anwender:innen eingebettet sind, steht 3.) auch die *Veralltäglichsung* von Interfaces angeleiteter digitaler Praktiken neu zur Debatte. Jedenfalls von Anbieter:innenseite wird fortwährend propagiert, dass ein Wandel zu auditiven Bedienparadigmen gerade der Naturalisierung und Habitualisierung von Medienpraktiken im Umgang mit dem Digitalen diene (so beispielsweise in der Rede von sog. „natural user interfaces“ (Wigdor & Wixon, 2011). Mit der erwünschten „everydayness“ (Wirth, 2016) geht 4.) aber auch eine problematische Unsichtbarmachung von *Interfacepolitiken* einher, die bestehende Macht- und Wissensasymmetrien verschärft. Es darf nicht vergessen werden, dass die globale Industrie der großen Digitalkonzerne ein vitales Interesse an der Verdichtung gesprochener Sprache hat und im Rahmen des dominanten Plattformkapitalismus neuerliche digitale Einhegungen (Andrejevic, 2007; Hörl, 2018) zu erwarten sind. Schließlich erlaubt es eine derart ausgerichtete Analytik akustischer Interfaces 5.), auch die *Environmentalität*, also die Umweltbedingtheit und Situativität von Interfaces, zu adressieren. Beispielsweise wird Sonar unter Wasser verwendet, weil die Übertragung von Schallwellen im Wasser günstigeren physikalischen Gesetzmäßigkeiten folgt als in der Luft, was als ökologische Affordanz Angebot und Zwang zugleich ist (Miyazaki, 2012; Borbach, 2022; Stock, et al., 2025). Insbesondere in der dominanten Erscheinungsform assistiver Medien (Smart Speaker und andere Voice Machine Interfaces) wird die Umweltlichkeit akustischer Interfaces besonders deutlich: Ihre Nützlichkeit bewährt sich im Gelingen ihres Situationsbezugs und sie sind eingebettet in Gebrauchsweisen, die sich nicht unabhängig vom Kontext sinnvoll erfassen und beschreiben lassen.

⁵Vergleiche hierzu den Beitrag von Markus Spöhrer im vorliegenden Band.

3 Sound Studies und Wissensgeschichte akustischer Interfaces

Aus Perspektive der medienhistorisch argumentierenden Sound Studies betrachtet, ist der tatsächliche wie fiktionale Dialog zwischen Menschen und Maschinen weitaus älter, als es die aktuelle Relevanz von Smart Speakern vermuten lässt. Musikhistorisch gesehen waren es zudem die Tastaturen musikalischer Instrumente, die zuerst als akustische Interfaces fungierten. Bemerkenswert aus epistemologischer Perspektive ist bei jenen Tastaturen zudem, dass ihre Geschichte bis in gegenwärtige Interfacekulturen hineinreicht, obgleich sie ein Wandel der phänomenologischen Ordnung charakterisiert: Die akustischen Interfaces musikalischer Instrumente wurden im Verlauf ihrer Genese ebenso in rein symbolische, d. h. stille, Medienpraktiken eingebunden. Im Folgenden werden diese Dimensionen akustischer Interfaces aus Perspektive der Sound Studies und der Wissensgeschichte nachvollzogen.

Medienhistorisch betrachtet artikuliert sich in literarischen Schriften seit spätestens dem frühen 19. Jh. ein Wunsch nach einem Dialog mit Maschinen, einem gesellschaftlichen „Mitspracherecht“ von Automaten und damit implizit nach akustischen Interfaces. Die Epoche der Romantik fiel mit einer Blütezeit des Automatenbaus zusammen, in welcher mechanische Flötenspieler:innen, Tänzer:innen und erste sprechende Maschinen konstruiert wurden. So stellte in der romantischen Literatur der Dualismus Mensch-Maschine ein fortwährendes Motiv dar, dessen Grenzen vornehmlich artikulatorisch verwischten, geleitet von der protowissenschaftlichen Frage, ob Maschinen sprechen können (Borbach, 2017). Wurde die Mechanisierung intellektueller Tätigkeit bereits für möglich erachtet, war es die Eventualität maschineller Stimmen, die eine besondere Faszination auslöste, aber stets als unheimlich angesehen und kritisch bewertet wurde. So wurde dem vermeintlichen Schachautomaten Wolfgang von Kempelens von Zeitzeug:innen Authentizität bescheinigt, seiner Sprechmaschine (von Kempelen, 1791) hingegen Betrug unterstellt, obgleich es faktisch umgekehrt war. Einen zentralen Topos nehmen sprechende Maschinen bei E.T.A. Hoffmann ein. In *Die Automate* von 1814 werden beide Maschinen von Kempelens zum literarischen Motiv, wobei es die Figur Ludwig ist, die das Unheimliche einer sprechenden Maschine präzisiert: „Mir sind“, sagte Ludwig, „alle solchen Figuren, die dem Menschen nicht sowohl nachgebildet sind, als das Menschliche nachäffen, diese wahren Standbilder eines lebendigen Todes oder eines toten Lebens, im höchsten Grade zuwider““ (Hoffmann, 1958 [1814], S. 414). Und auch in seiner Erzählung *Der Sandmann* von 1816 ist es bezeichnend, dass die Figur der rätselhaften Olimpia, die sich schließlich als Automat entpuppt, eine bezaubernde Singstimme aufweist.

