

6. ›Ver-rückt‹ durch einen Algorithmus

Immersive Audio: Mediation und Hörbeziehungen

Joseph Klett

Du sitzt in einem Zimmer. Auf deinem Kopf trägst du scheinbar handelsübliche Kopfhörer. Eine Markenbezeichnung ist auf den Ohrmuscheln eingeprägt: *Mantle*. Die Kopfhörer sind an eine Blackbox-Konsole angeschlossen, die ihrerseits mit einer Nutzer-Oberfläche auf deinem Computer verbunden ist. Am Bildschirm wählst du »Benutzer kalibrieren« aus. Das System generiert eine Reihe von Tönen und fordert dich auf, jeden einzelnen Ton in Beziehung zu deinem Avatar auf dem Bildschirm zu verorten: ein Ton scheint direkt hinter deiner rechten Schulter zu entstehen; ein anderer Ton erklingt direkt vor dir, vielleicht eher ein bisschen von links. Du lokalisierst eine gewisse Anzahl von Tönen – vielleicht zwölf insgesamt – und dann wartest du, bis das System dein Profil erstellt hat. Du bekommst ein Signal, dass das System jetzt bereit für dich ist.

Immer noch im Sitzen wählst du aus deiner Programmbibliothek das Lied »Deja la Vida Volar (REMASTERED)« von Victor Jara aus. Ein Mediaplayer mit den üblichen Schaltflächen PLAY, PAUSE, ADVANCE und REVERSE erscheint auf dem Bildschirm. Neu ist ein weiteres Menü LOCATION; du bleibst erst einmal bei der Voreinstellung »Studio«. Du drückst auf »Play«. Jaras Akustikgitarrenklänge perlen vor dir von einer Seite zur anderen, während die Holz-Perkussion dumpf zu deiner Rechten stampft. Deine Augen folgen dem Klang der Andenflöte, der fast zum Greifen nah linker Hand schmachtet. Als Jaras Stimme ertönt, drängt sie in den Vordergrund, direkt vor dir – doch sie klingt merkwürdig hohl, als käme sie aus einer Höhle. Alles klingt so wie aus einem halbwegs anständigen Kopfhörer in digitalem Stereo.

Dann bewegst du den Kopf. Ein Gyroskop im Kopfhörerbügel registriert deine Bewegung. In der Konsole übersetzt ein Mikroprozessor die Information, ein Algorithmus passt sofort das Ausgangssignal an deine Kopfbewegungen an. Während du nach links blickst, kannst du die Stimme, die eben noch vor dir war, nun hinter deiner rechten Schulter hören. Die Flöte ertönt jetzt grade außerhalb der Reichweite deiner *rechten* Hand. Das Schlagzeug ist hinter dir. Du schwenkst den Kopf langsam von rechts nach links und der Klang passt sich an, gleicht deine Richtungswechsel mit seinem virtuellen Raum ab. Du bewegst den Kopf schneller und der Algorithmus schwenkt um; eine Kompressionssequenz verhindert hörbares *Audio Clipping*. Dank des Zerstreungskegels der Wahrneh-

mung im menschlichen Ohr nimmst du gar nicht wahr, dass das Playback vorübergehend weniger Samples enthält.

Du wendest dich wieder der Nutzeroberfläche und deren Menü zu. Du schaltest bei LOCATION von »Studio« auf »CBGB«. Der Sound verwandelt sich automatisch von der Studioschachtel, in der Jara aufnahm, in die langgestreckte Akustik des berühmten Kult-Clubs. Es spielt keine Rolle, dass der New Yorker Club überhaupt erst zwei Monate nach Jaras tragischem Tod in Chile eröffnet wurde; ein anderer Online-Jara-Fan hat sich die Mühe gemacht, eine verlustfreie MP3-Aufnahme in dem objekt-basierten Format, das du jetzt hörst, zu remastern. Dieses Format übersetzt die Aufnahme in einzelne Klang-Objekte, die anhand vorher aufgezeichneter akustischer Parameter tatsächlicher Räumlichkeiten wiedergegeben werden. In Verbindung mit dem Algorithmus kannst du nun mithilfe von Kopfhörern so hören, als wärest du tatsächlich an verschiedenen Orten. Dies ist ein fiktiver Bericht. Aber du könntest diese Geschichte auch von den Technikern hören, die die Algorithmen für *Immersive Audio* erstellen.

»Was machen Algorithmen?« Solon Barocas, Sophie Hood und Malte Ziewitz (Barocas u.a. 2013) werfen diese Frage auf, um die funktionale Geschlossenheit und Eindimensionalität von Algorithmen infrage zu stellen. Algorithmen sind keine immateriellen Formeln, sondern praktische Ausdrücke, die Auswirkungen auf die phänomenale Welt von Menschen haben. Daher muss man diese phänomenalen Modifizierungen auf ihre Relevanz als algorithmische Vorgänge untersuchen (Gillespie 2014). Statt Algorithmen für selbstverständliche Mechanismen zu halten, könnten wir diese technischen Verfahren ›in freier Wildbahn‹ aufsuchen, um so besser zu verstehen, wie sie Erfahrungen vermitteln. Algorithmen werden buchstäblich und im übertragenen Sinn von handelnden Personen ›kodierte‹, die Bedeutungen und Werte vor Ort in technische Prozeduren übersetzen. Ein Beispiel dafür finden wir, wenn wir uns die digitale Technik von *Immersive Audio* näher anschauen.

In diesem Kapitel nehme ich Sie mit in ein Forschungs- und Entwicklungslabor (*R&D*), in dem Tontechniker mit psychoakustischen Regeln und objekt-orientierten Berechnungen experimentieren, um Klänge in einer virtuellen Hörumgebung zu reproduzieren. Ein Mikroprozessor (und der darin eingebettete Code) vermittelt zwischen live ertönenden oder aufgezeichneten Quellen und einem Kopfhörer für die Wiedergabe. Der Klang wird in seiner Beziehung zwischen Wahrnehmungscharakteristika und Raumausrichtung des Hörers und dem Klangverhalten einer Umgebung umgestaltet. *Immersive Audio* wurde ursprünglich für Anwendungen im Film, in der Musik und in Videospiele vermarktet – obwohl die Entwicklung dieses neuen Audio Formats (Sterne 2012) nicht gegen einen Einsatz in verschiedenen Bereichen erweiterter Realität (*Augmented Reality* = *AR*) spricht, wie etwa auf Videokonferenzen oder bei virtuellen Führungen. In der Tat betonen die zeitgenössischen Entwickler, dass die Zukunft von *AR* nicht davon abhängt, welche Informationen mithilfe

fe dieser Technologie übermittelt werden, sondern von den Zugriffsmöglichkeiten dieser Technologie auf alle Arten von Information. Deshalb kodieren Tontechniker im Digitalbereich *Immersive Audio* weniger für spezifische Anwendungen, sondern eher auf einen bestimmten Zustand von akustischen Bezügen hin. Diese akustischen oder Hörbezüge entstehen vermittelt durch Personalisierung, Desorientierung und räumlicher Verlagerung (Translozierung).

Im Folgenden werde ich ethnographisches Material nutzen, um zu zeigen, wie diese Vermittlung aus technischen Entscheidungen und Wertsetzungen der Techniker entsteht, während sie daran arbeiten, *Immersive Audio* zum Leben zu erwecken.¹ Im Rückgriff auf Theorien in *Sound Studies* (kulturwissenschaftliche Klangforschung) und Kulturosoziologie behaupte ich, dass die algorithmischen Transformationen durch *Immersive Audio* einen Effekt herstellen, den ich *baffling* nenne, also »baff-machend« oder »ver-rückend« – in der Doppelbedeutung des Wortes »baffle«.² Ich meine also sowohl im übertragenen als auch im technischen Sinn eine Desorientierung, Umlenkung, Verblüffung oder Ver-Rückung. *Baffling* isoliert den Zuhörer vom gemeinsamen Klangraum und arrangiert die Bedeutungen in seiner Wahrnehmung dieses Raumes neu. Dieser technische Akt der Vermittlung oder Mediation (im agentiven und enkulturierenden Sinn, wie ihn Appadurai 2015 entwickelte) fesselt die Wahrnehmung des Nutzers und blockt gleichzeitig physisch andere Klänge ab. Die Antwort auf die Frage von Barocas u.a., was die Algorithmen von *Immersive Audio* denn tun, lautet also, dass sie physische und symbolische Prozesse ins Werk setzen, die tatsächlich die akustischen Bezüge zwischen Hörern und ihren unmittelbaren sozialen Welten umstrukturieren.

