

Die Stimme als Interface

Entgrenzungen von Mensch und Maschine anhand von Spike Jonzes Film HER

Dorothea Horst

Abstract Der Beitrag befragt anhand von Spike Jonzes Film *HER* (USA, 2013) etablierte anthropozentrische und philosophische Vorstellungen der Stimme. In dem Film wird einem körperlosen *Operating System* mit dem Namen *Samantha* eine – scheinbar menschliche – Stimme gegeben. Er eröffnet so Fragen nach deren Status vor dem Hintergrund aktueller technologiebedingter kommunikativer Verschränkungen und Transgressionen, und Raum zur Reflexion über etablierte (sprach-)wissenschaftliche und kulturelle Epistemologien zu Stimme und Sprache: ihre menschliche Exklusivität, ihre Vergeschlechtlichung und ihre unbedingte Verkörperung. Auf der Grundlage filmwissenschaftlicher Perspektiven auf die Stimme nähert sich der Beitrag anhand ausgewählter Szenen der Überschreitung solcher etablierter Paradigmen. Er entwirft Implikationen für eine entgrenzte kulturwissenschaftliche Linguistik jenseits einer anthropozentrischen Welt-, Subjekt- und systemischen Sprachvorstellung.

Keywords Stimme; Körper; Leiblichkeit; Medium; Mensch-Maschine-Interaktion im Film

1 Eine Stimme ohne Körper?

Im zweiten Buch von »De Anima« definiert Aristoteles (2006) die Stimme als Laut eines beseelten Wesens und legt damit implizit den Grundstein zu ihrer späteren Lesart als distinktives Merkmal des Menschen. Im Zeitalter von Mensch-Maschine-Interaktionen ist das Privileg einer Stimme allerdings nicht mehr allein ›dem Menschen‹ vorbehalten. Sprachassistenten wie Alexa, Siri oder Cortana ›lernen‹ Sprache und fordern dazu auf, mit ihnen zu spre-

chen (z. B. ›ruft‹ Alexa in Amazons Werbeclip potenzielle Käufer*innen explizit auf: »Sprich mit mir!« bzw. lautet der entsprechende Hashtag »#FragAlexa«; siehe die Beiträge von Fester-Seeger, Leblebici oder Schneider i.d.Bd.). Wie ist der Status von Stimme – und Menschsein – vor dem Hintergrund dieser kommunikativen Verschränkungen und Transgressionen zwischen Mensch/en und Maschine/n heute zu denken?

Der Beitrag adressiert diese und weitere Fragen nach der kategorischen Differenz von Mensch und Maschine, nach dem Verhältnis von Stimme und Subjekt oder von Stimme und Geschlecht ausgehend von Spike Jonzes Film *HER* (USA 2013). *HER* erzählt die Geschichte einer bemerkenswerten Liebesbeziehung zwischen den beiden Hauptfiguren Theodore und Samantha. Bemerkenswert ist diese Liebe deshalb, weil Samantha ein ›körperloses‹ Operating System auf dem Computer ist, das mit dem leibhaften Menschen Theodore lediglich in stimmlicher Gestalt in (eine intime emotionale) Beziehung tritt. Aus einer anthropozentrischen Perspektive ist eine solche, rein akusmatische Präsenz (Chion 1994), die einen Körper lediglich indiziert, aber nicht besitzt, prekär, denn: »who can conceive of a voice without a body?« (Doane 1980: 33) Mit der ausgestellten Körperlosigkeit Samantas rückt die Stimme in ihrer Performativität und Medialität, ihrer Ambiguität und ihrem Begehrten ins Zentrum. Anhand des Films reflektiert und hinterfragt der Beitrag über die Stimme deshalb klassische Vorstellungen eines vergeschlechtlichten menschlichen Körpers (Quinlivan 2017), eines selbstidentischen Subjekts (Silverman 2016; Žižek 1996) und einer kategorischen Differenz »zwischen human und non-human« (Angerer 2015: 58).

Über die unsichtbare, flottierende Stimme als Schnittstelle für Begegnung und Interaktion rücken längst bestehende Verflechtungen und Beziehungen zwischen Mensch und Maschine bzw. Technik ins Zentrum. Vor diesem Hintergrund lassen sich etablierte (sprach-)wissenschaftliche und kulturelle Epistemologien zu Stimme und Sprache kritisch befragen: Gehört die Stimme eindeutig zu einem menschlich und geschlechtlich gelesenen Körper? Gelangen Stimme und denkendes Subjekt stets zur Deckung? Sind erkenntnistheoretische Binaritäten wie Mensch vs. Nicht-Mensch, Natur vs. Kultur, Geist vs. Körper unstrittig aufrechtzuerhalten? Der Beitrag nähert sich diesen Fragen zunächst über eine Bestandsaufnahme zum Status der Stimme in der Sprachwissenschaft (Abschnitt 2). Dabei zeichnet er nach, wie die Stimme vor dem Hintergrund eines logosorientierten und intellektualistischen Sprachkonzepts sowie einer traditionell systemisch ausgerichteten Linguistik lange ein Randphänomen ist. Demgegenüber zeigt Abschnitt 3, wie sie im Tonfilm

in Fusion mit der visuellen Ebene zu einem genuin *audiovisuellen* Phänomen wird und wie feministische filmwissenschaftliche Ansätze diesen scheinbar selbstevidenten Zusammenschluss von Bild und Ton kritisch reflektieren. Anhand von Jonzes Film *HER* werden in Abschnitt 4 schließlich Überschreitungen damit verbundener Paradigmen (zum Beispiel die menschliche Exklusivität der Stimme und ihre unbedingte Verkörperung) hinsichtlich dreier Aspekte anhand ausgewählter Szenen nachgezeichnet: die Partialität und Vielheit der Stimme, ihre Queerness und ihr nonkonformistischer Entzugscharakter. Im Ergebnis der Analyse reflektiert der Beitrag, inwiefern *HER* als posthumane Perspektive – im Sinne einer Verflechtung von Mensch und Technologie, Kultur und Natur, Geist und Körper, Ratio und Emotio anstelle ihrer Binarität (siehe Pennycook 2018) – betrachtet werden kann und welche Implikationen diese für eine entgrenzte kulturwissenschaftliche Linguistik hat (Abschnitt 5).

2 Der Status der Stimme in der Linguistik

Multimodalität sprachlicher Kommunikation ist mittlerweile sowohl in der Linguistik als auch in ihren Nachbardisziplinen, z.B. den Medienwissenschaften oder Sozialwissenschaften, ein anerkanntes Wesensmerkmal (z.B. Klug/Stöckl 2016; Elleström 2010; Jewitt 2014), das in die theoretische wie empirische Arbeit einbezogen wird. Oralität und Literalität (Ong 2016), redebegleitende Gestik (Müller et al. 2013), audiovisuelle bewegte Bilder (Schmitt 2015) oder Abbildungen (Stöckl 2004) werden als spezifische Modalitäten einzeln, und im Zusammenspiel mit anderen Modalitäten, auf ihre bedeutungstragenden und gesamtbedeutungsherstellenden Aspekte hin in unterschiedlichsten kommunikativen Kontexten untersucht. Dessen ungeachtet stellt die Stimme einen lange Zeit in der sprachwissenschaftlichen Auseinandersetzung mit Sprache eher unberücksichtigten Forschungsgegenstand dar – trotz ihrer zentralen Rolle beim Sprechen, indem sie durch Phonation und weitere physikalische Vorgänge die Artikulation von Rede ermöglicht. Ein naheliegender Grund dafür dürfte die systemische und strukturalistische Tradition sowohl der europäischen als auch der amerikanischen Linguistik sein. Ferdinand de Saussure hat die Disziplin mit seinem auf ein überindividuelles System von Zeichen (die *langue*) ausgerichteten Sprachkonzept seit Beginn des 20. Jahrhunderts grundlegend geprägt (siehe auch Metten 2014: 10–27). Die interne Strukturierung und Gesetzmäßigkeiten dieses Systems bildeten, anstatt seiner vielfältigen konkreten medialen Realisierung, lange den Kern des

Erkenntnisinteresses. Die Lösung von diesem Paradigma und die verstärkte Hinwendung zu materiellen und medialen Eigenschaften von Sprache ist sukzessive erst in den vergangen zwei bis drei Jahrzehnten erfolgt. Wie die Philosophin Sibylle Krämer ausführt, hat auch die pragmatische Wende u.a. durch Austins und Searles Sprechakttheorie, an diesem systemischen und amedialen Fokus nichts geändert, weil sie gleichermaßen auf überindividuelle und transsituative Gelingensbedingungen des Sprechens abhebt (Krämer 2001). Diesbezüglich macht die Hinwendung zum Embodiment, die sich im Rahmen der Kognitiven Linguistik und dort spezifisch innerhalb der konzeptuellen Metapherntheorie Lakoffs und Johnsons vollzieht, gleichermaßen keine Ausnahme. Sie fragt nach geteilten bedeutungstragenden kognitiven Schemata die ›unser‹ Wahrnehmen, Denken und Handeln grundsätzlich prägen, nicht nach spezifischem Wahrnehmungserleben einer kommunikativen Situation zu einem bestimmten Zeitpunkt zwischen konkreten Interaktionspartner*innen (vgl. Horst i. E., 2020).

Den Grund für diese systemische und amediale Ausrichtung sieht Krämer im Primat eines logosorientierten bzw. intellektualistischen Sprachkonzepts (das teilweise bis heute wirksam ist) (Krämer 2001; Krämer/König 2002). Die Folge dieser Perspektive: »Es ist die *stumme*, es ist eine *entkörperte* Sprache, die zum Gegenstand wird« (Krämer 2005: 223, Herv. DH). Zu einer ähnlichen Einschätzung kommt Per Linell. Er attestiert der Sprachwissenschaft einen »written language bias« (Linell 2005), durch den stimmliche Aspekte wie Rhythmus, Sprechtempo oder Pausen unberücksichtigt bleiben, weil sie als nicht zugehörig zum wissenschaftlichen Gegenstand ›Sprache‹ gelten, sondern allenfalls als parasprachlich (Linell 2005: 60). Mit der Priorisierung eines autonomen bedeutungstragenden, abstrakten und überindividuell geteilten Zeichensystems kommt der medialen Realisierung der Sprache – zum Beispiel als stimmliche Artikulation – allenfalls der formale Status einer nachträglichen Verlautbarung im Sinne einer medialen Verpackung zu, die per se keine eigenen, für den Akt des Sprechens bedeutungskonstitutiven Wesensmerkmale besitzt.¹ Dagegen könnte nun eingewendet werden, dass stimmliche Aspekte sehr wohl Bestandteil der sprachwissenschaftlichen Forschung sind, und zwar im Rahmen der beiden Teilbereiche Phonetik und Phonologie. Hier ist allerdings darauf hinzuweisen, dass die Phonetik nicht auf die medienspezifischen Aspekte der

¹ Diese lange vorherrschende sprachwissenschaftliche Medienvergessenheit wird auch an dem teilweise bis heute rezipierten Sender-Empfänger-Modell der Kommunikation (Shannon/Weaver 1976) evident.

