

3D Audio Sound Branding

Neue Dimension in der akustischen Markenkommunikation

Markus Sternbauer

Universität für Musik und darstellende Kunst Graz
Institut für Elektronische Musik und Akustik

FH JOANNEUM University of Applied Sciences
Institute of Design & Communication

Betreuer: Univ. Prof. Dipl.-Ing. Dr.techn. Alois Sontacchi

zur Erlangung des akademischen Grades
Master of Arts (M.A.)

September 2020

Eidesstattliche Erklärung

Ich erkläre ehrenwörtlich, dass ich die vorliegende Masterarbeit selbstständig angefertigt und die mit ihr verbundenen Tätigkeiten selbst erbracht habe. Ich erkläre weiters, dass ich keine anderen als die angegebenen Hilfsmittel benutzt habe. Alle aus gedruckten, ungedruckten oder dem Internet im Wortlaut oder im wesentlichen Inhalt übernommenen Formulierungen und Konzepte sind gemäß den Regeln für gutes wissenschaftliches Arbeiten zitiert und durch Fußnoten bzw. durch andere genaue Quellenangaben gekennzeichnet.

Die vorliegende Originalarbeit ist in dieser Form zur Erreichung eines akademischen Grades noch keiner anderen Hochschule vorgelegt worden. Diese Arbeit wurde in gedruckter und elektronischer Form abgegeben. Ich bestätige, dass der Inhalt der digitalen Version vollständig mit der gedruckten Version übereinstimmt. Ich bin mir bewusst, dass eine falsche Erklärung rechtliche Folgen haben kann.

Graz, am 25.08.2020

A handwritten signature in black ink, appearing to read 'Markus Sternbauer', written in a cursive style.

Markus Sternbauer

Danksagung

An dieser Stelle möchte ich mich bei meinem Betreuer Alois Sontacchi für die Hilfestellungen und die wertvollen Gespräche über die letzten zwei Jahre im Rahmen der Betreuung dieser Masterarbeit und darüber hinaus bedanken.

Außerdem möchte ich mich bei den Teilnehmer*innen der Hörversuche bedanken, die sich die Zeit nahmen und mir in konstruktiven Gesprächen Feedback gaben.

Ebenfalls möchte ich mich bei meinen Studienkolleg*innen bedanken, die durch eine ganz besondere Gruppendynamik die letzten zwei Studienjahre kurzweilig, unglaublich lustig und bereichernd gemacht haben.

Ganz besonders möchte ich mich bei meiner Familie bedanken, die mir trotz Umwegen Vertrauen und jegliche Form der Unterstützung zukommen ließ.

Zuletzt möchte ich mich bei Eva bedanken, die mit viel Geduld und Verständnis meinen Weg hierher begleitet hat und mir immer mit Rat und Tat zur Seite steht.

Abstract

Recent developments in 3D audio technology have not only attracted more attention from a broader audience, but also enable easier integration into everyday life. This applies to both binaural playback and the 3D audio speaker systems. Furthermore, it creates the opportunity for acoustic brand communication (sound branding) to establish a new touchpoint and experience for customers. This master thesis explores the creative possibilities and difficulties that emerge.

In the theoretical part the directional and distance perception of the human hearing system is explained in acoustic and psychoacoustic terms. The visual component in hearing perception is also addressed.

In the practical part the production process of several sound logos for the IKO, a 20-sided speaker system, is documented. Different movement patterns and sound objects were created using typical sound materials. These were examined by test persons for their comprehensibility and impact.

The listening tests confirm not only a spectral dependency in the localization but also differences in the perception of the spatial dimension. Besides certain design criteria, factors influencing the subjective assessment could also be found. In addition, the influence of visual cues on more complex auditory objects is confirmed.

Kurzfassung

Neuste Entwicklungen in der 3D Audio Technologie haben nicht nur für mehr Aufmerksamkeit eines breiteren Publikums gesorgt, sondern ermöglichen eine zunehmend einfachere Integration in den Alltag. Das betrifft sowohl die binaurale Wiedergabe als auch die 3D Audio fähigen Lautsprechersysteme. Daher entsteht für die akustische Markenkommunikation (Sound Branding) die Möglichkeit, einen neuen Kontaktpunkt mit Kunden zu etablieren. Welche gestalterischen Freiräume und Schwierigkeiten dabei entstehen, wird in dieser Masterarbeit ausgelotet.

Im theoretischen Teil wird die Richtungs- und Distanzwahrnehmung des menschlichen Gehörs in akustischer und psychoakustischer Hinsicht erläutert. Dabei wird ebenfalls die visuelle Komponente in der Hörwahrnehmung thematisiert.

Im praktischen Teil wird der Produktionsprozess mehrerer 3D Audio Sound Logos für den IKO, ein 20-seitiges Lautsprechersystem, dokumentiert. Unter der Verwendung gängiger Klangmaterialien wurden unterschiedliche Bewegungsmuster und Klangobjekte kreiert. Diese wurden durch Proband*innen auf ihre Prägnanz und Wirkung untersucht.

Die Hörversuche bestätigen neben einer spektralen Abhängigkeit in der Lokalisation auch Unterschiede in der Wahrnehmung der räumlichen Dimension. Es konnten neben gestalterischen Kriterien auch Einflussfaktoren auf die subjektive Bewertung gefunden werden. Außerdem bestätigt sich der Einfluss durch visuelle Anker auf komplexere auditive Objekte.

CONTENT

EIDESSTATTLICHE ERKLÄRUNG	2
DANKSAGUNG.....	3
ABSTRACT	4
KURZFASSUNG	5
1 EINLEITUNG.....	8
1.1 IMMERSION	9
2 GRUNDLAGEN DES RÄUMLICHEN HÖRENS.....	10
2.1 HORIZONTALE EBENE.....	11
2.2 VERTIKALE EBENE.....	13
2.3 DISTANZWahrnehmung	14
2.4 AUßENOHRÜBERTRAGUNGSfunktion.....	15
3 VISUELLER EINFLUSS.....	16
4 RÄUMLICHES HÖREN MEHRERER SCHALLQUELLEN	18
4.1 PRÄZEDENZEFFekt	18
5 BEAMFORMING.....	20
6 3D-AUDIO FORMATE	22
6.1 KANALBASIERt.....	22
6.2 OBJEktBASIERt.....	23
6.3 SZENENBASIERt	24
6.3.1 <i>Ambisonics</i>	24
6.4 MPEG-H.....	27
7 SOUND BRANDING	27
7.1 SOUND DNA	28
7.2 NEUE AKUSTISCHE INTERAKTIONEN	29
8 3D AUDIO UND EMOTIONEN	29
9 WERKSTÜCK	30
9.1 KLANGKONZEpt	31
9.2 IKOSAEDER (IKO).....	31
9.3 PRoDUKTIONSUMGEBUNG	33

9.4	ARBEITSUMGEBUNG	37
9.5	SOUND LOGOS	38
9.5.1	<i>Sound Logo 1</i>	39
9.5.2	<i>Sound Logo 2</i>	40
9.5.3	<i>Sound Logo 3</i>	41
9.5.4	<i>Sound Logo 4</i>	42
9.6	ZUSAMMENFASSUNG DER GESTALTUNGSMAßNAHMEN.....	44
9.6.1	<i>Harmonische Struktur und Tonhöhe</i>	44
9.6.2	<i>Lautstärke</i>	45
9.6.3	<i>Ambisonische Ordnung</i>	46
9.6.4	<i>Maskierung von Transienten</i>	46
9.6.5	<i>Simultane Klänge</i>	46
9.6.6	<i>Akzente</i>	47
10	HÖRVERSUCHE	47
10.1	PROBANDEN UND ABHÖRRAUM	48
10.2	METHODE	48
10.3	ABLAUF.....	50
10.4	RESULTATE	51
11	DISKUSSION.....	52
12	ZUSAMMENFASSUNG UND AUSBLICK.....	55
	LITERATURVERZEICHNIS	58
	ONLINE QUELLEN.....	67
	ABBILDUNGSVERZEICHNIS.....	69
	ANHANG – SOUND LOGOS.....	71
	FRAGEBOGEN HÖRVERSUCH.....	72

1 Einleitung

“Tomorrow belongs to those who can hear it coming”. Dieses berühmte Zitat von David Bowie erweist sich in der akustischen Markenkommunikation als zutreffender denn je, da besonders in Anbetracht zunehmender Interaktionen über Sprachassistenten der Fokus stärker auf die auditive Wahrnehmung rückt. Einhergehend mit diesem Trend, ermöglicht die zunehmend versierte Technik von 3D Audio-fähigen Wiedergabegeräten eine einfachere und kontinuierlich steigende Integration in den Alltag. Daher gilt es, frühzeitig die Gestaltungsmöglichkeiten zu erkennen und zu nutzen.

Das Ziel dieser Arbeit ist es, durch die Konzipierung und Umsetzung praktischer Beispiele Einblick in einen möglichen Entwicklungsprozess der akustischen Markenkommunikation durch eine 3D Klangwiedergabe zu bekommen. Dabei werden künstlerische Empfehlungen herausgearbeitet, wie mit Hilfe des Audioformats Ambisonics bestimmte Klangobjekte unterschiedlicher Komplexitätsstufen dynamisch im Raum bewegt werden können.

Im ersten Teil der Arbeit werden die theoretischen Grundlagen für die akustische und psychoakustische Richtungs- und Distanzwahrnehmung genauer erläutert. Ebenso wird in diesem Kontext auf den visuellen Einfluss in der auditiven Wahrnehmung eingegangen. Außerdem werden die gängigen Formate der 3D Audio Wiedergabe näher beleuchtet, einhergehend mit einer Erörterung der Vor- und Nachteile der jeweiligen Konzepte. Mit der Einführung einiger wichtiger Prinzipien des Sound Brandings wird der Übergang zum praktischen zweiten Teil geschaffen.

Darin wird ein künstlerischer Beitrag geleistet, in dem der Arbeitsprozess an fiktiven 3D Audio Sound Logos dokumentiert wird. Anhand dieses Entstehungsprozesses werden Möglichkeiten und Schwierigkeiten beschrieben und ausgehend davon Gestaltungsmaßnahmen abgeleitet. Anschließend werden die intersubjektive Wahrnehmung dieser Sound Logos in Hörversuchen evaluiert und allgemeine Erkenntnisse präsentiert.

Diese Arbeit soll durch die Erstellung eines künstlerischen Leitfadens zur Erarbeitung von 3D Audio Konzepten einen wesentlichen Beitrag zur Weiterentwicklung der akustischen Markenkommunikation leisten und durch neue Erkenntnisse weitere Forschungsfragen stellen.

1.1 Immersion

Der Begriff der Immersion wurde in den letzten Jahren in unserem Sprachgebrauch immer prominenter und führte je nach Kontext zu einer unterschiedlichen Auslegung. Da Immersion sowohl im Audibereich als auch im Kontext der Markenkommunikation ein gern verwendetes Schlagwort ist, wird der Begriff eingangs beleuchtet, um in weiterer Folge einer begrifflichen Verwirrung in dieser Arbeit vorzubeugen.

Mit zunehmender Reichweite von *Virtual Reality* und *Augmented Reality* wurde der Begriff häufig im Zusammenhang mit dem Eintauchen in digitale Umgebungen verwendet.¹ Dem entspricht auch die Definition von *Oxford Dictionary*²: „*generating a three-dimensional image which appears to surround the user*“.

Roginska und Geluso³ beschreiben die Immersion im Audibereich als ein Gefühl, dort zu sein und dass Immersion das Ziel verfolgt, eine Umhüllung in einer dreidimensionalen Umgebung zu erreichen. Zudem definiert Kendall⁴ Präsenz und Umhüllung als wichtige Eigenschaften für die Immersion. Daraus lässt sich bereits eine enge Verstrickung der Begriffe Immersion und Umhüllung erkennen. Mit steigender Popularität von 3D Audio wird in diesem Kontext gerne von immersiven Audioformaten gesprochen und führt daher zu einer synonymen Verwendung der Begriffe Immersion und Umhüllung.⁵

Laut Berg⁶ lässt sich Immersion nicht so einfach mit einer Umhüllung von Klängen erklären, da zusätzlich zur auditiven Wahrnehmung auch andere Einflüsse die Komplexität erhöhen. Diese Komplexität wird durch unterschiedliche Konzepte und Arten der Immersion von Zhang, Perkis und Arndt⁷ beschrieben. Ebenso werden in dieser Untersuchung ein kognitives Eintauchen durch eine emotionale Immersion (*Emotional Immersion*) und ein sensorisches Eintauchen durch eine räumliche Immersion (*Spatial Immersion*) gegenübergestellt. Dabei zeigt sich, dass die emotionale Komponente der ausschlaggebendere Faktor ist, um eine Immersion zu erreichen.⁸

¹ Miller, Digital Storytelling.

² Stevenson, immersive.

³ Roginska und Geluso, Immersive Sound.

⁴ Kendall, „Spatial Perception and Cognition in Multichannel Audio for Electroacoustic Music.“

⁵ Agrawal, Simon, und Bech, „Defining Immersion: Literature Review and Implications for Research on Immersive Audiovisual Experiences.“

⁶ Berg, „The Contrasting and Conflicting Definitions of Envelopment.“

⁷ Zhang, Perkis, und Arndt, „Spatial immersion versus emotional immersion, which is more immersive?“

⁸ Zhang, Perkis, und Arndt.

Daher ist es für die Markenkommunikation wichtig, ein ganzheitliches, multisensorisches, aber auch kognitives Konzept zu entwickeln, um Kunden auf möglichst vielen Ebenen binden zu können. Dazu kann die auditive Komponente, besonders 3D Audio durch eine sensorische Einhüllung beitragen. In dieser Arbeit ist mit Immersion dieses eben beschriebene ganzheitliche, kognitive und multisensorische Konzept gemeint, hingegen wird bei der auditiven Immersion von einer Umhüllung gesprochen.

2 Grundlagen des räumlichen Hörens

Da sich das menschliche Gehör nicht vor den Umgebungsgeräuschen verschließen lässt, findet eine ständige Evaluierung der Klangumgebung statt. Die Fähigkeit, einzelne Schallquellen im Raum lokalisieren zu können, wird dadurch ständig geschärft.

Damit man mit 3D-Audio eine möglichst überzeugende Illusion der Wirklichkeit entwerfen kann, muss nicht nur die sehr komplexe Funktionsweise der räumlichen Wahrnehmung grundlegend verstanden werden, sondern auch jede klangliche Erweiterung perfekt in das natürliche Klangbild passen. Die Lokalisation einer Schallquelle im Raum findet dabei durch eine Richtungswahrnehmung auf der horizontalen und der vertikalen Ebene statt, sowie durch eine Entfernungswahrnehmung. Wenn man von der Lokalisierung bzw. Platzierung von Klängen spricht, werden die Begriffe Azimut für die horizontale und Elevation für die vertikale Ebene verwendet.⁹ Wie in Abbildung 1 zu sehen ist, wird für die genaue Bestimmung der Richtungswahrnehmung ein polares Koordinatensystem verwendet, wobei die Blickrichtung mit 0° Azimut und 0° Elevation beschrieben wird.¹⁰

⁹ Roginska und Geluso, Immersive Sound.

¹⁰ Roginska und Geluso.

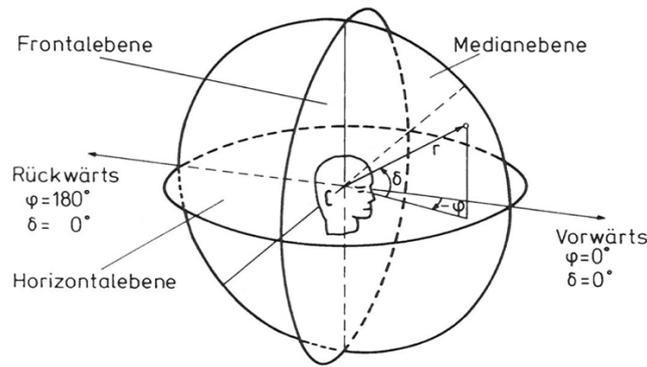


Abbildung 1: Koordinatensystem für die Beschreibung der Richtungswahrnehmung durch den Azimut auf der Horizontalebene und die Elevation auf der Medianebene.¹¹

In den nachfolgenden Kapiteln werden sowohl die Funktionsweise der Richtungswahrnehmung und der Distanzwahrnehmung für eine einzelne Schallquelle genauer beschrieben als auch die psychoakustischen Einflüsse bei der auditiven Wahrnehmung von mehreren Schallquellen in einem geschlossenen Raum, entsprechend den Untersuchungen von Blauert¹².

2.1 Horizontale Ebene

Der Richtungswahrnehmung auf der horizontalen Ebene liegt maßgeblich die von Lord Rayleigh beschriebene *Duplex-Theorie* aus dem Jahr 1907 zugrunde.¹³ Dabei sind für die Lokalisation der Schallquelle die Informationsunterschiede der beiden Ohren in Bezug auf Pegelunterschiede und der zeitlichen Verzögerung ausschlaggebend. Man spricht dabei von interauraler Pegeldifferenz (*Interaural Level Differences, ILD*) und interauraler Laufzeitdifferenz (*Interaural Time Differences, ITD*).

