

Maximilian KOCK

# WIE SOUNDDESIGN WIRKT

Von der Psychoakustik zum  
auditiven Medien- und Produktdesign



**Im Internet:**  
Video- und Tonbeispiele

HANSER

2., vollständig überarbeitete Auflage

### **Disclaimer zur Barrierefreiheit**

Der Carl Hanser Verlag unternimmt große Anstrengungen, um seine Produkte barrierefrei zu machen. Dazu gehört auch, dass Bilder oder Tabellen für blinde und sehbehinderte Menschen zugänglich gemacht werden. Dies geschieht durch zusätzliche beschreibende Texte (Alternativtexte), die in den Daten integriert sind. Die Alternativtexte können von assistiven Technologien (z. B. Screenreadern) vorgelesen werden. Bei der Erstellung dieser Texte kommt eine KI zum Einsatz. Die inhaltliche Verantwortung liegt weiterhin bei den Lektor:innen und Autor:innen.

Kock

## **Wie Sounddesign wirkt**



Maximilian Kock

# **Wie Sounddesign wirkt**

Von der Psychoakustik zum auditiven Medien-  
und Produktdesign

2., vollständig überarbeitete Auflage

HANSER

Dieses Buch ist der ersten Auflage unter dem Titel „Wie der Ton zum Bild passt: Wege zu effektivem Sounddesign“, 2019, bei Schiele & Schön GmbH erschienen.

Der Autor:

Prof. Dr. Maximilian Kock ist Professor für Audioproduktion in den Studiengängen Medienproduktion und Medientechnik an der Ostbayerischen Technischen Hochschule (OTH) Amberg-Weiden. Gleichzeitig ist er im Verband Deutscher Tonmeister (vdt) als Leiter des Referats für Aus- und Weiterbildung tätig.



Print-ISBN: 978-3-446-48300-2

E-Book-ISBN: 978-3-446-48301-9

E-Pub-ISBN: 978-3-446-48583-9

Die allgemein verwendeten Personenbezeichnungen gelten gleichermaßen für alle Geschlechter.

Alle in diesem Werk enthaltenen Informationen, Verfahren und Darstellungen wurden zum Zeitpunkt der Veröffentlichung nach bestem Wissen zusammengestellt. Dennoch sind Fehler nicht ganz auszuschließen. Aus diesem Grund sind die im vorliegenden Werk enthaltenen Informationen für Autor:innen, Herausgeber:innen und Verlag mit keiner Verpflichtung oder Garantie irgendeiner Art verbunden. Autor:innen, Herausgeber:innen und Verlag übernehmen infolgedessen keine Verantwortung und werden keine daraus folgende oder sonstige Haftung übernehmen, die auf irgendeine Weise aus der Benutzung dieser Informationen – oder Teilen davon – entsteht. Ebenso wenig übernehmen Autor:innen, Herausgeber:innen und Verlag die Gewähr dafür, dass die beschriebenen Verfahren usw. frei von Schutzrechten Dritter sind. Die Wiedergabe von Gebrauchsnamen, Handelsnamen, Warenbezeichnungen usw. in diesem Werk berechtigt also auch ohne besondere Kennzeichnung nicht zu der Annahme, dass solche Namen im Sinne der Warenzeichen- und Markenschutz-Gesetzgebung als frei zu betrachten wären und daher von jedermann benützt werden dürfen.

Die endgültige Entscheidung über die Eignung der Informationen für die vorgesehene Verwendung in einer bestimmten Anwendung liegt in der alleinigen Verantwortung des Nutzers.

Bibliografische Information der Deutschen Nationalbibliothek:

Die Deutsche Nationalbibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliografie; detaillierte bibliografische Daten sind im Internet unter <http://dnb.d-nb.de> abrufbar.

Dieses Werk ist urheberrechtlich geschützt.

Alle Rechte, auch die der Übersetzung, des Nachdruckes und der Vervielfältigung des Werkes, oder Teilen daraus, vorbehalten. Kein Teil des Werkes darf ohne schriftliche Einwilligung des Verlages in irgendeiner Form (Fotokopie, Mikrofilm oder einem anderen Verfahren), auch nicht für Zwecke der Unterrichtsgestaltung – mit Ausnahme der in den §§ 53, 54 UrhG genannten Sonderfälle –, reproduziert oder unter Verwendung elektronischer Systeme verarbeitet, vervielfältigt oder verbreitet werden.