MEDIATION

Medientechnologien sind heutzutage allgegenwärtige Vermittler von Kultur. Das Auditive (kurz: Audio) beispielsweise ist Mittel und Medium zur Wahrnehmung symbolischen Materials und symbolischer Handlung. Im Kopfhö-

1 | Das empirische Material in diesem Kapitel stammt aus einem größeren Forschungsprojekt zur kulturellen Wahrnehmungsgestaltung durch Tontechnik (Klett, in Kürze erscheinend). Dieses Forschungsprojekt wurde von der *National Science Foundation* gefördert (SES-1128288).

2 | Das Wortspiel ist im Deutschen nicht genau so möglich: »to baffle« ist im Englischen der Vorgang, jemanden zu verblüffen, »baffled« entspricht dem deutschen »verblüfft«, »perplex«, »baff«. Ein *Baffle* ist aber auch eine Vorrichtung in der Akustik, mit der Schallwellen umgelenkt werden, z.B. eine Schallwand in Lautsprechern, die Töne umlenkt. In diesem Doppelsinn wird jemand durch *Immersive Audio* desorientiert, klanglich ver-rückt und verblüfft (A.d.Ü.).

rer konstruieren wir eine neue klangliche Beziehung zur Welt, die Bedeutung überlagert (Beer 2007; Bijsterveld 2010; Bull 2012; Hosokawa 2012). Das mechanische Verhalten dieser Technologie ist bedeutsam, weil es den Symbolgehalt, den es vermittelt, selbst färbt. Diesbezüglich thematisieren Jonathan Sterne und Tara Rodgers (2011) die Semiotik algorithmischer Signalverarbeitung und beziehen Metaphern des »Rohen« und des »Gekochten« auf die Klangfarbe der Audio-Wiedergabe. Techniker benutzen diese Kategorien, um »unberührte« Klänge von verarbeiteten zu unterscheiden; solche symbolischen Kategorien erlauben es ihnen – im Vakuum des Mediums – absolute oder »reine« Klänge zu spezifizieren, die entweder noch verarbeitet oder unbearbeitet belassen werden.

Über hörbare Signale hinausgehend kodieren Tontechniker auch Algorithmen für akustische Eigenschaften von Klang. Sterne (2015) beschreibt die Bedeutung, die »dry« und »wet« Audio für Tontechniker haben, wenn sie vorgegebene *Signale* mit algorithmischem *Affekt* kombinieren. Es ist Symbolarbeit, wenn Tontechniker den situativen Effekt des Halls separieren, so als habe er eine unabhängige Beziehung zum absoluten Klang-Objekt – sie imaginieren faktisch, dass ein Klang und der Raum dieses Klangs effektiv voneinander getrennt, aus ihrer akustischen Situation extrahiert und anderswo neu kombiniert werden können. Nach dieser Logik ist ein Niesen »an sich« immer der gleiche isolierte Klang – in den Gewölben einer modernen Bibliothek, in den Dünen eines Strands oder auf der sehr kurzen Strecke zwischen Kopfhörern und Gehörgang. Aus der Perspektive der Tontechnik ändert sich dabei lediglich, wie diese Räume die Wahrnehmung des Niesens modulieren. Für einen Toningenieur mag dies eine attraktive Einstellung sein: Wenn Nieser ontologische Objekte sind, die getrennt von ihrer Erkenntnis in einem Wahrnehmungs-Raum existieren, dann würde eine unabhängige Verbindung zwischen beiden es Algorithmen erlauben, die Wahrnehmung des Klangraums bei der Wiedergabe zu manipulieren. Lass einen beliebigen Klang durch einen Filter laufen und schon klingt es so, als käme der Klang aus einer völlig anderen Räumlichkeit. Solche Algorithmen rekonstruieren also den hörbaren »Raum«, indem sie innerhalb einer Audio-Produktion ein Set zuvor festgelegter akustischer Charakteristika als Referenz nehmen. In dieser Hinsicht, bestätigt Sterne, produzieren Algorithmen die Repräsentation eines Klangraums, die sich aus, bestenfalls, unvollständigen Informationen über die Räumlichkeiten speist, die sie zu repräsentieren vorgibt. »Künstlicher Hall repräsentiert Raum und konstruiert ihn dabei gleichzeitig«, so Sterne (2015: 113).

An dieser Stelle sollte ich unterstreichen, dass die symbolischen Kategorien, die Techniker verhandeln, nicht getrennt von der Arbeit des Aushandelns tatsächlicher Signale existieren. Diese pragmatische Einstellung schützt uns vor dem Drang, mit Symbolen so umzugehen, als wären sie auf einer höheren Ebene angesiedelt als ihr physikalischer Ausdruck. In diesem Sinn warnt die

Medienhistorikerin Lisa Gitelman (2004) uns vor »Entmaterialisierung« in Theorien medialer Effekte: Nur wenn wir Medien in einem Vakuum denken, als wären sie nie in Kontakt mit praktischer Nutzung, können wir uns diese Klänge als wirklich »virtuell« vorstellen, als wäre Erfahrung so flexibel und austauschbar wie Text auf einer leeren Seite. Eine derartige Verdinglichung von Symbolen-als-Text verhindert jedoch zu verstehen, warum ein bestimmter Akt der Mediation verführerischer sein kann als ein anderer. Audio-Technologie ist immer da besonders überzeugend, wo statische Signale durch multistabile Systeme, die zusätzlich zu den Signalen noch einen Kontext liefern, ersetzt werden (Langsdorf 2006). In der Praxis sind Kontexte Situationen, die sich aus einer Mischung aus materiellen und immateriellen Interaktionen zusammensetzen. Wie wir eine Situation verstehen – wenn wir sie überhaupt erkennen – hängt von unserer Wahrnehmung dieser Interaktionen ab. Laut Martina Löw muss »aus der Fülle des Wahrnehmbaren eine Auswahl getroffen werden, weshalb diese Wahrnehmung nicht direkter Natur ist. Sie vermittelt lediglich den Eindruck von Direktheit, ist aber ein hochgradig selektiver und konstruktiver Prozess« (Löw 2008: 41). Was wir für die Transparenz der neuen Medien halten, ist vielmehr ein Produkt der sinnlichen »Atmosphären« (Löw 2008) oder »Hüllen« (Rawes 2008: 74), die durch den Kontakt zwischen enkulturierten Hörern und auditiven Situationen hergestellt werden.

Als materielle Kultur ist Klang kein passives Medium, das die Verbindung zwischen zwei Subjekten herstellt und unterbricht, z.B. in einem Dialog. Klang ist eher ein aktives Ereignis in einer »hyper-relationalen Welt«, die die auditiven Bezüge und Nicht-Bezüge jener Subjekte konstituiert (Revill 2015). Wenn wir beispielsweise die schlecht isolierten Ohrknöpfchen benutzen, die wir in der Regel mit unseren digitalen Geräten mitgeliefert bekommen, sind es die anderen, die den Exzess an Klangenergie hören, der aus unseren Gehörgängen in die gemeinsame Klang-Umwelt durchsickert. Zur Illustration dieses Beziehungsaspekts von Klang verfolgt Heike Weber (2010: 346) die öffentliche Geschichte des Hörens mit Kopfhörern von der stationären Nutzung zur mobilen, zeichnet den Diskurs des »respektvollen Hörens« nach, der das einohrige Hören mit einem Kopfhörerknopf umgab; respektvoll, weil es sich nur auf das Ohr des Individuums bezog, aber auch, weil es das andere Ohr der nicht-mediatisierten Welt überließ. Tragbare Stereo-Kopfhörer »privatisierten« diese Erfahrung dann, indem sie sie in den intimen Raum des Ohres (Weber spricht von »Kokon«) verbrachten, während der geheime sinnliche Zugang, dessen sich der Kopfhörernutzer erfreute, demonstrativ öffentlich sichtbar war (Hosokawa 2012). Wie diese Studien zeigen, lässt uns Klang auf verschiedene Weisen Teil einer sozialen Situation sein, von denen uns nicht alle bewusst werden. Je mehr Zeit wir im Zustand der Audio-Nutzung verbringen, desto weniger stehen wir diesen anderen Situationen zur Verfügung, auch nicht den

anderen Zuhörern, die diese Situationen in akustischen Bezügen zu uns mitkonstruieren.