Wenn eine Schallquelle einen Azimut von 0° besitzt, treffen die Schallinformationen sowohl am rechten als auch am linken Ohr identisch ein. Sobald sich aber die Schallquelle auf der horizontalen Ebene auf eine Seite bewegt, kommt auf Grund der verlängerten Distanz die Schallinformation am gegenüberliegenden Ohr zeitlich verzögert an. Durch die zeitliche Verzögerung kommt es zu einer Phasenverschiebung, die ausschlaggebend für eine präzise

¹¹ Weinzierl und Verband Deutscher Tonmeister, Handbuch der Audiotechnik.

¹² Blauert, Spatial Hearing.

¹³ Rayleigh, „XII. On Our Perception of Sound Direction“.

Lokalisation der Schallquelle ist. Die präzise Bestimmung über die Laufzeitunterschiede ist zusätzlich an den Frequenzbereich der Schallquelle gekoppelt, da die halbe Wellenlänge größer sein muss als die maximale Wegstrecke um den Kopf.¹⁴ Daher ist die Richtungswahrnehmung durch die Laufzeitdifferenz nur bis zu einem Frequenzbereich um 1500Hz möglich.^{15 16}

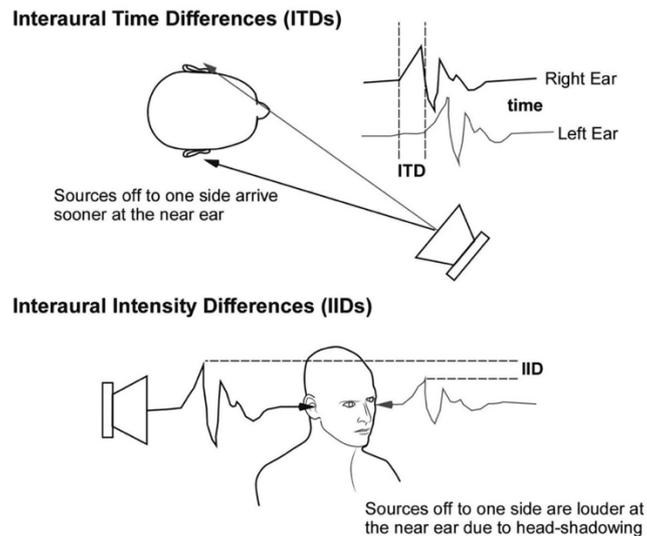


Abbildung 2: Veranschaulichung der Richtungswahrnehmung auf der horizontalen Ebene durch Pegel- und Laufzeitunterschiede¹⁷

Um den Frequenzbereich von 1500Hz wird je nach spektraler Information und Raumeinflüssen die Laufzeitdifferenz zu ungenau und die Richtungswahrnehmung erfolgt über die interaurale Pegeldifferenz.^{18 19} Da der Kopf in der Schallausbreitung ein Hindernis darstellt, verhalten sich die Schallwellen ab diesem Frequenzbereich auf Grund ihrer Wellenlänge unterschiedlich. Schallwellen unterhalb von 1500Hz können sich wegen der größeren Wellenlängen um den Kopf beugen und bewirken dadurch kaum eine Änderung der Druckamplitude. Höhere Frequenzen mit kürzeren Wellenlängen werden allerdings reflektiert

¹⁴ Sontacchi, „Dreidimensionale Schallfeldreproduktion für Lautsprecher- und Kopfhöreranwendung“.

¹⁵ Roginska und Geluso, Immersive Sound.

¹⁶ Sontacchi, „Dreidimensionale Schallfeldreproduktion für Lautsprecher- und Kopfhöreranwendung“.

¹⁷ Roginska und Geluso, Immersive Sound.

¹⁸ Brughera, Dunai, und Hartmann, „Human Interaural Time Difference Thresholds for Sine Tones“.

¹⁹ Rakerd und Hartmann, „Localization of Sound in Rooms. V. Binaural Coherence and Human Sensitivity to Interaural Time Differences in Noise“.

und gedämpft.²⁰ Dadurch tritt auf der gegenüberliegenden Kopfseite der sogenannte *Head-Shadow Effect*²¹ auf, wie er in Abbildung 3 visualisiert wird.

Somit verwendet das menschliche Gehör für die Richtungswahrnehmung auf der horizontalen Ebene, abhängig vom Frequenzbereich, eine Kombination aus interauralen Laufzeit- und Pegelunterschieden.

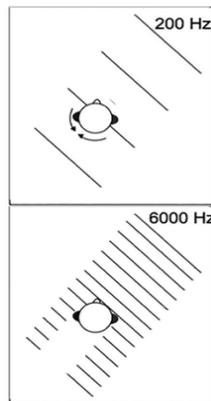


Abbildung 3: Head-Shadow Effect unterhalb und oberhalb des Grenzbereiches um 1500Hz.²²

2.2 Vertikale Ebene

Die interauralen Pegel- und Laufzeitunterschiede werden bei der Richtungswahrnehmung auf der vertikalen Ebene obsolet, da sie keine Information über die Elevation enthalten. Befindet sich der Klang auf einer beliebigen Position auf der Medianebene, ergibt sich ein identisches Signal auf beiden Ohren und sogenannte *monaurale*²³ Eigenschaften liefern vertikale Richtungsinformationen. Es handelt sich dabei um spektrale Interferenzen, die durch Auslöschung und Überhöhung der Schallreflexionen durch das Außenohr und durch den menschlichen Torso entstehen.²⁴ ²⁵ Durch die Mischung aus dem direkten Signal und den individuellen Reflexionen entsteht ein Kammfilter-Effekt²⁶, der je nach Elevation einem

²⁰ Sontacchi, „Dreidimensionale Schallfeldreproduktion für Lautsprecher- und Kopfhöreranwendung“.

²¹ Roginska und Geluso, Immersive Sound.

²² Roginska und Geluso.

²³ Blauert, Spatial Hearing.

²⁴ Sontacchi, „Dreidimensionale Schallfeldreproduktion für Lautsprecher- und Kopfhöreranwendung“.

²⁵ Weinzierl und Verband Deutscher Tonmeister, Handbuch der Audiotechnik.

²⁶ Beim Kammfiltereffekt kommt es zu Verzerrungen, indem durch Reflexionen und gebeugte Schallwellen einzelne Frequenzen ausgelöscht und angehoben werden, Friesecke, Die Audio-Enzyklopädie.

bestimmten Muster aus Spitzen und Einkerbungen (*peaks* und *notches*) folgt.²⁷ Diese Muster werden anschließend im Gehirn ausgewertet und einer Position zugeordnet.

Wie die Untersuchungen von Blauert²⁸ gezeigt haben, gibt es neben den individuellen Reflexionsmustern auch allgemeine frequenzabhängige Lokalisierungsmuster auf der Medianebene. Die sogenannten richtungsbestimmenden Frequenzbänder zeigen, dass beispielsweise ein Klang im Bereich von 8kHz tendenziell von oben und ein Klang mit 1kHz eher von hinten wahrgenommen wird.

Generell ist aber die Auflösung auf der Medianebene, im Speziellen über dem Kopf, deutlich unschärfer gegenüber der Richtungswahrnehmung auf der horizontalen Ebene.²⁹

2.3 Distanzwahrnehmung

Um eine möglichst realistische und einhüllende Umgebung zu kreieren, ist neben der Richtungswahrnehmung besonders die Wahrnehmung von Distanz und den natürlichen Reflexionen ein entscheidender Faktor, um Klängen einen entsprechenden Kontext zu geben. Für die Distanzwahrnehmung sind besonders die Lautstärkenverhältnisse und die Vertrautheit des Gehörten von Bedeutung.³⁰ Vorrangig wird die Lautstärkenintensität und die interpretierte Lautheit des Schallreizes verwendet, um die Distanz einschätzen zu können.³¹ Hierzu wurden von Blauert Eigenschaften formuliert, die in Abhängigkeit von der Entfernung den Klang beeinflussen.³²

- Bei einer mittleren Distanz von 3 – 15m bestimmt der Schalldruckpegel des Eingangssignals an den Ohren die Distanz zum Schallereignis. Im Freifeld nimmt dieser um 6dB bei jeder Verdoppelung der Distanz ab.
- Bei einer Entfernung von mehr als 15m wird durch den Luftwiderstand das Spektrum des Klanges verändert, indem die hohen Frequenzen stärker gedämpft werden als die niedrigeren Frequenzen.

²⁷ Middlebrooks und Green, „Sound Localization by Human Listeners.“

²⁸ Blauert, Spatial Hearing.

²⁹ Sontacchi, „Dreidimensionale Schallfeldreproduktion für Lautsprecher- und Kopfhöreranwendung.“

³⁰ Sontacchi.

³¹ Begault, „3-D Sound for Virtual Reality and Multimedia.“

³² Blauert, Spatial Hearing.

- Bei einer Entfernung von weniger als 3m kommt es auf Grund der Krümmung der Schallwellen am Kopf zu linearen Verzerrungen und somit ebenfalls zu spektralen Veränderungen.

Man bezieht sich vor allem auf diese Intensitätswahrnehmung bei weniger vertrauten Schallquellen, während in bekannten Klangumgebungen durch eine räumliche Einordnung der Klänge die reine Intensitätsbeurteilung durch Erfahrung überstimmt wird.³³ Ein bekanntes Beispiel bietet die Interpretation unterschiedlicher Sprechformen. Entsprechend unserer Erfahrung wird ein Flüstern unabhängig des Schallpegels näher und intimer wahrgenommen als die normale Sprechstimme.³⁴

Ein weiteres wichtiges Maß zur Einschätzung der Entfernung einer Schallquelle in einer reflektierenden Umgebung ist das Verhältnis zwischen Direktschall und den Schallreflexionen.³⁵ Je größer die Distanz zur Schallquelle wird, desto geringer wird der Anteil des Direktschalls gegenüber den diffusen Reflexionen. Sobald die Distanz des Hallradius³⁶ erreicht wird, überwiegt der Pegel des entfernungsunabhängigen Diffusanteils.³⁷ Außerdem liefert der zeitliche Versatz zwischen dem Direktschall und den ersten Reflexionen (*Initial Time Delay Gap*) einen weiteren Anhaltspunkt für die Abschätzung der Distanz.³⁸ Je größer dieser zeitliche Versatz ist, desto kürzer ist die Distanz zur Schallquelle gegenüber dem zurückgelegten Weg der Reflexionen.

2.4 Außenohrübertragungsfunktion

Bei der Außenohrübertragungsfunktion oder der *Head-Related Transfer Function* (HRTF) werden die oben erwähnten Faktoren der Richtungswahrnehmung für den dreidimensionalen Raum zusammengefasst.³⁹

³³ Begault, „3-D Sound for Virtual Reality and Multimedia.“

³⁴ Roginska und Geluso, Immersive Sound.

³⁵ Roginska und Geluso.

³⁶ Der Hallradius ist der Abstand zur Schallquelle, bei dem Direktschall- und Diffusschallenergiedichte gleich groß sind.

Weinzierl und Verband Deutscher Tonmeister, Handbuch der Audiotechnik.

³⁷ Sontacchi, „Dreidimensionale Schallfeldreproduktion für Lautsprecher- und Kopfhöreranwendung.“

³⁸ Füg, Werner, und Brandenburg, „Controlled Auditory Distance Perception Using Binaural Headphone Reproduction – Algorithms and Evaluation.“

³⁹ Boren und Roginska, „The Effects of Headphones on Listener HRTF Preference.“

Bei einer binauralen Reproduktion von 3D-Audio über Kopfhörer versucht man Klänge mit einer möglichst realistischen Nachahmung dieser Übertragungsfunktion als virtuelle Schallquellen im Raum zu platzieren. Da vor allem die spektrale Filterung sehr sensibel auf die Physis einer einzelnen Person abgestimmt ist und besonders für die Medianebene bzw. die vorne/ hinten Unterscheidung wichtig ist, müsste man, um eine möglichst realistische Illusion zu bekommen, für jede Person eine eigene HRTF anfertigen.⁴⁰ Da sich diese allerdings noch nicht in einem schnellen Verfahren individuell anpassen lässt, wird versucht einen Standard zu finden, der für viele Personen zu einer möglichst realistischen Wiedergabe führt.

Ein Hauptaugenmerk liegt dabei auch auf der Distanzwahrnehmung der Schallquelle und den raumakustischen Eigenschaften, da ansonsten der Klang im Kopf (*Inside-the-head locatedness, IHL*) platziert wird und durch eine fehlende Externalisierung die Illusion einer natürlich Klangumgebung nicht überzeugend funktioniert.^{41 42}

3 Visueller Einfluss

Wie bei der Distanzwahrnehmung (siehe Kapitel 2.3) bereits angemerkt, gibt es Diskrepanzen zwischen der Verarbeitung der akustischen Reize in isolierten Untersuchungen und der alltäglichen Erfahrung. Dabei spielt vor allem die multimodale, respektive die visuelle Wahrnehmung eine wichtige Rolle. So entspricht die Distanzwahrnehmung einem lebenslangen Lernprozess audiovisueller Beobachtungen von Intensitätsunterschieden in Korrelation mit der physikalischen Bewegung einer Klangquelle.⁴³ Auch zeitliche Muster zwischen dem Visuellen und dem Auditiven werden vom Menschen sehr sensibel wahrgenommen und führen laufend zu einer gegenseitigen Evaluierung und Korrektur.⁴⁴ So werden die Stärken und Schwächen der beiden Sinneswahrnehmungen ausgeglichen.⁴⁵ Um eine optimale Verarbeitung zu erreichen, sollten beide Informationen bestmöglich miteinander kombiniert werden, besonders in Hinsicht auf die Lokalisierung von Klängen.

⁴⁰ Boren und Roginska, „The Effects of Headphones on Listener HRTF Preference.“

⁴¹ Roginska und Geluso, Immersive Sound.

⁴² Leclère, Lavandier, und Perrin, „On the Externalization of Sound Sources with Headphones without Reference to a Real Source.“

⁴³ Begault, „3-D Sound for Virtual Reality and Multimedia.“

⁴⁴ Bregman, „Auditory Scene Analysis: The Perceptual Organization of Sound.“

⁴⁵ Bregman.

Entsprechend konnte festgestellt werden, dass sich der Lokalisierungsfehler mit verbundenen Augen gegenüber dem Hören mit visuellem Feedback sowohl in der horizontalen als auch in vertikalen Ebene signifikant erhöht.⁴⁶ Blinde Menschen jedoch bilden eine sensiblere Verarbeitung der spektralen Eigenschaften aus, um die visuellen Einflüsse zu kompensieren.⁴⁷ Dieser Lokalisierungsfehler tritt ebenso bei bewegten Klängen auf. Dabei konnte festgestellt werden, dass vor allem durch die Abwesenheit eines visuellen Ankers ein Versatz der wahrgenommenen Position in Bewegungsrichtung stattfindet.⁴⁸

Außerdem konnten differenziertere Erkenntnisse über die Auswirkung von visuellen Stimuli auf die auditive Wahrnehmung festgestellt werden.⁴⁹ So wurde ein Zusammenhang zwischen dem visuellen Einfluss und Attributen wie Realismus und Raumgefühl festgestellt, allerdings nicht für spektrale Balance und Umhüllung.⁵⁰

Um an das Thema HRTF (siehe Kapitel 2.4) anzuschließen, zeigt die Untersuchung von Udesen, Piechowiak und Gran⁵¹, dass visuelle Anker einen signifikanten Einfluss auf die Externalisierung von Klängen haben. Somit kann eine hohe Kongruenz in der audiovisuellen Wiedergabe helfen, die Schallereignisse außerhalb des Kopfes zu lokalisieren.

In Anbetracht der 3D Audio Wiedergabe über Lautsprechersysteme konnte ein visueller Einfluss auf die Echounterdrückung (*echo suppression*) festgestellt werden.⁵² In halligen Umgebungen kann die Filterung der Richtungsinformationen resultierend aus Reflexionen, vor allem im späteren Verarbeitungsprozess, unterstützt werden (siehe Kapitel 4.1).

Das audiovisuelle Zusammenspiel ist nicht nur für die Wahrnehmung unserer Umgebung und für die Lokalisierung ausschlaggebend, sondern beeinflusst maßgeblich wie wir etwas wahrnehmen. Daher konnte in unterschiedlichen Versuchsanordnungen ein Zusammenhang zwischen dem Visuellen und der rezipierten auditiven Qualität festgestellt werden.^{53 54} So zeigte ein Versuch von Rummukainen et. al. in einer VR Anwendung, dass eine Reduzierung der visuellen Qualität signifikant zu einer Reduzierung der auditiven Qualität führt.⁵⁵

⁴⁶ Tabry, Zatorre, und Voss, „The Influence of Vision on Sound Localization Abilities in Both the Horizontal and Vertical Planes.“

⁴⁷ Voss u. a., „Relevance of Spectral Cues for Auditory Spatial Processing in the Occipital Cortex of the Blind.“

⁴⁸ Getzmann und Lewald, „Localization of Moving Sound.“

⁴⁹ Woodcock, Davies, und Cox, „Influence of Visual Stimuli on Perceptual Attributes of Spatial Audio.“

⁵⁰ Woodcock, Davies, und Cox.

⁵¹ Udesen, Piechowiak, und Gran, „Vision Affects Sound Externalization.“

⁵² Bishop, London, und Miller, „Neural Time Course of Visually Enhanced Echo Suppression.“

⁵³ Platz und Kopiez, „When the Eye Listens.“

⁵⁴ Hammerschmidt und Wöllner, „Audio-Visual Quality Perception in Musical Performance Videos.“

⁵⁵ Rummukainen u. a., „Influence of Visual Content on the Perceived Audio Quality in Virtual Reality.“

Chion prägte im Kontext dieser audiovisuellen Beeinflussung den Begriff *Visu-Audition*⁵⁶, wo der Fokus der Wahrnehmung auf dem Auditiven liegt, aber in einen visuellen Kontext eingebettet ist. Dieser Kontext kann durch die Einflussnahme auf der Metaebene zu einer Unterstützung, Verstärkung oder auch einer Verzerrung der Wahrnehmung führen.