Wir behalten uns auch eine Nutzung des Werks für Zwecke des Text und Data Mining nach § 44b UrhG ausdrücklich vor.

© 2025 Carl Hanser Verlag GmbH & Co. KG, München

Vilshofener Straße 10 | 81679 München | [info@hanser.de](mailto:info@hanser.de)

[www.hanser-fachbuch.de](http://www.hanser-fachbuch.de)

Lektorat: Frank Katzenmayer

Herstellung: Eberl & Koesel Studio GmbH, Kempten

Coverkonzept: Marc Müller-Bremer, [www.rebranding.de](http://www.rebranding.de), München

Covergestaltung: Max Kostopoulos

Titelmotiv: © Max Kostopoulos, unter Verwendung von Grafiken von © [gettyimages.de/edwardolive](http://gettyimages.de/edwardolive) und [firefly.adobe.com](http://firefly.adobe.com)

Satz: le-tex publishing services GmbH, Leipzig

Druck: Elanders Waiblingen GmbH, Waiblingen

Printed in Germany

---

# Inhalt

---

|  |           |
|--|-----------|
| <b>Vorwort .....</b>   | <b>IX</b> |
| <b>1    Einleitung .....</b>   | <b>1</b>  |
| <b>2    Sounddesign: Wahrnehmung und qualitative Beurteilung ..</b>                              | <b>9</b>  |
| 2.1    Audiotechnik, Audiopostproduktion und Sounddesign .....                                   | 10        |
| 2.2    Akustik, Psychoakustik und audiovisuelle Wahrnehmung .....                                | 15        |
| 2.3    Musikpsychologie und Neurowissenschaften .....  | 24        |
| 2.4    Sounddesign in den audiovisuellen Medien .....  | 45        |
| 2.5    Sounddesign in der Industrie .....  | 50        |
| 2.6    Exkurs: Sounddesign und künstliche Intelligenz .....                                      | 56        |
| 2.7    Sounddesign: Wie kann man seine Wirkung messen? .....                                     | 69        |
| <b>3    Sounddesign in der Anwendung .....</b>   | <b>73</b> |
| 3.1    Warum man Sounddesign in der Industrie<br>und in den audiovisuellen Medien benötigt ..... | 73        |
| 3.2    Visuelle Rezeption versus auditive Rezeption: Beispiele aus der Praxis ..                 | 80        |
| 3.3    Soundeffekte (Geräusche), Musik und Sprache .....   | 83        |
| 3.4    Die besondere Funktion der Sprache bei der Entstehung des Tonfilms ...                    | 85        |
| 3.5    Diegetischer und nicht-diegetischer Ton .....   | 89        |
| 3.6    Exkurs: Die deutsche Sprache als Tonsprache .....   | 94        |
| 3.7    Eine mögliche Methodik zur Effizienzmessung von Sounddesign .....                         | 98        |