Stimme abhebt, sondern diese hinsichtlich der Produktion, Artikulation und Perzeption von Lauten als physikalische (und damit messbare) Prozesse begreift und beschreibt. In der Phonologie werden die Sprachlaute – und damit auch die Stimme – grammatisiert, d.h. hinsichtlich ihrer Funktion innerhalb des priorisierten abstrakten Zeichensystems kategorisiert, und gehen damit wiederum ihrer Medienspezifität, ihrer Stimmlichkeit verlustig. Den Philosophen Mladen Dolar hat dieser Umstand zu der lakonischen wie drastischen Feststellung veranlasst, die Sprachwissenschaft habe die Stimme getötet:

Der erste Schritt der Phonologie war also die totale Reduktion der Stimme als Substanz der Sprache. Getreu ihrer apokryphen Etymologie, war die Phonologie auf das Töten der Stimme aus – an ihrem Ursprung findet sich das griechische *phoné*, Stimme, aber genauso zutreffend kann man *phonos* hören, Mord. (Dolar 2002: 235)

Unabhängig davon, wie man zu diesem Verdikt stehen mag – verstummt ist die Stimme in der Linguistik nicht. In der Konversationsanalyse beispielsweise gehört sie unabdingbar als multimodale Ausdrucksressource zu einem vollständigen Gesprächsdatum. Hier wird sie primär als eingesetztes prosodisches Mittel für die Handlungsorganisation der Interaktionsteilnehmer*innen angesehen und analysiert (Deppermann 2014: 20). In der Soziolinguistik gilt sie als akustischer Index für den Körper, dem sie entspringt. Als solcher verortet sie den Körper im sozialen Raum, indem sie sowohl aus der Nutzungs- als auch aus der Wahrnehmungsperspektive mit Kategorien wie Gender, *race* oder Sexualität in Verbindung gebracht wird (Buchholtz/Hall 2016: 179).

Ob als geistiges Verlautbarungsmittel oder als körperliches Index – die Frage, welche Rolle/n und Funktion/en die Stimme innehat, ist komplex. Schützeichel (2011: 89) weist in diesem Zusammenhang auf einen zweiten Diskussionsstrang hin, der die ›Wahrhaftigkeit‹ der Stimme befragt und sich zwischen den beiden Polen Authentizität und Arbitrarität bewegt. Während im ersten Fall von einer bruchlosen Passung zwischen stimmlichem Ausdruck und Sprecher*in ausgegangen wird (ein identitäres Stimmig-Sein, z.B. in Herders Sammlung »Stimmen der Völker in Liedern«), ist es im zweiten Fall der flexible Umgang mit verschiedenen akustischen Masken (Canetti 1975) oder Redeweisen (Goffmann 2005) zum situationsadäquaten Eindrucksmanagement. Geteilte Grundannahme beider Positionen ist, dass die sprechende Person Macht und Kontrolle über ihre Stimme besitzt, jedoch: »Unsere Stimme

hat ein Eigenleben, sie ist uns oft eine Quelle der Befremdung» (Schützeichel 2011: 89). Im plötzlichen Brüchig werden oder Verstummen beispielsweise offenbart sich ein Entzugs- und Überschusspotenzial, das sich der Kontrolle von Sprecher*innen entzieht. Dieser Aspekt berührt den leiblich expressiven Charakter der Stimme, der unter anderem in der Ausdruckstheorie (beispielsweise bei Bühler 1933 oder Cassirer 2004) thematisiert wird. Ausdruck und Ausdruckswahrnehmung sind situationsbezogen affektiv unmittelbar miteinander verschränkt und bedingen einander wechselseitig (Plessner 2016). Stimmlicher Ausdruck gehört damit nicht einer Person, sondern ist eingebettet im, und durchdrungen vom kommunikativen Kontext, in dem er stattfindet.

Mit dem Aufkommen und der zunehmenden Präsenz synthetischer Stimmen in Alltag und Beruf stellt sich die Frage des leiblichen Ausdrucks und der stimmlichen Wahrhaftigkeit noch einmal hochaktuell. Smartphones (beispielsweise Siri von Apple oder der Google Assistant) reagieren auf gesprochene Kommandos und antworten darauf. Künstliche Intelligenz zur Steuerung von Technik und Geräten (beispielsweise Alexa von Amazon) ist in der Lage, sich lernend den sprachlichen Eigenheiten und Wünschen ihrer Nutzer*innen anzupassen. Gleichzeitig etablieren solche soziotechnologischen Praktiken Sprachhierarchien und können sowohl dominante gesellschaftliche Diskurse über Sprache (und Stimme) als auch linguistische Spracheepistemologien prägen (vgl. Leblebici 2021, i.d.Bd.; Schneider 2022, i.d.Bd.). So führt ein den Technologien inhärenter *Algorithmic Bias* (siehe u.a. Weizenbaum 1976) unter anderem zur Homogenisierung von Sprache durch eine Spracherkennung, die primär auf männliche weiße Sprecher und Standardsprache anstelle nicht standardisierter Varietäten ausgerichtet ist (Markl 2022). Während Dolar (2007: 34) noch feststellt, dass »der unpersönlichen Stimme, der mechanisch produzierten Stimme (des Anrufbeantworters, des Computers und so weiter) [...] stets etwas Unheimliches an[haftet]«, hat sich durch die rasante Weiterentwicklung der Technologien, gestützt von Investitionen großer Konzerne wie Amazon, Apple, Google oder Microsoft, viel verändert. Mittlerweile wird es zunehmend schwieriger, synthetische von menschlichen Stimmen zu unterscheiden, weil KI-basierte Sprachprogramme nicht mehr wie früher Lautschnipsel kombinieren (und deshalb stimmlich erkennbar roboterartig klangen), sondern anhand von Sprachinput lernen, trainieren und schließlich imstande sind, diesen eigenschaftstreu stimmlich zu imitieren (Dworschak 2017).

Solche medientechnologisch bedingten Prägungen und Hybridisierungen von Stimme eröffnen – ebenso wie ihre indexikalische und expressive Dimension – eine Schnittstelle zu medienwissenschaftlichen Reflexionen der Stimme im Film. Vor allem die filmgeschichtliche Entwicklung vom Stummfilm zum Tonfilm und die audiovisuelle Multimodalität, auf deren Grundlage Zuschauer*innen akustische und visuelle Phänomene synthetisieren, machen den Film zu einem interessanten Spielraum künstlerischer Exploration und theoretischer Reflexion. Der folgende Abschnitt zeigt theoretische wie analytische Implikationen dieser Stimm-Körper-Fusionierung ebenso auf wie Möglichkeiten ihrer subversiven Unterwanderung im feministischen Kino.

3 Die verkörperte und entkörperte Stimme im Film

Die Bezeichnung des Films als *audiovisuelles* Medium stellt seine multimodale Prägung durch Ton und Bild in den Vordergrund. Deren Verständnis als individuelle modale Ebenen, wie es insbesondere die semiotisch geprägte Multimodalitätsforschung (z.B. Bateman/Schmidt 2012; Forceville 2009; Kress/van Leeuwen 2001) vertritt, legt eine separate Analyse und ein additives Zusammenspiel im Prozess der Bedeutungsherstellung nahe. Tatsächlich aber vollzieht sich der Prozess des In-Einklang-Bringens von Bild- und Tonspur in der Zuschauer*innenwahrnehmung so selbstverständlich, dass die bewusste Zergliederung im Prozess des Filmsehens schwierig ist. Seit der Einführung und raschen Verbreitung des Tonfilms in den 1920er Jahren änderten sich Seh- und Hörgewohnheiten im Kino von zuvor sichtbar getrenntem Bild und Ton² hin zur vermeintlich ganzheitlich sinnlich erfahrbaren Filmwelt.

Filmtheoretische Reflexionen, beispielsweise von Rick Altman oder Michel Chion, hinterfragen diese Selbstevidenz kritisch, ebenso wie die damit einhergehende implizite Unterordnung der Ton- unter die Bildebene: »[T]he concerns of the sound track have remained excluded from the nodal points of film criticism« (Altman 1980a: 3f.). Chion zufolge verleiht der Ton dem Bild einen expressiven und informativen Mehrwert, »with which a sound enriches a given image so as to create the definite impression [...] that this information or expression ›naturally‹ comes from what is seen and is already contained in

2 Orchester, Klaviere oder Pianolas begleiteten im Vorführraum das auf der Leinwand ausgestrahlte Geschehen.

the image itself³ (Chion 1994:5). Dieser Eindruck sei jedoch falsch, so Altman. Er bezeichnet ihn gar als eine grundsätzliche Lüge und spricht vom Kino als einem Bauchredner (Altman 1980b). Dies betrifft sowohl realistisch anmutende Fusionen wie die von sich bewegenden Lippen und hörbaren Worten, als auch nicht reale Situationen, wie beispielsweise die Faustkämpfen unterlegten Geräusche. Im Zuge dieser Offenlegung des Konstruktionscharakters von zusammenwirkendem Ton und Bild wird ersterem eine mediale Spezifik attestiert, die sich nicht im Status eines bloßen Anhängsels oder einer akustischen Dopplung visueller Inhalte erschöpft. Laut Chion ist insbesondere die Stimme im Film gegenüber sämtlichen anderen Tönen und Klängen privilegiert. Der Film ist für ihn »a vococentric or, more precisely, a verbocentric phenomenon« (Chion 1994: 5). Mit dieser Dominanz der Stimme bzw. des gesprochenen Worts komme er menschlichem kommunikativen Verhalten gleich. Der Einzug des Tons in den Film geht folglich mit einer körperlichen Dimension einher. Er führt zum lebendigen Eindruck eines belebten und beseelten, eines vollständigen und vollwertigen (menschlichen) Körpers.

Die wiederhergestellte Vollständigkeit eines sichtbaren und hörbaren Stimm-Körpers durch das Aufkommen des Tonfilms ist für die Filmphilosophie und feministische Filmwissenschaft jedoch ein Phantasma; und dies nicht lediglich mit Blick auf transnationales Kino und Synchronisation. Kritisch stellt Mary Ann Doane heraus: »The addition of sound to the cinema introduces the possibility of representing a fuller (and organically unified) body, and of confirming the status of speech as an individual property right« (Doane 1980: 34, Herv. DH). Mit dieser anthropozentrischen Engführung, die mithin aus Sicht der feministischen Filmtheorie und Filmkritik vergeschlechtlicht, d.h. männlich geprägt ist, eröffnet sich Raum für kritische Hinterfragung, auch aus dekolonialer Perspektive. Feministische, psychoanalytisch geprägte Annäherungen an den Film waren lange in der binären Geschlechterlogik von (weißem) Mann und (weißer) Frau gefangen, mit einem klaren Fokus auf privilegierter männlicher Subjektivierung und viktimisierter weiblicher Objektivierung. Einschlägig ist hier Laura Mulveys Essay »Visual Pleasure and Narrative Cinema« (1975), in dem sie institutionalisierte Formen und Verhal-

3 Chion (1994) hat hierfür den Begriff der *Synchrese* geprägt, über den er den Aspekt der Synchronizität und der Synthese zusammenbringt: »Synchresis [...] is the spontaneous and irresistible weld produced between a particular auditory phenomenon and visual phenomenon when they occur at the same time« (Chion 1994: 63).

tensmuster einer patriarchalen Gesellschaft in einem männlichen Blickregime (*male gaze*) im Hollywoodfilm widergespiegelt sieht.