Die Entwicklung tatsächlicher Sprechmaschinen ist ebenso auf die erste Hälfte des 19. Jhs. zu datieren. Zu nennen sind hier beispielsweise die aus Leichenteilen konstruierte Sprechmaschine des Physiologen, Biologen und Anatomen Johannes Müller (1839) oder die Apparatur „Euphonia“ von Joseph Faber, die er erstmals im Jahr 1840 in Wien und kurze Zeit später in den USA öffentlich vorführte. Nach einer Konjunktur der Technikfiktion

sprechender Roboter in der Hochzeit der Science-Fiction-Literatur⁶ konkretisierte sich die Vision einer stimmbasierten Mensch-Computer-Kommunikation im bekannten Aufsatz „As We May Think“ von Vannevar Bush (1945). Dort visioniert Bush zusätzlich zu den inzwischen vertrauten Desktopanwendungen und -metaphoriken eine Interaktion zwischen Mensch und Computer, die sich – nach dem Vorbild des in den Bell Labs entwickelten Vorders – auditiv realisierte. Hier zeigt sich ein kultureller Phonologozentrismus, nach welchem es nur folgerichtig ist, Kommunikation zwischen Menschen und Technologien stimmlich zu realisieren (Borbach, 2016).

Medientechnologisch und -kulturell hat sich, wie eingangs skizziert, erst in den letzten Jahren das auditive Interfaceparadigma der maschinellen Sprachgenerierung und stimmlichen Steuerung (insbesondere durch die anhaltende Konjunktur rezenter Smart Speaker-Technologien) auf dem Massenmarkt etablieren können. Dieses nimmt zunehmend Abstand von alphanumerischer und bildschirmbasierter Voice Machine Interaction und privilegiert den akustischen Kanal. Durch die technologischen Fortschritte algorithmischer Sprachsynthese, deren Kehrseite immer schon die Sprachanalyse und damit zugleich die Spracherkennung war, ist verbale Mensch-Maschine-Kommunikation im 21. Jh. in das häusliche Umfeld eingekehrt und findet nicht mehr in exklusiven Umgebungen wie Laboren von Industrieforschungsunternehmen statt (Haensch, 2021). Auf Grundlage technischer Entwicklungen in den Bereichen der maschinellen Erzeugung, Erkennung und Prozessierung von Stimmen (Story, 2019) konsolidiert sich sprach- und tonbasierte HCI nunmehr in verschiedenen privaten Anwendungsfeldern – von Smartphones, Smart TVs und Spielekonsolen bis hin zu den sprechenden und tönenden Navigationssystemen der Automobilität. Derartige auf Sprachsteuerung basierende Interfaces werfen eine Reihe neuer Fragen auf, die von der Veralltäglichsung und Habitualisierung ubiquitärer User Interfaces bis zur kritischen Auseinandersetzung mit rezenten Techniken KI-gestützter Voice Analytics reichen, die wiederholt die Differenz von Signal und Rauschen nachrichtentechnisch zu formalisieren suchen und Fragen der Signaldiskriminierung (Chun, 2021) virulent werden lassen. Damit ist der Frage nach akustischen Interfaces eine immanent politische (Bendel, 2018) sowie mitunter postkoloniale Dimension und Signalökonomie eingeschrieben.⁷

Rezente akustische Interfaces wie beispielsweise Smart Speaker begründen eine umweltliche Beziehung zwischen User und Technologie. Wurde dem Radio seit geraumer Zeit von Medien- und Kulturwissenschaftler:innen attestiert, zum „Nebenbeimedium“ geworden zu sein, da es Hörenden in der Alltagspraxis vornehmlich um Hintergrundbeschallung statt um aktive Mediennutzung gehe (Hickethier, 1997), liegt in jenem Status des „Hintergrundmediums“ bei Smart Speakern ihr Potenzial und ihre Affordanz. Da sich Kommunikation mit ihnen auditorisch artikuliert, realisiert sich in den Umgebungen, in denen sie agieren, aufgrund ihres „Always-on“- und „Always-on-us“-Charakters (Turkle, 2008) idealiter eine Form der auditiven Nahtlosigkeit in der HCI. Die menschliche Stimme

⁶Vergleiche den Beitrag von Liz Faber in diesem Band.