**4 Methode einer empirischen Untersuchung ..... 101**

4.1 Empfindungsgrößen Immersion und Spannung ..... 101

4.2 Versuchsmethode ..... 104

4.3 Inhalt der Testvideos ..... 106

4.4 Produktion der Testvideos ..... 110

4.5 Die emoTouch-Applikation ..... 110

4.6 Der Fragebogen ..... 113

4.7 Beschreibung und Auswahl der Probanden ..... 114

4.8 Versuchsablauf ..... 115

4.9 Rohdaten und deren statistische Auswertung ..... 117

4.10 Nachbetrachtung ..... 119

**5 Interpretation der Ergebnisse des Hörversuchs ..... 121**

5.1 Auswertung der Hörversuche: Signifikanzprüfung ..... 121

5.2 Visualisierung der Ergebnisse in Boxplots ..... 124

5.2.1 Realfilm Katakomben mit Soundeffekten (F/X) ..... 125

5.2.2 Realfilm Katakomben mit Musik ..... 126

5.2.3 Realfilm Katakomben mit Musik und F/X ..... 127

5.2.4 Animationsfilm Goldenberg mit F/X ..... 128

5.2.5 Animationsfilm Goldenberg mit Musik ..... 130

5.2.6 Animationsfilm Goldenberg mit Musik und F/X ..... 131

5.3 Mittlerer Verlauf der Effekte über die Filmlaufzeit ..... 132

5.4 Auswertung der Fragebögen ..... 136

5.5 Visualisierung der Versuchsergebnisse im Vergleich zueinander ..... 138

**6 Diskussion ..... 141**

6.1 Diskussion der Ergebnisse der einzelnen Filmversionen ..... 141

6.1.1 Realfilm Katakomben mit Soundeffekten (F/X) ..... 145

6.1.2 Realfilm Katakomben mit Musik ..... 145

6.1.3 Realfilm Katakomben mit Musik und F/X ..... 146

6.1.4 Animationsfilm Goldenberg mit F/X ..... 146

6.1.5 Animationsfilm Goldenberg mit Musik ..... 147

6.1.6 Animationsfilm Goldenberg mit Musik und F/X ..... 148

6.2 Diskussion der Ergebnisse der Fragebögen ..... 148

6.3 Angemessenheit der Methode und praktische Relevanz der Studie ..... 150

6.4 Vergleich der Ergebnisse mit denen anderer Studien ..... 153

6.5 Zusammenfassung ..... 155



|          |   |            |
|----------|---|------------|
| <b>7</b> | <b>Empirisch fundierte Grundsätze für die Audiogestaltung<br/>in der Anwendung .....</b>    | <b>157</b> |
| <b>8</b> | <b>Glossar der verwendeten Fachbegriffe .....</b>   | <b>167</b> |
| <b>9</b> | <b>Verzeichnisse .....</b>  | <b>177</b> |
| 9.1      | Literaturverzeichnis .....  | 177        |
| 9.2      | Filmographie in alphabetischer Reihenfolge nach deutschen Titeln .....                      | 182        |
| 9.3      | Links zu den Versuchsvideos und zu den Hörbeispielen<br>Psychoakustik und Sounddesign ..... | 183        |
|          | <b>Index .....</b>  | <b>187</b> |



---

# Vorwort

---

Sounddesign ist das tonale Erzählen von Geschichten, die im dreidimensionalen Raum erlebt werden können. Wir nehmen diese Geschichten über unseren Hörapparat wahr. Unsere Ohren sind unser Zeitorgan, beim Hörerleben fallen Vergangenheit und Zukunft in der Gegenwart zusammen. Damit lässt uns unser auditiver Fernsinn – in Ergänzung zu unserem Sehorgan – unsere Umwelt in ihren drei Raumdimensionen wesentlich intensiver erleben, als es Vielen bewusst sein mag.

Dieses Buch erzählt unter anderem die über hundertjährige Geschichte des Sounddesigns für die Medien- und Produktindustrie und wie diese Industrien unsere auditive Wahrnehmung und unsere Kultur des Hörens verändert haben.

Die Basis für ein bewusstes Hör-Erleben und Hör-Verstehen ist das Wissen, wie unser Hörsinn eigentlich funktioniert und wie sich unser Hörgedächtnis im Laufe unseres Lebens formt und somit unser alltägliches Handeln beeinflusst.

Mithilfe eines nach aktuellen psychoakustischen Erkenntnissen gestalteten und durchgeführten Hörversuchs mit 240 Probanden werden im letzten Buchabschnitt Grundsätze für ein immersives und wirkungsvolles Sounddesign für Produkte und Anwendungen sowie für die Tonspurgestaltung audiovisueller Medien formuliert.

Viele der kursiv gesetzten Fachausdrücke werden in einem Glossar in Kapitel 8 erläutert und zudem werden einige Begriffe, die für ein besseres Verstehen einen auditiven Zugang einfordern, für ein aktives Hör-Erleben mit vertonten Videos verbunden. Die Video-Links und die näheren, jeweiligen Erläuterungen dazu sind in Abschnitt 9.3 zu finden.

München, im September 2025

Maximilian Kock



---

# 1

## Einleitung

---

*„Das Auditive hat gegenüber dem Visuellen den Vorteil, dass es Raum lässt für Gedanken, für Fantasie, für Nebentätigkeiten. Das Auditive lässt uns in der Welt sein.“*

Bernhard Heckler [Heckler, 2025]

Das Hören erschließt uns die Welt um uns herum, sie wird durch unsere beiden Ohren dreidimensional erlebbar und damit im Wortsinne begreifbar. Aufgrund akustischer Signale und deren Reflexionen können wir uns im dreidimensionalen Raum orientieren, die Richtungen von Signalen einordnen und die Entfernungen von Tonquellen abschätzen.