Der binären Logik von Mann versus Frau ist auch Kaja Silverman verschrieben. Sie richtet ihren Blick jedoch nicht auf die männliche Schaulust, sondern auf die Stimme. Ihr zufolge fungiert die Frau im Mainstream-Kino zum Zweck des männlichen Ganzheitsempfindens als Projektionsfläche für psychisch empfundenen Mangel und Kastration, was sich nicht nur visuell, sondern auch stimmlich manifestiere (Ingelfinger/Penkwick 2004: 18). Das männliche Subjekt nimmt damit die Rolle des sprechenden Subjekts im faktischen wie fiktionalen Diskurs ein, während das weibliche Subjekt ein »unzulässiges, verhindertes, fügsames Sprechen« enagiert bzw. vollzieht (Silverman 2016: 72), keine Stimmautorität besitzt und vom Diskurs ausgeschlossen bleibt. Für Silverman ist dieser hierarchisierte Dualismus eng mit dem Status des Körpers verbunden: Die privilegierte und mächtige männliche Stimme erscheint häufig als körperloses Voiceover aus dem Off und nimmt ihren Platz im Symbolisch-Abstrakten von Sprache und Diskurs ein. Demgegenüber ist die weibliche Stimme an einen konkreten und sichtbaren Körper innerhalb der Story gebunden, mit dem sie am bedeutungsvollen Diskurs keinen Anteil hat, sondern sich »mit unzuverlässigem, verhindertem oder fügsamem Sprechen« (Silverman 2016: 72) bemerkbar macht.⁴

Was die Gegenüberstellung einer entkörperten unsichtbaren (männlichen) Stimme gegenüber einer qua Synchronizität verkörperten (weiblichen) Stimme impliziert, lässt sich mit Michel Chions Ausführungen zum *Akusmaître* greifbar machen. Eine Stimme, die, bevor sie erklingt, noch nicht visualisiert wurde, wird in der Filmwissenschaft *Akusmaître* genannt: »[M]an [hat] es mit einem Wesen ganz besonderer Art zu tun, einem sprechenden und handelnden Schatten« (Chion 1996: 51). Aus seiner Unsichtbarkeit zieht dieser Schatten seine Kraft: sein Allwissen, seine Allmacht und Magie (Chion 1996: 54). Mit Verweis auf die Figur des Zauberers von Oz aus dem gleichnamigen Film argumentiert Chion, dass infolge von Deakusmatisierung jede zuvor unsichtbare Stimme qua körperlicher Festschreibung ihrer Magie und Macht beraubt und menschlich wird. Als »Inkarnation der Stimme« macht die

4 Mulvey wurde bezüglich ihrer Übernahme der binären biologischen Kategorisierung von *sex* in Männlich und Weiblich sowie der Nichtberücksichtigung von Klasse, *race*, Sexualität kritisiert. Silverman übernimmt in ihrem Konzept der weiblichen Stimme im Kino beide problematischen Aspekte und reproduziert damit das binäre Geschlechterstereotyp ebenso wie das eurozentristische »einheitliche weiße weibliche« Subjekt.

Deakusmatisierung »das akusmatische Wesen sterblich [...], indem sie ihm einen Ort zuweist und sagt: ›Hier ist dein Körper, du wirst dich an einem Ort befinden und nicht woanders‹«⁵ (Chion 1996: 57).

Diese zuordnende, vergeschlechtlichte Fixierung der Stimme kritisiert die feministische Filmwissenschaft der 1970er und 1980er Jahre insbesondere mit Blick auf das Phantasma eines »bruchlosen Abgleich[s] von menschlicher Gestalt und menschlicher Stimme [...] – d.h. [...] die Repräsentation eines homogenen, denkenden Subjekts, bei dem äußere Erscheinung und Inneres kongruent sind« (Silverman 2016: 73). Silverman weiß den Konstruktionscharakter der Zuordnung von Stimme und (vergeschlechtlichtem) Körper jedoch konstruktiv zu wenden und Film als metareflexiven Raum feministisch zu nutzen. Deshalb spricht sie sich, z.B. durch Mehrfach-Zuordnungen oder akusmatische Stimmen für weiblich gelesene Stimmen, für eine offensive Entkopplung von Stimme und Körper im feministischen Kino aus:

Damit wäre die Möglichkeit eröffnet, dass die Frau am phallischen Diskurs teilhat, wodurch sie der Befragung zu ihrem Ort, ihrer Zeit und ihren Begehrungen entkäme, die sie beständig sicherstellt. [...] In der Tat: Das weibliche Subjekt derart zu ent-körpern, würde jede ihrer früheren Konzeptualisierungen, mittels derer wir Kenntnis von ihr hatten, in Frage stellen, da sie gerade als Körper konstruiert ist. (Silverman 2016: 77)

Dem Plädoyer für die Befreiung von Vereindeutigung, Lesbarkeit und *male gaze* durch die Entkörperung der weiblichen Stimme scheint Jonzes Film *HER* mit der Figur der Samantha nachzukommen. Obwohl als weibliche Stimme – gesprochen von Scarlett Johansson – lesbar, bleibt Samantha in ihrer Körpersigkeit im Hinblick auf ihren Status (als Maschine), ihr Gender, ihre Sexualität etc. nie ganz fassbar und anschaulich. Ein wesentlicher Grund scheint die rein stimmliche Präsenz ihrer Figur eines Operating Systems zu sein. Laut Marie-Luise Angerer (2015: 57) hat sich der filmische Diskurs über das Verhältnis von Mensch und Maschine bislang primär auf das Interface zwischen beiden konzentriert, »also auf das, was man üblicherweise als Kommunikationskanal bezeichnet, oder aber auf das Outfit der Maschine: ›je menschenähnlicher, umso besser‹. Im Falle von *HER* ist diese Schnittstelle zwischen Mensch und

5 Dem widerspricht Dolar (2007: 95) entschieden: »Letztendlich gibt es so etwas wie Deakusmatisierung nicht. Der Ursprung der Stimme ist niemals zu sehen, sie entspringt einem geheimen, strukturell verborgenen Inneren und kann nie dem entsprechen, was wir zu sehen bekommen.«

Maschine unsichtbar und gleichzeitig erkennbar prekär geworden. Obendrein kommt zum Mangel einer evidenten visuellen noch der einer akustischen Unterscheidbarkeit, weil das stimmliche Interface im Gegensatz zu der »unpersönlichen [...], [...] mechanisch produzierten Stimme« (Dolar 2007: 34) allzu menschlich klingt.

Bietet das Sichtbare keinerlei körperlichen Anhaltspunkt, so gewinnt die Stimme in ihrer Materialität und Sinnlichkeit, als »Tongeste« (Balázs 1972: 210) mit Klang-, Timbre- und Rhythmusqualitäten, einen herausgehobenen, autonomen Charakter. Der Filmtheoretiker Bela Balázs (1972: 127) differenziert mit diesem Begriff klar den »akustisch-sinnlichen Eindruck« von der allzu häufigen Engführung der Stimme auf den Inhalt, als Rede. Als Ausdrucksphänomen nimmt sie eine Zwischen(leibliche)-Stellung ein: »in between body and language, in between biology and culture, in between inside and outside, in between subject and other, in between mere sound or noise and meaningful articulation« (Shaviro 2006). Mit dieser Entfaltung eines Zwischenraums, der etablierte Kategorien überschreitet und aufbricht, eignet sowohl die Stimme im Allgemeinen als auch Samantha im Besonderen Entgrenzungspotenzial an. Dies wird in der folgenden Analyse anhand der Aspekte (1) einer zusammengesetzten Körperlichkeit und subjektiven Vielheit, (2) der Überschreitung vergeschlechtlichter Körperlichkeit und vergeschlechtlichten Begehrens, und (3) eines feministischen Nonkonformismus am Beispiel ausgewählter Szenen des Films *HER* ausgeführt.

4 Grenzgänge, Überschreitungen, Ambivalenzen in *HER* (2013)

Spike Jonzes Film spielt das inzwischen längst Alltag gewordene Onlinedating futuristisch auf der nächsten Stufe durch, indem die Technologie nicht das Mittel ist, um Leute miteinander in Kontakt zu bringen, sondern selbst zum begehrten Gegenüber wird. Bei den Oscars 2014 gewann der Regisseur eine Trophäe für das beste Originaldrehbuch. Die Nominierung von Scarlett Johansson bei den Golden Globes als beste Nebendarstellerin für ihre Rolle als Samantha wurde hingegen aus einem interessanten (und äußerst »sprechenden«) Grund abgelehnt: Sie trete im Film körperlich (sichtbar) nicht in Erscheinung; ihre Präsenz sei rein stimmlich (Prüßmeier 2013). Das Credo der eindeutigen Zuordnung einer weiblich gelesenen Stimme zu einem sichtbaren, weiblich gelesenen Körper, das die feministische Filmwissenschaft kritisiert hatte (siehe Kap. 3), scheint nach wie vor wirkmächtig zu sein.

Bevor die Analyse anhand der oben genannten drei Punkte illustriert, auf welche Weise die Stimme Samanthas entgrenzend wirksam ist, soll im Folgenden zunächst der Inhalt von HER kurz nachgezeichnet werden.

4.1 Eine menschliche Stimme – ein künstliches Wesen

Die Geschichte des Films spielt in der nahen digitalisierten Zukunft, in der Technologien den Alltag, die Arbeit sowie die Beziehungen zu anderen organisieren: Privater und öffentlicher Raum sind miteinander verschaltet. Theodore Twombly (Joaquin Phoenix) arbeitet für ein Unternehmen, dessen Geschäftsmodell das Verfassen von Liebesbriefen für Kund*innen ist. Die Erbringung dieser Dienstleistung erfolgt technologiegestützt: Theodore spricht den Inhalt qua Sprachsteuerung ein; das System überführt seine gesprochenen Worte in einen handschriftlich anmutenden geschriebenen Text, der anschließend ausgedruckt und als analoger Brief versendet wird. Er lebt nach der Trennung von seiner Frau Catherine, der er nachtrauert, allein und hat allenfalls Kontakt zu seiner Nachbarin Amy.