⁷Vergleiche den Beitrag von Christoph Klauke in diesem Band.

wird zur Referenz für die technologische Gestaltung von Medien, mithin selbst Teil von Interfaces. Dass Smart Speaker hierbei keine neutralen Technologien darstellen, ist wiederholt betont worden. Die aktuelle Forschung zu sprachinteraktiven Interfaces hat verschiedentlich darauf hingewiesen, dass die als gehorsam konnotierte Stimme von Siri normative Genderrollen widerspiegelt und verstärkt (Hennig & Hauptmann, 2019). Heather Suzanne Woods legt zudem dar, dass gerade jene weiblich gehörten Stimmen Formen der Überwachung im Plattformkapitalismus verschleiern (2018).

Darüber hinaus bleibt festzuhalten, dass das *Interface* als integrierte Komponente der digitalen Signal- und Datenverarbeitung im Kontext der computerisierten Akustik – das „Audiointerface“ – seit den 1980er-Jahren über eine explizite Begriffstradition verfügt. Audiointerfaces wie das MIDI (das „Musical Instrument Digital Interface“, das 1982 erstveröffentlicht wurde) realisieren Analog-Digital-Umwandler, die aus analogen klanglichen Signalen digitale Signale transduzieren, und umgekehrt. Zudem sind es die Tastaturen musikalischer Instrumente, die im musikalischen Diskurs als Bedingung eines „Musizierens“ und damit als Interfaces gelten. Dass diese wiederum formatierende Effekte in der musikalischen Praxis zeitigten, wurde bereits aufgearbeitet. Mit ihrem Fokus auf das „keyboard interface“ problematisiert beispielsweise Emily Dolan die klassische Klaviatur als technologisches Problem u. a. des 18. Jhs.: Durch das Aufkommen neuer Techniken der Tonerzeugung stellte sich zugleich das Problem der Tonsteuerung, d. h. der Formatierung des technischen Spielens jener Instrumente. Hier bot sich die Klaviatur als geeignete, aber auch limitierende Schnittstelle an, so Dolan: „The ‚problem‘ became a universal solution: novel ways of producing sound could be efficiently instrumentalized using the keyboard“ (2012, S. 6).

Dieser Prozess der Zurichtung musikalischer Instrumente durch akustische Interfaces, den Dolan „keyboardification“ nennt, ist historisch nicht singulär. Wie Carsten Wernicke zeigt, sieht sich die Einführung neuer digitaler Musikinstrumente fortwährend mit dem Problem konfrontiert, adäquate Interfaces zu entwickeln (Wernicke, 2022), und greift daher oft auf erprobte Designs zurück. Bereits Robert Moog, Entwickler des Moog-Synthesizers, hatte auf diesen Aspekt der kulturellen Habituation verwiesen: „The keyboards were always there ... for some reason or other it looks good if you’re playing a keyboard. People understand then you’re making music“ (Pinch, 2001, S. 386). Im Gegensatz zu Moog sah Synthesizerpionier Don Buchla daher seit den 1960er-Jahren von der Verwendung einer Klaviatur zugunsten von Drehreglern und Touchpads zur Bedienung seiner Synthesizer ab, um, mit Theodor Adorno und Max Horkheimer gesprochen, eine Reproduktion des Immergleichen zu vermeiden, so Buchla: „When you’ve got a black and white keyboard there it’s hard to play anything but keyboard music“ (Buchla zit. n. Pinch & Trocco, 2002, S. 44). Die Intention einer intuitiven Instrumentenbedienung via Klaviaturen – allgemeiner gesprochen: Tasten (Weber, 2009) – wird damit ebenso also Paradigma totaler Kontrolle lesbar, die sich über akustische Interfaces realisiert: Klaviaturen ermöglichen nicht nur ein Musizieren, sondern sie formatieren und limitieren es auch.