Neben dem Sehen ist das Hören der zweite Fernsinn des Menschen. Die weiteren drei Sinne Tasten, Riechen und Schmecken sind dagegen unsere Nahsinne. Erst wenn man mit einer (fremden) Person oder einer neuen Situation vertraut ist, lässt man diese sprichwörtlich nahe an sich herankommen, um diesen Menschen oder dieses Ereignis mit all seinen Sinnen – besonders den Nahsinnen – hautnah erfahren zu können. Eine Annäherung an ein Lebewesen, einen Gegenstand oder an eine neue Situation setzt daher zunächst eine visuelle und akustische Abklärung und Absicherung in einiger Entfernung durch unseren Seh- und Hörapparat voraus, um sich dann auch nach einer weiteren Annäherung sicher fühlen zu können.

Die Fern- oder Distanzsinn Hören und Sehen sind im heutigen Zeitalter der in den elektronischen Medienkanälen proliferierenden computergenerierten Virtualität unsere elementaren Fenster zur Welt geworden. Durch diese Fenster wird die Welt um uns herum von der Mehrheit der Menschen betrachtet. Aber leider wird sie seltener er-hört, da viele Menschen die Erfahrung des Sehens der des Hörens vorziehen.

Denn das (Zu-)Hören erfordert manchmal einige wenige Minuten an *Zeit*, die viele von uns heutzutage nicht mehr übrig zu haben scheinen. Die derzeitige Dominanz des Sehens zeigt sich auch darin, dass wir alle nahezu ständig über Schrift- und (auf-

gezeichneten) Sprachnachrichten zeitversetzt untereinander kommunizieren, aber das Realzeit-synchrone Telefonieren für einige Betroffene inzwischen eine psychische Belastung darstellt, nicht selten sogar Angstzustände auslösen kann [Fischer, 2023].

Unser Gehör ist primär ein Zeitorgan, das Hören lässt uns Menschen das Gegenwärtige und somit den Ablauf von Zeit erleben und folglich unsere Lebenszeit gestalten. Wir Menschen erfahren Zeit, indem wir unser akustisches Umfeld er-hören, indem wir Töne und Klänge in einem für uns sinnvollen Zeitgerüst von Ursache und Wirkung anordnen und uns so im dreidimensionalen Raum und im eindimensionalen Zeitablauf orientieren (siehe Abschnitt 2.6). Denn wir können mit unseren Ohren jeden Raum aufgrund der Schallreflexionen seiner Wände, die zeitlich verzögert bei uns eintreffen, in seiner Größe und Ausstattung auditiv erfassen.

Unser Fernsinn Sehen hilft uns dabei weniger, da das für diesen Sinn notwendige Licht wegen seiner enorm hohen Ausbreitungsgeschwindigkeit uns Menschen eine (falsche) Gleichzeitigkeit suggeriert.

Das auditive Erleben und Einordnen des ständigen Tönens um uns herum benötigt folglich einen gewissen Erlebenszeitraum. Dieser Erlebenszeitraum kann, im Gegensatz zum Betrachten eines Bildes oder beim Lesen von Texten, nicht durch ein schnelles Überfliegen verkürzt werden. Denn gesprochene Texte werden in dem Fall einer verkürzten Wahrnehmung nicht mehr verstanden und Musik verliert ihre Aussage und ihre emotionale Wirkung. Das kann jede(r) feststellen, wenn sie oder er Musik oder *Sprache* in überhöhter Abspielgeschwindigkeit wiedergibt.

Ein umfassendes auditives Wahrnehmen benötigt folglich Realzeit und am besten ein Erleben in zeitsynchroner Präsenz. Unser Hörorgan hilft uns so, in der Gegenwart zu leben, weil Töne – und besonders Musik – im Jetzt passieren. Auch viele psychologische Ratgeber empfehlen, dass ein aktives Zuhören in Gesprächen besser bei völliger geistiger und körperlicher Präsenz gelingt und erst so die anerkennende Wertschätzung meines Gegenübers möglich ist.

Denn Telefonierende bemerken sofort eine fehlende (geistige) Präsenz der Gesprächspartnerin am anderen Ende der Verbindung, sie nehmen die abnehmende Aufmerksamkeit über die Veränderung der Klangfarbe der anderen Stimme wahr.