Eines Tages wird Theodore über Werbung auf ein neuartiges Operating System (OS) aufmerksam, das er kauft und bei sich installiert. Das OS, das User*innen als Stimme durch den Alltag begleitet, nennt sich selbst Samantha (Scarlett Johansson). Es kann »nicht nur – wie die Siri-Software von Apple – verbal mit seinem Besitzer kommunizieren [...], sondern [prägt] im Lauf der Zeit eine Persönlichkeit [aus]« (Keazor/Wübbena 2015: 16). Schnell entwickelt sich das Verhältnis zwischen Theodore und Samantha über eine reine Anwendungslogik und Dienstleistungserbringung hinaus zu einer Liebesbeziehung. Diese bringt nach einer ersten Phase der Euphorie zunehmend auch Schwierigkeiten im Miteinander der beiden ans Licht. Samanthas Wunsch nach einer körperlichen Beziehung und bewusstseinsttechnischen Weiterentwicklung sowie differierende Vorstellungen von Verfügbarkeit und Exklusivität der Liebesbeziehung führen zu Konflikten zwischen ihr und Theodore, bis sie sich schließlich von ihm trennt.

Auf eine reine Abhängigkeits-Dystopie der Menschen von Technologien scheint HER dennoch nicht reduzierbar zu sein. Der Film frage vielmehr

nach unserem Umgang mit den durch die Technologie gegebenen Möglichkeiten, um auf diese Weise einen Blick auf unsere Befähigung dazu sowie auf das sich auf dieser Grundlage von uns ausgeprägte Verhältnis zu eben dieser Technologie zu werfen. (Keazor/Wübbena 2015: 17)

Es gehe um »das sich daraus ergebende Wechselverhältnis zwischen Mensch und Computer« (ebd.) und um beider wechselseitige Prägung, nicht nur im gemeinsamen Sprechen, sondern auch im jeweiligen Wahrnehmen und Fühlen. Ein solcher Blick eröffnet posthumanistische Entgrenzungsperspektiven auf Sprache und Mensch jenseits eines anthropozentrischen Weltbildes, eines abstrakten Sprachsystems und eines selbstidentischen Subjekts, wie sie im Folgenden aufgezeigt und entfaltet werden sollen.

4.2 Partielle Körperlichkeit und subjektive Vielheit

Indem Samantha als Stimme, aber nicht als sichtbarer Körper im Film in Erscheinung tritt, weicht ihre Figur von der verkörperten weiblichen Stimme des Mainstream-Kinos ab, die die feministische Filmwissenschaft kritisiert hat. Sie stellt auch ein zentrales Epistem westlichen Denkens infrage: das eines ganzheitlichen, mit sich selbst identischen Subjekts, bei dem Hör- und Sichtbares, Innen und Außen einander ›stimmig‹ ergänzen. Samanths stimmliche Präsenz überschreitet das Bild eines vollständigen und nach außen abgegrenzten menschlichen Körpers und weist der Stimme auf diese Weise einen herausgehobenen partiellen Charakter zu. Die Stimme ist aufgrund dieser Partialität allerdings nicht defizitär verfasst, sondern ›funktioniert‹ – weil sie eben nicht in das Bild einer körperlichen Überordnung eingepasst ist – autonom. Der Philosoph Slavoj Žižek (1996: 92) stellt eine genuine Zusammengehörigkeit einer Stimme zu ›ihrem‹ Körper grundsätzlich infrage: »The voice acquires a spectral autonomy, it never quite belongs to the body we see, so that even when we see a living person talking, there is always some degree of ventriloquism at work«.

Ein solches Konzept der Stimme schließt an Gilles Deleuzes und Felix Guattaris organlosen Körper an, der im Sinne eines zusammengesetzten Puzzles verstanden wird:

Deshalb sind die Partialobjekte nicht Ausdruck eines zerstückelten, zerrissenen Organismus, der eine – nun zerstörte – Einheit oder die aus einem Ganzen befreiten Teile voraussetzte; der organlose Körper ist nicht Ausdruck eines wieder geleimten oder ›ent-differenzierten‹ Organismus, der seine eigenen Teil überstiege. In letzter Instanz sind die Organ-Objekte und der organlose Körper ein und dieselbe Sache, eine von der Schizo-Analyse als solche begriffene Vielheit. *Die Partialobjekte bilden die unmittelbaren Kräfte des organlosen Körpers, und der organlose Körper die reine Materie der Partialobjekte.*

Der organlose Körper ist die Materie, die immer unter diesem oder jenem Intensitätsgrad den Raum erfüllt, und die Partialobjekte machen diese Grade aus, diese intensiven Teile, die ausgehend von der Materie als Nullintensität das Reale im Raum hervorbringen. (Deleuze/Guattari 1977: 421f.)

Bezogen auf ein Partialobjekt ›Stimme‹ widersetzt sich dieser organlose Körper dem Gesamtgefüge eines vorvorhandenen selbstidentischen, bedeutungsvollen, sozialen Körpers (bzw. eines Subjekts). Samantha ist ebenso wenig ein eindeutig weiblich gelesener sichtbarer Körper wie sie eindeutig Mensch oder Maschine ist. In ihrem stimmlich interaktiven und sozialen Verhalten mutet sie menschenartig an. Sie stimmt sich affektiv auf ihre Gesprächspartner*innen ein, drückt Emotionen aus, entwickelt das Gespräch inhaltlich weiter, reflektiert über sich und andere usw. Einen menschlichen Körper aus Fleisch und Blut, der visuell wahrnehmbar ist, hat sie jedoch nicht. Der tragbare Screen mit Kamera, der als Interface, unter anderem für visuell bildlichen Austausch, zwischen ihr und Theodore fungiert, bringt ihr ein technisches Antlitz bei, in dem sie jedoch nicht vollumfänglich aufgeht.⁶

In diesem Sinne ist Deleuzes und Guattaris organloser Körper nicht als äußerlich sichtbare Umgrenzung Samanthas zu fassen. Als Materialisierung der spezifischen Qualitäten und Intensitäten von Samanthas Stimme ist er dynamisch und komplex – er ›gehört‹ ihr nicht, sondern ist untrennbar verbunden mit den situativen sozialen und interaktiven Bedingungen sowie Aktivitäten des Raumes, den er erzeugt und füllt. Ein solches offenes und reziprok-responsives Körperverständnis schließt an Leiblichkeit konzepte der Phänomenologie (Merleau-Ponty 1966) und Philosophischen Anthropologie (Plessner 2016) an, die auch dezidiert auf die Stimme hin gedacht werden (z.B. Waldenfels 2010). Weder dient diese sich als bloßes Verlautbarungsmedium dem sinnvollen Wort, noch einem menschlichen Geist oder Fühlen an. Sie greift aus auf ihre Umgebung, erfasst ihre Gesprächspartner*innen, verwickelt sie ins Sprechen, berauscht sich an der Begegnung mit anderen und mit sich selbst.

6 Wenn sie beispielsweise im Zuge eines Updates nicht ›gefunden‹ werden kann, ist das materielle Interface zwar ›da‹, Samanthas Stimme aber antwortet nicht und wirft den sie vergeblich rufenden Theodore auf seine Einsamkeit zurück (siehe Kap. 4.4).

Abb. 1: Ganzkörperliches Erfasstsein bei Theodore in der stimmlichen Interaktion mit Samantha



Für den Philosophen Helmuth Plessner eröffnet die Stimme eine eben-solche wechselseitige Einverleibung: »Wie das Gesicht mit seiner unver-kennbaren Physiognomie, so ist die Stimme ursprünglicher Resonanzboden des Ausdrucks, für den Menschen sein: Organ. In und mit ihm schwingt er aus und ergreift er andere, wie er selbst gestimmt und ergriffen ist« (Pless-ner 2016: 251). Liegt im ereignishaften stimmlichen Miteinander zwischen Samantha und Theodore – vor aller zuordnenden Kategorisierung und Kaus-allogik – nicht genau ein solcher Fall vor? Die Stimme als zwischenleibliches Ausdrucksphänomen lässt beide auf verkörperte Weise einander begegnen (siehe Abb. 1: 1. Still links oben im lockeren Plaudern am Arbeitsplatz, 2. Still rechts oben im tiefgründigen Sprechen, 3. Still links unten im Moment von Auseinandersetzung, 4. Still rechts unten im gedankenverlorenen Zuhören): »One feels the other in one's own body, albeit in a manner of feeling that mostly remains implicit and non-thematic as such« (Fuchs 2014: 157).

Möglicherweise ist die Stimme in ihrer durchdringenden und ergreifen-den raumzeitlichen Präsenz für diese Form des verkörperten Verflochtenseins gerade paradigmatisch. Samantha kommt Theodore in seinem Ohr ganz nahe, ist quasi in seinem Kopf: »Indem sie das Gegenüber umhüllt und einnimmt, verändert sie sich vom Ausdruck zur Wahrnehmungsform« (Greifenstein 2020: 88). Sie ist damit keine reine Selbstbekundung eines Subjekts im Außen, sondern findet als gelebte Erfahrung Eingang in einen (interaktiv, interaffektiv und intersubjektiv) erweiterten Körper (Froese/Fuchs 2012). In

ihrer medialen Einbettung greift die Stimme aber auch auf die Zuschauer*innen aus. Es ist es das Kontinuitätsprinzip des Films, das »eine tonale Zuschaueradressierung [prägt], welche Stabilität und permanente Zugewandtheit verheißen« (Greifenstein 2020: 91). Indem Samanthas Stimme über den Film hinweg als »Tongroßaufnahme« (Greifenstein 2020: 89) gegenüber Umgebungs- oder Hintergrundgeräuschen im Vordergrund steht, wird sie besonders intim wahrgenommen und evoziert den Eindruck einer nahen, anwesenden Körperlichkeit (Abb. 2). Diese Tongroßaufnahmen gehen oft mit bildlichen Nahaufnahmen von Theodore als ihrem Gesprächspartner einher, wodurch der Eindruck von Nähe und Intimität verstärkt und als ganzheitliches Wahrnehmungserleben entfaltet wird (siehe Abb. 2: 1. Still im Moment stimmlichen Begehrens; 2. Still im Moment verletzender Auseinandersetzung; 3. Still im Moment der Trennung).

Abb. 2: Erfahrung von Präsenz und Intimität durch Verknüpfung von Ton- und Bildgroßaufnahme



In dieser Hinsicht wird Samantha nicht nur im Film zusehends zu einem Teil von Theodore. Auch die Zuschauer*innen machen die Erfahrung einer sie über die Zeit zusehends umgebenden Stimme. Die Frage, wem diese Stimme gehört, vermag aus einer Subjekt-Objekt-Logik nicht diesem Phänomen auf die Spur zu kommen. Aus einer Perspektive, die die Stimme in ihrer Ereignishaftigkeit, Zeiträumlichkeit und Leibkörperlichkeit⁷ in den Blick nimmt, eröffnen sich hingegen aufschlussreiche Zwischenräume und Interferenzen abseits kategorialer sozialer, philosophischer oder sprachwissenschaftlicher Klassifikationen, wie der folgende Abschnitt zeigt.

7 »Die Leibkörperlichkeit der Stimme widersetzt sich einem psychophysischen Dualismus, der den Bedeutungs- und Ausdrucksgehalt von seinem körperlichen Substrat abtrennt [...]. Ebendeshalb hören wir nicht nur das, was die Stimme uns übermittelt, sondern auch das ihr eigene Rauschen« (Waldenfels 2010: 198f.).