Besonders musikalische Interfaces dienen entsprechend einer Kontrolle auf Distanz, die zu Steuerndes und Steuerung räumlich voneinander entkoppelt und damit eine

Virtualisierung des Instrumentenspiels begründet. Wie Michael Harenberg festhält, hat diese Entwicklung eine lange musikhistorische Tradition:

„Ab 1400 wurde es durch verschiedene technische Verbesserungen möglich, die Klangerzeugung in größerer Entfernung vom Spieler zu installieren, was die Trennung vom Spieltisch als mediales Interface und dem klingenden Instrument, der Pfeifenlade, ermöglichte. Die Bedienungseinheit samt Spieler wird historisch erstmalig unabhängig vom Klangkörper gedacht und inszeniert.“ (Harenberg, 2012, S. 175–176)

Grund dafür, dass sich das Spielen nicht in den Unberechenbarkeiten des Realen ausgestaltete, war jene „Bedienungseinheit“, die ein diskretes Format aufweist: die Klaviatur. Dieses Verständnis von Tastaturen *als* standardisierende Schnittstellen der musikalischen Praxis ist zentral für eine *Epistemologie* akustischer Interfaces. Tasteninstrumente, die es seit rund 500 Jahren historisch verbürgt gibt, sind die ersten diskreten bzw. konkreter: die ersten diskretisierenden akustischen Interfaces, insofern sie das klangliche Kontinuum in konkrete Tonhöhen einteilen. Die musikalische Tastatur, so Maren Haffke, ist damit „das erste Push-Button-Interface überhaupt“ (2019, S. 8).

Um 1600 avancierten Tasteninstrumente zu den Universalinstrumenten musikalischer Praxis. Als derartiger Standard hatten Klaviaturen wiederum weitreichende Medieneffekte. Im Laufe des 18. Jhs. etablierte sich die noch heute vertraute Klaviertastatur mit ihren zwölf Halbtonschritten, mit je zwei bzw. drei schmalen versetzten Tasten, die seitdem die Programmatik westlicher Musikvorstellung präsentierte und präfigurierte. Die Klaviatur als akustisches Interface machte das „wohltemperierte Klavier“ – also die Standardisierungen der Intervallordnung – erforderlich und wurde irreduzible Bedingung der Soundästhetik des Pianos und der westlichen Musik. Zudem, in einer eher metaphorischen Wendung des Begriffs, ist die diskrete Klaviatur das Interface, das der westlichen Notenschrift als ebenso diskreter Ordnung entspricht. Klangästhetik ist hier also dezidiert auf die Formatierung akustischer Interfaces zurückzuführen. Damit wird deutlich, dass Fragen der Formatierung von Interfaces Politiken repräsentieren: Sie sind Ergebnis von Aushandlungsprozessen über die Tonhöhenverteilung auf Klaviaturen, mithin Resultat von Frequenzpolitik.

Klaviaturen entfalteten nicht allein musikalische Effekte, die sich in gegenwärtigen Medienkulturen fortschreiben. Trotz alternativer Entwürfe von Klaviaturen, die sich nie längerfristig stabilisieren konnten, ist epistemologisch nicht die Formatierung der Tastaturen *en detail* entscheidend, sondern vielmehr, dass sich via das Interface diverser musikalischer Instrumente ein Prinzip der Steuerung etablierte, das sich über *Tasten* realisierte. Dieses wurde breit praktiziert und festigte sich dadurch habituell und kulturell – und konnte kulturtechnische Wirkmächtigkeit in anderen Praxisbereichen entfalten. Auch die „Komposition von Druckseiten mit beweglichen Typen Mitte des 19. Jhs. geschieht unter Rückgriff auf ein pianistisches Tastenmodell“, schreibt Haffke (2019, S. 64). Zudem wurden die Steuerungen von Schreibmaschinen in direkter Inspiration von Klaviaturen entwickelt. Tatsächlich verwendeten diverse Prototypen von Schreibmaschinen im 19. Jh. Tastaturen, die an die Tastaturen von Klavieren erinnerten. Der Grund dafür, dass sich Kla-

viere und Schreibmaschinen, sogar manche Telegrafen, im historischen Kontext dasselbe Interface teilten, lag in der langen Tradition der Klaviertasten, die eine körpertechnische Vertrautheit mit dem Instrument befördert hatte und in routinierten Praktiken resultierte, die auf jeweils neue Apparaturen und Medientechniken appliziert werden konnte. Mit der Schreibmaschine wurden nunmehr Symbole statt Töne in Anschlag gebracht.