Das kann eventuell auch die oben genannten Angstzustände beim zeitsynchronen Telefonieren erklären: Die für dieses Kommunikationsmedium notwendige ständige geistige Präsenz – auch aufgrund der Erfordernis einer in Telefongesprächen unmittelbar nötigen, zeitnahen (mündlichen) Reaktion – kann für einen Telefonierenden eine so große psychische Belastung werden, dass das Telefon als Kommunikationskanal generell gemieden wird.

Die Nahsinne Fühlen, Riechen und Schmecken haben in unserer digitalen Erlebenswelt bedauerlicherweise keine Bedeutung mehr. Oder sie haben noch keine Bedeutung für uns vornehmlich digital kommunizierende Menschen. Unsere damit auf le-

diglich zwei Fernsinne (Mediensinne) reduzierte Lebenserfahrung lässt uns auf diese Weise emotional verarmen: Wir leben in virtuellen Filterblasen, in die nur Bilder und Töne zeitversetzt eindringen können.

Wir verharren somit in einer künstlichen Verkapselung – mit Bildschirmen umgeben – allein mit uns selbst: im Auto, im Büro, zu Hause, mit Mobilgeräten, mit VR-Brillen und Kopfhörern. Um uns herum sind fast ständig digitale Bild- und Tonerzeuger, mit denen wir immer mehr Zeit alleine verbringen und dabei zunehmend vereinsamen.

Es ist also etwas völlig anderes, einen fremden Menschen, ein Ton- oder ein Bildereignis im persönlichen realen Nahfeld und in Realzeit (live) mit allen fünf Sinnen und in Präsenz zu erleben oder diese Ereignisse nur distanziert über elektronische Medienkanäle in Ton und Bild zeitversetzt zu rezipieren.

Ein gutes praktisches (analoges) Beispiel ist das Anschauen eines Filmes in einem Kino: Hier erlebe ich vor der Leinwand andere zuschauende, mitfiebernde Menschen, ich kann sie sehen und hören, unter Umständen auch riechen und fühlen. Zu Hause vor meinem eigenen Bildschirm bin ich dagegen häufig allein.

Auch die zeitversetzte Wiedergabe einer digitalen Aufzeichnung eines kulturellen Ereignisses ersetzt nicht das Erleben in Präsenz, denn ein elektronisches Erleben ist kein Erfahren mit allen unseren fünf menschlichen Sinnen. Dies könnte eine naheliegende Erklärung sein, die nicht selten auftretende Abgestumpftheit und Gleichgültigkeit der Digital Natives gegenüber ihren Mitmenschen in der realen Umwelt zu deuten und auch eventuell besser zu verstehen. Auch die immer häufiger auftretenden Angststörungen, könnten eine Folge dieser proliferierenden digitalen, asynchronen Kommunikation sein.

Elementar für unser tatsächliches Real-Erleben ist auch – was in den folgenden Abschnitten dieses Buchs noch detaillierter dargelegt werden wird – die inner-synchrone Wahrnehmung eines auditiven und visuellen Ereignisses. Sounddesigner bezeichnen dies als *Sound-Diegeese*: „See a dog, hear a dog.“ Ein nicht diegetisches und synchrones Wahrnehmen von Hör- und Sehereignissen – und das ist bei den digitalen Medien nicht selten der Fall – irritiert uns und lässt uns an der Realität, an der Übertragungstechnik oder gelegentlich sogar an unseren eigenen Sinnen (ver-)zweifeln. Dieses Thema wird unter anderem in Abschnitt 2.6 „Sounddesign und künstliche Intelligenz“ näher erörtert.

Mit Beginn der 1990er-Jahre bestimmten bei der Produktion von audiovisuellen Medien und der audiovisuellen Produktgestaltung immer mehr digitale computerbasierte Systeme (die sogenannten Workstations) den Produktionsprozess. Die Arbeitsgänge des Kreierens von Sound – *Geräusche*, Klangebene und Tonatmosphären – sowie auch von passender Musik vereinfachten und beschleunigten sich wesentlich. Durch die Digitalisierung, beginnend in den 1980er-Jahren, und zusätzlich seit den späteren 2010er-Jahren durch die *künstliche Intelligenz* wurde der Einsatz von Sound-

design in den audiovisuellen Medien und bei der Produktgestaltung so populär, dass auch Marketingfachleute sich mit der Wirkung von *Soundeffekten* in der Tonspur und bei der Gestaltung (digitaler) Produkte intensiv zu beschäftigen begannen.