Wenn wir die materiellen und symbolischen Dimensionen vermittelten Hörens wirklich erfassen wollen, sollten wir uns das Zuhören nicht als binäre (ein/aus) Praxis vorstellen, sondern eher als eine kontinuierliche Beziehung zu Klang im Raum. Zu diesem Zweck argumentierte Tia DeNora (2000), dass Audio ein »Grundgerüst« (*scaffolding*) bereitstellt, in das das Selbst eingeklinkt werden kann und wodurch es gestützt wird, auf dem ein Handelnder aktiv an Rhythmen angekoppelt und affektiv von Klängen umgestimmt werden kann; Befragte berichten über die Nutzung von Audio im Alltag als pragmatische Technik und als Quelle der Selbsterkenntnis. Gordon Waitt, Ella Ryan und Carol Farbotko (2014) nennen dies die »Bauch-Politik« des Klangs: Die praktische Erfahrung eines Diskurses durch Klang (in ihrem Beispiel auf einem Demonstrationszug gegen den Klimawandel) liefert eine Resonanz zwischen Bedeutung und Schall. Und obwohl die meisten Hörerfahrungen sich auf Musikstücke beziehen, die Hörer aktiv auswählen, fand Mack Hagood (2011) heraus, dass die Hörer sich dabei gleichzeitig auch dafür entscheiden, die Umgebungsgeräusche außerhalb ihrer Audio-Geräte zu unterdrücken. Im Fall von Geräuschunterdrückungstechnologien gehört zu den Umgebungsgeräuschen eben auch der invertierte »noise« dazu, der als verworfenes Störsignal entsteht, um das herum Hörer sich weiter in ihre persönlichen Klanglandschaften zurückziehen. Diese eher pragmatisch orientierten Untersuchungen zeigen, dass Wahrnehmung in zwei Richtungen funktioniert: Die Fokussierung der auditiven Aufmerksamkeit hat notwendigerweise das Ausblenden anderer wahrnehmbarer Informationen in der Umgebung zur Folge.

Das gesellschaftliche Leben ist eine Bibliothek von Darbietungen, nach Genres organisiert, und Wahrnehmung ist von entscheidender Bedeutung, wenn es darum geht, diese Genres zu erkennen. In diesem Sinn betreffen Hörbeziehungen unmittelbar, was Ari Adut (2012) einen »allgemeinen Sinneszugang« zu einem gemeinsamen Satz an Erfahrungen genannt hat. Unsere phänomenale Erfahrung qualitativer Information hängt davon ab, dass Wahrnehmung innig mit Erkenntnis verknüpft, nicht ihr untergeordnet ist (McDonnell 2010; Martin 2011; Klett 2014). Dies bedeutet, dass ich in direkten und unmediatisierten Begegnungen vielleicht neue Wahrnehmungsgewohnheiten entwickle, die sich mit meinen Denkgewohnheiten in Einklang bringen wollen. Je mehr eine Gewohnheit zur Routine wird, desto eher wird sie meine Wahrnehmung diktieren. Aber wenn ich einer persönlichen Audio-Technologie lausche, werde ich perzeptuell individuiert, nicht sozial. Wenn wir Kopfhörer mit aktiver Geräuschunterdrückung benutzen, leuchtet für andere sozusagen eine rote Lampe auf, die signalisiert, dass wir trotz physischer Anwesenheit für andere nicht zugänglich sind. Und doch bin ich anderen gegenüber immer noch für meine Anwesenheit und mein Verhalten verantwortlich

– mein Husten, mein Rascheln, mein Kichern. In dem Sinne, dass ich meine Aufmerksamkeit (aber nicht mich) der unmittelbaren Situation zu entziehen vermag, setzen Vermittlungsprozesse meine Prioritäten als Handelnder nicht außer Kraft, sie formen sie lediglich um. Fragt sich nur: Woher stammen diese Prioritäten? Wenn uns Klangforscher sagen, dass die Mediation, die Vermittlung von Hörbeziehungen, ein Produkt kultureller Codes in ihrer Anwendung als Algorithmen sind, wer schreibt dann diese Codes?

IM FORSCHUNGS- UND ENTWICKLUNGSLABOR VON MANTLE (MANTLE R&D)

Mantle ist ein internationaler Audio-Hersteller mit Büros in verschiedenen Ländern auf der ganzen Welt. *Mantle R&D* ist das Labor der Firma im Silicon Valley. Das Labor besteht aus einem Direktor, acht angestellten Technikern und einer Handvoll Praktikanten, die jeweils nur für eine Saison dabei sind. Spezialisiert sind sie unter anderem auf Elektro- und Computertechnik, Sound Design und natürlich auf das Schreiben und Kodieren von Algorithmen.

Der größte Teil der technischen Arbeit im Labor vollzieht sich in einer Mischung aus digitalem und elektro-akustischem Experimentieren. Zwar mag ein Begriff wie »Forschungs- und Entwicklungslabor« ein Bild von weißen Kitteln und Apothekerschränken heraufbeschwören; man sollte sich aber eher das Großraumbüro eines Software-Entwicklers oder technischen Startup-Unternehmens vorstellen: lange Tische, graue Polstermöbel, normale Bürotüren und -fenster, Poster an den Wänden und diverse Spielzeuge überall verstreut. Das Ganze auf etwa einer Größe von knapp 112 m².

Mantle R&D wurde eigens zu dem Zweck gegründet, digitale Signalprozessoren (DSP) für verschiedene Formate der Lautsprechertechnik zu entwickeln. Mit DSP kann man elektronische Klänge auf vielen verschiedenen Ebenen bearbeiten und dabei die gesamte Substanz des Klanges erzeugen. Zwar besteht die eigentliche Aufgabe von *Mantle R&D* in der Entwicklung praktischer Anwendungen für DSP-Technik, aber weiter gehen die Anweisungen ›von oben‹ auch nicht. Dass das Büro mitten im Silicon Valley angesiedelt ist, deutet darauf hin, dass man sich seitens des Unternehmens einiges an ›Immersion‹ davon verspricht, eine Handvoll Forschungsingenieure in eine Brutstätte für digitale Technologie zu setzen. Während meines Aufenthalts im Labor von Mai bis August 2011 begannen die Ingenieure mit experimenteller Forschung zur Nutzung von DSP für *Immersive Audio*.

BAFFLING, IN DREI SCHRITTEN

Um zu verstehen, was die Algorithmen von *Immersive Audio* eigentlich tun, studiere ich die technische Bearbeitung, die der abschließenden Kodierung des Prozessors vorausgeht. Zu dieser Arbeit gehört eine Reihe von rekursiven Projekten, die ich über mehrere Monate hindurch durchführte. Geleitet wird das Projekt vom Cheftechniker Adam und seinem Assistenten Stefan, der in seinen Ferien von der Ingenieurschule hier als Praktikant arbeitet.

Während Adam und Stefan damit beschäftigt sind, einen Versuchsaufbau zu konstruieren, Daten zu sammeln, wissenschaftliche Texte heranzuziehen und Codes zu schreiben, navigieren sie sich zu ihrem Ziel *Immersive Audio* auf einem Floß voller Detailfragen durch, die ihrer Zielerreichung ständig praktischen Widerstand leisten. Regelmäßig stehen sie vor Entscheidungen, für die technisches Know-How nur begrenzt hilfreich ist; die symbolischen Codes werden eher im Fortgang des Experiments zugeteilt und ausgewertet – und dabei haben diese Entscheidungen direkte Auswirkungen auf die Funktionsweise der Technologie.

Ingenieure digitaler Audio-Produkte gestalten die ästhetischen Eigenschaften eines Klangs: Klangfarbe, Klangtemperatur und -textur. Außerdem kodifizieren sie bestimmte algorithmische Prozesse, die die Hörbezüge zwischen Hörern, Situationen und akustischen Umgebungen gestalten. Diese Bezüge geben eine Reihe lokaler Bedeutungen wieder, die nach Ansicht der Ingenieure am besten durch eine Reihe von symbolischen Transformationen des Hörerlebnisses repräsentiert werden. Schauen wir uns diese drei Vermittlungen (Mediationen) einmal an:

- *Personalisierung* definiert den Hörer durch eine Reihe von Audio-Charakteristika. So kann Audio auf eine große Bandbreite von Hörer-Körpern zugeschnitten werden. Während der Körper zu einem Teil des Audio-Systems wird, wird ihm jedoch gleichzeitig eine unabhängige Verbindung zu den Reproduktionsmitteln, d.h. zu den Lautsprechern in seinen Kopfhörern, zugewiesen.
- *Desorientierung* betrifft eine Bewegungsfreiheit der Körperausrichtung innerhalb des Signals des Audiogeräts. Eine willkürliche Beziehung zwischen Hörer und Lautsprecher schafft den Fokuspunkt konventioneller Audio-Felder ab und gewährt dem Hörer einen größeren Grad an Freiheit innerhalb des Feldes.
- *Verlagerung (Translozierung)* transformiert Klang nach vorgefertigten Charakteristika verschiedener akustischer Profile. Wenn der Hörende verlagert wird, brauchen die akustischen Bedingungen des eingenommenen Raumes nicht mehr zu gelten, es sei denn sie überlagerten sich mit dem Signal von *Immersive Audio*.