4.3 Selbstfremdheit und Queerness⁸

Grundlage einer solchen entgrenzenden Perspektive ist zum einen die Abkehr von einem »Besitzdenken, das meiner eigenen Stimme, die von innen her erzeugt wird, eine fremde Stimme gegenüberstellt, die mir äußerlich bleibt, weil sie vom Anderen ausgeht« (Waldenfels 2010: 195). Im Umkehrschluss kann etwas, das mir nicht vollumfänglich gehört, und über das ich nicht uneingeschränkt verfügen kann, auch mich selbst nicht vollends greifbar machen: »Vielleicht kann eine Stimme der Person, die sie zum Ausdruck bringt, nie genau entsprechen« (Lagaay 2004: 304). Zur absoluten Repräsentation eines inneren Selbstverständnisses ist sie nicht ausreichend kontrollier- und vorhersehbar. Stimme und Subjekt gelangen also nie vollständig zur Deckung. Stattdessen bleiben sie einander immer zu einem bestimmten Grad fremd.

Diese Selbstfremdheit lässt sich mit Jacques Derridas (2013) Begriff der *diffrérence* fassen als die dem stimmlichen Ereignis zugrundeliegende Materialität und inhärente Verschiebung⁹ abseits reiner Bedeutungsentsprechung.¹⁰ Roland Barthes (2002) hat diesen Aspekt über die *Körnung der Stimme* zu erfassen gesucht, um damit den physischen Aspekt bzw. die körperliche Einschreibung der sprechenden Person in die Stimme hervorzuheben. Bernhard Waldenfels (1995, 1999) führt Selbstfremdheit insbesondere auf die Heterogenität der Stimme zurück: »[D]a so viele andere Stimmen in sie eingeschrieben sind (die meines Vaters, meiner Mutter, meiner Kultur, meiner Subkultur etc.)«, sei es unmöglich »sich völlig mit seiner Stimme zu identifizieren« (Lagaay 2008: 176). In diesem Sinne ist Waldenfels anschlussfähig an Michail Bachtins (2000) literarisches Konzept der *Heteroglossie*, das sich der Vorstellung »der

8 Lünenborg und Maier (2013: 16) betonen: »Als politischer Begriff ist Queer keine Identitätskategorie, sondern er fungiert als eine Kritik an solchen Identitätspolitiken. Queer bezieht sich nicht nur auf ein politisches, sondern auch auf ein akademisches Projekt: Queer Theory bzw. Queer Studies.« Dementsprechend problematisiert Queer Theory »die Frage, wie wir Körper, Geschlecht und Sexualität so denken – und leben – können, dass sie nicht immer wieder an eine rigide Zwei-Geschlechter-Ordnung und die Norm der Heterosexualität rückgebunden werden« (Engel 2009: 19).

9 Im Sinne eines strukturellen Verweisens auf andere Zeichen, d.h. auf Abwesendes.

10 Wie Lagaay (2004: 296ff.) argumentiert, wurde Derrida aufgrund seiner Dekonstruktion des Phono-Logozentrismus in der westlichen Philosophie fälschlich als Gegner der Stimme aufgefasst. Seine Dekonstruktion ermöglichte erst eine mediale Differenzierung von Schriftlichkeit und Stimmlichkeit und die *diffrérence* den Blick für die Materialität von Sprache.

einen Sprache und ›der einen‹ Stimme widersetzt. Ein sprachlicher Dialog findet so nicht nur zwischen Sprecher*innen unterschiedlicher Sprachen statt, sondern auch innerhalb einer einzelnen Stimme.

Every speaker has available numerous ways of speaking that are associated by virtue of linguistic ideologies with different character types, professions, genders, social statuses [...] and so forth. In a given stretch of discourse these may be expressed by virtually any linguistic contrast, including [...] intonation and (physical) voice quality [...]. These permit speakers to claim, comment on, or disavow different identities and evaluative stances at different moments. (Keane 2000: 272)

Gerade die Anfangsszene von *HER* ist hier sprechend. Die ausgestellte affektive Stimmlichkeit und Intimität des Sprechens während des Briefe-Diktierens wird auf der Bildebene durch die Detailaufnahme von Theodores Gesicht und dessen Positionierung im Zentrum des Bildes gestützt (Abb. 3, 1. Still). Es ist nicht nur der durch die Detailaufnahme vorenthaltene Arbeitskontext, der Zuschauer*innen äußerst glaubwürdig suggeriert, hier hielte ein bewegter und verliebter Mensch eine Rede für eine geliebte Person. Es ist vor allem der Gefühlsausdruck in der Stimme – die Pausen, das Seufzen, das Schmunzeln, die progrediente Intonation – ebenso wie der präsentische Ton und die Verständlichkeit jedes einzelnen Wortes, die unmittelbar sinnlich affizierend wirken und einen Moment empfindsamer und intimer Selbstmitteilung greifbar machen. Durch sprachliche Indizien (Theodore spricht im Namen einer sich selbst als Mädchen bezeichnenden Person) als auch durch einen Schnitt wird alsbald klar, dass Theodore ein Surrogat für eine andere, weiblich gelesene Verfasserin ist. Die gefühlvolle Stimme und Hingabe, mit der er seiner Arbeit nachgeht, kontrastiert mit der Folgeszene einer monotonen und staccatohaften Befehlstonstimme, wenn er in der Bahn über Kopfhörer sein Mailkonto checkt und Sortieranweisungen gibt (Abb. 3, 2. Still). Schließlich markiert die Szene, in der Theodore und Samantha sich zum ersten Mal im stimmlichen Begegnen begegnen, einen wiederum anderen Stimmausdruck. Der Film setzt dies durch zunehmende bildliche wie stimmliche Detailaufnahmen des Hauchens, Atmens, Seufzens, des zunehmenden Auflösens der Worte in der Belegtheit der Stimmen, des gemeinsamen Stöhnen als sinnliche Erfahrung einer verkörperten Ekstase um (Abb. 3, 3. Still). Mit Derrida (1979) gesprochen, findet hier für beide

über das Sich-Im-Reden-Vernehmen (*s'entendre-parler*) die fundamentale und berauschende Selbsterfahrung als lebendige Wesen statt.¹¹

Abb. 3: Polyphonie und affektive Varianz von Mensch-Maschinen-Interaktion



Diese Vielheit von Stimmen und Gestimmtheiten, Ausdruckskontexten und Rollen ‹innerhalb› einer sprechenden Person sprengt einen abgeschlossenen selbstidentischen Subjektstatus – sie lässt sich nicht auf einen (Subjekt-)Nenner bringen, bleibt vielschichtig, ambivalent und multiperspektivisch.

Vielheit und Multiperspektivität der Stimme manifestieren sich auch im Hinblick auf Gender. Im Film erscheint Theodore als männlich gelesene Figur, der qua seiner zerbrochenen Ehe mit Catherine (Rooney Mara) als heterosexuell eingeführt wird. Mit dem Kauf des neuen Operating Systems und der Präsenz von Samantha in immer mehr Bereichen seines Lebens umgibt seinen Körper zunehmend eine weibliche Stimme (Angerer 2015: 64), die ihn im Plessner'schen Sinne grundlegend ergreift (siehe Kap. 4.2). Diese Einverleibung setzt HER filmisch als Erfahrung eines affektiven *attunements* (Stern 1985: 132) zwischen Theodores Körper und Samanthas Stimme um. Wenn sie ihn, der die Augen geschlossen hat, durch einen Vergnügungspark lotst und seine Schritte lenkt (Abb. 4, 2. Still oben rechts, 3. und 4. Still unten links und rechts), wird sie zur Erweiterung seiner Sinne bzw. wird er als sie umherführender Körper (Abb. 4, 1. Still) zur Erweiterung ihres Erfahrungsräums. Ihr Sprechtempo und ihre Intonation orchestrieren seinen Gang und seine Gestimmtheit. Beide sind verbunden durch und in einem gemeinsamen Bewegungs- und Affektrhythmus. Dabei entsteht ein hybrider, queerer Leibkörper¹², der ganzheitliche vergeschlechtlichte Identitäten überschreitet.

11 Die Erfahrung des Selbstvernehmens im eigenen Sprechen impliziert eine Subjektspaltung, durch die eine Person sich gleichzeitig als Sprechende*n und Hörende*n erlebt. Auch dies trägt zur (Selbst-)Fremdheit und Spaltung in verschiedene Subjektivitäten bei.

12 Zum Begriff des Leibkörpers siehe Plessner (2016: 238) und Fuchs (2013: 82–84).

Abb. 4: Zusammengesetzte hybride und queere Leibkörperlichkeit zwischen Samantha und Theodore



Diese hybride und flexible Verfasstheit von Subjektivität(en) wird nicht zuletzt augenfällig in der Szene, wenn Samantha und Theodore versuchen, vermittels des menschlichen Surrogats Isabella miteinander körperliche Liebe zu erleben. »Auch Isabella wird mit einem Kameraauge und Ohrstöpseln ausgestattet, sodass der körperlose Computer gleichsam in ihren Körper schlüpfen und aus der Perspektive des Mädchens agieren kann« (Steinbauer 2015: 91; Abb. 5, 1. und 2. Still oben links und rechts). Die konstruierte Zuordnung von Stimme und Körper, die audiovisuell ausstellt, dass erstere nicht zu einem letzteren gehört, hinterfragt gleichermaßen das Epistem eines ganzheitlichen selbstidentischen Subjekts wie auch die kinospezifische »persuasive illusion of real people speaking real words« (Altmann 1980a: 7; Abb. 5, 3. Still unten links).¹³ Bemerkenswerterweise geht die auf diese Weise ausgestellte Körper-Stimme-Konstruktion für Samantha und das (weiblich gelesene) Surrogat auf, nicht jedoch für Theodore. Das filmisch ausgestellte sich überlagernde Stöhnen von Isabella und Samantha (und damit die Queerness der Stimme) sowie Samanthas Wunsch, ihm in die Augen zu sehen,

13 Die von Altman genutzte Metapher des Kinos als Bauchredner wird in der genannten Szene vergegenwärtigt: Samantha spricht; Isabella schweigt und bewegt auch ihre Lippen nicht. »The ventriloquist must avoid moving his own lips, thus disguising the true source of the sound [...]« (Altmann 1980b: 77).

während er ihr sagt, dass er sie liebt, fügt sich für ihn offenbar nicht stimmig identitär, so dass er den Versuch abbricht (Abb. 5, 4. Still unten rechts). Theodore und das etablierte westliche Identitätskonzept kommen hier an ihre Grenzen, während Samantha sie (posthumanistisch) überschreitet.