Denn auch aufgrund dieser vor mehr als vierzig Jahren in den klassischen Kommunikationsmedien startenden technologischen Entwicklung wird Sounddesign heutzutage nicht nur in den audiovisuellen Medien eingesetzt, sondern fast überall praktiziert: in der Industrie (z. B. bei Automobilen und Maschinen), im Produktdesign (z. B. für Haushaltsgeräte, Verpackungen für Lebensmittel) und in der Mobilitätsbranche (z. B. beim Reisen: Navigationssysteme, Bahnhofs- und Zugansagen, mobile Kommunikation).

Folgende zentrale Fragen kamen hinsichtlich der Produktion und des effektiven Einsatzes von *Sounddesign* aufseiten der Gestalter und Nutzer auf:

- Wie wirkt Sounddesign genau auf den Zuhörer? Was löst es für Emotionen, Erinnerungen und Reaktionen aus?
- Was ist wirkungsvoller: Sounddesign (Soundeffekte) oder Musik (auch hinsichtlich des Mischungsverhältnisses beider Elemente)?
- Wie authentisch kann und soll Sounddesign sein? Was ist überhaupt ein authentisches Sounddesign?
- Wie viele Soundeffekte sind sinnvoll und notwendig? Wann ist die Tonspur oder ein Produkt mit Sounddesign überfrachtet?

Mit dem Erfolg der *Digitalen Audio-Workstations (DAWs)* geriet der Ton als Gestaltungselement zumindest bei den audiovisuellen Medien damit zum vierten Mal in seiner mittlerweile über hundertjährigen Geschichte in den Fokus der Konsumenten:

- Zum ersten Mal nach dem kommerziellen Erfolg des Tonfilms Ende der 1920er-Jahre,
- zum zweiten Mal bei der Etablierung des Stereo-Kinotons in den 1950er-Jahren und
- zum dritten Mal durch die Einführung des analogen Mehrkanal-Kinotons in den 1970er-Jahren.

Besonders in der TV-Produktion war die Tonspur bis in die 1990er-Jahre eher als ein Fernsehbegleitton auf ihre bloße Funktion der Sprachübertragung begrenzt worden. Das eigentliche gestalterische Potenzial des *Audiodesigns* wurde bis dahin nur selten in vollem Umfang genutzt.

Dabei war jedem (TV-)Produzenten schon damals mehr oder weniger bewusst, dass das audiovisuelle Medium eine einzigartige Chance bot, einem Zuschauer eine Botschaft optimal zu übermitteln. Kein anderes Medium ermöglicht es, eine Information gleichzeitig über zwei der wichtigsten Kommunikationskanäle des Menschen, das Sehen und das Hören, zu verbreiten.



Die Übertragung von Video ist allerdings in ihrer geometrischen Dimension und in ihrer Wirkung immer limitiert. Der Bildschirm ist nur ein Fenster, selbst eine große Kinoleinwand ist in ihren Ausmaßen immer begrenzt. Zudem kann der Zuschauer als Empfänger die Augen schließen, seinen Kopf abwenden und damit die Wahrnehmung und Aufnahme einer visuellen Botschaft blockieren.

Die Audioübertragung dagegen vermag es, den Empfänger *räumlich* zu umhüllen: Ein gutes Design der Tonspur wirkt immersiv. Dem Ton können die Zuhörer nicht ausweichen, das Hörorgan ist immer empfangsbereit.

Audio kann damit in seiner Wirkung räumliche Grenzen überwinden. Lediglich die Aufmerksamkeit des Rezipienten, sein Achtgeben sollte nach bisherigen Erfahrungen von Sounddesignern durch eine möglichst interessant gestaltete Tonspur gewährleistet sein [Sonnenschein, 2001; Holman, 2010; Raffaseder 2010].

Dass die Audioebene die Bildebene manchmal in ihrem Aufmerksamkeitspotenzial sogar übertreffen kann, verdeutlicht die folgende, typische Situation in einem Fernsehbeitrag eines TV-Reporters: Eine Zuschauerin kann dessen Bericht nicht oder nur mühsam folgen, weil im Off, also außerhalb des Bildbereichs eine weitere gleichlaute Stimme zu hören ist. Die aufmerksame Zuschauerin fragt sich die ganze Zeit, wer dort (für sie unsichtbar) redet. Dabei verpasst sie den eigentlich sie interessierenden Tonbericht des Korrespondenten.