4 Räumliches Hören mehrerer Schallquellen

Um die Grundlagen des räumlichen Hörens zu verstehen, wurden viele Erkenntnisse aus isolierten Untersuchungen herangezogen. Da sowohl in einer natürlichen Umgebung mit reflektierenden Oberflächen als auch bei der Reproduktion von Klängen in einem mehrkanaligen System nicht nur eine Schallquelle existiert, sondern eine Vielzahl an zusätzlichen räumlichen Informationen zum Ohr gelangen, werden die Einflüsse von mehreren Schallquellen im nachfolgenden Kapitel näher besprochen.

4.1 Präzedenzeffekt

In „natürlichen Räumen“ und der Annahme eines akustischen Freifeldes entspricht die erste Wellenfront, die am Ohr eintrifft, dem direkten Weg zur Schallquelle. In einer halligen Umgebung treffen neben dem Direktschall über unterschiedliche Wege zahlreiche Reflexionen ein und müssen vom menschlichen Gehirn interpretiert werden. Grundlegend handelt es sich bei einer Reflexion um eine gedämpfte, örtlich getrennte und verzögerte, aber dennoch zusammenhängende Kopie des Direktschalles.⁵⁷ Durch die spektrale Ähnlichkeit der beiden Signale treten allerdings in Abhängigkeit von der Zeitverzögerung einige psychoakustische Phänomene auf, zu denen ebenfalls der sogenannte Präzedenzeffekt gezählt wird.⁵⁸

⁵⁶ Chion, Audio-Vision.

⁵⁷ Litovsky u. a., „The Precedence Effect.“

⁵⁸ Litovsky u. a.

Dieses Phänomen wurde erstmals von Wallach 1949⁵⁹ beschrieben und ist seither zentraler Aspekt weiterführender Untersuchungen unter anderen bei Blauert⁶⁰, Brown⁶¹, Litovsky⁶², Donovan⁶³, Rakerd und Hartman^{64 65}.

In verschiedenen Versuchsanordnungen, unter anderen von Blauert⁶⁶, wurde die Verarbeitung unmittelbar aufeinander eintreffender Schallereignisse anhand von zwei kurz nacheinander über Lautsprecher abgespielter Klänge untersucht. Wie in Abbildung 4a veranschaulicht, stellt das verzögerte Signal (*Lag*) die Reflexion der originalen Schallquelle (*Lead*) dar.

In der Abbildung 4b wird die Auswirkung einer zeitlichen Verzögerung zweier identischer Signale in einer symmetrischen Lautsprecheranordnung, mit jeweils einem Lautsprecher im Winkel von 45° zur Mitte dargestellt. Ohne zeitliche Verzögerung kommt es zu einer Summenlokalisation (*Fusion*), wo ein Signal als Phantomschallquelle (*Phantom Source*) zwischen den Lautsprechern wahrgenommen wird. Sobald eine zeitliche Verzögerung eintritt, kommt es zum Präzedenzeffekt. Bis zu einer Zeitverzögerung von 1ms bewegt sich die Wahrnehmung hin zum *Lead*-Signal (*localization dominance*). Ab einer Zeitverzögerung von 1ms wird das *Lag*-Signal in Lokalisation nicht mehr wahrgenommen (*discrimination suppression*),⁶⁷ bis die Signale ab der Echschwelle (*Echo Threshold*) eigenständig wahrgenommen werden.⁶⁸ Die Echschwelle beschreibt die kürzest mögliche *Lead-Lag* Verzögerung, wo zwei eigenständige Signale ausgemacht werden können.⁶⁹

⁵⁹ Wallach, Newman, und Rosenzweig, „A Precedence Effect in Sound Localization.“

⁶⁰ Blauert, „Localization and the Law of the First Wavefront in the Median Plane.“

⁶¹ Brown, Stecker, und Tollin, „The Precedence Effect in Sound Localization.“

⁶² Litovsky u. a., „The Precedence Effect.“

⁶³ Donovan, Nelson, und Takahashi, „The Contributions of Onset and Offset Echo Delays to Auditory Spatial Perception in Human Listeners.“

⁶⁴ Rakerd und Hartmann, „Localization of Sound in Rooms, II.“

⁶⁵ Rakerd und Hartmann, „Localization of Sound in Rooms, III.“

⁶⁶ Blauert, *Spatial Hearing*.

⁶⁷ Brown, Stecker, und Tollin, „The Precedence Effect in Sound Localization.“

⁶⁸ Donovan, Nelson, und Takahashi, „The Contributions of Onset and Offset Echo Delays to Auditory Spatial Perception in Human Listeners.“

⁶⁹ Brown, Stecker, und Tollin, „The Precedence Effect in Sound Localization.“

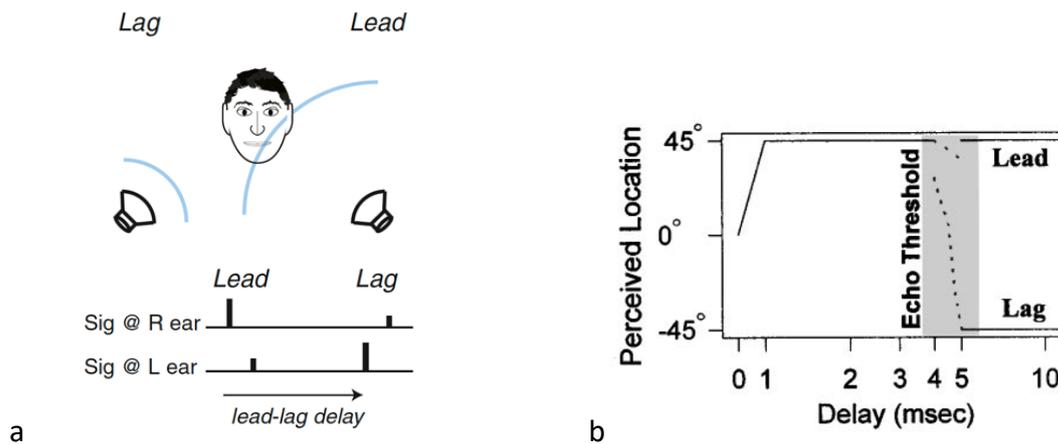


Abbildung 4: a: Versuchsanordnung mit zwei Lautsprechern und einem Lead- und Lag-Signal⁷⁰

b: Verlauf der Lokalisierung durch die zeitliche Verzögerung zweier Signale in Millisekunden.⁷¹

Die Lokalisierung beim Präzedenzeffekt ist über die zeitliche Verzögerung hinaus noch von weiteren Parametern abhängig, die von Brown zusammengefasst wurden.⁷² Je schmaler das Spektrum eines Klanges ist, desto schwächer ist auch die *localization dominance* ausgeprägt. Je breitbandiger das Signal ist, desto robuster kann der Klang lokalisiert werden. Ebenso wird der Präzedenzeffekt durch die Länge eines Signals beeinflusst, indem mit zunehmender Dauer auch die Dominanz des jeweiligen Signals gegenüber einem kürzeren verstärkt wird. Da die menschliche Wahrnehmung für ITD in einem Frequenzbereich von 500 – 1000Hz am sensibelsten ist, wird in diesem Frequenzbereich jede zeitliche Verzögerung noch präziser erfasst und überlagert andere Informationen über die wahrgenommene Position.

5 Beamforming

Beamforming beschreibt eine starke Bündelung von Klängen, die durch bestimmte Lautsprecheranordnungen gezielt in eine Richtung abgespielt werden können.^{73 74} Dabei werden einzelne Lautsprecher so gewichtet, dass in die gewünschte Richtung konstruktive

⁷⁰ Brown, Stecker, und Tollin.

⁷¹ Litovsky u. a., „The Precedence Effect.“

⁷² Brown, Stecker, und Tollin, „The Precedence Effect in Sound Localization.“

⁷³ Zotter und Frank, „Investigation of Auditory Objects Caused by Directional Sound Sources in Rooms.“

⁷⁴ Zotter, „Analysis and Synthesis of Sound-Radiation with Spherical Arrays.“

Interferenzen entstehen, also eine Überhöhung der Schallenergie und destruktive Interferenzen, also Auslöschungen von Schallwellen in die Richtung, wo keine Schallenergie abgestrahlt werden soll⁷⁵. Durch diese gerichtete Schallwirkung können gezielt Klänge an die Decke, auf den Boden oder an die Wände projiziert werden, damit sogenannte Spiegelquellen entstehen, wie in Abbildung 5 zu sehen ist.⁷⁶ Mit Hilfe von Spiegelquellen kann die Richtungs- und Distanzwahrnehmung so manipuliert werden, dass sich Klänge von der Schallquelle wegbewegen lassen und eine 3D- Audio Wiedergabe von einem Punkt aus im Raum möglich ist. Auf diesem Prinzip basiert die 3D-Audio Wiedergabe von modernen Soundbars, von 3D-Audio Speakern und unter anderem auch dem IKO; dieser wird bei der Produktionsumgebung für den praktischen Teil noch näher beschrieben.

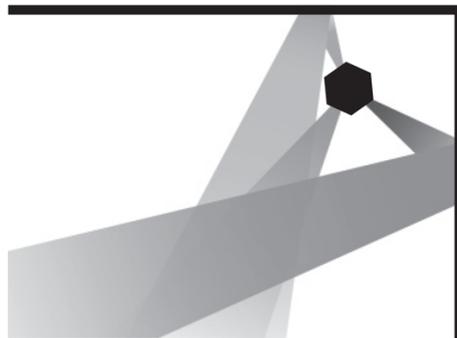


Abbildung 5: Grundprinzip des Beamformings, durch die gezielte Verwendung von Wandreflexionen mit Hilfe von gebündelten Schallstrahlen⁷⁷

Für die gebündelten Klänge können nicht nur die Abstrahlwinkel statisch eingestellt werden, sondern sie können auch frei bewegt werden.⁷⁸ Somit kann nicht nur die Lokalisation der Klänge beeinflusst werden, sondern auch die Distanzwahrnehmung, wie Laitinen et. al⁷⁹ gezeigt haben. Dafür wird das Verhältnis zwischen dem direkten Schall und dem diffusen Nachhall (D/R) verändert, indem die Lautstärke und die Richtung des Beamforming Musters entsprechend eingestellt wird. Auf Grund der Entfernung zur Schallquelle in herkömmlichen Wohnräumen wird angenommen, dass die Steuerung des Verhältnisses zwischen dem

⁷⁵ Sharma, „Komponieren mit skulpturalen Klangphänomenen in der Computermusik.“

⁷⁶ Sakamoto und Haneda, „Sound Localization of Beamforming-Controlled Reflected Sound from Ceiling in Presence of Direct Sound.“

⁷⁷ Zotter u. a., „A Beamformer to Play with Wall Reflections.“

⁷⁸ Sharma, „Komponieren mit skulpturalen Klangphänomenen in der Computermusik.“

⁷⁹ Laitinen u. a., „Controlling the Perceived Distance of an Auditory Object by Manipulation of Loudspeaker Directivity.“

direkten Schall und dem diffusen Nachhall am besten für die Steuerung der Distanzwahrnehmung ist.⁸⁰ Das Manipulieren der Richtungs- und Distanzwahrnehmung ermöglicht somit neue kreative Wege, auditive Szenarien zu entwerfen.

6 3D-Audio Formate

3D-Audio ermöglicht gänzlich neue Wege, die Zuhörer in die Geschichte miteinzubeziehen, da gegenüber der Surround-Wiedergabe bei 3D-Audio auch die Höhen abgebildet werden. Dadurch können immer bessere Illusionen der Wirklichkeit entstehen. Die Audioformate beschreiben dabei, wie die Übertragung der Audioproduktion in einer sogenannten *Digital Audio Workstation* (DAW) auf die jeweilige Wiedergabesituation übertragen wird. Man unterscheidet dabei zwischen den drei Ansätzen kanalbasiert, objektbasiert und szenenbasiert.

„3D-Audio ist die konsequente Fortsetzung der Evolution von Mono, Stereo und Surround Sound.“⁸¹ Diese logische Fortsetzung bringt auch eine entsprechende Komplexität gegenüber den vorhergehenden Formaten mit, weshalb ein grundlegendes Wissen über das Prinzip und die Vor- und Nachteile dieser 3D-Audioformate relevant ist. Zusammenfassend werden diese in den nächsten Kapiteln näher beschrieben.

6.1 Kanalbasiert

Das konventionelle System für die Audiowiedergabe ist ein kanalbasiertes System wie Stereo, 5.1, 9.1 uvm. Dabei gibt es nicht nur eine fixe Anzahl an Lautsprechern, sondern auch die genaue Position dieser muss in der Produktion berücksichtigt werden, um etwaige Abbildungsfehler in der Reproduktion zu verhindern. Es ist daher essentiell, die Wiedergabekonfiguration zu kennen.

⁸⁰ Laitinen u. a., „Controlling the Perceived Distance of an Auditory Object by Manipulation of Loudspeaker Directivity.“

⁸¹ Nipkow, Zielinsky, und Ammermann. „Die Bedeutung von 3D bei Immersive Audio.“

In einem kanalbasierten System wird für die genaue Anzahl an Lautsprechern gemischt und entsprechende Metadaten wie die Lautstärke, *Panning*⁸², *EQ*⁸³, Nachhall oder die Laufzeitunterschiede werden zwischen den Lautsprechern übertragen und auf den vorgegebenen Kanälen abgespeichert⁸⁴. Daher ist es wichtig, dass es vor allem im Bereich des *Broadcastings* einen Standard gibt, der die genauen Abstände und Winkel der einzelnen Lautsprecher vorgibt. Diese Standards wurden von der *International Telecommunication Union* (ITU) festgelegt.⁸⁵

Bei einer Produktion in einem Stereo-Setup funktioniert das kanalbasierte System sehr gut, während mit zunehmender Kanalanzahl auch Schwierigkeiten in der Produktionskette entstehen. Ein Vertreter der 3D-Audio Wiedergabe über ein kanalbasiertes System ist *Auro-3D*.

6.2 Objektbasiert

Bei der objektbasierten Audiowiedergabe wird nicht mehr ein Audiokanal auf einen Lautsprecher gerendert, sondern es werden einzelne Klangobjekte abgespeichert und mit Metadaten versehen. Diese Metadaten beinhalten räumliche Informationen zu den Objekten, die anschließend am Endgerät durch einen eingebauten Renderer für die Wiedergabesituation gemischt werden.⁸⁶ Somit lassen sich sowohl individuelle Lautsprecheranordnungen als auch die von der *ITU* festgelegten Standards wie 5.1 und 7.1 realisieren.

Für die Wiedergabe von Klangobjekten werden einzelne Monosignale verwendet, die in den Metadaten über genaue Informationen zur räumlichen Position verfügen.⁸⁷ Dadurch wird den Algorithmen der Renderer eine besonders wichtige Rolle zuteil, da diese je nach Wiedergabeanordnung die Objekte entsprechend der räumlichen Wahrnehmung der

⁸² Beim Panning wird ein Signal stufenlos durch Intensitäts- und Laufzeitunterschiede zwischen zwei gleichen Schallquellen positioniert, Friesecke, Die Audio-Enzyklopädie.

⁸³ EQ oder Equalizer haben die Aufgabe, den Klang eines Signals zu gestalten oder Fehler im Frequenzgang zu korrigieren, Friesecke.

⁸⁴ Rumsey, „Spatial Audio.“

⁸⁵ International Telecommunication Union (ITU), “Multichannel sound technology in home and broadcasting applications.”

⁸⁶ Rumsey, „Spatial Audio.“

⁸⁷ Roginska und Geluso, Immersive Sound.

Menschen im Raum platzieren müssen. Die Vorteile von objektbasiertem Audio sind neben der flexiblen Wiedergabeordnung vor allem die Personalisierung und das interaktive Eingreifen des Endverbrauchers. So lassen sich leicht unterschiedliche Sprachversionen realisieren bzw. kann man auch die Lautstärke einzelner Objekte bearbeiten oder sogar gänzlich entfernen, da die Mischung im Renderer am Endgerät passiert.⁸⁸

Zusätzlich gibt es auch hybride Formate, wo Objekte einerseits einem bestimmten Lautsprecher zugeordnet werden, andererseits auch kanalunabhängig im Raum platziert werden können. Die kanalbasierten Objekte, die sogenannten *Beds*, werden unter anderem für Atmos verwendet. Der bekannteste Vertreter dieses hybriden Formats ist Dolby Atmos⁸⁹.

6.3 Szenenbasiert

Bei den szenenbasierten Audioformaten handelt es sich um eine physikalische Rekonstruktion eines Schallfeldes.⁹⁰ Die prominentesten Vertreter dieses Prinzips sind die Wellenfeldsynthese (WFS) und Ambisonics, letzteres wird nachfolgend ausführlicher beschrieben, da die Produktion des Werkstücks und die Wiedergabe der Hörversuche auf diesem Format beruhen.