Diese drei nebeneinander bestehenden Mediationen erschaffen das gewünschte Erlebnis der Immersion, also das Eintauchen oder die Versenkung in einen Klangraum. Im Dialog mit den aufstrebenden Techniken der erweiterten Realität (AR) und der Virtuellen Realität (VR) strebt Immersion danach, die akustischen Grenzen der phänomenalen Hörsituation zu transzendieren. Dieser akustische Effekt ist jedoch von einer untergeordneten sozialen Funktion abhängig, die Ingenieure nicht unbedingt vorhersehen. Ich nenne diesen Prozess »baffling«. *Baffling* ermöglicht dem Hörer eine dynamische Beziehung zum vermittelten Klang, indem Wahrnehmung von den sozialen Situationen des unmittelbaren Hörens abgekoppelt wird.

Personalisierung

Forschungen zur Wahrnehmung deuten darauf hin, dass körperliche Unterschiede von Hörern die Klangwahrnehmung beeinflussen. Schließlich ist Hören ein Prozess, der zwischen Veränderungen in hörbarer Energie (Evens 2005) und der Sinnesenergie von Körpern an bestimmten Orten unterscheidet. Für den Tontechniker gehören zu diesen Variablen: Größe und Form der Ohren, Position der Ohren am Kopf und Schnitt der Schultern. Diese messbaren Unterschiede zwischen Hörern werden häufig als »individuelle Voraussetzungen« für das Hören bezeichnet. Die Ingenieure von *Mantle R&D* versuchen, diese individuellen Voraussetzungen durch Algorithmen zu modellieren. Darum geht es beim Prozess der *Personalisierung*.

Im *R&D Labor* erwuchs das Interesse an Personalisierung aus einem Treffen mit einem Professor, der die psychoakustischen Effekte individueller Voraussetzungen studiert. Auch wenn die technische Anwendung eine Neuheit ist, beruht die wissenschaftliche Arbeit des Professors auf einer ziemlich alten Wahrnehmungstheorie, die unter dem Namen *Head-Related Transfer Function* (HRTF, »Kopfbezogene Übertragungsfunktion« oder »Außenohrübertragungsfunktion«) bekannt ist. Kurz zusammengefasst handelt es sich bei HRTF um eine algebraische Repräsentation dessen, wie ein Klang erst vom einen, dann vom anderen Ohr gehört wird und wie das Gehirn dadurch die Klangquelle triangulieren kann. Durch Modellierung der HRTFs aus einer Stichprobe verschiedener Individuen glauben die Ingenieure von *Mantle R&D* den Code für eine Vorlage (Template) schreiben zu können, die für den einzelnen Benutzer nur noch modifiziert werden muss. Für die Erhebung der HRTF-Daten kommt eine Versuchs-Apparatur namens *AURA* zum Einsatz.

AURA besteht aus zwölf Lautsprechern, die an einem Gerüst aus zwei waagrechten Ringen befestigt sind. Sechs Lautsprecher stehen auf Ständern in etwa 120 cm Höhe und sechs hängen in etwa 300 cm Höhe von der Decke. Alle Lautsprecher sind auf einen Stuhl in der Mitte ausgerichtet. Auf dem Stuhl sitzt ein Hörer mit Mikrofonen in den

Ohren, die den Klang aufnehmen sollen, sobald dieser kurz hinter dem Außenohr den Gehörgang erreicht. Für die Messung spielen die Lautsprecher nur einen Ton, um einen gleichmäßigen Klangeindruck zu erzielen. Die Aufnahmen werden an Stefans Computer übermittelt und dort als visuelles Schaubild wiedergegeben. Mithilfe der so dargestellten Kurven soll eine audiologische Landkarte davon erstellt werden, wie der Klang den Hörer erreicht.

Die meisten Tontechniker wären sich wohl einig darin, dass Hörer leibliche Individuen mit einzigartigen Wahrnehmungsapparaten sind. Aber nicht alle stimmen zu, dass jeder Hörer gleich behandelt werden sollte. Vor der Entstehung von *Immersive Audio* gab es für Ingenieure keinen Anlass, individuelle Voraussetzungen jenseits ihrer eigenen zu berücksichtigen. Stereo – der vorherrschende Standard in der Audio-Technologie – entsteht aus zwei komplementären Soundkanälen, die in ihrer Gesamtheit nach objektiven Qualitätskriterien beurteilt werden, zum Beispiel aufgrund des Signal-Rausch-Abstands. Das Fehlen von Rauschen, um mal nur eine Variable zu nennen, zeichnet ein System im Vergleich zu anderen als objektiv besser aus. In Bezug auf diese Geräte von objektiver Qualität, so argumentierten professionelle und selbsternannte Audiophile, könnten sich aber auch Hörer qualitativ unterscheiden, nämlich darin, wie sie ihrerseits den Klang aus diesen Geräten hörten. Im Unterschied dazu behandeln Stefan und die anderen Tontechniker bei *Mantle R&D* individuelle Voraussetzungen als wesentlich für das Hörerlebnis und daher als unabdingbar für ein erstrebenswertes digitales Produkt.

Nachdem AURA im Labor montiert wurde, macht Stefan erste Probeaufnahmen, um zu schauen, wie die Daten aussehen. Dadurch dass das Gerüst für AURA offen ist, werden unweigerlich auch Geräusche von außerhalb zusammen mit den kontrolliert abgegebenen Tönen aufgezeichnet. Das Labor ist ein Großraumbüro, in dem andere Ingenieure ihre eigenen Forschungen betreiben, was nichtsdestotrotz zur Klanglandschaft des Labors beiträgt.

Am Computer müht sich Stefan damit ab, in den Aufzeichnungen die absichtlichen Klänge von AURA von den Störgeräuschen zu unterscheiden. Ohne festen Referenzpunkt kann er nicht klar abgrenzen, was der Hörer als Störgeräusch wahrnimmt und was Umgebungsgeräusch ist.

Sich eine neue Epistemologie des Hörens zu eigen zu machen, birgt Herausforderungen – vor allem, wenn man das Hören als essentiell subjektiv betrachtet. Im Prozess der Personalisierung gibt es keine vorher festgelegte oder ideale Wiedergabe, weil am Ende der Klang immer das subjektive Erlebnis des Benutzers ist. Wenn man irgendein Klang-Artefakt als »unerwünscht« einordnet, über-reduziert man willkürlich Teile des Signals. Stattdessen enthält nach diesem Ansatz das Signal sein eigenes »gutes Rauschen« als hörbaren lokalen

Kontext des Hörers (Klett 2014). *Immersive Audio* stellt gleichzeitig Figur und Grund zu Verfügung, da beides zum räumlichen Hören dazu gehört – wenngleich diese grundsätzlich subjektive Hörerfahrung den Technikern ungefähre Entscheidungen abverlangt, wo näherungsweise die Erfahrung einer anderen Person anfängt und wo sie endet.

Desorientierung

Audiotechnologie funktioniert mittels der Übertragung von Elektrizität in hörbare Schwingungen eines Zwerchfells oder einer Lautsprechermembrane. Deshalb hat man sich bisher den Lautsprecher, also die Stelle, wo sich System und Nutzer berühren, meist als statisches Gestell vorgestellt, in dem Audio-Sound faktisch mündet. Ingenieure der Stereo-Ära kodifizierten den Lautsprecher als objektiven Ausdruck von Audio und den Hörer als das Subjekt, das diesen objektiven Ausdruck aufsucht. Aber die Orientierung eines subjektiven Hörers an einer objektiven Klangquelle verhindert ein wirkliches Eintauchen in den Klang, eine wirklich immersive Hörerfahrung, in der Sound im Ohr des Hörers mündet. Daher kodieren die Techniker von *Mantle R&D* den Lautsprecher um, machen ihn eher zur Vermittlungsinstanz als zu einem objektiven Orientierungspunkt. Darum geht es beim Prozess der *Desorientierung*.