Hintergrund der vier weiter oben in dieser Einleitung gestellten Fragen nach der Wirkung von Sounddesign ist auch der Wunsch nach einer Effektivitätssteigerung bei der Konzeption und Produktion audiovisueller Botschaften. Lässt sich durch ein ausgeklügeltes Sounddesign ein höherer Aufmerksamkeitsgrad bei den Rezipienten bewirken? Verbessern gute Soundeffekte eventuell sogar grundsätzlich die Qualität bei der Kommunikation und bei der Beurteilung von Produkten?

In den letzten drei Jahrzehnten der fortschreitenden Digitalisierung wurde allerdings eine grundsätzliche Frage nicht beantwortet: Was sind die wesentlichen, die Rezeption des Zuschauers beeinflussenden Aspekte eines guten Sounddesigns für das Bewegtbild und in der Produktgestaltung? Gibt es dort eine Gestaltungsgrundlage wie zum Beispiel bei der visuellen Gestaltung das Gesetz des *Goldenen Schnitts*? Wirken die einzelnen Tonelemente, also Sprache, Musik und Soundeffekte auch alleine oder nur in der richtigen Kombination (in passender Mischung und korrekter zeitliche Anordnung zueinander) erfolgreich auf die Rezipienten ein?

Dass dabei die *Sprache* als Tonelement eine Sonderstellung einnimmt, ist naheliegend. Ein gesprochener Text muss bei einer Tonspur und bei einem kommunizierenden Gegenstand im Gegensatz zu Musik und Sounddesign zu verstehen sein. Hier sind besonders im Laufe der fast einhundertjährigen Geschichte des Tonfilms elementare Mischstandards entstanden [Raffaseder, 2010, S. 240 ff.; Tauchnitz, 1990, S. 8].

Aber wie ist das mit den beiden anderen Tonelementen Musik und *Soundeffekte*? Sind sie für einen größeren Aufmerksamkeits- und Wiedererkennungsgrad der Zuschauere-

rinnen und Zuhörer notwendig? Oder lenken sie eventuell eher von der Sprachebene ab?

Ein Ziel dieses Buchs ist die Entwicklung von (empirischen) Grundsätzen und die Bestätigung impliziten Wissens aus dem Sounddesign für den Fall gleichzeitiger Rezeption von Musik und Soundeffekten. Mithilfe von speziell konzipierten Hörversuchen mit Probanden soll eine wissenschaftliche Grundlage für die passende Gestaltung und Mischung der Tonelemente Soundeffekte und Musik formuliert werden.

Das zweite Kapitel wird einen Überblick über den derzeitigen Stand der Forschung der audiovisuellen Wahrnehmung in den Disziplinen des Sounddesigns (bei den audiovisuellen Medien und in der Industrie), der Psychoakustik, der Musikpsychologie, der Neurowissenschaften und der Filmwissenschaften geben. Auch der Einfluss künstlicher Intelligenz auf das Sounddesign und dessen mögliche qualitative Verbesserung durch sie wird in diesem Abschnitt thematisiert.

Im dritten Kapitel wird die generelle Problematik der adäquaten und damit immersiven Gestaltung der Tonspur beschrieben. Grundlage ist dabei die Annahme, dass eine stärkere *Immersion* einen stärkeren Eindruck, d. h. ein intensiveres Erleben der auditiv wahrgenommenen Informationen, beim Rezipienten bewirken kann.

Die Basis für das zweite und dritte Kapitel bilden

- die Quellen aus der Literatur des Sounddesigns und der Tongestaltung,
- das Maßstäbe setzende Sounddesign einiger Filme,
- das Design von Produkten, die unseren Umgang mit Technologie verändert und geprägt haben (Computermaus, Touchscreen, Assistenzsysteme, Haushaltsgeräte etc.).
- die Lehren aus der Gestaltpsychologie, aus der Musikpsychologie und aus den Neurowissenschaften,
- die Literatur über die Geschichte des Tonfilms und seiner Entstehung in den 20er-Jahren des letzten Jahrhunderts,
- sowie die über dreißigjährige Berufserfahrung des Autors im Audiodesign.