6.3.1 Ambisonics

Das Konzept von Ambisonics wurde bereits in den 70er Jahren von Michael Gerzon⁹¹ entwickelt. Ambisonics wurde durch einen stärkeren Fokus auf 3D-Audio Anwendungen, vor allem durch die Entwicklungen im Bereich der 360° Videos und Virtual Reality, wieder mehr Aufmerksamkeit zuteil.^{92 93} Es basiert auf der mathematischen Zerlegung eines holophonen Klangfelds in einzelne Kugelflächenfunktionen. Diese Funktionen bilden die Basis für die räumlichen Spektren, mit denen ein Schallfeld nachgebildet werden kann.⁹⁴ Diese sind in Abbildung 6 und 7 visualisiert.

⁸⁸ Roginska und Geluso, Immersive Sound.

⁸⁹ Roginska und Geluso.

⁹⁰ Sontacchi, „Dreidimensionale Schallfeldreproduktion für Lautsprecher- und Kopfhöreranwendung.“

⁹¹ Gerzon, „Periphony“.

⁹² Zotter und Frank, Ambisonics.

⁹³ Hafsati u. a., „Sound Source Separation in the Higher Order Ambisonics Domain.“

⁹⁴ Zotter, Frank, und Sontacchi, „The Virtual T-Design Ambisonics-Rig Using VBAP.“

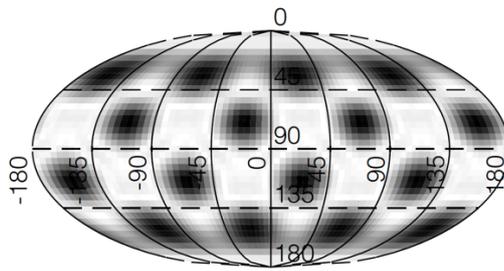


Abbildung 6: Approximative Nachbildung eines räumlichen Schallfeldes durch Kugelflächenfunktionen.⁹⁵

In einer endlichen Anzahl an Kugelflächenfunktionen wird so das Schallfeld reproduziert. Diese Anzahl wird durch die ambisonische Ordnung ausgedrückt und entspricht der räumlichen Auflösung und der klanglichen Genauigkeit des Schallfelds. Die Ordnung N entscheidet über die Anzahl der erforderlichen Kanäle auf dem Ambisonics Bus⁹⁶ und der minimalen Anzahl der Lautsprecher für die Wiedergabe.⁹⁷ Die Berechnung der benötigten Anzahl der Kanäle für die jeweilige Ordnung erfolgt durch die Formel $(N+1)^2$.⁹⁸ Man unterscheidet dabei zwischen der ersten Ordnung *First Order Ambisonics* (FOA) oder einer höheren ambisonischen Ordnung *High Order Ambisonics* (HOA).

Auf der Produktionsseite wird zuerst ein Encoder benötigt, um das verwendete Audiomaterial samt den Metadaten in das virtuelle Schallfeld zu bringen. Dabei ist die Anzahl der verwendeten Objekte uneingeschränkt und unabhängig von der ambisonischen Ordnung.⁹⁹ In diesem Schallfeld können dann die einzelnen Sound Objekte mit Hilfe von frei verfügbaren Plugins¹⁰⁰ im Raum platziert und auch bewegt werden.

Dieses virtuelle Schallfeld wird anschließend im *ambix*-Format abgespeichert. Da in diesem Format keine Informationen über die Wiedergabekonfiguration enthalten ist, sondern lediglich das Schallfeld übertragen wird, benötigt man auf der Endnutzerseite einen entsprechenden Decoder.

⁹⁵ Zotter und Frank, Ambisonics.

⁹⁶ Auf dem Ambisonics Bus werden alle Kanäle in einer DAW zusammengeführt, die in das ambisonische Format encodiert wurden.

⁹⁷ Frank, Zotter, und Sontacchi, „Producing 3D Audio in Ambisonics.“

⁹⁸ Frank, Zotter, und Sontacchi.

⁹⁹ Frank, Zotter, und Sontacchi.

¹⁰⁰ <http://www.matthiaskronlachner.com/?p=2015>
<https://plugins.iem.at>

Der Decoder muss dabei an die vorhandene Wiedergabekonfiguration angepasst werden, damit die virtuellen Klangobjekte als Phantomschallquellen¹⁰¹ im Raum platziert werden können. Dabei bietet der *Allrad Decoder*¹⁰² mit einer Mischung aus mathematisch errechneten Amplituden Panning und dem *VBAP* (Vector-Base Amplitude Panning)¹⁰³ eine flexible Decodierung für unterschiedliche Lautsprecheranordnungen.¹⁰⁴

Da Ambisonics ein virtuelles Schallfeld darstellt, bietet es die Möglichkeit einer sehr einfachen und glatten Rotation des Klangbildes ohne großen Rechenaufwand und ist daher für Head-Tracking für *Virtual Reality* und *Augmented Reality* Anwendungen prädestiniert.¹⁰⁵¹⁰⁶ Durch sehr effiziente Algorithmen zur binauralen Decodierung ermöglicht *HOA* eine hochaufgelöste auditive Wiedergabe über Kopfhörer.¹⁰⁷

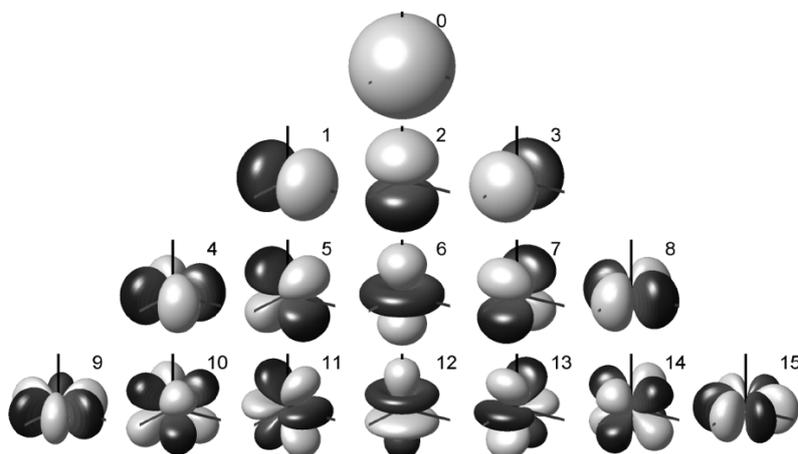


Abbildung 7: Darstellung der Kugelflächenfunktionen in einem polaren Koordinatensystem. Die Reihen zeigen die Anzahl der Kanäle von der 0 bis zur 3. Ordnung, sowie die Anordnung der Kugelflächenfunktionen für die mathematische Reproduktion des Schallfeldes.¹⁰⁸

¹⁰¹ „Unter einer Phantomschallquelle versteht man die Lokalisation eines Schallereignisses aus einer Richtung, in der sich kein Schallwandler befindet.“

Friesecke, Die Audio-Enzyklopädie.

¹⁰² Zotter und Frank, „All-Round Ambisonic Panning and Decoding.“

¹⁰³ Pulkki, „Virtual Sound Source Positioning Using Vector Base Amplitude Panning.“

¹⁰⁴ Zotter und Frank, „All-Round Ambisonic Panning and Decoding.“

¹⁰⁵ Shivappa u. a., „Efficient, Compelling and Immersive VR Audio Experience Using Scene Based Audio / Higher Order Ambisonics.“

¹⁰⁶ Altmann u.a., „Immersive Audio for VR.“

¹⁰⁷ Shivappa u. a., „Efficient, Compelling and Immersive VR Audio Experience Using Scene Based Audio / Higher Order Ambisonics.“

¹⁰⁸ Zotter und Frank, Ambisonics.

6.4 MPEG-H

Mit dem 3D-Audioformat MPEG-H der *Moving Picture Experts Group* (MPEG) etabliert sich ein flexibles Kontainerformat das gleichermaßen kanalbasierte, objektbasierte und szenenbasierte Audioformate unterstützt.¹⁰⁹ Dabei sind in den MPEG-H Metadaten alle wichtigen Informationen gespeichert, um eine Reproduktion unabhängig der Wiedergabekonfiguration zu ermöglichen. Mit dem *USAC-3D* (Unified Speech and Audio Coding) Decoder werden die jeweiligen Kanäle zugeordnet und die Metadaten der Audioobjekte und *HOA* Inhalte an die richtige räumliche Position gerendert.¹¹⁰

Ebenfalls können Formate miteinander kombiniert werden. So lassen sich Objekte ergänzend zu einer kanalbasierten oder *HOA* Wiedergabe hinzufügen. Der große Vorteil liegt dabei ebenfalls in der Personalisierung der Audioinhalte, indem aktiv bei der Wiedergabe eingegriffen werden kann und je nach eigenen Vorstellungen Audioobjekte angesteuert und angepasst werden können.¹¹¹

7 Sound Branding

Sound Branding beschreibt den Einsatz von akustischen Reizen im Rahmen der Markenkommunikation und die generelle Idee sich klanglich bemerkbar zu machen. Es wird dabei das Ziel verfolgt, mit Musik oder Klängen eine Verbindung zwischen den Menschen und einer Marke oder einem Unternehmen herzustellen.¹¹² Obwohl in diesem Zusammenhang viele semantisch unterschiedliche Begriffe wie Audio Branding, Sonic Branding, Acoustic Branding, Corporate Sound uvm. existieren, ist der Kern dahinter der gleiche, einem Unternehmen eine akustische Identität zu verleihen.¹¹³

Entsprechend dem visuellen Pendant versucht man auch beim Auditiven die Markenidentität, also ein Selbstbild der Marke, und das Markenimage, also die Wahrnehmung der Marke auf der Konsumentenseite, in ihrer Gesamtheit zu erfassen und die Essenzen

¹⁰⁹ Herre u. a., „MPEG-H Audio“.

¹¹⁰ Herre u. a., „MPEG-H Audio“.

¹¹¹ <https://www.iis.fraunhofer.de/en/ff/amm/broadcast-streaming/mpeg-h.html>

¹¹² Groves, Commusication.

¹¹³ Kleinjohann, Marketingkommunikation mit Acoustic Branding.

herauszuarbeiten.¹¹⁴ Es sollte so ein kongruentes Stimmungsbild zwischen der Marke und den assoziierten Emotionen entstehen, das auf möglichst vielen Sinneskanälen gemeinsam transportiert wird. Ähnlich wie beim visuellen Erscheinungsbild gibt es bestimmte Designelemente, die mit den Markenattributen angereichert werden und in möglichst viele Kontaktpunkte mit dem Kunden integriert werden. Diese akustischen Designelemente werden in der *Sound DNA*¹¹⁵ zusammengefasst.

7.1 Sound DNA

Die Sound DNA beschreibt die wichtigsten Bausteine einer akustischen Identität einer Marke. Dazu gehören Melodien, Akkorde, Instrumente und ganz bestimmte Klänge, die eindeutig mit einer Marke in Verbindung gebracht werden können. Das Prinzip der Sound DNA entspricht der Leitmotiv-Technik, die vor allem durch Richard Wagner sehr bekannt wurde und auch heute unter anderem sehr intensiv im Film eingesetzt wird.¹¹⁶

Um möglichst viele Kontaktpunkte zwischen der Marke und dem Kunden mit der eigenen akustischen Identität zu bespielen, liefert die Sound DNA einen Pool an Klängen, die mit dem Markenattributen in Verbindung gebracht werden und während eines akustischen Designprozesses entwickelt werden.¹¹⁷ Für die Sound DNA gelten dabei ähnliche Anforderungen^{118 119 120 121} wie für Sound Logos. Die Sound DNA spiegelt idealerweise möglichst viele Markenattribute wider (Marken Fit) und das prägnant, damit möglichst effizient eine Wiedererkennung stattfindet und die assoziierten Emotionen hervorgerufen werden (Prägnanz). Außerdem ist es wichtig, sich klanglich entsprechend von Konkurrenten abheben zu können (Unverwechselbarkeit).

Ein besonders wichtiger Aspekt ist die Anpassungsfähigkeit (Flexibilität), sowohl aus inhaltlicher Sicht an die verschiedenen Werbeinhalte als auch aus technischer Sicht an die möglichen klanglichen Interaktionen.

¹¹⁴ Steiner und Joachimsthaler, Sound Branding.

¹¹⁵ Auch hier gibt es unterschiedliche Bezeichnungen, die allerdings in ihrer Bedeutung ebenfalls synonym verwendet werden.

¹¹⁶ Flückiger, Sound design.

¹¹⁷ Groves, Commusication.

¹¹⁸ Anzenbacher, Audiologos.

¹¹⁹ Steiner und Joachimsthaler, Sound Branding.

¹²⁰ Groves, Commusication.

¹²¹ Ringe, „Audio Branding Guide.“

7.2 Neue akustische Interaktionen

Um sich dauerhaft in der Wahrnehmung von Kunden zu etablieren, muss sich der Klang der Marke weiterentwickeln und sich technisch und gesellschaftlich modernisieren, ohne dabei die Identität zu verlieren.¹²² Es ist daher wichtig, die Konzipierung des Sound Brandings vorausschauend und zukunftsorientiert zu gestalten und dabei sämtliche Kontaktpunkte in den kreativen Prozess miteinfließen zu lassen.

In den letzten Jahren zeichnete sich bereits eine Veränderung in der Markenkommunikation ab, die zunehmend den Fokus auf die auditive Interaktion rückt und maßgeblich durch die Entwicklung von Sprachassistenten beschleunigt wird.¹²³ Einhergehend mit dieser Neuorientierung entsteht auch in akustischer Hinsicht ein Raum, sich mit kreativen Ansätzen zu profilieren. Dabei ermöglicht die Integrierung von 3D Audio-fähigen Smart Speakern und binauralen Wiedergabegeräten neue akustische Interaktionen im Alltag.

Es gibt bereits erste Marken die 3D-Audio als Werkzeug verwenden, um Markenerlebnisse zu kreieren und so die Kunden emotional an die Marke zu binden.¹²⁴

8 3D Audio und Emotionen

Um eine emotionale Bindung an eine Marke zu erreichen, muss eine Immersion entstehen, die es den Menschen ermöglicht, auf einer sensorischen und kognitiven Ebene eintauchen zu können. In einer Untersuchung von Hahn¹²⁵ konnte festgestellt werden, dass durch eine Wiedergabe von 3D Audio, im Vergleich zu einer Stereo und 5.1 Wiedergabe die Emotionen zusätzlich getriggert werden. Für die Untersuchung wurden die Proband*innen in die drei Gruppen Tonmeister*innen, Musiker*innen und ungeübte Hörer*innen unterteilt. Anhand zweier unterschiedlicher Ausschnitte eines Werkes aus der klassischen Musik wurden die Parameter *emotionale Erregung* und *Gefallen* bewertet. Dabei konnte sich der emotionale Mehrwert am deutlichsten bei der Gruppe der ungeübten Hörer*innen feststellen lassen. (siehe Abbildung 8)

¹²² Steiner und Joachimsthaler, Sound Branding.

¹²³ Ringe, „Audio Branding Guide.“

¹²⁴ <https://blog.dearvr.com/overview/lufthansa-3d-audio-podcast-by-zdf-digital>

¹²⁵ Hahn, „Musical Emotions Evoked by 3D Audio.“

Daraus lässt sich schließen, dass 3D Audio nicht nur eine emotionale Verstärkung in der akustischen Markenkommunikation bewirken kann, sondern durch eine starke Wirkung auf ungeübte Hörer*innen ebenfalls eine breite und kommerzielle Nutzung möglich ist.

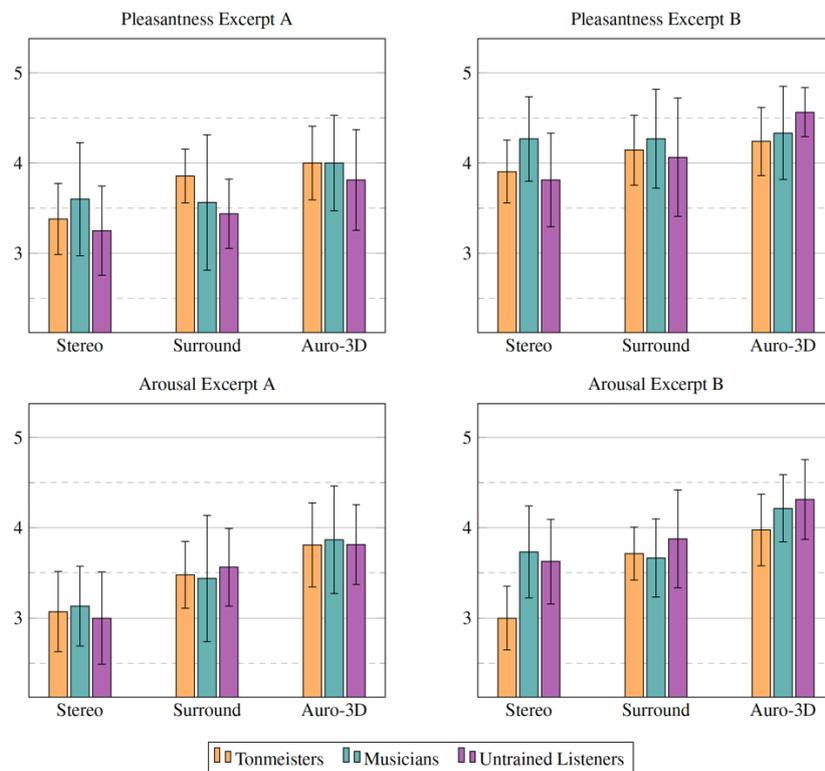


Abbildung 8: Ergebnisse der Untersuchung von Hahn in Bezug auf die emotionale Erregung und Gefallen.¹²⁶

9 Werkstück

Um das Thema nicht nur theoretisch zu beleuchten, sondern auch die praktische Umsetzung zu untersuchen, wurden prototypisch vier Sound Logos (kurzes akustisches Erkennungsmerkmal) für die akustische Markenkommunikation über ein 3D Audio Lautsprechersystem erarbeitet. Dafür wurden für den IKO (siehe Kapitel 9.2) unterschiedliche Klangobjekte entwickelt, um mit variierenden Komplexitätsstufen die gestalterischen Möglichkeiten auszuloten.