Die Kunsthistorikerin Caroline Jones (2005) bezeichnete »Hi-Fi« Stereo-Anlagen als Gipfel des Sounds der Moderne: Stereo ist statisch, zweidimensional und idealerweise direkt vor dem Hörer – wo sich in der Sprache der Stereoverkäufer der »sweet-spot« befindet. Stereo ist ein »kanal-basiertes« Format in dem Sinne, dass Stereo-Signale in einem einzelnen Kreislauf, der im Lautsprecher endet, eingefangen, gespeichert und wiedergegeben werden. Man geht von einer stabilen, direkten und unveränderten Wiedergabe von Sound-Daten aus. Von der Aufnahme bis zur Wiedergabe geht der rechte Kanal in den rechten Kanal und der linke Kanal in den linken Kanal. Dass im Diskurs von Stereo der Sound fest mit einem Paar Lautsprecher verdrahtet ist, bedeutet, dass Audio in erster Linie nach modernen Wertmaßstäben der Objektivität beurteilt wird, von einem stabilen und unbeweglichen Standpunkt aus, der immer außerhalb des zu beurteilenden Systems liegt. Weil das System beurteilt wird, als befände es sich in einem Vakuum, ist die kanalbasierte Beziehung letztlich eine Beziehung des Aufeinanderfolgens, nicht des Raumes.

Um es in den Worten eines Ingenieurs von *R&D* auszudrücken: Es war schwer, die Weisheit des modernen Stereoformats infrage zu stellen, weil die biologische Assoziation von Stereo mit zwei-ohrigem Hören die Idee, dass zwei Kanäle die optimale Wiedergabeweise seien, quasi als naturgegeben vermittelt hat. In der Psychoakustik beschreibt HRTF jedoch eine Beziehung zwischen Tönen in Zeit *und* im Raum. Zum Beispiel können wir in kakophonischen Situationen den Klang einer Stimme ›unter‹ oder ›zwischen‹ vielen anderen heraus-

hören – der sogenannte »Cocktail-Party Effekt«. Dieser kognitive Prozess legt nahe, dass räumliche Beziehungen mehr sind als nur eine Aufeinanderfolge.

Zusätzlich zu den von AURA kontrolliert ausgestoßenen Klängen werden in der offenen Struktur des AURA Gerüsts noch andere Geräusche aus dem Laborraum aufgenommen. Adam markiert auf der Kurve der Daten die Schwelle, ab der Umgebungsgeräusche die Töne zu überlagern beginnen, und empfiehlt Stefan, die Daten aggressiv zu »kappen«, auch wenn dabei Mikrosekunden der aufgenommenen Daten wegrasiert werden. Durch diese Trunkierung könnte man präziser wiedergeben, wie der Raum innerhalb der AURA-Anordnung aufgenommen wird. Doch geht die Präzision hier auf Kosten der Sorgfalt. Die digitale Manipulation der Klangdaten weicht vom menschlichen Wahrnehmungshandeln vor Ort ab, das sie doch gerade nachbilden möchte.

Adam schreibt ein Programm, das die aufgenommenen Klangdaten in eine algebraische Beschreibung überführen soll, die durch einen Algorithmus repliziert werden kann. Da er ein Verhältnis zwischen Subjekt und Umgebung beschreibt, braucht er mindestens eine bekannte Variable. Adam weist Stefan an, er solle nur die Töne bekannter Frequenzen aufnehmen. Diese Frequenzen werden als Visualisierungen der aufgezeichneten Daten kartographiert – in der Erwartung, dass sich die Form, die diese Frequenzen annehmen, extrahieren lässt und nur noch die Hörbeziehung zwischen Hörer und Umgebung, die sie ja darstellen wollen, übrig bleiben. Wenn sie erst diese Beziehung beschreiben können, könnten sie auch die Hörer-Wahrnehmung auf jegliche Orientierung im Raum hin manipulieren.

Die Daten, die man mit AURA sammeln kann, liefern keine stabilen Modelle, sondern dynamische räumliche Beziehungen, die zur Reproduktion immer Verarbeitung voraussetzen. In diesem Sinn behandelt AURA den Hörer als Informationsverarbeitungs-Instanz und Klanginformationen (in räumlichen Begriffen konzipiert) als »Reflexionen«: Während sich Signale von der Quelle zum Wahrnehmenden bewegen, bewegt sich überschüssige Klangenergie weiter, wird von Oberflächen zurückgeworfen und zum Hörer zurückgesendet. Dies gibt dem Klang ein Gefühl von Richtung und Abfolge – Dimensionen eines Signals, die man so manipulieren kann, dass sie unser Gefühl davon, wo sich ein Klang in Bezug auf einen anderen befindet, umorientieren.

Um das »gute Rauschen« im Audio-Material der Umgebung einzufangen, konzipieren die Techniker AURA so, dass auch Daten über den Hörer in Bezug auf »den Raum« als akustischen Raum gesammelt werden. Durch Rekonstruktion der Art und Weise, wie wir räumliche Orientierung wahrnehmen, konstruieren Ingenieure einen neuen Maßstab für Klangtreue (*fidelity*). Klangtreue oder Wiedergabegüte bezog sich früher auf die Wirklichkeitsnähe einer Aufnahme oder Geräteleistung. Wirklichkeitsnähe setzt einen stabilen, objektiven Standpunkt voraus, von dem aus Leistungen beurteilt werden können. *Immersive Audio* verlagert den Fokus von der Interpretation auf die räumliche Orien-

tierung und bezieht sich so auf die Wiedergabe eines Klangraums in all seinen akustischen Charakteristika. Sound ist in erster Linie eine Sache von Zeit und Raum und hat keinen einzelnen Ausgangspunkt. Nun fragen nicht mehr nur Kognitionswissenschaftler, sondern auch Computer-Programmierer: »Kann man die Gestalt eines Raums hören?« (Dokmanic' u.a. 2011)

Stefan verbringt mehrere Achtstundentage mit der Ausrichtung und Kalibrierung der AURA-Lautsprecher. Weil er noch keine Daten sammelt und es ein mühseliger Prozess ist, kommt eine Kopf- und Oberkörper-Attrappe aus Gummi zum Einsatz, der man Mikrophone ins Ohr gesetzt hat. Dies erlaubt ihm sorgfältigere und präzisere Messungen (weil Attrappen nicht viel herumhampeln).

Stefan nimmt einen kontrollierten Test-Ton auf – ein schrilles »Biiiiiiijuuuuu!« – und spielt ihn durch die auf die Attrappe gerichteten Lautsprecher ab. Diese Aufnahmen wiederholt er den ganzen Tag hindurch, auch zu Zeiten, wenn im Labor weniger los ist und die Geräuschkulisse dadurch gleichförmiger. Diese sich ständig wiederholenden Klänge – von den Ingenieuren ›gutes Rauschen‹ genannt – zeichnen sich in den Daten leichter ab und helfen dadurch, das akustische Territorium zu markieren.

Da man inzwischen Zugang zu billigeren und schnelleren Mikroprozessoren für DSP, zu Gyroskopen und zu Globalen Positionsbestimmungssystemen (GPS) hat, beginnen Toningenieure, die nicht aus der Stereo-Tradition kommen, damit, mithilfe von Algorithmen die Grenzen nicht beim Lautsprecher anzusetzen, sondern beim Zuhörer direkt außerhalb des Lautsprechers. Wenn man die Beziehung zwischen den aufgenommenen Tönen und dem Zuhörer kontinuierlich und in Echtzeit kalibrieren könnte, würde das System eine ›unabhängige Verbindung‹ zwischen den Tönen und den Quellen (d.h. den Lautsprechern), aus denen sie vibrieren, herstellen. In diesem Bezugsfeld könnte sich ein Subjekt frei in einer 360-Grad-Ausrichtung auf den Klang in der Horizontalen bewegen. So wie die Einbeziehung des Hörerkörpers den Eindruck, in die eigene Klangwelt eingetaucht zu sein, steigert, bedeutet die mediatisierte Emanzipation des Hörers von den statischen Bedingungen einer lokalen Situation, dass das Ohr sich nicht mehr an einer verdinglichten Klangquelle im Lautsprecher orientieren muss. Unter solchen Bedingungen der Möglichkeit täuscht ein Algorithmus vielleicht sogar das Gefühl vor, in einer ganz anderen Klangumgebung zu sein – ein Ideal von Virtualität, das sich bisher der Audio-technik entzogen hatte.