Das vierte Kapitel stellt die Versuchsmethodik und die Versuchsdurchführung vor. Das fünfte Kapitel nennt die Versuchsergebnisse, deren statistische Auswertungen und Visualisierungen. Im sechsten Kapitel werden die Ergebnisse aus den Hörversuchen diskutiert. Mit den empirisch fundierten Grundsätzen für ein immersiveres und auch damit kommunikativeres und eventuell beeindruckenderes Sounddesign schließt dieses Buch.

Wir leben momentan in einer Zeitenwende, in der sich die von uns geschaffenen automatisierten Systeme immer mehr verselbstständigen, siehe Abschnitt 2.6. Gerade in dieser Zeitenwende des immer bedeutsamer werdenden Einsatzes von Algorithmen der künstlichen Intelligenz (KI) im Alltag des Suchens, Findens und des automatisierten (und des dadurch leider häufig zu gleichförmigen) Gestaltens bleibt das gegenseitige Zuhören unter uns Menschen elementar.

Künstliche Intelligenz fördert und unterstützt, wie jede von uns geschaffene Maschine, zunächst die Bequemlichkeit des Menschen: Das mühsame selbstständige Suchen, Finden und Kreieren – auch im Sinne einer tieferen und nachhaltigeren persönlichen Erkenntnis – wird ersetzt durch das einfachere Eingeben von Suchanfragen und sogenannte Prompts mit fast gleichzeitiger, häufig oberflächiger und standardisierter Beantwortung bzw. Umsetzung.

Der Einsatz von künstlicher Intelligenz führt zunächst einerseits zu einer Reduktion von Fehlern und einer Vermeidung von Fehlentscheidungen.

Andererseits basiert jede menschliche Erkenntnis zunächst auf Fehlern, auf Zweifeln und auf Fehlversuchen (siehe Abschnitt 2.6). Diese Wege des Zweifels und der Fehler werden beim Nutzen intelligenter Algorithmen aber nicht mehr beschritten. Jegliches leise Zweifeln wird durch die prompte Reaktion der Algorithmen erstickt.

Ein Forschen, ein Nachfragen, ein Bezweifeln einer Antwort ist nahezu unmöglich, denn eine KI-Maschine ist prinzipiell ohne Selbstzweifel. Auch weil diese automatisierten, rekursiven Systeme (noch) kein Selbstbewusstsein aufweisen. Und erst ein Zweifeln ist eine Basis für ein gesundes, d. h., sich auch infrage stellendes Selbstbewusstsein. Und dieses Zweifeln – neben unserer Begabung, intuitiv Lösungen zu finden – unterscheidet unsere menschlichen Denkprozesse (noch) von denen der intelligenten Algorithmen, siehe den Exkurs in Abschnitt 2.6.

Eine Selbstbewusstwerdung, ein Prozess über das eigene Zweifeln und Fragen hinweg, der schließlich zu einem gesunden Selbstbewusstsein führt, macht uns Menschen erst menschlich. Ohne unsere täglichen gesunden Selbstzweifel werden wir zur Maschine und damit zu einem unmenschlichen Tyrannen. Es ist daher zu fragen, ob KI-Systeme uns Menschen wirklich helfen und uns nicht eher mit ihren schnellen und oberflächigen Standardantworten in unserer Entscheidungsfreiheit einengen, tyrannisieren oder – im Gegensatz dazu – ruhigstellen.

Wir dürfen nicht unterschätzen, dass ein Erkenntnisgewinn meistens auf zuvor auftretenden Fehlern oder Fehlanwendungen und Irrwegen beruht: Wer keine Fehler macht, lernt nichts dazu.

KI-Systeme arbeiten vornehmlich bild- und textbasiert, selten tonbasiert. Die Erklärung dafür mag darin liegen, dass das Hören im Gegensatz zum Sehen ein Sinn des Zweifels und des Nachfragens ist: „Ich habe mich wohl verhört“, sagt man eher als „Ich habe mich wohl ver-sehen“.

Zuhören, d. h., das alltägliche Aufnehmen und Einordnen der Signale aus unserem klingenden dreidimensionalen Umfeld, macht uns folglich erst zum Menschen; zu einem Menschen, der im gegenwärtigen Jetzt lebt und sich unaufhörlich (in) Frage(n) stellt.

Dieses Buch möchte nach den Vorstellungen und Wünschen seines Autors ein Plädoyer für das Hören und das Hör-Erleben sein, auch um das bewusstere Wahrnehmen und ganzheitliche Erleben unserer akustischen (Um-) Welt zu fördern.