In einer medienepistemologischen Wendung waren es akustische Interfaces, die das emblematische Prinzip einer Kommunikation zwischen Apparaturen und Menschen oder zumindest der unilateralen, weitgehend intuitiven Bedienung präsentierten. Haffke zitiert diesbezüglich den Musikwissenschaftler Ivan Raykoff, der eine umfangreiche Studie zum Klavier zur Zeit der Romantik anfertigte:

„In the early nineteenth century, the piano keyboard provided both a conceptual and a practical model for new communication devices such as the typewriter and the telegraph. Some early models of these writing machines utilized a stretch of piano keys as their keyboard, offering users a familiar interface for transmitting written language. Placed in their historical context alongside the piano, the telegraph and typewriter can be seen as comparable technologies of the fingers.“ (Raykoff zit. n. Haffke, 2019, S. 65)

Mechanische Tastaturen von Musikinstrumenten stellten ein ideales (haptisch-basiertes) Funktionsschema bereit, wie eine gedruckte symbolische Ordnung via Schreibmaschinen zu realisieren sei. Mit dieser „Stillstellung“ der Tastatur durchlebten Klaviaturen genealogisch einen Wechsel der phänomenologischen Ordnung: vom Akustischen hin zum Diskret-Symbolischen getippter Sprache auf Papier. Als verstummte, nicht länger im Sonischen operierende Interfaces, bargen Klaviaturen ein Potenzial, das sich nicht auf die Bedienung musikalischer Dinge zu begrenzen brauchte. Dabei inspirierten Schreibmaschinen als symbolische Leitmedien des 19. Jhs. wiederum grundlegend die Formatierung der symbolischen Schnittstelle eines Leitmediums des späten 20. Jhs.: die „Klaviaturen“ von Digitalcomputern.

Der Erfolg der Tastatur als akustisches Interface der originär musikalischen Kontrolle verdankt sich also der produktiven Einnischung in jeweils neue Technologien. Neue Instrumente und technische Medien konnten sie adaptieren, um Schnittstellen zu verwirklichen, die an bereits vertraute Kultur- und Körpertechniken der haptischen Kontrolle mittels Tasten anschlossen bzw. sich in diese integrierten. Stillgestellte akustische Interfaces, so ließe sich resümieren, haben sich in unsere Medienkultur jenseits des Akustischen eingeschrieben, da Tastaturen integraler Bestandteil rezenter und ubiquitärer Medienpraktiken sind. Im Sinne einer „longue durée“ akustischer Interfaces gilt es entsprechend festzuhalten, dass die akustischen Interfaces von Tasteninstrumenten zunächst im 19. Jh. in Telegraphen und Schreibmaschinen einkehrten und – von diesen ausgehend – im 20. Jh. in die Tastaturen von Computern und Laptops sowie schließlich in deren digitale Darstellung auf den Touchscreens von Smartphones im 21. Jh. Wir hätten es also gleichsam mit einer interfacehistorischen Variante des „Vergessen[s] hinein in die Struktur“ (Winkler, 1997, S. 148) zu tun, die Hartmut Winkler als charakteristisch für Prozesse der Konventionalisierung und Traditionsbildung im Feld der Medien allgemein ausmacht.

Aus Perspektive der Sound Studies muss zudem betont werden, dass sich das epistemische Potenzial akustischer Interfaces nicht in der Realisierung eines Dialogs zwischen Mensch und Maschine oder im Kontrollparadigma via Klaviaturen erschöpft. Bereits durch die Konjunktur erster elektroakustischer Medien in der zweiten Hälfte des 19. Jhs. – zu nennen ist hier vor allem das Telefon – erfuhren hörende Zugriffe auf die Welt eine radikale Veränderung. Beispielsweise war es mit der von René Théophile Hyacinthe Laennec begründeten diagnostischen Praxis der Auskultation (Laennec, 1819) lediglich möglich, über Hörtechniken einen formatierenden Zugriff auf Klangphänomene zu erlangen, die bereits akustisch vorlagen. Das Potenzial der Elektrifizierung bestand hingegen darin, originär nichtklangliche Phänomene in Akustik zu transduzieren⁸ und Wirklichkeiten auditorisch zuzurichten. Jene sonifizierenden Techniken nebst den klanglichen Ereignissen, die sie produzierten, realisierten akustische Interfaces, die vornehmlich in der medizinischen und physiologischen Forschung zur Anwendung kamen, insofern sie direkt an körperimmanente biologische Nervenströme oder Organe angeschlossen werden konnten. Das Telefon in der physiologischen Praxis wurde dabei u. a. dafür verwendet, Muskelreizungen von Tieren sprachlich-telefonisch vorzunehmen oder, andersherum, Muskelreizungen als akustische Telefonsignale anzuzeigen. „So bringt das Telefon zu Gehör, was nie zum Hören bestimmt gewesen war und ermöglicht eine auditive Ausforschung vormals unzugänglicher Phänomenbereiche“, wie Axel Volmar jenen Techniken bescheinigt (2007, S. 110).