Verlagerung (Translozierung)

Unter dem Einfluss von Psycholinguistik und dem Niedergang der Moderne als Design-Ideologie erforschen Tontechniker jetzt einen Modus der Klangreproduktion, der grundlegend nutzerzentriert ist. Die Art und Weise, wie

Ohren hören, ist dreidimensional, sehr ähnlich dem, wie Augen sehen: Wir nehmen Klänge von rechts und links kommend wahr, von oben und unten, von fern und von nah. Zu diesem Zweck nähern sich Algorithmen diesen Dimensionen an, um ein Gefühl von Hall ›innerhalb‹ der Räume, in denen Klangenergie tatsächlich vom Ohr wahrgenommen wird, zu erzeugen. Bei *Mantle R&D* stellt man sich dieses Verpflanzen von akustischen Charakteristika von einem Raum auf einen anderen so vor, dass die Nutzer von *Immersive Audio* tatsächlich so hören können, als befänden sie sich in einer komplett anderen Hörumgebung. Darum geht es beim Prozess der *Verlagerung (Translozierung)*.

Um diese Empfindung des Hörens in einer anderen physischen Umgebung zu bewerkstelligen, müsste der Algorithmus ein Signal anhand eines Modells akustischer Resonanzen auf einen tatsächlichen Raum – von Ingenieuren auch ›der Raum‹ genannt, selbst wenn er nicht umschlossen ist – arrangieren, gebildet aus Referenzaufnahmen eines kontrollierten Tons in einer bestimmten Umgebung. Sterne erklärt dies so:

»Wir können uns Klangwellen so vorstellen, als schwärmten sie aus und erforschten den Raum, bis sie die Außenwände erreichen und dann zum Mittelpunkt des Raumes zurückkehren. Dabei zeichnen sie das Territorium sozusagen nach. Diese Resonanzen werden dann auf das trockene Signal aufgepfropft, so als befände sich dieses Signal in diesem Raum.« (Sterne 2015: 125)

Im Reich der Tontechnik erfolgt dieses »Nachzeichnen des Territoriums« anhand verräterischer Klang-*Reflexionen*. Reflexionen, zurückgeworfene Wellen, fangen die komplexen Beziehungen zwischen Klängen, Umgebungen und Nutzer ein. Theoretisch sind Reflexionen Nebenprodukte von Signalen. Aber Ingenieure wissen, dass beim Hören von Signalen – ganz gleich ob live oder technisch vermittelt – Reflexionen höchstwahrscheinlich gar nicht zu vermeiden sind und auch nicht vermieden werden müssen. Ingenieure, die von sich behaupten, dass sie sich in Akustik auskennen, sagen, die Physik der Reflexion sei zwar reichlich mysteriös, aber nicht prinzipiell unbegreiflich.

Innerhalb von AURA ergießt sich das kontrollierte Signal um den Körper in der Mitte der Versuchsanordnung herum. Aber Klang bewegt sich schnell fort. Wie die Energie auf den Körper im Inneren auftrifft, prallt sie auch jenseits der offenen Lautsprecheraufhängung auf die harten Oberflächen des Raumes. Dieser Klang kehrt in den Experimentalraum zurück und wird als Reflexion des Original-Signals mit aufgenommen. Visuelle Repräsentationen des innerhalb von AURA aufgenommenen Klangs können so auseinandergezogen werden, dass man den Punkt erkennen kann, an dem das Signal ankommt, noch bevor irgendwelche Reflexionen ankommen. Aber wenn Stefan seine Messungen vornimmt, zeigen die Bilder, die er zurate zieht, dass die Daten gespickt mit Anomalien sind. Er bittet Dave um Hilfe, der mit Tontechnik mehr Erfahrung hat.

Dave schlägt vor, eine Reihe von Referenzmessungen zu machen, um die Quelle oder die Quellen dieser vermutlich sehr frühen Reflexionen zu lokalisieren. Er holt mich zu Hilfe und weist auf verschiedene Stellen der Apparatur, von der diese zusätzlichen Daten ausgehen könnten. Er und ich halten an diese Stellen ein Stück gepolsterte Verkleidung, während Stefan den Testton abspielt. Wir machen mehrere Aufnahmen an verschiedenen Stellen, um diese physischen Stellen mit den Stellen auf den visuellen Aufzeichnungen zu korrelieren.

Wenn die Physik der Reflexion als Gleichung ausgedrückt werden kann, kann ein Algorithmus sie imitieren. Um Reflexionen für eine abbildende Darstellung nachzeichnen zu können, entwerfen die Ingenieure ein Modell des akustischen Raumes, der den Körper umgibt, als Gegenmodell zu dem verkürzten Raum zwischen Ohr und Kopfhörer. In Annäherung an eine akustische Umgebung werden die Experimentalbedingungen von AURA nicht am Nullsummen-Ideal von Stille gemessen, sondern durch eine visuelle Abbildung aufgezeichneter Reflexionen in graphischen Daten. Indem sie die im Laufe der Zeit kartographierte Verteilung von Klangwellen studieren und in den tatsächlichen akustischen Raum innerhalb der AURA-Anordnung rückübersetzen, erahnen Dave und Stefan, welche Klänge wünschenswert und welche unerwünscht sind. Wenngleich letztlich unpräzise, hilft AURA dennoch, sich erstrebenswerten Reflexionen als Beziehungsinformationen anzunähern, während es letztere als außerhalb des virtuellen akustischen Territoriums befindlich herauschneidet.

Anders als Stereo, das Klang zwischen seinen Lautsprechern wie ein Porträt ›rahmt‹, nutzt *Immersive Audio* Reflexionen, um durch das Arrangement von Klängen eine räumliche Beziehung zu konstruieren. Als fingen sie das Territorium in einem Globus ein, benutzen die Ingenieure digitale Parameter, um eine unbegrenzte Menge an Territorien, die auf ähnliche Weise kartographiert wurden, zu überlagern, und erzeugen so den Eindruck, man lausche bestimmten Klangbeziehungen, die in einer völlig anderen Umgebung aufgefunden wurden, ohne vorher definierte Orientierung und in einzigartiger Bewegungsfreiheit. Dieses System ist ein Zusammenfluss von Mediationen – Personalisierung, Desorientierung und Translozierung – von denen in der Abfolge des Hörens keine Vorrang hat. Stattdessen sind diese Prozesse in den Algorithmen von *Immersive Audio* grundsätzlich immer im Gange.

ALGORITHMEN UND HÖRBEZIEHUNGEN

Sozialstudien von Algorithmen neigen dazu, Code als Strukturregeln herauszustellen, die die Welt als Text montieren – Zahlen, Buchstaben, Symbole. Auf die Frage, wie diese Regeln im materiellen Sinn Beziehungen herstellen,

wird weniger Aufmerksamkeit verwendet. Aber Code gibt es nur verwirklicht in phänomenalen Medien, als Beziehungen zwischen den Objekten, die sie repräsentieren, und den Subjekten, die sie wahrnehmen (Coleman 2010). Studien zur Auseinandersetzung mit digitalen Technologien zeigen, wie weit die technische Bedienung eines Apparates in kognitiver, sinnlicher und praktischer Beziehung eines Subjekts bedarf (Schüll 2012). In diesem Sinne verfügt Medientechnologie gar nicht über unbegrenztes interaktives Potenzial. Eher sind es Medien wie *Immersive Audio*, die eine bestimmte Beziehung während der Nutzung mit-produzieren.

Für die Techniker von *Mantle R&D* sind Körper und Raum, die Klangreflexionen produzieren, wichtig, denn mit ihnen werden Klänge in allererster Linie klanglich. Im Einvernehmen mit der objektorientierten Programmierung virtueller und erweiterter Wirklichkeiten entwickelt sich im Umkreis relationaler Modelle der Wahrnehmung eine Kultur der Tontechnik – inspiriert von der Psychoakustik, der Kognitionswissenschaft und der Audiologie. Diese Kultur ist in erster Linie algorithmisch, denn nur mithilfe geschickten Hochleistungsrechnens kann unsere Wahrnehmung so effektiv mediatisiert werden. Wenn Klänge physische Entitäten mit definierten Grenzen sind, können Techniker Wahrnehmung besser auf einen Algorithmus modellieren, der uns die Arbeit des Hörens reflexiv abnimmt.³

Nach der Kalibrierung von AURA sammelt Stefan jetzt erste Hörerdaten, wobei er den Dummy-Torso als sein Subjekt nimmt. Eifrig schaltet er zwischen dem Computer-Bildschirm auf seinem Schreibtisch und dem Dummy-Subjekt im Zentrum des Experimentaufbaus hin und her.