¹²⁶ Hahn, „Musical Emotions Evoked by 3D Audio.“

Dem praktischen Konzept liegt die Frage zu Grunde, wie sich Marken in einer zunehmend audiophileren Kommunikation über Voice-Systeme¹²⁷ bemerkbar machen können. Dafür ergeben sich nicht nur für das Marketing eines Unternehmens neue Fragen, sondern auch für den Designprozess der Sound DNA, da sich die klassischen Kontaktpunkte^{128 129} erweitern. Diese müssen auch im Entwicklungsprozess mitbedacht werden, besonders in Anbetracht einer zunehmenden Bedeutung von 3D Audio, damit durch technische oder akustische Limitationen nicht die klangliche Qualität in Mitleidenschaft gezogen wird. Ein möglicher Designprozess wird anhand der Sound Logos konstruiert.

9.1 Klangkonzept

Damit die Sound Logos klanglich möglichst zu verallgemeinern sind, wurden die klanglichen Konzepte nicht an eine Marke angepasst, sondern orientieren sich an bekanntem akustischem Material. Dafür wurde als Vorlage ein Ranking¹³⁰ der besten akustischen Branding Konzepte für den klanglichen Findungsprozesses verwendet.

Aus diesem Ausgangsmaterial wurden fiktive Sound Logos (siehe Kapitel 9.5) erstellt und ebenfalls mit einer sehr basalen visuellen Identität ausgestattet. Diese visuellen Ankerpunkte entsprechen unter anderem der Morphologie der Klangobjekte, die durch eine 3D Audio Wiedergabe im Raum dargestellt werden.

9.2 Ikosaeder (IKO)

Als Wiedergabesystem wurde der IKO der Firma *Sonible*¹³¹ und des Instituts für Elektronische Musik und Akustik¹³² verwendet. Es handelt sich dabei um ein kompaktes Lautsprechersystem in der geometrischen Form eines Ikosaeders, also eines 20-seitigen Würfels. Wie in Abbildung

¹²⁷ Voice-Systeme oder auch CUI (Conversational User Interfaces) umfasst die Kommunikation zwischen Mensch und Maschine mit natürlicher Sprache. Der Begriff der natürlichen Sprache umfasst sowohl das gesprochene Wort, als auch das geschriebene Wort.

Kabel, Dialog Zwischen Mensch und Maschine Conversational User Interfaces, Intelligente Assistenten und Voice-Systeme.

¹²⁸ Steiner und Joachimsthaler, Sound Branding.

¹²⁹ Anzenbacher, Audiologos.

¹³⁰ <http://ampsoundbranding.com/best-audio-brands-ranking/>

¹³¹ <https://iko.sonible.com/en.html>

¹³² <https://iem.kug.ac.at/projects/osil/about-the-iko.html>

9 dargestellt, ist auf jeder dieser Seitenflächen ein Lautsprecherchassis eingebaut und ermöglicht so die Reproduktion eines kugelförmigen Schallfeldes, die für die *High Order Ambisonics* (HOA) Wiedergabe grundlegend ist. Der IKO ist in der Lage, Ambisonics bis zu dritter Ordnung wiederzugeben.



Abbildung 9: IKO, ein 20-flächiges ikosaederförmiges Lautsprechersystem der Firma Sonible und des Instituts für Elektronische Musik und Akustik Graz.¹³³

Durch die große Anzahl an Lautsprechern und die besondere Form, lassen sich die mit Hilfe von speziellen Algorithmen gebündelten Schallwellen (*Sound Beams*) frei und auch dynamisch in alle Raumrichtungen wiedergeben. Dabei können diese *Sound Beams* nicht nur frei im Raum bewegt werden, sondern auch die Breite lässt sich beliebig variieren. Durch die Möglichkeit, mehrere *Sound Beams* simultan abspielen zu können, lassen sich komplexere Reflexionsmuster realisieren.¹³⁴

Als Audiointerface wurde ein RME MADiface USB verwendet. Zusätzlich ergänzte der *d:24*¹³⁵, ein mehrkanaliger Verstärker der Firma *Sonible* das Setup für den IKO, siehe Abbildung 9.

¹³³ <https://iko.sonible.com/en.html>

¹³⁴ Wendt u. a., „Perception of Spatial Sound Phenomena Created by the Icosahedral Loudspeaker.“

¹³⁵ <https://www.sonible.com/d24/>

9.3 Produktionsumgebung

Die Produktion der Sound Logos gliedert sich in zwei Abschnitte. Zuerst wurden sie in einem klassischen Stereoformat in der *Digital Audio Workstation (DAW) Logic Pro X*¹³⁶ entwickelt. Dafür wurden *Sample Libraries*, verschiedene digitale Synthesizer und Klangaufnahmen verwendet und den ausgearbeiteten Klangkonzepten entsprechend durch Veränderung des Spektrums, der Tonhöhe oder des zeitlichen Verlaufes manipuliert.

Im zweiten Schritt wurden die einzelnen Ebenen und klanglichen Elemente der Sound Logos in Mono- und Stereodateien exportiert und wiederum für die räumliche Anordnung der Klänge in die DAW *Reaper* importiert. Da für die Anfertigung der 3D Klangobjekte mit Ambisonics dritter Ordnung gearbeitet wurde, bietet sich *Reaper* dahingehend an, dass insgesamt 64 Kanäle pro Spur¹³⁷ möglich sind. Entsprechend der Formel zur Berechnung der ambisonischen Kanäle (siehe Kapitel 4.3.1) würde diese Kanalanzahl eine Produktion bis zur siebten Ordnung ermöglichen.

Für die Platzierung der Klänge im Raum wurden die frei verfügbaren Plugins der *IEM Plug-in Suite*¹³⁸, *ambiX – Ambisonics plug-in suite*¹³⁹ und der *mcfx – multichannel plug-in Suite*¹⁴⁰ verwendet. Dafür wurde zuerst eine *Reaper* Datei eingerichtet und der Abbildung 14 entsprechend geroutet.

Zuerst wurde der Ambisonics Bus, wo alle einzelnen Spuren zusammengefasst werden, erstellt und für die dritte Ordnung mit 16 Kanälen eingestellt. Anschließend wurde der *mcfx_convolver* eingefügt, eine mehrkanalige Faltungsmatrix, die durch die Entzerrung des Signals die Klangabstrahlung¹⁴¹ des IKOs kontrolliert.¹⁴² Dafür konnte eine bereits vorhandene Vorlage verwendet werden.

¹³⁶ <https://www.apple.com/de/logic-pro/>

¹³⁷ <https://www.reaper.fm/about.php#technical>

¹³⁸ <https://plugins.iem.at>

¹³⁹ <http://www.matthiaskronlachner.com/?p=2015>

¹⁴⁰ <http://www.matthiaskronlachner.com/?p=1910>

¹⁴¹ Ein convolver verwendet eine gemessene Impulsantwort eines Raumes oder eines Lautsprechers. Diese Impulsantwort wird wie ein Filter auf das Eingangssignal gelegt und angeglichen. Durch einen mathematischen Prozess können so Raumeigenschaften und Abstrahleigenschaften angepasst werden.

<https://www.izotope.com/en/learn/the-basics-of-convolution-in-audio-production.html>

¹⁴² Kronlachner, „Spatial Transformations for the Alteration of Ambisonic Recordings.“

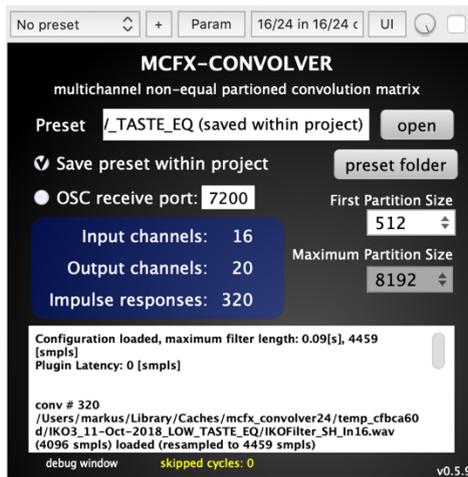


Abbildung 10: Screenshot des `mxfc_convolver` mit der Filtervorlage für den IKOs

Zusätzlich wurde auf dem Ambisonics Bus der *EnergyVisualizer* eingefügt, damit die Intensität der Schallabstrahlung in der Klangsphäre nachvollzogen werden kann. Wie in der Abbildung 11 zu sehen ist, kann sowohl die Lautstärke als auch die Verteilung des Signals durch die grafische Oberfläche schnell eruiert werden.

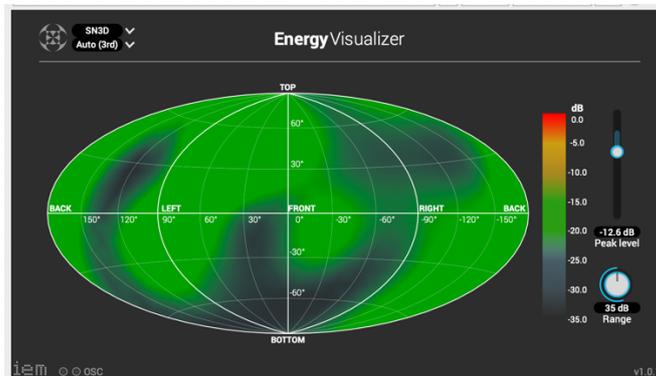


Abbildung 11: Screenshot des *EnergyVisualizer*.

In der graphischen Darstellung wird die Verteilung des Signals auf einer aufgefächerten Kugel abgebildet. Die Lautstärke wird dabei farblich veranschaulicht.

Auf den Ambisonics Bus wurden die einzelnen Mono- und Stereodateien der Sound Logos eingefügt und geroutet. Um die Signale in die Ambisonics Domäne zu bringen, wurden sie mit Hilfe von *StereoEncoder*¹⁴³ und *ambix_encoder_o3* enkodiert. Diese beiden Encoder sind in den Abbildungen 12 und 13 gegenübergestellt. Neben der unterschiedlichen Benutzeroberfläche, wo durch ein *Graphical User Interface* (GUI) die Parameter verändert werden

¹⁴³ <https://plugins.iem.at/docs/plugindescriptions/>

können, besitzen beide Encoder gewisse Vor- und Nachteile, die jeweils von der Funktion der Klänge abhängig sind. So kann bei *StereoEncoder* die Breite eingestellt und automatisiert werden, indem das Stereosignal geteilt wird. Zudem ist die Einstellung der Rotation in dem Plugin möglich. Dieser Encoder wurde in den meisten Fällen für statische Klänge verwendet, die durch eine fixierte Position das Klangpanorama verbreitern. In *mcfx_encoder_o3* ermöglicht der Parameter *Size* (siehe Abbildung 12) eine stufenlose Veränderung der ambisonischen Ordnung. Durch diese Veränderung lässt sich ein dynamisches Bewegungsmuster mit einer Tiefenstaffelung einstellen und somit eine glaubwürdige Distanzänderungen realisieren (Details siehe Kapitel 9.6.3).

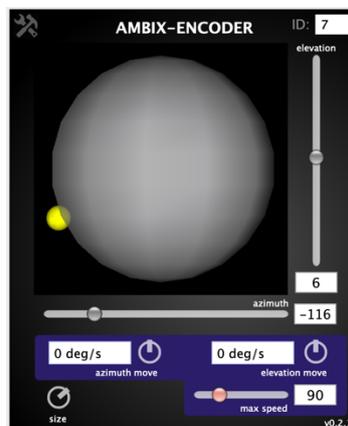


Abbildung 12: Screenshot des *mcfx_encoder_o3*.

Der Encoder stellt das Schallfeld aus der Vogelperspektive dar. Die gelbe Kugel zeigt das Signal und lässt sich auf der graphischen Oberfläche frei im Schallfeld positionieren. Die Position kann ebenfalls durch die Schieberegler verändert werden. Links unten lässt sich die ambisonische Ordnung stufenlos durch den Parameter *Size* mit einem Drehregler verändern.

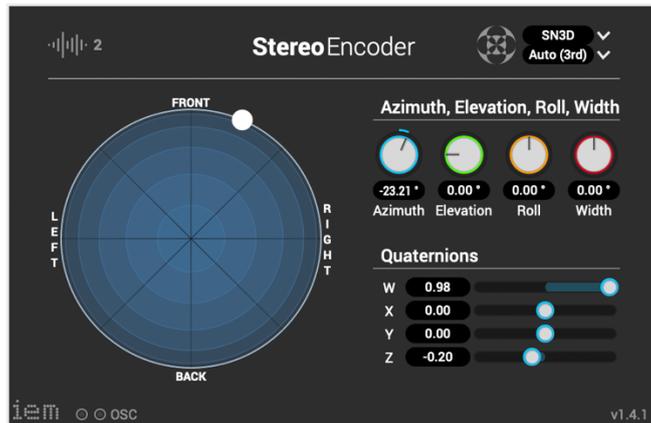


Abbildung 13: Screenshot des Stereo Encoder.

Ebenfalls aus der Vogelperspektive wird das Signal als weiße Kugel dargestellt. Das Signal lässt sich durch die graphische Oberfläche frei im Schallfeld bewegen. Ebenso lassen sich die Parameter Azimut, Elevation, Rotation und Breite des Stereosignals durch Drehregler verändern.

Weiters wurde auch *MultiEncoder*¹⁴⁴ verwendet, um mehrere Klänge gleichzeitig in Ambisonics zu enkodieren und auch gemeinsam in parallelen Bewegungsmustern zu bearbeiten (Siehe Sound Logo 3).

Für eine optimale Mischung der Klänge trotz der räumlichen Entfernung wurde *FdnReverb*¹⁴⁵ verwendet. Durch das Beifügen von Hall sollte die Lokalisierung der Klänge diffuser werden und somit ein kompakteres und dichteres Klangfundament für die dynamischen Klänge entstehen (siehe Kapitel 8.6.4).

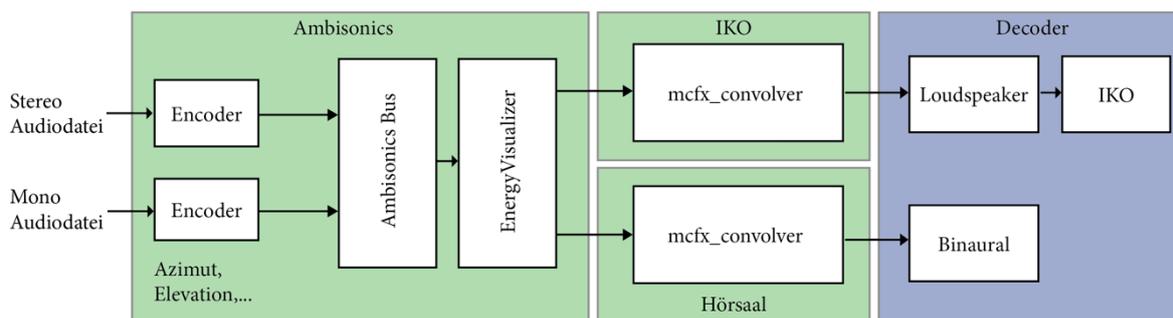


Abbildung 14: Routing der Spuren in Reaper sowohl für den IKO, als auch für die virtuelle Umgebung über eine binaurale Wiedergabe.

¹⁴⁴ <https://plugins.iem.at/docs/pluginDescriptions/>

¹⁴⁵ Ist ein künstlicher Nachhall, der durch eine Verzögerung des Feedback-Signals erreicht wird. Ein Feedback ist eine akustische Rückkoppelung.

<https://plugins.iem.at/docs/pluginDescriptions/>

9.4 Arbeitsumgebung

Bei der Konstruktion komplexer Klangobjekte ist eine sehr präzise Arbeit mit den Wandreflexionen und den räumlichen Gegebenheiten notwendig. Da es allerdings nicht immer möglich ist die künstlerische Arbeit vor Ort oder auch mit dem IKO durchzuführen, muss die reale Wiedergabesituation im Entwicklungsprozess imitiert werden. Daher wurde für die Vorarbeit zu den Hörversuchen eine virtuelle Simulation des IKOs und der räumlichen Gegebenheiten verwendet, der sogenannte *VIKO* (virtual IKO).¹⁴⁶ So ist man nicht an die Räumlichkeiten und an den IKO selbst gebunden und kann über eine binaurale Wiedergabe die Raumeigenschaften simulieren und eine entsprechende räumliche Anordnung vornehmen.

Dafür wurden von Mitarbeitern des *IEM (Institut für Elektronische Musik und Akustik)* Raumimpulsantworten des Hörsaales des *IEM* mit dem 32-kanaligem Mikrofon *Eigenmike EM32* an der gewünschten Abhörposition aufgenommen.¹⁴⁷ Da für die Hörversuche sowohl derselbe Hörsaal als auch dieselbe Abhörposition verwendet wurden, konnte diese virtuelle Umgebung für die Vorbereitung genutzt werden.