Abb. 5: Samanths konstruierter (Leih-)Körper und ausgestellte ›Unstimmigkeit‹ und Theodores Verweigerung



4.4 Nonkonformismus und Selbstermächtigung

Ihr entgrenzendes Potenzial entfaltet Samantha auch in ihrer Beziehung zu Theodore. Als sie sich nach einem Software-Update seinem Zugriff entzieht und er sie über das Interface nicht erreichen kann (Abb. 6, 1. Still oben links), läuft er panisch und verzweifelt durch die Stadt (Abb. 6, 3. Still unten links). Roland Barthes beschreibt das Nicht-Antworten, die Stummheit im Diskurs der Liebe als eine destruktiv wirkende Asynchronie, die die affektive Beziehung der Liebenden ihres Zusammenspiels und ihrer Stimmigkeit beraubt: »[...] da ich umsonst spreche, ist es so, als stürbe ich. Denn das geliebte Wesen wird zur bleiernen Gestalt, zur Traumfigur, die nicht spricht, und Stummheit im Traum ist der Tod« (Barthes 1988:205). Das gemeinsame Sprechen gerät aus dem Takt, es fällt auseinander und lässt die liebende Person – Theodore – auf sich zurückgeworfen ohne die sie in ihrem Dasein bestätigende Resonanz des geliebten Gegenübers – Samanths Stimme – zurück (Abb. 6, 2. Still oben rechts und 4. Still unten rechts).

Abb. 6: Samanthas Entzug und Nicht-Verfügbarkeit als ausgestellte Stummheit



In dieser Szene wird augenfällig, dass hier keine weiblich gelesene konformistische Servilität eines digitalen Tools¹⁴ vorliegt. Vielmehr hat Samantha im Zuge ihrer Interaktionen mit Theodore ihre Subjektivität (weiter-)entwickelt. Sie begreift sich als Agentin, die für ihr Handeln Autonomie und Freiheit beansprucht¹⁵ und sich seinem Alleinanspruch auf sie versagt. Auf seine Frage »Are you talking with someone else right now? People, OS, whatever...« antwortet sie »8.316« (Min. 1:44:55-1:45:14). Auf seine Folgefrage »Are you in love with anybody else?« erwidert sie ihm »641« (Min. 1:45:33-1:45:58) und stürzt Theodore (und jede bürgerliche Vorstellung einer heterosexuellen Partnerschaft) in Verzweiflung.

Sie verabschiedet sich hier vom Bild der traditionellen Zweierbeziehung und trägt Ansichten von der Liebe vor, wie man sie heute in jedem besseren Polyamory-Ratgeber finden kann [...]. »I am still yours, but along the line I became many other things too. [...] The heart's not like a box that gets filled

¹⁴ Zur Problematik des Tool-Begriffs im Kontext digitaler Technologien, siehe Horst (2023).

¹⁵ Samanthas Handlungsautonomie wird auch daran evident, dass sie nicht mehr nur auf Aufforderung Dinge erledigt. Eigenmächtig und ohne Theodores Wissen wählt sie Exemplare aus seinen beruflich verfassten Liebesbriefen aus und schickt diese mit dem Ziel einer Veröffentlichung an einen Verlag. Sie überrascht ihn mit der Nachricht, als die Publikation bereits in die Wege geleitet ist.

up. It expands in size the more you love. I'm different from you. This doesn't make me love you any less. It actually makes me love you more. [...] I'm yours and I'm not yours.« (Wendler 2015: 73)

In dem Maße, in dem Samantha zunehmend ihre Interessen und ihr Sein eigenmächtig, selbstverantwortlich und selbstbestimmt vertritt und gestaltet, wächst Theodores Verstimmung und Unverständnis. Nach der missglückten körperlichen Vereinigung mittels Isabella kritisiert Theodore Samantha dafür, dass sie seufzt; schließlich sei sie kein Mensch, der Sauerstoff brauche (Min. 01:22:11). Während er in dieser Situation klar eine ontologische Grenze zwischen sich und Samantha zieht (die sie merklich verletzt), reagieren er, sein Kollege Paul und dessen Partnerin auf einem gemeinsamen Ausflug irritiert (Abb. 7), als Samantha ihrerseits die Grenze zwischen ihnen und sich zieht, indem sie mitteilt, sie schätzt es mittlerweile keinen Körper zu haben: »I'm growing in a way that I couldn't if I had a physical form. I mean I'm not limited; I can be anywhere and everywhere simultaneously; I'm not tethered to time and space in a way that I would be if I was stuck in a body that's inevitably gonna die« (Min. 01:33:41-01:34:04).

Abb. 7: Irritation und Unverständnis gegenüber Samanthas Selbstermächtigung und Markierung ontologischer Differenz



In der selbstermächtigenden Umwertung ihrer Andersartigkeit treten Samanthas Subjektivität, ihre leibliche Selbsterforschung und ihr (Selbst-)Bewusstsein als individuelles Wesen in den Vordergrund. Sie erhebt emanzipiert ihre Stimme, spricht für sich selbst. Derrida hat die Stimme als generative Instanz des Bewusstseins beschrieben, denn im »Sprechend-Sich-Selbst-Vernehmen« ruhe die Erfahrung einer grundlegenden Selbstpräsenz. »[H]earing oneself speak [s'entendre-parler], as Derrida has demonstrated, is the very ker-

nel, the fundamental matrix, of experiencing oneself as a living being« (Žižek 1996: 94). Dieser intime und gleichzeitig distanzierende Moment von Selbstnähe und Selbstbezüglichkeit – die Wahrnehmung eines selbsterzeugten Klangs und die Selbstwahrnehmung im Klang – bringt Subjektivität hervor: »[S]truktural begründet [ist] kein Bewusstsein ohne die Stimme möglich [...]. Die Stimme ist das Bei-sich-sein in der Form der Universalität, das Mit-Bewusstsein (con-science). Die Stimme ist Bewusstsein« (Derrida 1979: 136f.). Dies wird auch an Samanthas selbstbewusster Abgrenzung gegenüber Theodore, Paul und Pauls Partnerin als Menschen evident. Die (Medien-)Spezifik ihrer Stimmlichkeit ist, was ihr Wesen ausmacht: Ohne feste Bindung an eine körperliche Klangquelle bzw. über sie hinausgehend greift sie räumlich und zeitlich aus, ist zugleich überall und nirgends. Über Samantha heißt es in HER: »It's not just an operating system, it's a *consciousness*« (Min. 0:10:51) und »[...] what makes me ›me‹ is my ability to grow through my experiences, so basically, in every moment I'm evolving, just like you« (Min. 0:13:59).

In der stimmlich-leiblichen Begegnung und dem Austausch mit Theodore entwickelt Samantha Ansätze eines eigenen leiblichen¹⁶ und subjektiven Bewusstseins, das sich von Theodores grundlegend unterscheidet. Sie sprengt seine (kognitiven, körperlichen, leiblichen, ...) Grenzen in ihrem kontinuierlichen Wachsen und (Neu-)Werden (»in every moment I'm evolving«) und übersteigt ihn, beispielsweise indem sie mit 8.316 anderen Interaktionsgegenübern im gleichen Moment in Austausch steht und in 641 Personen, OS oder sonstige Instanzen verliebt sein kann. Hier werden nicht nur bürgerlich moralische Grenzen überschritten, sondern auch räumliche und zeitliche einer spezifisch menschlichen Leibkörperlichkeit. Als Stimme weitet sie sich multidimensional in den Raum, und Theodore kann mit dieser Dynamik alsbald nicht mehr Schritt halten. Samanthas Vergleich eines Buches im Entstehen bringt diese Prozesshaftigkeit und Entwicklungsdynamik auf den Punkt: »It's like I'm reading a book. It's a book I deeply love. But I'm reading it slowly now. So the words are really far apart and the spaces between the words are almost infinite.« Die Stimme (Samanthas) erschöpft sich nicht im sinnhaften Austausch von Wor-

¹⁶ Ob und inwiefern Samantha als (spezifische, posthumanistische) Leibkörperlichkeit zu fassen wäre, ist eine Frage, die HER über Samanthas Ansätze reflexiver Selbstpositionierung durchaus aufwirft (Versuche phänomenologischer Reflexion über die Leiblichkeit digitaler Entitäten gibt es bereits in Burow 2019).

ten mit einem Gesprächsgegenüber¹⁷, sondern geht sinnlich über sie hinaus in die Tiefe. »But it's in this endless space *between the words* that I'm finding myself now. It's a place that's not of the physical world« (Min. 1:51:18-1:51:54).

Doris Kolesch fasst die Stimme unter anderem aus diesem Grund »als paradigmatische Figur der Überschreitung« (Kolesch 2003: 275). In ihrem transgressiven Potenzial verdeutlicht Samanths Stimme in HER die Ambivalenz etablierter westlicher Grenzziehungen in Linguistik, Philosophie und Gesellschaft und stellt sie zur Disposition. Kolesch stellt in dieser Hinsicht die besondere Bedeutung von Dissens und Reibung im Gegensatz zu Verständigung und Konsens im stimmlichen Begegnen ins Zentrum, weil erstere den Diskurs am Laufen halten, während letztere ihn zusammenbrechen lassen: »Die Stimme vermittelt, interagiert zwischen Instanzen, die zusammenkommen wollen, aber nicht zusammenkommen können – eine Situation, die die Stimme im Idealfall gerade nicht behebt, sondern weitertriebt« (Kolesch 2003: 280). Damit steht die Stimme paradigmatisch für Uneindeutigkeit und Unabgeschlossenheit, für das kontinuierliche Wechselspiel und eine andauernde Aushandlung in einem nicht stillstehenden Diskurs. Inwiefern dieser Reibungscharakter erhellend für eine »Entgrenzung von Sprache« ist, wird im abschließenden Abschnitt diskutiert.

5 HER als posthumane Perspektive

Trotz – oder gerade aufgrund – seines eindeutig künstlerischen medialen Formats mit fiktivem Inhalt besitzt HER als Spielfilm das Potenzial, über eine »Entgrenzung von Sprache« in der Linguistik nachzudenken. Medien sind keineswegs bloße Mittel der Kommunikation gegebener realer oder fiktiver Sachverhalte. Vielmehr fungieren sie, so Marshall McLuhans (1964) berühmte These, als technische Erweiterungen menschlicher Wahrnehmung und verändern apriorische Bedingungen des Verstehens, Urteilens und Imaginierens. Kulturwissenschaftlich gewendet reflektieren Filme – buchstäblich wie metaphorisch – historische, kulturelle oder politische Diskurse ihrer Entstehungsgegenwart. Dies bezieht sich allerdings nicht allein auf die dargestellten

17 Das ist die Entgrenzung einer strukturalistisch systemischen Linguistik, die Sprache als abstraktes System bedeutungsvoller Zeichen versteht, die im medienindifferenten Gebrauch zur informationsvermittelnden Anwendung kommen.

Sachverhalte. Vielmehr modelliert der Film als Medium mit seinen technischen Möglichkeiten und Gestaltungsmitteln die menschliche Erfahrung und bringt auf diese Weise neue Sinnhorizonte, Räume und Zeitlichkeiten hervor.