Mike, der gerade kommt, um Leute zum Lunch abzuholen, macht sich an den Dummy heran und fragt: »Na, wie geht's dem Jungen?« Noch beim Tippen hebt Stefan den Kopf und blickt auf den Oberkörper, der sich leicht schief auf einem Schreibtischstuhl lümmelt. »Ein bisschen instabil, aber okay, solange ihn keiner anfasst.«

3 | Man sollte beachten, dass die Bedeutung von »subjektiv« in diesem Zusammenhang sehr spezifisch definiert ist. Die Körper und Räume, die mithilfe von *Immersive Audio* modelliert werden, sind statisch, passiv. Die individuelle Voraussetzung eines Subjekts bestehen im Wesentlichen darin, über einen Oberkörper mit Schultern erkennbarer Breite und über einen Kopf mit fleischigen Ohren zu verfügen. Andere Wahrnehmungsunterschiede werden zu anderer Zeit wiedergegeben oder als für die Wahrnehmung von Audio unnötig verworfen. In Anlehnung an Sternes (2003) Geschichte der Anfänge von Audio kann man sagen, dass die Klang-Erfahrungen in sogenanntem *Immersive Audio* oft bis zur Indexikalität überreduziert sind. In diesen Technologien »schwingen unsere Ohren in sympathischer Resonanz mit Maschinen, die für uns hören«.

Studierende der *Sound Studies* ermuntern Theoretiker dazu, Audio nicht als eine alternative Wirklichkeit, sondern als eine alltägliche kulturelle Praxis zu verstehen. Es gibt wenig Tabus, was die Benutzung von Audio in gesellschaftlichen Räumen betrifft. Selbst Gruppenaktivitäten mit Familie und Freunden schließen die Benutzung von Kopfhörern nicht aus. So steht es in der heutigen algorithmischen Kultur um die Hörbeziehungen. Wie Brandon LaBelle (2006) argumentiert: Hörbeziehungen sind irreduzibel und unwandelbar – aber sie sind verpflanzbar. Das phänomenale Wesen des Klangs macht unser Gefühl aus, einen Raum in der Zeit innezuhaben. Indem wir diese Elemente manipulieren, können wir vielleicht auch die sozialen Beziehungen zwischen denen, die an bestimmte Technologien gewöhnt sind, und jenen, die andere Gewohnheiten haben, umgestalten.

Algorithmen sind wie alle Technologien Einschreibungen sozialer Beziehungen. Statt die Medien der Algorithmen aus der Ferne zu lesen, schaue ich mir an, wie sie bei der Arbeit der Tontechnik gelesen, geschrieben und umgeschrieben werden. Diese sehr sinnliche Praxis enthüllt, wie die mechanistischen Prozeduren von *Immersive Audio* in der kontingenten, aber wertbesetzten Arbeit einer kleinen Gruppe von Individuen ihren Anfang nehmen. Die Toningenieure benutzen Algorithmen nicht nur, um die dynamischen Effekte spezifischer akustischer Bedingungen zu repräsentieren, sondern auch, um den Beziehungseffekt dessen, was es heißt, als individueller Körper mit einzigartigen Eigenschaften einen Raum einzunehmen, zu reproduzieren. Durch die Vermittlung von Personalisierung, Desorientierung und Translozierung lenkt *Immersive Audio* die Wahrnehmung von Hörern von realen Situationen weg. Statt einen generischen Hörraum für alle Hörer zu betreffen, verarbeitet *Immersive Audio Subjektivität in unabhängiger Beziehung zu einem nicht-lokalen Raum*. Das ›Objekt‹ von *Immersive Audio* ist dann wesentlich subjektiv und doch stets der Perspektive der Toningenieure zugeneigt.

Ohne etwas zu den sozioökonomischen Ungleichheiten, die den Zugang zu Audiotechnologie umgeben, zu sagen, argumentiere ich, dass ›Ver-Rückung‹ unser Verständnis von ›Wir-heit‹ in sozialen Situationen infrage stellt. Sinnliche Erfahrung ist in dem Maße bedeutungsvoll, in dem sie teilbar ist (Adut 2012), und Audio fungiert als »kultureller Filter« (Blessner/Salter 2006), der bestimmte Erfahrungen blockiert, andere favorisiert. Die Algorithmen von *Immersive Audio* repräsentieren eher persönliche Unterschiede als eine geteilte Erfahrung gemeinsamer Phänomene. Wenn diese kulturellen Filter robuster werden, verliert die unmittelbare Erfahrung lokaler Interaktionen an Macht, die Situation zu definieren. Ver-rückte Hörer durchlaufen so eine doppelte Privatisierung: eine objektive Isolierung, indem sie die Kopfhörer auf die Ohren pressen, und eine subjektive Isolierung, indem sie maßangefertigte Klänge interpretieren.

Durch die Mediationen von *Immersive Audio* wird die Fähigkeit des Subjekts, gemeinhin Klangeigenschaften zu definieren und zu unterscheiden – geschweige denn zu bewerten – durch Personalisierung rekonstruiert und wirkt stattdessen auf den Hörer als Objekt zurück, um die Hörsituation umzudefinieren. Das soll nicht heißen, dass Leute durch die Empfindungen von *Immersive Audio* verwirrt oder irregeleitet werden. Eher ermutigt diese Technologie Menschen dazu, aus unmittelbaren Erfahrungen auszusteigen und sich auf höhere Ebenen individueller Unterschiede zu begeben. Dies schafft neue Herausforderungen für politische Anerkennung in komplexen und stratifizierten Gesellschaften, wo Menschen sich kognitiv aus den Situationen, in denen sie sich befinden, »ausklinken« (Beer 2007: 858).

Angesichts der ungleichmäßigen Verteilung von Geräuschbelastung in Gesellschaften und der wachsenden Märkte für digitale Audio-Produkte, wird Stille rasch zu einem Luxusgut in städtischer Umgebung (Keizer 2010; Hagoood 2011; Stewart/Bronzafst 2011; Biguenet 2015). Aber in diesem Fall heißt »Stille« nicht Ruhe »da draußen«, sondern bedeutet eher Kontrolle über das eigene Hören »hier drin«. Zum Beispiel ist das *HERE Active Listening* ein Gerät im Ohr, das in Verbindung mit einem Smartphone Schallverarbeitung in Echtzeit anbietet. Das Gerät verfügt über einen Equalizer und eine Reihe von Preset-Filtern, mit denen der Nutzer die eigene Erfahrung von Live-Musik »kuratieren« kann (www.hereplus.me). Als eine Art von Kontrolle durch Klassifizierung suggeriert »Kuration«, dass Individuen aus einer festgelegten Menge von Eigenschaften ihre Auswahl treffen. Das ist für die kulturwissenschaftliche Forschung zu Algorithmen interessant, aber nicht wegen dem »Suchen Sie sich was aus«-Teil. Eher verspricht die *fest umrissene Menge an Eigenschaften* die größte theoretische Zugkraft.

Gewahrwerden und Erkennen von, Verantwortlichkeit für Situationen – hier fordern Algorithmen unsere Vorstellungen von Technologie als Epistemologie heraus. In der praktischen Anwendung von *Immersive Audio* sind wir mit einem Überschuss an Objekten auf Kosten immer weniger Subjekte konfrontiert. Solch ein Zustand an Hörbeziehungen kann unsere darauffolgenden Zustände des In-der-Welt-Seins beeinflussen. Ohne nun über den affektiven Stand der Empathie bei Generationen von »ver-rückten« Hörern spekulieren zu wollen, können wir die grundlegende Neuorganisation der Hörbeziehungen würdigen, die in den Verfahren von *Immersive Audio* auftritt. Selbst wenn wir keinen Mangel an Empathie innerhalb der Gesellschaft erfahren, bleibt da immer noch das eher praktische Problem des Verlusts an Fürsorge. Und Fürsorge oder Zuwendung ist ein Test, dem wir unsere Algorithmen unterziehen können.