In der realen Wiedergabesituation wurde durch den rechteckigen Grundriss des Raums der IKO in den Ecken platziert, da so sehr gezielt mit den Reflexionen zweier Wände gearbeitet werden kann.¹⁴⁸ Eine konkave Reflektoren-Anordnung hinter dem IKO würde zwar vielseitigere Möglichkeiten bieten, wurde allerdings auf Grund einer kaum realisierbaren Wiedergabesituation für herkömmliche 3D Audio Lautsprecher nicht berücksichtigt. Zusätzlich hat Sharma¹⁴⁹ festgehalten, um eine gute Balance zwischen den Reflexionen und dem Direktschall zu bekommen, sollte der Abstand des IKOs zum Publikum/zur Abhörposition mindestens so groß sein, wie zwischen dem IKO und den Wänden. Der Grundriss des Raumes und die jeweiligen Distanzen sind in Abbildung 15 dargestellt. Die Lautsprecher des IKOs befanden sich bei den Hörversuchen in einer Höhe von 1,6 Metern. Die Nachhallzeit im Hörsaal beträgt 0,6s.¹⁵⁰ Im Hörsaal befindet sich ein Parkettboden und die Decke besteht aus akustischen Fliesen. Von der Abhörposition aus rechts befindet sich eine Glasfassade,

¹⁴⁶ Zaunschirm, „The (Virtual IKOsahedral Loudspeaker Array) VIKO“.

¹⁴⁷ Zaunschirm.

¹⁴⁸ Wendt u. a., „Perception of Spatial Sound Phenomena Created by the Icosahedral Loudspeaker“.

¹⁴⁹ Sharma, „Komponieren mit skulpturalen Klangphänomenen in der Computermusik“.

¹⁵⁰ Sharma.

während die restlichen drei Wände aus Gips bestehen. Es befinden sich auf der Rückseite Melaminharzabsorber und an der Vorderseite zwei Tafeln. Auf der linken Seite der Abhörposition ist die Gipswand von Glasfenster und einer Glastür durchbrochen.

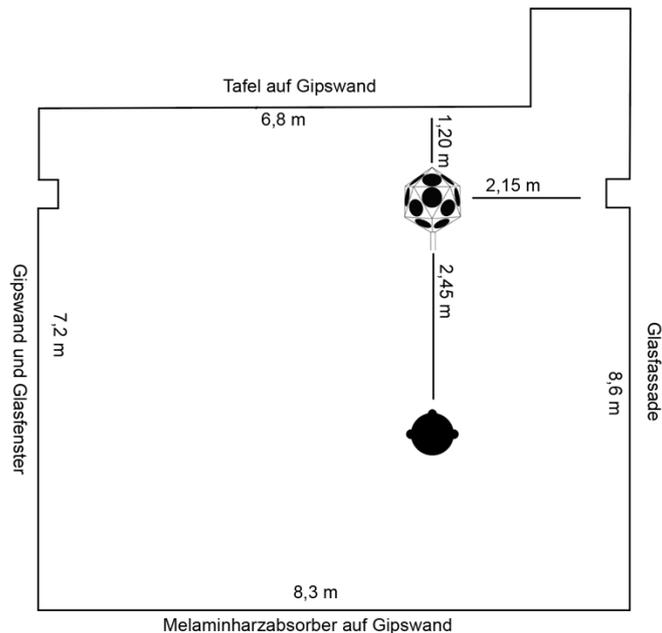


Abbildung 15: Grundriss des Hörsaals des Instituts für Elektronische Musik und Akustik inkl. der Raummaße und Wandmaterialien. Der Hörsaal hat einen Parkettboden und akustische Fliesen an der Decke.

9.5 Sound Logos

Ein Sound Logo ist ein prägnantes akustisches Erkennungsmerkmal, das sehr eng mit einer Marke verbunden ist.¹⁵¹ Dabei sollte innerhalb weniger Töne ein Wiedererkennungseffekt stattfinden und Assoziation mit den gewünschten Markenattributen transportiert werden. Es handelt sich dabei um einen der wichtigsten Träger der Sound DNA, da es besonders durch die Kürze sehr flexibel eingesetzt werden kann.

Insgesamt wurden für den praktischen Teil vier Sound Logos erstellt, die sich aus statischen Klängen und dynamischen Klangbewegungen zusammensetzen. Für die Klangbewegungen wurden sowohl einfache geometrische Formen, als auch erfundene visuelle Markenidentitäten, wie zum Beispiel Buchstaben verwendet.

¹⁵¹ Anzenbacher, Audiologos.

9.5.1 Sound Logo 1

Das Fundament dieses Sound Logos ist eine E-Bass Melodie, die auf Grund der omnidirektionalen¹⁵² Abstrahlung der tiefen Frequenzen für eine Einhüllung sorgt. Die Melodie wird von einer bestimmten Akkordfolge begleitet. Beide Komponenten sind ein wichtiger Bestandteil der Sound DNA. Sie werden im Raum verteilt, um ein möglichst breites Klangpanorama zu erreichen. Darüber hinaus enthält das Sound Logo einige perkussive Klänge, welche die Bass Melodie auf Grund ihrer Tonhöhe im oberen Frequenzbereich unterstützen. Durch den hohen Anteil an transienten¹⁵³ Signalen werden diese sehr nahe am IKO wahrgenommen.

Das Bewegungsmuster wird von einem rauschhaften Klang mit ansteigender Lautstärke vollzogen. Im Beispiel A bewegt sich der Klang in Äquatorhöhe (auf der Höhe der Ohren) im Halbkreis von rechts nach links (siehe Abbildung 16). Durch die geringe Lautstärke am Beginn des Klanges verharrt er verhältnismäßig länger auf der rechten Seite, bis er lokalisiert werden kann und vollendet anschließend die Trajektorie¹⁵⁴ nach links. Ein weiterer Grund für die anfängliche Verzögerung der Bewegung liegt unter anderem in einer tendenziellen Vorwärtsverschiebung der Wahrnehmung von Trajektorien in der Anfangsphase, besonders ohne visuelle Hilfestellung.¹⁵⁵

Das Bewegungsmuster B vollzieht ebenfalls eine halbkreisförmige Bewegung über den IKO hinweg auf der vertikalen Ebene. (siehe Abbildung 16). Dabei wurde der Klang möglichst auf den zur Wand gerichteten Lautsprechern abgebildet, damit auf Grund der reduzierten Anzahl der Lautsprecher im oberen Kreis der Klang durch Deckenreflexionen lokalisiert wird und nicht vom Direktschall überlagert wird.

Bewegungsmuster C befindet sich auf der sagittalen Ebene und macht auf Äquatorhöhe einen Distanzfade¹⁵⁶ durch den IKO hin zur Abhörposition. (siehe Abbildung 16) Dabei entspricht das Crescendo einem Amplitudenanstieg durch die Verringerung der Distanz. Zusätzlich

¹⁵² Im Bassbereich unter 100Hz wird das Signal über alle Lautsprecher abgespielt. Daher kann der Bassbereich besser abgebildet werden.

Sharma, „Komponieren mit skulpturalen Klangphänomenen in der Computermusik.“

¹⁵³ Wikipedia. „Transient (acoustics).“

¹⁵⁴ Der Weg den ein Flugkörper zurücklegt.

Clapham und Nicholson, trajectory.

¹⁵⁵ Getzmann und Lewald, „Localization of Moving Sound.“

¹⁵⁶ Eine kontinuierliche Veränderung der Entfernung zur Schallquelle.

wurde das Bewegungsmuster C durch die Veränderung der ambisonischen Ordnung von einem diffusen entfernten Signal zu einem präzisen Signal näher an der Abhörposition.

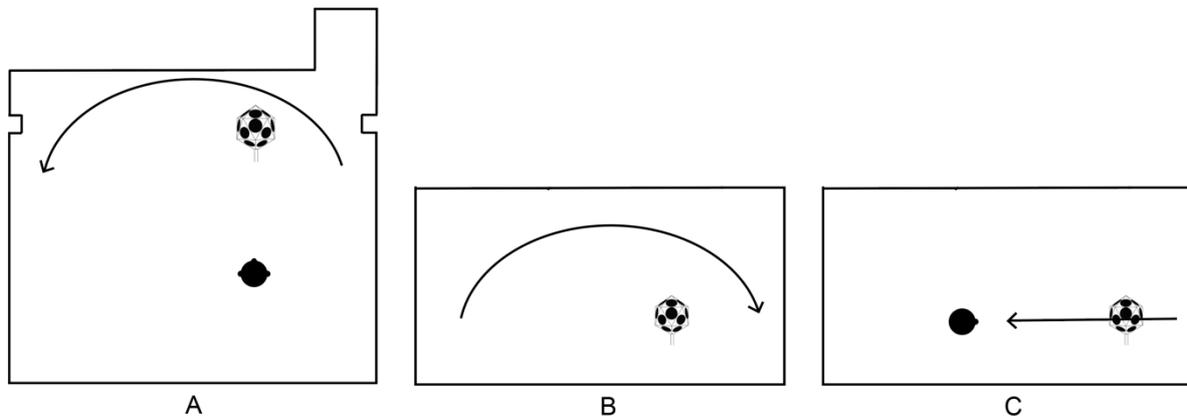


Abbildung 16: Bewegungsmuster des gesuchten Klanges des Sound Logos 1. Das Bewegungsmuster A befindet sich auf der horizontalen Ebene, das Bewegungsmuster B auf der vertikalen Ebene und das Bewegungsmuster C befindet sich auf der sagittalen Ebene.

9.5.2 Sound Logo 2

Das zweiteilige Sound Logo besteht aus einer synthetisch nachgebauten Imitation des Herzschlags und einem ebenfalls synthetischen flächigen Klang, der für das Bewegungsmuster verwendet wird. Der flächige Klang besitzt eine unregelmäßige spektrale Textur. Als visuelle Markenidentität für das Bewegungsmuster wurde der Buchstabe „N“ gewählt. Durch eine Amplitudenmodulation wird das dreiteilige Klangobjekt unterstützt, da Richtungsänderungen vorgenommen werden, um das „N“ im Raum abzubilden. Das Bewegungsmuster startet links hinter der Abhörposition und bewegt sich mit annähernd drei geraden Linien in die rechte Ecke des Raumes. (siehe Abbildung 17) Um eine möglichst gerade Linie zu ermöglichen, wurden die Automationen des Azimuts, der Elevation und der Lautstärke entsprechend angepasst. Je nach Distanz zur Abhörposition wurde auch hier die ambisonische Ordnung verändert, damit durch die Unschärfe des Klanges mehr Tiefe erreicht wird.

Bei den Bewegungsmustern B und C handelt es sich jeweils um Kreisbewegungen auf der horizontalen und auf der sagittalen Ebene. Bei durchgehenden Trajektorien wirkt sich die Amplitudenmodulation negativ auf die Nachvollziehbarkeit des Bewegungsmuster aus.

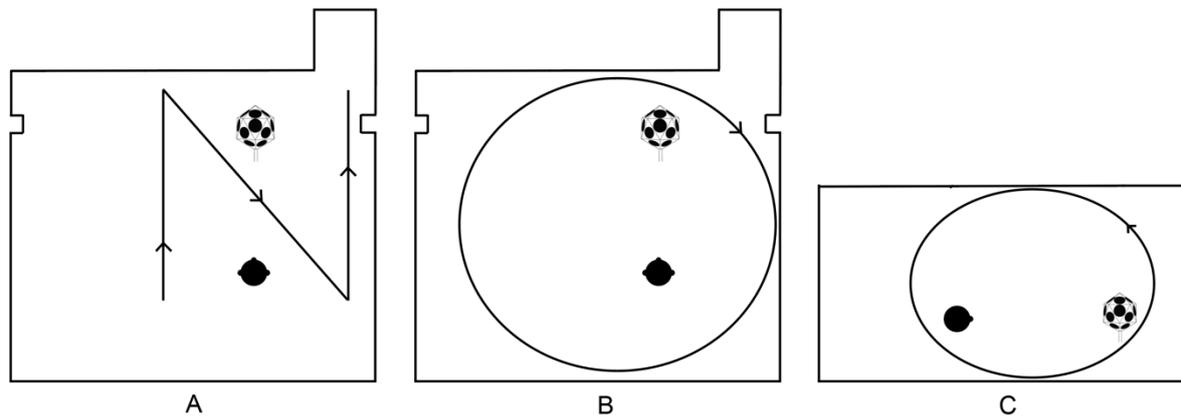


Abbildung 17: Bewegungsmuster des gesuchten Klanges des Sound Logos 2. Die Bewegungsmuster A und B befinden sich auf der horizontalen Ebene und Bewegungsmuster C befindet sich auf der sagittalen Ebene.

9.5.3 Sound Logo 3

Dieses Sound Logo gewinnt zunehmend an Komplexität, da es neben dem gesuchten Bewegungsmuster nicht nur statische Klänge, sondern auch dynamische Bewegungen im Raum gibt. Das Sound Logo beginnt mit einem Arpeggio von Streichinstrumenten. Dieser Klang lässt sich durch das obertonreiche Spektrum sehr leicht lokalisieren, weshalb sich auch die kreisförmige Bewegung gut nachvollziehen lässt. Als Fundament dienen ein tiefer Streicherklang und ein ausgedehnter tiefer Schlag. Abgerundet wird das Sound Logo mit einem helleren Nachklang.

Das Bewegungsmuster A und B bilden jeweils die Form eines „T“ auf der horizontalen Ebene nach. (siehe Abbildung 18) Dabei handelt es sich auch hier nicht um ein durchgehendes Bewegungsmuster, sondern es gliedert sich in zwei Teile. Im ersten Teil findet die Bewegung eines Klaviertones weg von der Abhörposition in Richtung IKO statt. Auch hier wurde mit einer Kombination aus Lautstärkenänderung und Änderung der ambisonischen Ordnung eine räumliche Tiefe erzeugt. Das Klangobjekt schließt im Bewegungsmuster A eine Gerade von links nach rechts ab. Um einen möglichst geradlinigen Verlauf zu erreichen, führt der Klang hinter dem IKO vorbei. Außerdem nimmt der Klang bis auf die Höhe des IKOs an Lautstärke und ambisonischer Ordnung ab und erhöht sich zum Ende hin wieder. So kann vermieden werden, dass die Reflexionen der Rückwand gegenüber den seitlichen Reflexionen dominieren und der klangliche Verlauf verzerrt wird.

Im Bewegungsmuster B bleibt der erste Teil identisch wie bei A, lediglich der Verlauf von links nach rechts wird ersetzt durch ein Aufspannen nach links und rechts ausgehend von der Position des IKOs. Somit bewegen sich zwei ähnliche Signale simultan in entgegengesetzte Richtungen. Durch die Position des IKOs im Raum musste die Ausbreitungsgeschwindigkeit der beiden Signale der zurückzulegenden Distanz angepasst werden, weshalb der Verlauf des rechten Klanges langsamer ist als sein Gegenüber.

Das Bewegungsmuster C beginnt mit einem Distanzfade aus einer räumlichen Tiefe hin zur Abhörposition. Im zweiten Teil spannt sich entsprechend der Abbildung 18 ein vierstrahliger Stern auf. Dabei verlaufen die vier Strahlen simultan vom IKO in die Ecken des Raumes.

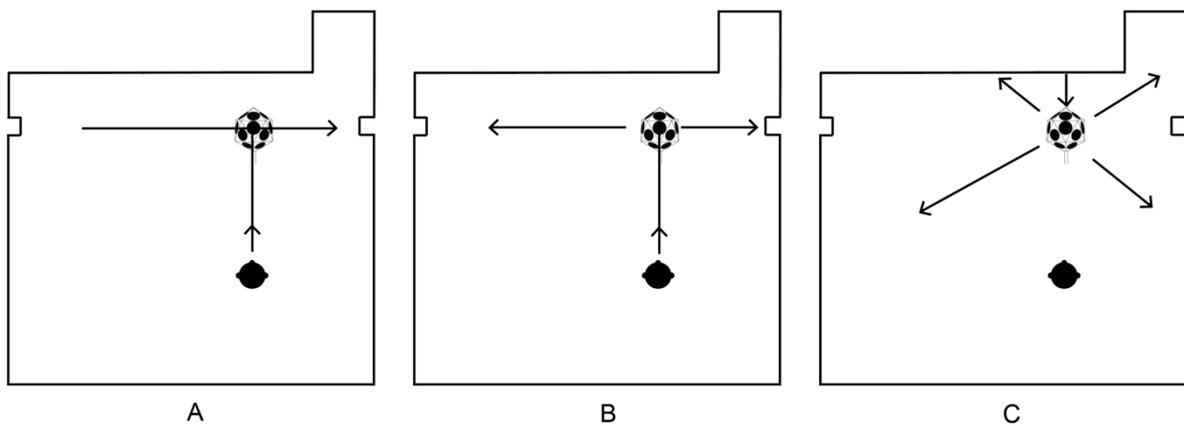


Abbildung 18: Bewegungsmuster des gesuchten Klanges des Sound Logos 3. Alle drei Bewegungsmuster finden auf der horizontalen Ebene statt.

9.5.4 Sound Logo 4

Das Sound Logo 4 startet mit einer Repetition eines Marimba Tones. Dem liegt ein tiefes Dröhnen zu Grunde, das die perkussiven Marimba Klänge leicht maskieren sollte (siehe Kapitel 9.6.4), um die Lokalisierung vom IKO zu entkoppeln.¹⁵⁷ Dieses Bewegungsmuster kann als Vorspiel für das eigentliche Sound Logo gesehen werden. Durch die Repetition des Klanges und ohne melodischen Verlauf, soll der erste Teil des Sound Logos gezielt komplexere Klangobjekte visueller Logos nachahmen, unter anderem einfache Pfeilbewegung, ähnlich einem *Marcato*.