Akustische Interfaces eröffnen mithin einen relationalen Raum der Kontrolle, aber auch der Interaktion, Kommunikation und Kooperation zwischen menschlichen und nicht-menschlichen Entitäten, der über die Sphäre stiller, symbolischer Codes hinausgeht und der Suprematie des Blicks in westlichen Kulturen eine auditive Dimension auf Basis flüchtiger Signale beifügt. Darüber hinaus, das zeigt die Medienpraxis der diagnostischen Medizin und der Physiologie ebenso wie des Echolots (Borbach, 2024, S. 147–178) oder des Sonars (Shiga, 2013), resultieren akustische Interfaces aus spezifischen Unsichtbarkeitsproblemen, in welchen klangliche Verfahren einen andernfalls unmöglichen Zugang zu Gegenstands- und Phänomenbereichen realisieren.

4 Interfaces und Behinderung

Im Bereich der medienkulturwissenschaftlichen Disabilityforschung (Ellis & Kent, 2018; Hartwig, 2020; Waldschmidt, 2022) sind im Hinblick auf akustische Interfaces verschiedene relevante Bereiche auszumachen, die Praktiken unterstützter Kommunikation, sprechbasierte Praktiken blinder Menschen sowie Regulationen akustischer Umgebungen bei Menschen mit Hörbehinderungen betreffen. Bevor diese Bereiche jeweils kurz skizziert werden, soll zunächst das Interface aus der Perspektive der kritischen Disabilityforschung näher betrachtet werden.

⁸Vergleiche hierzu den Beitrag von Jan Claas von Treeck in diesem Band.

In Form von engagierten Untersuchungen befassen sich Forschende der Disability Studies vor allem auch in kritischer Intention mit Interfaces und der Art, wie sie implizite Normen instanzieren, die auf unterschiedliche Weise zu Behinderungen beitragen. Sasha Costanza-Chock legt in *Design Justice* (2020) die engen Verbindungen zwischen der Geschichte der Behindertenrechtsbewegung und Fragen der Gestaltungsgerechtigkeit dar (Costanza-Chock, 2020, S. 52). Sie zeigen, dass digitalen Interfaces ein „machine bias“ (Costanza-Chock, 2020, S. 42) inhärent sei, der mehrheitlich strukturelle Ungleichheiten fortschreibt, während konventionelle Interfacekonzepte oft „idealtypische“ Nutzer:innen voraussetzen: Kanonische Texte wie Donald Normans *The Design of Everyday Things* (1990) oder Steve Krugs *Don't make me think* (2000) würden nicht darüber nachdenken, welche normative Körperlichkeit Interfaces implizieren: Norman „ignores race, class, gender, disability, and other axes of inequality ... the imagined user is ‚unmarked‘ and universalized. Terms like *race*, *class*, and *gender* never appear“ (Costanza-Chock, 2020, S. 37, 55). Auch die Frage der Multilingualität, die besonders für akustische Interfaces und deren sprachliche Normen eine Rolle spielt, werde wenig beachtet (Costanza-Chock, 2020, S. 55).