In ihrem Buch *Reclaiming Conversation* fordert Sherry Turkle (2015) eine neue Ebene zwischenmenschlicher Verantwortung bei der Herstellung von Technologien. Insbesondere empfiehlt sie, dass das interaktive Design digita-

ler Technologie durch die bewusste Einbeziehung von Werten weitreichend umstrukturiert werden sollte. Dies würde zweifellos ernsthaften Widerstand gegen das Ver-Rücken, das aus den latenten Mediationen von *Immersive Audio* resultiert, bedeuten. Aber genau darin liegt die Herausforderung zu verstehen, was Algorithmen ›tun‹: Wie können wir nicht nur die explizite technische Funktion, sondern auch die impliziten sozialen Funktionen, die in digitalem Code zum Ausdruck kommen, verstehen? Wie Jonathan Franzen (2015) in seiner Rezension von Turkles Buch schreibt: »Aber was wird aus der großen Masse an Menschen, die zu verunsichert oder zu einsam sind, um dem Lockruf der Technik zu widerstehen, die zu arm oder zu überarbeitet sind, um aus dem Teufelskreis auszubrechen?« *Immersive Audio* braucht kaum mehr soziale Solidarität als die Bereitschaft des Benutzers, die Schnittstelle zu kalibrieren und den Kopfhörer aufzusetzen.

Es gibt vielfältige Hörbeziehungen, die im Diskurs der Ingenieure ungeklärt bleiben. An keinem Punkt ihrer Entscheidungsprozesse brachten die Techniker von *Mantle R&D* ein Interesse am ›Ver-Rücken‹ des Individuums zum Ausdruck. Eher streben sie eine besondere Erfahrung auditiver Transzendenz an. Der Effekt, den Verblüffung oder Ver-Rückung auf Hörbeziehungen haben könnte, ist einfach keine Größe – wenngleich dadurch eine bestimmte Beziehung zwischen Hörer und sozialer Situation reproduziert wird. Kann die Erforschung algorithmischer Kulturen soziale Konzepte identifizieren und operationalisieren, um neue Technologien zu entwickeln? Unbedingt. Aber insofern algorithmische Kulturen kontinuierliche, sich anpassende Prozesse sind, können wir nicht davon ausgehen, dass Technologien nicht auch weiterhin verblüffen oder ver-rücken werden oder dass Nutzer sich nicht danach sehnen, verblüfft und ver-rückt zu werden.

Übersetzt von Dagmar Buchwald.

LITERATURVERZEICHNIS

- Adut, A. (2012): »A Theory of the Public Sphere«, *Sociological Theory* 30 (4), S. 238-262.
- Appadurai, A. (2015): »Mediants, Materiality, Normativity«, *Public Culture* 27 (76), S. 221-237.
- Barocas, S./Hood, S./Ziewitz, M. (2013): »Governing Algorithms: A Provocation Piece«, Social Science Research Network, online: http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2245322 (zuletzt aufgerufen am 12.12.2015).
- Beer, D. (2007): »Tune Out: Music, Soundscapes and the Urban Mise-en-scene«, *Information, Communication & Society* 10 (6), S. 846-866.

- Biguenet, J. (2015): *Silence*, New York: Bloomsbury.
- Bijsterveld, K. (2010): »Acoustic Cocooning: How the Car Became a Place to Unwind«, *The Senses and Society* 5 (2), S. 189-211.
- Blessner, B./Salter, L.R. (2006): *Spaces Speak, Are You Listening? – Experiencing Aural Architecture*, Cambridge, MA: MIT Press.
- Bull, M. (2012): »iPod Culture: The Toxic Pleasures of Audiotopia«, in: *The Oxford Handbook of Sound Studies*, hg. v. T. Pinch und J. Bijsterveld, New York: Oxford University Press, S. 526-543.
- Coleman, E.G. (2010): »Ethnographic Approaches to Digital Media«, *Annual Review of Anthropology* 39, S. 487-505.
- DeNora, T. (2000): *Music in Everyday Life*, Cambridge, UK: Cambridge University Press.
- Dokmanić, I./Yue, M.L./Vetterli, M. (2011): »Can One Hear the Shape of a Room: The 2-D Polygonal Case«, Vortrag auf der Konferenz: *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2011.
- Evens, A. (2005): *Sound Ideas: Music, Machines, and Experience*, Minneapolis, MN: University of Minnesota Press.
- Franzen, J. (2015): »Sherry Turkle's »Reclaiming Conversation««, *New York Times*, 28.9., online: www.nytimes.com/2015/10/04/books/review/jonathan-franzen-reviews-sherry-turkle-reclaiming-conversation.html (zuletzt aufgerufen am 1.11.2015).
- Gillespie, T. (2014): »The Relevance of Algorithms«, in: *Media Technologies: Essays on Communication, Materiality, and Society*, hg. v. Tarleton Gillespie, Pablo J. Boczkowski und Kirsten A. Foot, Cambridge, MA und London: MIT Press, S. 167-194.
- Gitelman, L. (2004): »Media, Materiality, and the Measure of the Digital; Or, The Case of Sheet Music and the Problem of Piano Rolls«, in: *Memory Bytes: History, Technology, and Digital Culture*, hg. v. Tarleton Gillespie, Pablo J. Boczkowski und Kirsten A. Foot, Cambridge, MA und London: MIT Press, S. 199-217.
- Hagood, M. (2011): »Quiet Comfort: Noise, Otherness, and the Mobile Production of Personal Space«, *American Quarterly* 63 (3), S. 573-589.
- Hosokawa, S. (2012): »The Walkman Effect«, in: *The Sound Studies Reader*, hg. v. Jonathan Sterne, Abingdon, UK und New York: Routledge, S. 104-116.
- Jones, C. (2005): *Eyesight Alone: Clement Greenberg's Modernism and the Bureaucratization of the Senses*, Chicago, IL: University of Chicago Press.
- Keizer, G. (2010): *The Unwanted Sound of Everything We Want: A Book about Noise*, New York: PublicAffairs.
- Klett, J. (2014): »Sound on Sound: Situating Interaction in Sonic Object Settings«, *Sociological Theory* 32 (2), S. 147-161, doi: 10.1177/0735275114536896.

- Klett, J. (im Erscheinen): »Organizing Sound: An Ethnographic Investigation into the Making of Listening Subjects and Sounding Objects«, unveröffentlichtes Manuskript.
- Labelle, B. (2006): *Background Noise: Perspectives on Sound Art*, New York: Continuum.
- Langsdorf, L. (2006): »The Primacy of Listening: Towards a Metaphysics of Communicative Interaction«, in: *Postphenomenology: A Critical Companion to Ihde*, hg. v. E. Selinger, Albany, NY: SUNY Press, S. 37-47.
- Löw, M. (2008): »The Constitution of Space: The Structuration of Spaces through the Simultaneity of Effect and Perception«, *European Journal of Social Theory* 11 (1), S. 25-49.
- McDonnell, T.E. (2010): »Cultural Objects as Objects: Materiality, Urban Space, and the Interpretation of AIDS Campaigns in Accra, Ghana«, *American Journal of Sociology* 115 (6), S. 1800-1852.
- Martin, J.L. (2011): *The Explanation of Social Action*, Oxford: Oxford University Press.
- Rawes, P. (2008): »Sonic Envelopes«, *The Senses and Society* 3 (1), S. 61-78.
- Revill, G. (2015): »How Is Space Made in Sound? Spatial Mediation, Critical Phenomenology and the Political Agency of Sound«, *Progress in Human Geography*, online: 22.2.2015, doi: 10.1177/0309132515572271.
- Schüll, N.D. (2012): *Addiction by Design: Machine Gambling in Las Vegas*, Princeton, NJ: Princeton University Press.
- Sterne, J. (2003): *The Audible Past: Cultural Origins of Sound Reproduction*, Durham, NC: Duke University Press.
- Sterne, J. (2012): *MP3: The Meaning of a Format*, Durham, NC: Duke University Press.
- Sterne, J. (2015): »Space within Space: Artificial Reverb and the Detachable Echo«, *Grey Room* (60), S. 110-131.
- Sterne, J./Rodgers, T. (2011): »The Poetics of Signal Processing«, *Differences* 22 (2-3), S. 31-53.
- Stewart, J./Bronzaft, A.L. (2011): *Why Noise Matters: A Worldwide Perspective on the Problems, Policies and Solutions*, Abingdon, UK: Earthscan.
- Turkle, S. (2015): *Reclaiming Conversation: The Power of Talk in a Digital Age*, New York: Penguin Press.
- Watt, G., Ryan, E./Farbotko, C. (2014): »A Visceral Politics of Sound«, *Antipode* 46 (1), S. 283-300.
- Weber, H. (2010): »Head Cocoons: A Sensori-Social History of Earphone Use in West Germany, 1950-2010«, *The Senses and Society* 5 (3), S. 339-363.