¹⁵⁷ Linke u. a., „How Masking Affects Auditory Objects of Beamformed Sounds.“

Somit kann eine Schnittstelle zwischen dem Visuellen und dem Auditiven etabliert werden. Ein mögliches Einsatzszenario wäre eine Interaktion mit einem Voice Assistenten, wobei die visuell konnotierte Einleitung als Startsignal gesehen werden kann und das Sound Logo als Abschlussfloskel.

Damit die einzelnen Marimba Klänge als einheitliche Gestalt wahrgenommen werden, findet in unserem Gehirn ein Prozess der Gruppierung statt, der von Bregman¹⁵⁸ als *Auditory Streaming* bezeichnet wird.

Bewegungsmuster A bildet, wie in Abbildung 19 zu sehen ist, ein Quadrat ab. Damit die Bewegung nicht als Kreis wahrgenommen wird, sondern die Ecken ebenfalls nachempfunden werden können, wurden jeweils die Eckpunkte mit Akzenten versehen.

Die Beispiele B und C haben ein identisches Bewegungsmuster, unterscheiden sich allerdings durch die klangliche Komponente, indem die Marimba durch Staccato Klänge von Streichinstrumenten ersetzt wird. Dieses identische Muster bewegt sich auf der vertikalen Ebene. Es startet von der Abhörposition rechts oben, wandert zum Boden und setzt die Bewegung diagonal nach links oben fort.

Das Sound Logo 4 besitzt gegenüber den anderen noch ein Bewegungsmuster D. Dieses Bewegungsmuster verläuft nicht nur eindimensional, sondern findet sowohl in der horizontalen als auch in der vertikalen Ebene statt. Während auf der horizontalen Ebene eine Bewegung zur Hörposition hin passiert, wird vertikal die Spitze eines Pfeils abgebildet.

¹⁵⁸ Bregman, „Auditory Scene Analysis: The Perceptual Organization of Sound.“

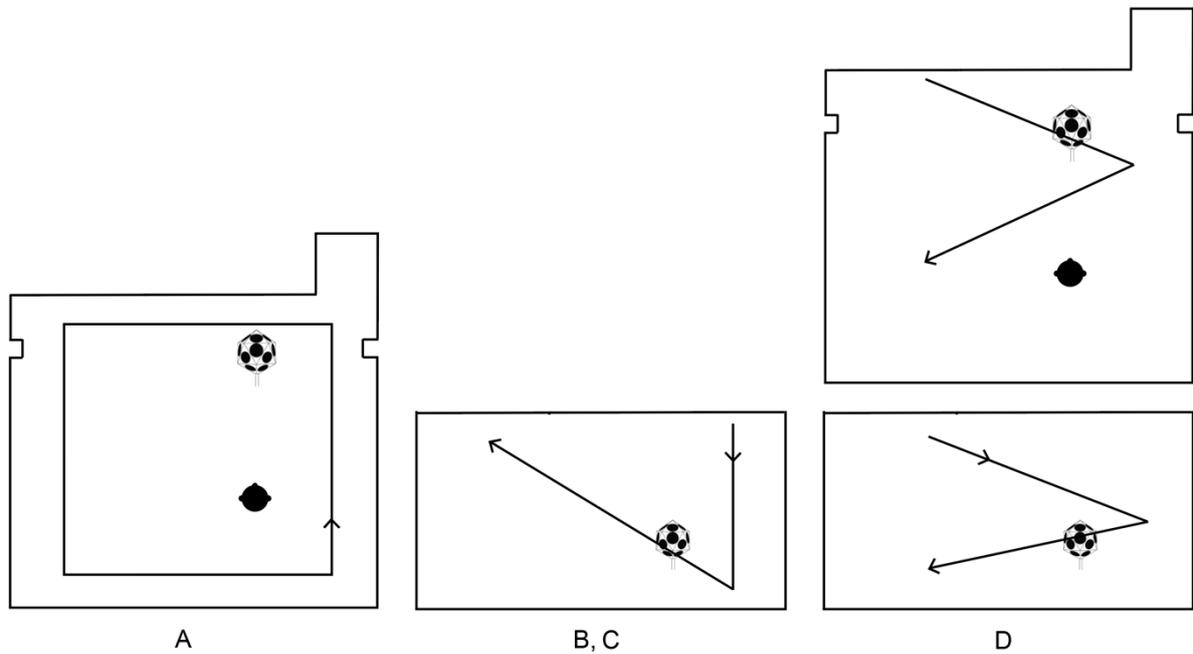


Abbildung 19: Bewegungsmuster des gesuchten Klanges des Sound Logos 4. Während das Bewegungsmuster A auf der horizontalen Ebene und die Bewegungsmuster B und C auf der vertikalen Ebene stattfinden, bewegt sich das Sound Logo D gleichzeitig auf der horizontalen, als auch auf der vertikalen Ebene.

9.6 Zusammenfassung der Gestaltungsmaßnahmen

Bei dem Entwurf und der Umsetzung der Sound Logos wurde ein iterativer Arbeitsprozess verfolgt, um etwaige Fehlerquellen und mögliche Lösungsansätze in einer konstanten Reflexion in die künstlerische Arbeit mit einfließen zu lassen. Ausgehend von der Arbeit von Sharma¹⁵⁹ wurden bereits vorhandene Erkenntnisse als Ausgangspunkt für weiterführende Untersuchungen verwendet und so die Grenzen des Umsetzbaren abgesteckt. Dabei werden nachfolgend einige wichtige Erfahrungen zusammengefasst.

9.6.1 Harmonische Struktur und Tonhöhe

Wie bereits bei der räumlichen Wahrnehmung beschrieben, spielt das Obertonspektrum eine wichtige Rolle für die Lokalisierung des Klanges. Demnach verfügen Sinustöne über wenig brauchbare Informationen, um die Position im Raum wahrnehmen zu können.¹⁶⁰ Daher

¹⁵⁹ Sharma, „Komponieren mit skulpturalen Klangphänomenen in der Computermusik.“

¹⁶⁰ Rakerd und Hartmann, „Localization of Sound in Rooms, III.“

wurden sinusähnliche Töne für Klangflächen verwendet, um ein möglichst einhüllendes Gefühl, besonders im tieferen Frequenzbereich zu bekommen.

Je mehr Obertöne in einem Klang enthalten sind, desto mehr Richtungsinformationen gelangen zu den Ohren und umso besser funktioniert die Lokalisierung. Das wurde in einem direkten Vergleich im Sound Logo 4 praktisch umgesetzt, wo dasselbe Klangobjekt einmal mit einer Marimba und einmal mit einem *Spiccato* Klang eines Streichinstruments in den Raum gezeichnet wird.

Ein weiterer wichtiger Faktor ist die Tonhöhe und die daraus resultierenden Lokalisierungsunterschiede. Diese entsprechen den frequenzabhängigen Bändern nach Blauert.¹⁶¹ Konkret kann dieses Phänomen in dem fünftönigen Marimba Motiv des Sound Logos 4 wahrgenommen werden, indem der Oktavsprung innerhalb der Melodie zu einer Veränderung der wahrgenommen Position auf der horizontalen Ebene führt.

9.6.2 Lautstärke

Um die Position eines Klanges feststellen zu können, braucht es nicht nur genügend Informationen für den Hörapparat, sondern diese müssen auch deutlich abgebildet werden, weshalb die Lautstärke eine wichtige Rolle spielt. Vor allem im Sound Logo 3A und 3B musste für den zweiten Teil des „T-förmigen“ Bewegungsmusters an den äußeren Positionen die Lautstärke deutlich angehoben werden, da auf Grund der zurückgelegten Distanz und der Reflexion das Signal zunehmend unschärfer wird. Diese Lautstärkenkompensation beträgt für das Sound Logo 3A knapp 10dB. Vor allem wenn zusätzliche Klänge die Dominanz des Bewegungsmusters reduzieren, ist ein entschiedener Start durch eine erhöhte Lautstärke notwendig.

Wie bereits bei der Distanzwahrnehmung erwähnt, ist die Intensität des Schallereignisses ausschlaggebend, um die Entfernung zur Schallquelle zu bestimmen. Es müssen somit Reflexionen und deren zurückgelegter Weg für die Balance der Lautstärke berücksichtigt werden. Daher wurden die Bewegungsmuster an der Position des IKOs leiser gemacht, da diese Reflexionen, abgesehen vom Direktschall, als erste an der Abhörposition eintreffen.

¹⁶¹ Blauert, Spatial Hearing: The Psychophysics of Human Sound Localization.

9.6.3 Ambisonische Ordnung

Neben der Lautstärke des Signals ist auch die ambisonische Ordnung entscheidend für die Distanzwahrnehmung. Durch eine stufenlose Veränderung der Ordnung kann die Verbreiterung des Signals so vorgenommen werden, dass ein geradliniger Verlauf entsteht und die Unschärfe erhöht werden kann. Bei einer niedrigeren Ordnung sind mehr Lautsprecher in die Klangreproduktion involviert, weshalb die Gefahr eines unerwünschten Direktschalls steigt. Es handelt sich bei der Imitation der Tiefe um einen Balanceakt zwischen der Ordnung und einer Lautstärkenkompensation. Wenn ein Distanzfade erreicht werden soll, aber auf Grund der Komplexität des Klangobjekts präzise Reflexionen notwendig sind, sollte eine möglichst hohe ambisonische Ordnung verwendet werden und die Intensität über die Lautstärke reguliert werden.

9.6.4 Maskierung von Transienten

Mit der Variation der ambisonischen Ordnung entstehen diffusere Schallfelder, die einer genauen Lokalisation hinderlich sind. Dieses diffuse Schallfeld kann aber für die Lokalisation anderer Klänge von großer Bedeutung sein, besonders für Klangobjekte mit einem hohen Anteil an Transienten.¹⁶² Daher wurde versucht, mit den einzelnen Klängen ein möglichst breites und dichtes Panorama zu erzeugen, damit Bewegungsmuster mit einer kurzen Einschwingzeit besser nachvollzogen werden können. Dieses diffuse Schallfeld wurde besonders durch den Einsatz von *FdnReverb* unterstützt. Auch hier ist die Balance der Lautstärke zwischen dem Bewegungsmuster und dem maskierenden Schallfeld wichtig.

9.6.5 Simultane Klänge

Die auditive Abbildung eines Klangobjekts mit mehreren gleichzeitig auftretenden Bewegungsmustern eines ähnlichen Klanges ist nur sehr schwer möglich. Bei zwei diametralen Bewegungen ausgehend vom IKO, entsteht trotz leichter Abweichung der beiden Signale, ähnlich der Stereowiedergabe, eine Art Phantomschallquelle. Man kann zwar ein

¹⁶² Linke u. a., „How Masking Affects Auditory Objects of Beamformed Sounds.“

breiteres Panorama feststellen, aber der Zuhörer tendiert dazu, den Klang am IKO zu lokalisieren. Dies erschwert die Wahrnehmung der exakten Bewegung. Bei vier gleichen und simultanen Bewegungen wie in Abbildung 18 dargestellt, ist es nicht mehr möglich, das Klangobjekt wahrzunehmen, allerdings erhöht sich dadurch die Umhüllung.

Wenn es sich um zwei oder mehr statische Klänge handelt, die auf unterschiedliche Positionen im Raum projiziert werden, können leichte Veränderungen der Richtungswahrnehmung die Lokalisierung unterstützen. Dabei wird der Winkel des Azimuts minimal hin und her bewegt. Dadurch verändert sich die Position des Klanges nicht, aber durch die zusätzlichen Informationen wird die Richtungswahrnehmung unterstützt.

9.6.6 Akzente

Durch einzelne Schwerpunkte in dynamischen Klangobjekten kann die Nachvollziehbarkeit von Bewegungsmustern unterstützt werden. Besonders wenn charakteristische Konturen auditiv gezeichnet werden sollen, kann eine stärkere Betonung einen Ankerpunkt liefern. Das betrifft vor allem Richtungswechsel wie sie für das Sound Logo 4 verwendet wurden.

Wenn es sich allerdings um durchgehende Trajektorien handelt, gilt es, abrupte Änderungen sowie Modulationen (Siehe Sound Logo 2) zu vermeiden, da dies bei komplexeren Klangobjekten zu einer zusätzlichen Verwirrung führen kann.

10 Hörversuche

Nach der Konzipierung der 3D Audio Sound Logos und deren praktischer Umsetzung mit IKO müssen die im Entstehungsprozess gewonnenen Erkenntnisse nicht nur aus der Sicht des Sound Designers beurteilt werden, sondern ebenfalls von Personen, die nicht intensiv in den Entwicklungsprozess involviert waren. Es gilt intersubjektive Diskrepanzen in der Wahrnehmung herauszufiltern. Dabei spielt bei einer Untersuchung im Rahmen des Sound Brandings nicht nur die Lokalisierung des Klanges eine wichtige Rolle, sondern auch die Wahrnehmung des Klangobjektes und der Effekt der 3D Audio Klangwiedergabe.

10.1 Probanden und Abhörraum

An dem Hörversuch nahmen insgesamt sieben Personen teil, die zum Zeitpunkt der Durchführung alle zwischen dem 20 und 30 Lebensjahr waren. Die Gruppe setzte sich aus vier männlichen und drei weiblichen Teilnehmer*innen zusammen. Davon bezeichneten sich drei Personen als professionelle Musiker*innen, Tontechniker*innen oder Sound Designer*innen. Zwei Personen ordneten sich der Gruppe der Hobbymusiker*innen zu und zwei Proband*innen bezeichnen sich als ungeübt. Somit ergibt sich eine Ausgewogenheit in der musikalischen Erfahrung der Versuchspersonen.

Bei einer genaueren Erhebung der Vorkenntnisse hatten vier Personen bereits Erfahrung mit 3D Audio und lediglich eine Person Erfahrung mit dem IKO. Eine Person hatte Erfahrung mit Sound Branding bzw. war mit den Grundprinzipien des Sound Brandings vertraut.

Die Hörversuche wurden im Hörsaal des *IEM* durchgeführt, somit entspricht der Aufbau der Hörversuche jenen in Abbildung 15 dargestellten räumlichen Parametern.

Die Wiedergabelautstärke der Sound Logos betrug gemittelt 72,6dB SPL A-bewertet¹⁶³. Der Geräuschpegel im Abhörraum betrug 29,9dB SPL A-bewertet.

10.2 Methode

Der Hörversuch setzt sich aus einer perceptiven und einer affektiven Erhebung der Daten zusammen. In einer indirekten perceptiven Erhebung mussten die Proband*innen das gesuchte Bewegungsmuster selbst per Hand in eine vorgefertigte Skizze einzeichnen. Dabei wurden für jedes Sound Logo und jedes Bewegungsmuster Skizzen der drei räumlichen Dimensionen (siehe Abbildung 20) vorgegeben. Der gesuchte Klang wurde zuerst einzeln und ohne eine räumliche Bewegung vorgespielt. Man konnte auch während der Aufgaben immer wieder den Beispielklang anhören. Da sich alle Bewegungsmuster lediglich in einer räumlichen Dimension (abgesehen von einer Ausnahme, die explizit als solche erwähnt

¹⁶³ Als A-Bewertung bezeichnet man einen Filter, wo die frequenzabhängige Lautstärkenwahrnehmung an das menschliche Ohr angepasst ist. Gegenüber einer linearen Messung des Schalldruckpegels werden Frequenzbereiche entsprechend gedämpft. Diese Filterung des Schalldruckpegels wird meist mit SPL (A) abgekürzt.
Friesecke, Die Audio-Enzyklopädie.

wurde) bewegten, bestand die erste Aufgabe in der Bestimmung der Dimension. Erst im nächsten Schritt sollte der klangliche Verlauf in der jeweiligen räumlichen Darstellung eingezeichnet werden (siehe Abbildung 20).

Die Sound Logos konnten beliebig oft abgespielt werden bzw. konnte auch während dem Hörversuch die Reihenfolge nach Belieben geändert werden. Somit sollte eine Verzerrung bei den ersten Beispielen durch eine Eingewöhnungsphase verhindert werden. Es wurde allerdings empfohlen, dass zumindest im ersten Durchgang die Reihenfolge auf Grund der zunehmenden Schwierigkeit beibehalten wird. Die Steuerung der Wiedergabe erfolgte über *Reaper*, indem mit Hilfe von Markern die Sound Logos durch den Ziffernblock angesteuert werden konnten.

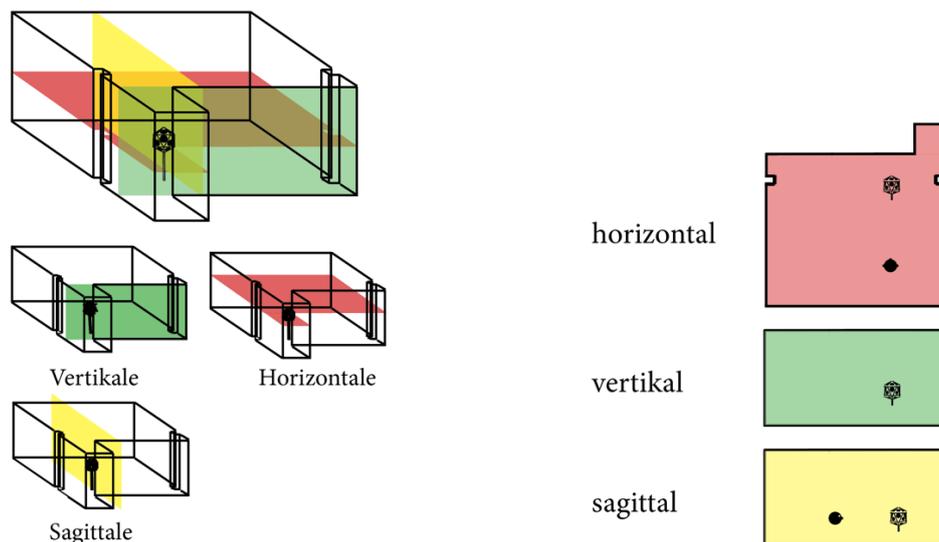


Abbildung 20: Links: graphische Darstellung der räumlichen Dimensionen.