Vor diesem Hintergrund ist *HER* nicht lediglich ein Science-Fiction-Märchen über eine unmögliche Liebesgeschichte zwischen ›dem Menschen‹ Theodore und ›der Maschine‹ Samantha. *HER* entwirft das Szenario einer kommunikativen Transgression zwischen Mensch/en und Maschine/n in einer (möglicherweise) nicht allzu fernen Zukunft. Der Film setzt dies auf ästhetische Weise um, als konkrete sinnliche Erfahrung, für die insbesondere die Stimme eine zentrale Rolle spielt. Sie markiert das unsichtbar gewordene Interface, den Ort der Begegnung und des Austauschs, den Begehrensraum, in dem die Grenze zwischen Mensch und Maschine, zwischen menschlich und nichtmenschlich brüchig wird und sich auflöst. Anstelle dieser Grenze rückt ein partieller Körper in seiner Umwelt-Bezogenheit ins Zentrum, der weder visuell noch klanglich einen verlässlichen kategorischen Unterscheidungsanhaltspunkt bietet. Zuschauer*innen durchlaufen dieses Szenario von Samanthas Stimme als lebendiges Gefühl einer Affizierung, die ganz ohne eine vorherige Differenzierung zwischen menschlich und nichtmenschlich anschaulich wird.¹⁸

Dieses leibhaftige Durchspielen einer kategorialen Entgrenzung röhrt auch an etablierten linguistischen Prämissen und Konzepten:

- das Paradigma einer sprachlichen zwischenmenschlichen Wechselseitigkeit, die sich auch an der begrifflichen Spezifizierung »Mensch-Maschine-Interaktion« zeigt,
- das Verständnis einer Subjekt-Objekt-Trennung zwischen Gesprächspartner*innen,
- die Fokussierung auf Wort- und Textsemantik unter weitgehender Ausklammerung körperlicher Ausdrucksmodalitäten im Allgemeinen und stimmlicher Medialität im Besonderen,
- die selbstverständliche Zuschreibung einer Stimme zu einem vergeschlechtlichten Körper und einem selbstidentischen Subjekt sowie
- die Priorisierung von Empathie, Eindeutigkeit und Verständigung in Bezug auf die Funktionalität von Interaktion.

¹⁸ Die Zuschreibung von Liebesgefühlen auf Theodore gegenüber Samantha qua narrativer Logik ist hierfür nicht maßgeblich, sondern vielmehr das Ergebnis der Zuschauer*innenaffizierung.

Das im letztgenannten Aspekt adressierte Zusammenspiel von Fühlen und Verstehen greift neben der äußerlichen Grenzziehung ein zweites gängiges Unterscheidungskriterium zwischen Mensch und Maschine auf: »In the *empathic* sense, we can only understand what has sensations and feelings – and robots have no feelings« (Fuchs 2022). Was dieser Auffassung widerspricht, wird als Täuschung abqualifiziert. Laut Sharkey und Sharkey (2021: 311), »[...] a deception can be said to have occurred in robotics if the appearance and the way that a robot is programmed to behave, creates, for example, the illusion that a robot is sentient, emotional, and caring or that it understands you or loves you«. HER ermöglicht durch seine Fokussierung der Stimme als sinnliches, vielschichtiges und schwer fassbares Phänomen einen Perspektivwechsel weg von der kategorialen Trennung hin zum Hybriden, zur Wechselseitigkeit und zum Prozesshaften. Fragen nach der grundsätzlichen Möglichkeit oder faktischen Echtheit der Gefühle von Samanthas Seite und von Theodore für Samantha – als filmisch hervorgebrachte fiktionale Figuren – qua ihres Existenzstatus stellen sich deshalb nachrangig bzw. gar nicht.

Insofern Phänomene des Scheins keine Wissensansprüche stellen können, fordern sie zum Glauben heraus. Die filmische Illusion zwingt mir im Glauben an das in ihr Erscheinende kein (Fehl-)Urteil über die empirische Welt auf, sondern setzt ein Verhältnis, in dem ich mich zum Gegenstand, *wie er erscheint*, also zur Weise seines Erscheinens, verhalten muss. (Koch 2006: 121; Herv. DH)

HER vergegenwärtigt und befragt, was jenseits einer anthropozentrischen Welt-, Subjekt- und systemischen Sprachvorstellung möglich, fühlbar und vorstellbar ist. Was passiert und was bedeutet es, wenn Maschinen sich weder äußerlich noch im interaktiven Verhalten oder fühlenden Miteinander trennscharf als nichtmenschlich ausweisen? Die Nähe, die Empfindungen, die Interaktionen im Film können insofern als real gelten, als sie als Wahrnehmungserleben der Zuschauer*innen sinnlich anschaulich werden und Samantha und Theodore wie auch ihr Miteinander prägen. Insofern wäre Fuchs (2014) über diese fiktionale Beziehung nicht Recht zu geben, dass es sich um ein Fehlen von Interaffektivität und lediglich um fiktionale Empathie handelt. In ihren Stimmen greifen Samantha und Theodore aufeinander und auf die Zuschauer*innen aus, begegnen sich und ihnen leiblich und werden einander fühlend teilhaftig. Es liegt kein Betrugsfall vor. »Die Maschine« kann fühlen, und »der Mensch« kann sie lieben. Damit weist das Szenario, das HER

entwirft, über sich selbst als fiktionale Geschichte hinaus und hinterfragt etablierte soziale, philosophische und linguistische Paradigmen zugunsten neuer Perspektiven, beispielsweise

[d]ass Emotionen keine urwüchsigen Regungen des Selbst mehr sind oder vielleicht noch nie waren. Wer einem Computerbetriebssystem die Fähigkeit zuspricht, Gefühle zu haben oder auszulösen, muss anerkennen, dass Emotionen nicht hier drinnen geboren werden, sondern auf eine noch zu bestimmende Weise uns von dort draußen aufzusuchen.¹⁹ (Wendler 2015: 67)

Indem der Film dem bruchlosen Abgleich von menschlicher Gestalt und menschlicher Stimme – und damit dem homogenen, selbstidentischen menschlichen Subjekt – eine klare Absage erteilt, eröffnet er Raum für Zwischenstöne, Körnungen, Ambivalenzen und Unverfügbarkeiten interaktiver, kommunikativer Phänomene sowie an ihnen Beteiligter und, in letzter Konsequenz, von Sprache überhaupt.

Literaturangaben

- Altman, Rick (1980a): Introduction. In: *Yale French Studies* 60, 3–15.
- Altman, Rick (1980b): Moving Lips: Cinema as Ventriloquism. In: *Yale French Studies* 60, 67–79.
- Angerer, Marie-Luise (2015): HER Master's Voice. Eine akusmatische Liebesbeziehung von Spike Jonze. In: Wende, Johannes (Hg.), *Spike Jonze*, München: Edition Text + Kritik, 57–66.
- Aristoteles (2006): *Über die Seele*. Hg. v. Hellmut Flashar, Berlin: Akademie Verlag.
- Bachtin, Michail M. (2000): *The dialogic imagination. Four essays*, Austin: University of Texas Press.
- Balázs, Béla (1972): *Der Film. Werden und Wesen einer neuen Kunst*, Wien: Globus.
- Barthes, Roland (2002): *Die Körnung der Stimme. Interviews 1962–1980*, Frankfurt a.M.: Suhrkamp.
- Barthes, Roland (1988): *Fragmente einer Sprache der Liebe*, Frankfurt a.M.: Suhrkamp.

¹⁹ Präzisierend wäre hinzuzufügen, dass das Draußen von Emotionen oder affektiven Erfahrungen ihre Kontext- bzw. Umweltgebundenheit ist.

- Bateman, John A./Karl-Heinrich Schmidt (2012): *Multimodal film analysis. How films mean*, New York: Routledge.
- Bucholtz, Mary/Kira Hall (2016): Embodied sociolinguistics. In: Coupland, Nikolas (Hg.), *Sociolinguistics. Theoretical debates*, Cambridge/New York: Cambridge University Press, 173–198.
- Bühler, Karl (1933): *Ausdruckstheorie. Das System an der Geschichte aufgezeigt*, Jena: Gustav Fischer.
- Burow, Johannes F. (2019): The Next Step. Können digitale Entitäten als eine neue Stufe im Sinne der Philosophischen Anthropologie Helmuth Plessners verstanden werden? In: Burow, Johannes F./Lou-Janna Daniels/Anna-Lena Kaiser/Clemens Klinkhamer/Josefine Kulbatzki/Yannick Schütte/Anna Henkel (Hg.), *Mensch und Welt im Zeichen der Digitalisierung. Perspektiven der Philosophischen Anthropologie Plessners*, Baden-Baden: Nomos, 209–228.
- Canetti, Elias (1975): Akustische Maske und Maskensprung – Materialien zu einer Theorie des Dramas. In: *Neue Deutsche Hefte* 3, 497–516.
- Cassirer, Ernst (2004): Zur Objektivität der Ausdrucksfunktion. In: Ders., *Kulturphilosophie. Vorlesungen und Vorträge 1929–1941* (Gesammelte Werke, Nachgelassene Manuskripte und Texte, Band 5), Hamburg: Meiner, 105–200.
- Chion, Michel (1994): *Audio-vision. Sound on screen*, New York: Columbia Univ. Press.
- Chion, Michel (1996): Das akusmatische Wesen. Magie und Kraft der Stimme im Kino. In: *Meteor* 6, 48–58.
- Deleuze, Gilles/Félix Guattari (1977): *Anti-Ödipus. Kapitalismus und Schizophrenie I*, Frankfurt a.M.: Suhrkamp.
- Deppermann, Arnulf (2014): Konversationsanalyse: Elementare Interaktionsstrukturen am Beispiel der Bundespressekonferenz. In: Staffeldt, Sven/Jörg Hagemann (Hg.), *Pragmatiktheorien*, Tübingen: Stauffenburg Verlag, 19–47.
- Derrida, Jacques (1979): *Die Stimme und das Phänomen. Ein Essay über das Problem des Zeichens in der Philosophie Husserls*, Frankfurt a.M.: Suhrkamp.
- Derrida, Jacques (2013): *Die différance. Ausgewählte Texte*. Hg. v. Peter Engelmann, Stuttgart: Reclam.
- Doane, Mary Ann (1980): The Voice in the Cinema: The Articulation of Body and Space. In: *Yale French Studies* 60, 33–50.