Derartige Betrachtungen von Designgerechtigkeit und Disability Justice lassen sich historisieren. Im Kontext der Analysen der Akteur-Netzwerk-Theorie (ANT), in der etwa Akrich (1992) die Fragen des Skripts technischer Geräte und damit die Figur des idealen Users problematisiert hat, sind Untersuchungen von Ingunn Moser (2000), Moser und John Law (1999) sowie Michel Callon (2005) daran interessiert, die Relationen von behinderten Menschen und assistiven Technologien zu untersuchen. Dieser Fokus auf Menschen und Dinge rückt laut Oudshoorn und Pinch die Frage in den Mittelpunkt, „how technologies work to articulate subjectivities“ (Oudshoorn & Pinch, 2003, S. 11). Subjektpositionen sind ihnen zufolge als Effekte von Netzwerken und hybriden Kollektiven zu verstehen: „Subject positions such as disability and ability are constituted as effects of actor networks and hybrid collectives“ (Oudshoorn & Pinch, 2003, S. 11). Solche Ansätze der ANT sind häufig wegen ihrer fehlenden politischen Sensibilität in die Kritik geraten. Im Kontext der Disabilityforschung hat etwa Winance (2006) darauf aufmerksam gemacht, dass die Etablierung soziotechnischer Relationen in erheblicher Weise von zusätzlicher psychisch-körperlicher Arbeit abhängt, wodurch Chancen gesellschaftlicher Teilhabe eingeschränkt werden. Ein solcher Aufwand wird auch häufig bei begrenzt zugänglichen Mainstreamtechnologien notwendig. Nach wie vor sind viele solcher (digitalen) Technologien durch Skripte gekennzeichnet, die normal-körperliche Nutzer:innen voraussetzen. Forschende sprechen daher davon, dass Behinderung durch Technologien konstituiert werde (Goggin & Newell, 2003). Dabei ist zentral, dass die jeweiligen Skripte einen sog. „preferred user“ (Ellcessor, 2016, S. 77; Ellis et al., 2020) imaginieren. Diese Einschreibungen normativer Körperlichkeit und sensorischer Verfasstheit werfen die Frage auf, welche Körperlichkeit durch Interfaces letztlich entworfen wird, da sich mit ihnen zwar jeweils spezifische Körper- und Sinnestechniken verbinden, diese jedoch nicht immer offensichtlich und explizit werden. Dieser Form des eingeschränkten Zugangs (Ell-

cessor, 2016) – die sich nicht zuletzt auch dem Umstand verdankt, dass Endgeräte und Anwendungen häufig von Designer:innen gestaltet werden, die sich ihrerseits nicht als behindert identifizieren – begegnen Menschen mit Behinderungen durch Hacks und Workarounds (Schabacher, 2017),⁹ um nichtzugängliche Welten zu navigieren (Bieling et al., 2023; Dokumacı, 2023; Hamraie & Fritsch, 2019; Hamraie, 2017) und unzulängliche Interfaces durch kreative Praktiken umzugestalten – häufig mit hohem zeitlichen und organisatorischen Aufwand (Olsen, et al., 2022). Obgleich Fragen der Zugänglichkeit von Interfaces thematisiert werden, über Inklusion und die UN-Behindertenrechtskonvention debattiert wird und Access bzw. Diversity als digitalkapitalistische Vermarktungsstrategie (Costanza-Chock, 2020, S. 75) an Bedeutung gewinnt, ist Zugänglichkeit vielfach noch nicht gegeben oder in genügendem Maße gesetzlich verankert. Hier setzt Ellcessors Vorschlag an, Zugänglichkeit zu begreifen als „a relational, unstable phenomenon that both grants benefits and interpellates individuals into larger social systems that may be empowering, exploitative, or both“ (Ellcessor, 2016, S. 7). Als veränderbare Konstellation verweist „access“ auf „a variable relationship between numerous material, and cultural, social factors“ (Ellcessor, 2016, S. 12). Vor diesem Hintergrund sind auch akustische Interfaces in ihrer Materialität und technologischen Struktur mit Praktiken verbunden, die gewisse Körperlichkeiten einbeziehen, adressieren oder ausschließen. Dabei verweist die Diskussion um die Zugänglichkeit von Mainstreaminterfaces auch generell auf die Frage, was assistive Technologien (Macele et al., 2025) sind, die teils unter der Bezeichnung „alternative Benutzerschnittstellen“ (Weber, 2015) firmieren, und wie bzw. ob sie sich von Mainstreamtechnologien unterscheiden lassen.

Spracherzeugende Assistenten waren und sind von großem Interesse, sowohl im Alltag, in der Popkultur als auch in der Forschung. Weniger Beachtung als smarte Anwendungen finden dabei etwa Technologien, die bei nonverbal-kommunizierenden Menschen oder spracheingeschränkten Personen zum Einsatz kommen können. Für Letztere ist etwa Jonathan Sterne Buch *Diminished Faculties* (Sterne, 2021) relevant, das für eine politische Phänomenologie der Einschränkung („impairment“) argumentiert. Sterne verknüpft Medienwissenschaft und Disability Studies, wobei drei „verminderte“ Vermögen („faculties“) – Sprach-, Hör- sowie verkörpertes Wahrnehmungsvermögen – im Zusammenhang persönlicher Erfahrungen und aus akademischer Perspektive problematisiert werden (Stock, 2022a). Im Kapitel „Degrees of Muteness“ erläutert Sterne Erfahrungen seiner eigenen Krebserkrankung und damit verbundene Veränderungen der Stimme (Sterne, 2021, S. 39) entlang der Trias von Behinderung, Einschränkung und Krankheit. Zudem diskutiert Sterne mobile Sprachverstärker, die er wegen der Krebsbehandlung über einen gewissen Zeitraum verwendete: In der „weird gray area of voice and techne“ (Sterne, 2021, S. 43) changiert der „Spokesman“ zwischen Assistenztechnologie, (vernachlässigtem) Designobjekt und Mainstreamtechnologie. Das Sprechen mittels dieses