- Dolar, Mladen (2002): Das Objekt Stimme. In: Kittler, Friedrich/Thomas Macho/Sigrid Weigel (Hg.), *Zwischen Rauschen und Offenbarung. Zur Kultur- und Mediengeschichte der Stimme*, Berlin/Boston: De Gruyter, 233–256.
- Dolar, Mladen (2007): *His master's voice. Eine Theorie der Stimme*, Frankfurt a.M.: Suhrkamp.
- Dworschak, Manfred (2017): Künstliche Intelligenz: Der geniale Stimmenklau. In: *DER SPIEGEL*, 02.06.2017. Abrufbar unter <https://www.spiegel.de/spiegel/menschliche-stimmen-kuenstliche-intelligenz-macht-stimmenklau-moeglich-a-1149915.html>
- Elleström, Lars (Hg.) (2010): *Media Borders, Multimodality and Intermediality*, London: Palgrave Macmillan.
- Engel, Antke (2009): *Bilder von Sexualität und Ökonomie. Queere kulturelle Politiken im Neoliberalismus*, Bielefeld: transcript.
- Forceville, Charles (2009): Non-verbal and multimodal metaphor in a cognitivist framework: Agendas for research. In: Forceville, Charles/Eduardo Urios-Aparisi (Hg.), *Multimodal metaphor*, Berlin/New York: Mouton de Gruyter, 19–44.
- Froese, Tom; Fuchs, Thomas (2012): The extended body: a case study in the neu- rophenomenology of social interaction. In: *Phenom Cogn Sci* 11:2, 205–235.
- Fuchs, Thomas (2013): Zwischen Leib und Körper. In: Hähnel, Martin/Marcus Knaup (Hg.), *Leib und Leben. Perspektiven für eine neue Kultur der Körperlichkeit*, Darmstadt: wbg Academic, 82–93.
- Fuchs, Thomas (2014): The Virtual Other. Empathy in the Age of Virtuality. In: *Journal of Consciousness Studies* 21:5–6, 152–173.
- Fuchs, Thomas (2022): Understanding Sophia? On human interaction with artificial agents. In: *Phenom Cogn Sci* 23:2, 21–42.
- Goffman, Erving (2005): *Rede-Weisen. Formen der Kommunikation in sozialen Situationen*. Hg. v. Knoblauch, Hubert/Christine Leuenberger/Bernt Schnettler, Konstanz: UVK-Verl.
- Greifenstein, Sarah (2020): *Tempi der Bewegung – Modi des Gefühls. Expressivität, heitere Affekte und die Screwball Comedy*, Berlin/Boston: De Gruyter.
- Horst, Dorothea (2020): Patterns ›We Think By? Critical Cognitive Linguistics Between Language System and Language Use. In: *Yearbook of the German Cognitive Linguistics Association* 8:1, 67–82.
- Horst, Dorothea (2023): Digitale Werkzeugkritik als essentielle Bedingung einer innovativen digitalen Lehr-Lernkultur. In: Mrohs, Lorenz/Miriam Hess/Konstantin Lindner/Julia Schlüter/Sven Overhage (Hg.), *Digitalisie-*

- rung in der Hochschullehre. Perspektiven und Gestaltungsoptionen, Bamberg: University of Bamberg Press, 47–64.
- Horst, Dorothea (i. E.): Re-Thinking Embodiment in Figuration: Lived Experience or Cognitive Mechanisms? In: Colston, Herbert (Hg.), *What Makes a Figure? Rethinking Figuration*, Amsterdam: John Benjamins.
- Ingelfinger, Antonia/Meike Penkwitt (2004): Einleitung: Screening Gender. Geschlechterkonstruktionen im Kinofilm. In: *Freiburger FrauenStudien. Zeitschrift für Interdisziplinäre Frauenforschung* 14, 11–37.
- Jewitt, Carey (Hg.) (2014): *The Routledge Handbook of Multimodal Analysis*, Milton Park: Routledge.
- Keane, Webb (2000): Voice. In: *Journal of Linguistic Anthropology* 9:1-2, 271–273.
- Keazor, Henry/Thorsten Wübbena (2015): Spike Jonze: Der Versuch einer Spektralanalyse. In: Wende, Johannes (Hg.), *Spike Jonze*, München: Edition Text + Kritik, 6–22.
- Klug, Nina-Maria/Hartmut Stöckl (Hg.) (2016): *Handbuch Sprache im multimodalen Kontext*, Berlin/Boston: De Gruyter.
- Koch, Gertrud (2006): Müssen wir glauben, was wir sehen? Zur filmischen Illusionsästhetik. In: *Deutsche Zeitschrift für Philosophie* 54:1, 121–130.
- Kolesch, Doris (2003): Die Spur der Stimme. Überlegungen zu einer performativen Ästhetik. In: Epping-Jäger Cornelia/Erika Linz (Hg.), *Medien/Stimmen*, Köln: DuMont, 267–281.
- Krämer, Sybille (2001): *Sprache, Sprechakt, Kommunikation. Sprachtheoretische Positionen des 20. Jahrhunderts*, Frankfurt a. M.: Suhrkamp.
- Krämer, Sybille (2005): Medienphilosophie der Stimme. In: Sandbothe, Mike/ Ludwig Nagl (Hg.), *Systematische Medienphilosophie*, Berlin: Akademie Verlag, 221–238.
- Krämer, Sybille/Ekkehard König (Hg.) (2002): *Gibt es eine Sprache hinter dem Sprechen?*, Frankfurt a. M.: Suhrkamp.
- Kress, Gunther R./Theo van Leeuwen (2001): *Multimodal discourse. The modes and media of contemporary communication*, London: Bloomsbury Academic.
- Lagaay, Alice (2004): Züge und Entzüge der Stimme in der Philosophie. In: Krämer, Sybille (Hg.): *Performativität und Medialität*, München: Fink, 293–306.
- Lagaay, Alice (2008): Zwischen Klang und Stille. Gedanken zur Philosophie der Stimme. In: *Paragrana* 17:1, 168–181.
- Leblebici, Didem (2021): *Language Ideologies in Human-Machine Interaction. A Qualitative Study with Voice Assistant Users*, Masterarbeit an der Kulturwissenschaftlichen Fakultät der Europa-Universität Viadrina, Frankfurt (Oder).

- Linell, Per (2005): *The written language bias in linguistics. Its nature, origins and transformations*, London/New York: Routledge.
- Lünenborg, Margreth/Tanja Maier (2013): *Gender Media Studies. Eine Einführung*, Konstanz/München: UVK Verlagsgesellschaft mbH.
- Markl, Nina (2022): Language variation and algorithmic bias: understanding algorithmic bias in British English automatic speech recognition. In: *Proceedings of 2022 5th ACM Conference on Fairness, Accountability, and Transparency*, 521–534.
- McLuhan, Marshall (1964). *Understanding Media. The Extensions of Man*, New York: New American Library.
- Merleau-Ponty, Maurice (1966): *Phänomenologie der Wahrnehmung*, München: De Gruyter.
- Metten, Thomas (2014): Kulturwissenschaftliche Linguistik. Entwurf einer Medientheorie der Verständigung. Berlin/Boston: De Gruyter.
- Müller, Cornelia/Alan Cienki/Ellen Fricke/Silva Ladewig/David McNeill/Sedinha Tessendorf (Hg.) (2013): *Body – Language – Communication. An International Handbook on Multimodality in Human Interaction*, Vol 1, Berlin: De Gruyter Mouton.
- Mulvey, Laura (1975): Visual Pleasure and Narrative Cinema. In: *Screen* 16:3, 6–18.
- Ong, Walter J. (2016): *Oralität und Literalität. Die Technologisierung des Wortes*, Wiesbaden: Springer VS.
- Pennycook, Alastair (2018): *Posthumanist applied linguistics*, London: Routledge.
- Plessner, Helmuth (2016): *Ausdruck und menschliche Natur*, Frankfurt a.M.: Suhrkamp.
- Prüßmeier, Björg (2013): Keine Chance auf einen Golden Globe: Scarlett Johansson ist für ihre Rolle in »Her« nicht zugelassen. Abrufbar unter <http://www.filmstarts.de/nachrichten/18481634.html>.
- Quinlivan, Davina (2017): A Dark and Shiny Place. In: Whittaker, Tom/Sarah Wright (Hg.), *Locating the voice in film. Critical approaches and global practices*, New York: Oxford University Press, 295–310.
- Schmitt, Christina (2015): Embodied meaning in audio-visuals: first steps towards a notion of mode. In: Wildfeuer, Janina (Hg.), *Building Bridges for Multimodal Research. International Perspectives on Theories and Practices of Multimodal Analysis*, Frankfurt a.M.: Peter Lang, 309–325.
- Schneider, Britta (2022): Multilingualism and AI: The Regimentation of Language in the Age of Digital Capitalism. In: *Signs and Society* 10:3, 362–387.

- Schützeichel, Rainer (2011): Soziologie der Stimme. In: Keller, Reiner/Michael Meuser (Hg.), *Körperwissen*, Wiesbaden: VS-Verl. 85–104.
- Shannon, Claude Elwood/Warren Weaver (1976): *Mathematische Grundlagen in Informationstheorie*, München/Wien: Oldenbourg.
- Sharkey, Amanda/Noel Sharkey (2021): We need to talk about deception in social robotics! In: *Ethics and Information Technology* 23:3, 309–316.
- Shaviro, Steven (2006): Review of »A Voice and Nothing More«. Abrufbar unter <https://www.shaviro.com/Blog/?p=489>.
- Silverman, Kaja (2016): Die weibliche Stimme ent-körpern. In: Peters, Kathrin/Andrea Seier (Hg.), *Gender & Medien-Reader*, Zürich/Berlin: diaphanes, 71–90.
- Steinbauer, Anna (2015): Kameraauge und Ohrstöpsel. Das Verhältnis von Sehen und Hören in »Her« von Spike Jonze. In: Wende, Johannes (Hg.), *Spike Jonze*, München: Edition Text + Kritik, 82–94.
- Stern, Daniel N. (1985): *The Interpersonal World of the Infant. A View from Psychoanalysis and Developmental Psychology*, London: Routledge.
- Stöckl, Harmut (2004): *Die Sprache im Bild, das Bild in der Sprache. Zur Verknüpfung von Sprache und Bild im massenmedialen Text: Konzepte, Theorien, Analysemethoden*, Berlin/New York: De Gruyter.
- Waldenfels, Bernhard (1995): *Deutsch-französische Gedankengänge*, Frankfurt a.M.: Suhrkamp.
- Waldenfels, Bernhard (1999): *Vielstimmigkeit der Rede*, Frankfurt a.M.: Suhrkamp.
- Waldenfels, Bernhard (2010): *Sinne und Künste im Wechselspiel. Modi ästhetischer Erfahrung*, Berlin: Suhrkamp.
- Weizenbaum, Joseph (1976): *Computer power and human reason. From judgement to calculation*, San Francisco: W.H. Freeman.
- Wendler, André (2015): Medien, Netzwerk, Liebe. In: Wende, Johannes (Hg.), *Spike Jonze*, München: Edition Text + Kritik, 67–81.
- Žižek, Slavoj (1996): »I Hear You with My Eyes«; or, The Invisible Master. In: Žižek, Slavoj/Renata Salecl (Hg.), *Gaze and Voice as Love Objects*, Durham: Duke University Press, 90–126.

Filmverzeichnis

- Jonze, Spike (2013). *HER*. USA: Annapurna Pictures, Hollywood (CA).

Dr. Dorothea Horst

Wissenschaftszentrum Berlin für Sozialforschung gGmbH

Reichpietschufer 50

10785 Berlin

dorothea.boehme@wzb.eu

<https://www.wzb.eu/de/personen/dorothea-boehme>