⁹Die bei Schabacher als Workaround beschriebene Umgehungslösung wird bei Petra Löffler als „produktive Umwegigkeit“ bezeichnet (Löffler, 2017, S. 139).

akustischen Interface – zu Vocodern und künstlichem Kehlkopf siehe Mills (2010) – distribuiert die Stimme in spezifischer Form am Kopf/Mund wie auch bezüglich des Mini-lautsprechers am Körper (Sterne, 2021, S. 47), woraus insbesondere bei wissenschaftlichen Vorträgen eine performative Qualität resultiert, die gewissermaßen als Metakommentar zu Fragen der Sound Studies und Medialität der Stimme verstanden werden kann (Sterne, 2021, S. 55).

Eine andere Form des akustischen Interface hinsichtlich der Stimme bzw. sprachbasierter kommunikativer Praktiken ist im Bereich der unterstützten Kommunikation zu finden. Dort fungiert der Talker – ein elektronischer Sprachcomputer – als Interface zwischen nonverbaler Person und lautsprachlich kommunizierenden Menschen. Mit der Verwendung eines Talkers verbinden sich spezifische Körpertechniken und Formen der sozialen Interaktion, wie Andreas Wagenknechts (2024) empirisch-theoretische Untersuchung herausarbeitet. Nonverbal kommunizierende Personen nutzen Sprachcomputer, um sich auszudrücken. Per Knopfdruck und spezieller Software können die Geräte in Tabletform Wörter und Sätze durch Sprachsynthese ausgeben, die Alltagskommunikation erleichtern und somit für mehr Lebensqualität und Selbstbestimmtheit sorgen. Dabei ändert sich die Zeitlichkeit mündlicher Kommunikation (Paterson & Hughes, 1999, S. 606): Denn Sätze müssen zuerst durch Eingabe der gewünschten Informationen gebildet und dann maschinell vorgelesen werden. Die kommunikative Struktur lautsprachlicher Unterhaltungen verlangsamt und verschiebt sich hinsichtlich der sequenziellen und thematischen Struktur, etwa weil Gesprächswendungen, spontane Korrekturen oder Reden mehrerer Gesprächspartner:innen in ein anderes Tempo verlagert werden (Niediek, 2022).

Eine andere Facette von mediengesättigter bzw. Interface-basierter Kommunikation betrifft das Hören von Sprache, wobei hinsichtlich der hier diskutierten Problematik der akustischen Interfaces Hörgeräte im Mittelpunkt stehen. Hörgeräte haben eine wechselvolle Technikgeschichte (Hüls, 1999) und einen ambivalenten Status in der Deaf History (Werner, 2024), da sie an der Ausdifferenzierung von Schwerhörigkeit und Gehörlosigkeit bzw. Taubsein mitgewirkt haben (Ladd, 2019). Hörgeräte sind vielfach bezüglich ihrer Sichtbarkeit und Gestaltung untersucht worden (z. B. Zdrodowska, 2023). Eine Rolle spielt dabei bislang weniger, wie diese Geräte als akustische Mittler fungieren. Vielmehr geht es um die Sichtbarkeit von Behinderung und wie diese durch Miniaturisierung (Mills, 2011) – im Kontext der Einführung von Transistoren oder digitaler Technologien – oder Designpraktiken (Weber, 2010) zum Verschwinden gebracht werden soll. Der sozialen Stigmatisierung von Hörverlust wurde entsprechend mit „invisible aids“ (Mills, 2011, S. 38) begegnet.

Von Beginn aber sind Hörgeräte auch im Zusammenhang mit einer Logik der „Verschaltung“ (Kittler, 1988, S. 348) zu verorten, oder, um es präziser zu fassen, als „Technologien relationaler Verschaltung“ (Hörl & Ochsner, 2023) zu verstehen. Sie wirken als technische Mittler bzw. „Operatoren“ in einem medizinisch-technischen Wissensbereich, in dem sich um 1900 das Normalhören und die Schwerhörigkeit ausdifferenzieren, wobei sie eine (verlorene) akustische Verbindung schwerhöriger Personen zur hörenden Mehrheitsgesellschaft – in Kultur, Bildung und Arbeit sowie im Privaten – durch „Einzel-“ oder