Rechts: Skizzen, wo das Bewegungsmuster eingezeichnet werden musste.

Zuerst musste die richtige räumliche Dimension erkannt werden und anschließend musste das Bewegungsmuster in die zutreffende Ebene eingezeichnet werden.

Für die affektive Datenerhebung musste zuerst eine Reihung des Bewegungsmusters für jedes Sound Logo nach dem Parameter „Gefallen“ vorgenommen werden. Auf einer 7-teiligen Skala musste anschließend bewertet werden, wie sehr man sich von dem Sound Logo miteinbezogen gefühlt hat. Ebenfalls mussten die Anzahl der einzelnen Ereignisse (Klänge, Bewegungen etc.) und die generelle Nachvollziehbarkeit ebenfalls in einer 7-teiligen Skala

ingezeichnet werden. So sollte die subjektive Wahrnehmung der Proband*innen festgehalten werden und Zusammenhänge zwischen den Parametern ausgemacht werden können. Nach jedem Sound Logo gab es einen entsprechenden Platz, um etwaige Anmerkungen zu protokollieren. (siehe Anhang: Fragebogen Hörversuche) Bei Sound Logo 2 mussten zusätzlich die Wiederholungen des Bewegungsmusters erkannt werden.

Bei Sound Logo 4 änderte sich die Aufgabenstellung dahingehend, dass in diesem Beispiel jeweils vier vorgefertigte Bewegungsmuster angeführt waren und entsprechend der wahrgenommenen Klangbewegung eines ausgewählt werden musste (siehe Abbildung 21). So sollte der Zusammenhang zwischen komplexen Klangobjekten und einer visuellen Hilfestellung eruiert werden.

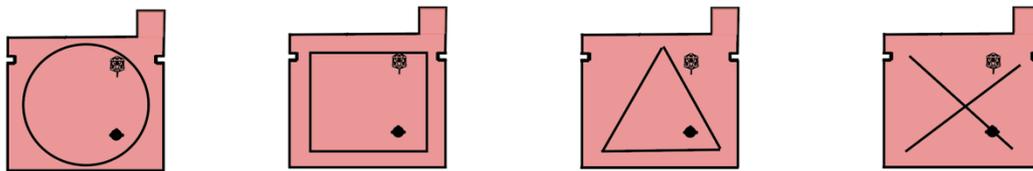


Abbildung 21: Auswahlmöglichkeiten für das Bewegungsmuster des Sound Logos 4.

10.3 Ablauf

Zuerst wurde den Proband*innen der inhaltliche Rahmen der Arbeit erklärt und der IKO vorgestellt. Anschließend wurden die räumlichen Dimensionen genau erklärt und mit Probebeispielen auditiv veranschaulicht. Da es in Testversuchen zu Unklarheiten bei den räumlichen Dimensionen gekommen war, wurde dafür ausreichend Zeit für präzise Erklärungen und Fragen eingeplant. Zusätzlich wurde eine Skizze angefertigt (siehe Abbildung 20), die auf jeder A4-Seite des Fragebogens als Hilfestellung abgebildet wurde. Anschließend wurden die Aufgabenstellungen durchbesprochen und die Wiedergabe über *Reaper* trainiert. Nach einer gewissen Vorbereitungszeit und der Beantwortung aller Fragen, konnten die Proband*innen alleine im Abhörraum ungestört die Hörversuch durchführen. Dabei gab es keinen zeitlich festgelegten Rahmen.

10.4 Resultate

Da ein großer Teil der Hörversuche auf einer graphischen Reproduktion der Hörwahrnehmung beruht, müssen die Ergebnisse sehr individuell betrachtet werden. Bei Sound Logo 1A konnten fünf von sieben Proband*innen das Bewegungsmuster erkennen, ebenso bei Beispiel C. Zusätzlich wurde bei dem Bewegungsmuster C die räumliche Dimension von allen Proband*innen richtig erkannt. Die Dimension des Beispiels B wurde zwar in vier von sieben Fällen richtig erkannt, der genaue Verlauf des Klanges konnte allerdings nur von zwei Personen nachvollzogen werden. Bei der Bewertung der Bewegungsmuster wurde mit einer Mehrheit von jeweils 57% A auf Platz 1 und B auf Platz 2 gewählt. Mit einer deutlicheren Mehrheit wurde das Bewegungsmuster C auf Platz 3 gewählt. Bei Sound Logo 2 konnten bei der Variante A und C jeweils die Dimension und das Bewegungsmuster nicht erkannt werden. Bei der Variante B konnten 85% die grobe Struktur des Klangobjektes erkennen, allerdings konnte es nur eine Person exakt erfassen. Bei Sound Logo 2 blieb eine Reihung nach dem Gefallen ohne Ergebnis, da sich keine eindeutigen Tendenzen ausmachen ließen.

Bei Sound Logo 3 konnten die komplexen Klangobjekte nur mehr fragmentarisch wahrgenommen werden. Bei der Bewertung haben sechs von sieben Personen die Variante A am besten bewertet.

Bei der Bewertung des Einbezogen-Seins, der Anzahl der Ereignisse und der Nachvollziehbarkeit (siehe Abbildung 22) kam es zu keinen merklichen Unterschieden zwischen den Sound Logos. Der Mittelwert der Bewertung des Einbezogen-Seins fällt bei allen drei Sound Logos gleich aus. Sound Logo 1 weist bei der Anzahl der einzelnen Ergebnisse den höchsten Wert auf. Je höher der Wert ist, desto mehr empfanden die Proband*innen die Anzahl als zu viel. Den zweithöchsten Wert hat Sound Logo 2, das lediglich aus zwei Klängen besteht, allerdings sehr ausgeprägte Klangbewegungen.

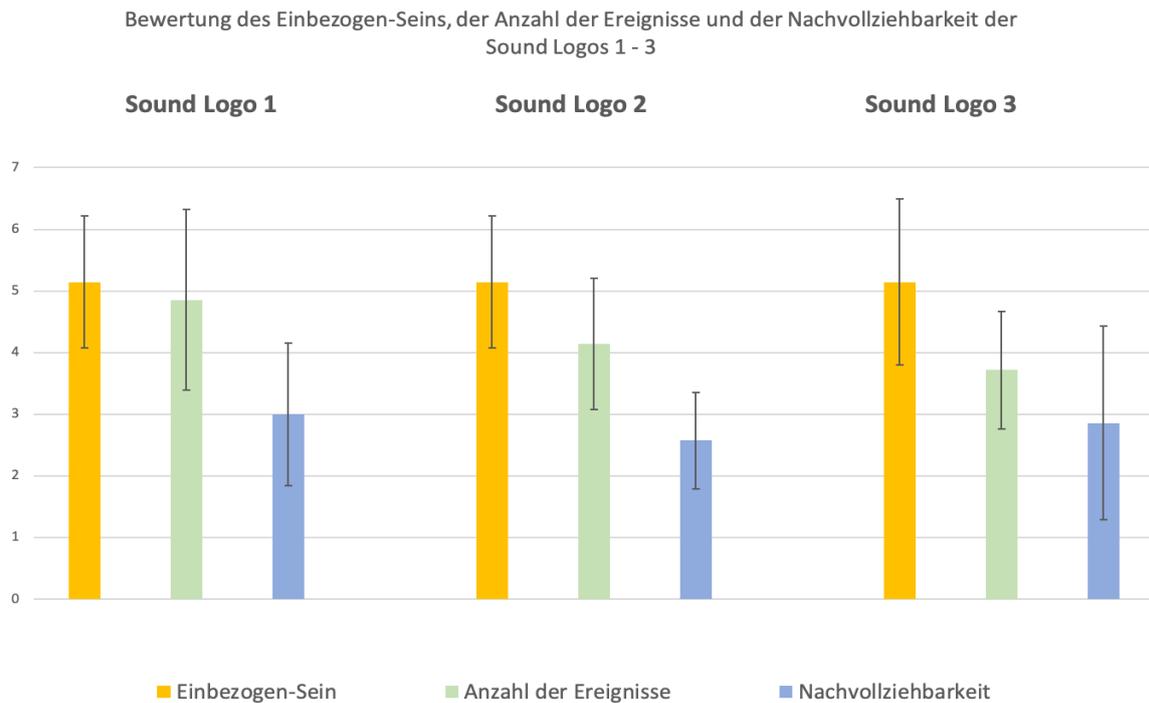


Abbildung 22: Mittelwert inklusive Fehlerindikation der Bewertung der Sound Logos in Bezug auf Einbezogen-Sein, Anzahl der Ereignisse und Nachvollziehbarkeit.

Bei Sound Logo 4A wurde in vier von sieben Fällen das richtige Klangobjekt erkannt. Bei der Variante B waren es nur drei und bei C lediglich 2. Bei der „zweidimensionalen“ Variante D konnten fünf von sieben Proband*innen das richtige Bewegungsmuster erkennen.

11 Diskussion

Die Ergebnisse der Hörversuche unterstreichen die große Anzahl an Einflussfaktoren, welche berücksichtigt werden müssen, um einer Verzerrung der klanglichen Abbildung vorzubeugen. Bei der Betrachtung der Ergebnisse von Sound Logo 1 lässt sich der Einfluss der räumlichen Dimension, in Anbetracht der in Kapitel 2 beschriebenen Faktoren für die Richtungswahrnehmung, deutlich veranschaulichen. So werden die Bewegungen des Klanges in der horizontalen Ebene präziser wahrgenommen, als jene auf der vertikalen Ebene wie in der Variante B.¹⁶⁴ Obwohl die Dimension durch eine Veränderung der Elevation von der

¹⁶⁴ Blauert, Spatial Hearing: The Psychophysics of Human Sound Localization.

Mehrheit der Proband*innen richtig erkannt wurde, konnte eine identische Bewegung wie in der horizontalen Ebene nicht in der vertikalen Ebene wahrgenommen werden.

Die gleiche Erkenntnis lässt sich auf die sagittale Ebene übertragen, wo die Distanzänderungen rein auf der Äquatorhöhe, wie im Sound Logo 3, deutlich besser wahrgenommen wurden. Mit zusätzlichen Richtungsinformationen auf der vertikalen Ebene, wie bei Sound Logo 2C, steigt jedoch der Lokalisierungsfehler. Auch bei Sound Logo 4C (siehe Abbildung 23 rechts) zeigt sich, wie sehr die Wahrnehmung von der räumlichen Dimension abhängt. In zwei von sieben Fällen wurde das richtige Bewegungsmuster erkannt, vier der falsch gelegenen fünf Proband*innen tippten allerdings auf eine einfache rechts – links Bewegung. (siehe Abbildung 23 links). Anhand dieses Beispiels erkennt man, dass vor allem die horizontalen Informationen für die Auswahl berücksichtigt wurden und die Richtungsinformationen auf der vertikalen Ebene verloren gingen.



Abbildung 23: Links: Auswahlmöglichkeit für das Bewegungsmuster des Sound Logos 4B, C.

Rechts: Bewegungsmuster des gesuchten Klanges des Sound Logos 4B, C.

Ebenfalls zeichnet sich eine Tendenz in der Wahrnehmung bestimmter Formen ab. Bregman beschreibt in der *Auditory Scene Analysis*¹⁶⁵ ausgewählte geometrischen Figuren als starke Formen. Als Beispiel wird von Bregman der Kreis angeführt, der in Untersuchungen auch trotz Unterbrechungen als geschlossene und zusammengehörige Einheit interpretiert wird. Dieser Zusammenhang zeichnet sich auch bei den Hörversuchen ansatzweise ab, indem durchgehende Trajektorien leichter wahrgenommen wurden. Das trifft besonders auf das Sound Logo 4A zu, wo trotz einzelner Marimba Schläge mit entsprechend schwerer Lokalisierbarkeit in sechs von sieben Fällen die grobe geometrische Form wahrgenommen werden konnte. Mit Hilfe von Akzenten, also Schwerpunkten auf einzelnen Klängen, konnten

¹⁶⁵ Bregman, „Auditory Scene Analysis: The Perceptual Organization of Sound.“

sogar noch schärfere Konturen gezeichnet werden. So konnten von den sechs Personen, welche die grobe geometrische Form erkannt hatten, sogar 66% ein Quadrat erkennen.

Bei Sound Logo 2B konnte die Kreisbewegung nicht erkannt werden, da ab einem gewissen Punkt die Kreisbewegung durch Direktschall unterbrochen wurde. Da durch den Präzedenzeffekt (siehe Kapitel 4.1) der Direktschall den Verlauf unterbricht, ist es sehr schwer, den Raum zwischen dem IKO und der Abhörposition zu bespielen. Wie die Ergebnisse des Hörversuchs zeigen, konnte auch die Rotationsgeschwindigkeit des Klanges die logische Fortsetzung des Klanges (*Auditory Scene Analysis*¹⁶⁶) nicht unterstützen, sondern die Auswertung der graphischen Darstellung der Proband*innen lässt auf den Effekt des *Spatial Beating* schließen.¹⁶⁷ Dabei wird durch eine schnelle Bewegung eines breitbandigen Signals nicht die Kreisbewegung wahrgenommen, sondern nur punktuelle Reflexionen. Es kann somit zu einem Pochen an den Wänden kommen bzw. auch eine klangliche Pendelbewegung stattfinden.

Sound Logo 3 zeigt, dass mit zunehmender Anzahl an simultanen Bewegungen das Klangobjekt nicht mehr erkannt werden kann. Auch wenn sich die Klänge leicht unterscheiden um Phantomschallquellen zu vermeiden, können sie dennoch nicht separiert werden. Die Erklärung ist hier ebenfalls bei der *Auditory Scene Analysis*¹⁶⁸ zu finden, da die einzelnen Bewegungen durch die klangliche Ähnlichkeit in sogenannte *Auditory Streams*¹⁶⁹ zusammengefasst werden und gruppiert weiterverarbeitet werden.

Wie bereits bei den Gestaltungsmaßnahmen erwähnt, hat das Klangspektrum einen entscheidenden Einfluss auf die Lokalisierung der Klänge. Bei Sound Logo 4D zeigt sich, dass ein Klang mit einem breiten Klangspektrum auch in einem mehrdimensionalen Bewegungsmuster nachvollzogen werden kann. Bei dem Versuchsaufbau von Sound Logo 4 zeigt sich auch, dass die visuelle Hilfestellung durch die vorgefertigten Skizzen die Nachvollziehbarkeit der Klänge deutlich erhöht. Während bei Sound Logo 4 trotz zunehmender Komplexität 50% aller Beispiele richtig erkannt wurden, lag der Durchschnitt bei Sound Logo 1 – 3 deutlich darunter.

¹⁶⁶ Bregman, „Auditory Scene Analysis: The Perceptual Organization of Sound.“

¹⁶⁷ Sharma, „Komponieren mit skulpturalen Klangphänomenen in der Computermusik.“

¹⁶⁸ Bregman, „Auditory Scene Analysis: The Perceptual Organization of Sound.“

¹⁶⁹ Auditory Stream beschreibt die neurale Gruppierung von auditiven Informationsströmen. Dafür werden im Gehirn Klänge mit ähnlichen spektralen Eigenschaften zu Clustern geformt. Mit dieser Ordnung wird eine Verbindung der physikalischen Welt und der mentale Repräsentation geschaffen. Das visuelle Pendant wäre das Objekt.

Wie in Kapitel 3 angemerkt, hat die visuelle Ebene einen direkten Einfluss auf die Bewertung der auditiven Qualität. Obwohl bei Sound Logo 1 – 3 keine visuellen Vorlagen gegeben waren, gibt es die Tendenz, dass einfachere Bewegungsmuster, die öfter erkannt wurden, auch dem Gefallen nach besser bewertet wurden. Daher stellt sich die Frage, wie sehr eine leichtere mentale Reproduktion des Bewegungsmusters des Sound Logos sich auf die Bewertung auswirkt.

Generell zeigt sich der visuelle Einfluss beim IKO selbst, da durch eine Fokussierung auf das Lautsprechersystem die Aufmerksamkeit für den Raumklang verloren geht.¹⁷⁰ Dieser Faktor rückt vor allem in den Fokus, weil nur eine der Testpersonen Erfahrung bzw. zuvor Kontakt mit dem IKO hatte. Für alle anderen Proband*innen war es eine neue Erfahrung, was vermuten lässt, dass die Aufmerksamkeit sehr stark auf den IKO gerichtet war und somit, entsprechend der Visu-Audition (siehe Kapitel 3), eine Verzerrung in der Wahrnehmung des Klangbildes nicht ausgeschlossen werden kann.

12 Zusammenfassung und Ausblick

Diese praktische Umsetzung unterschiedlicher Sound Branding Konzepte für 3D Audio-fähige Lautsprechersysteme zielt darauf ab, Gestaltungsmaßnahmen für den strategischen Einsatz von Klangobjekten in der akustischen Markenkommunikation abzuleiten. Dafür wurden in einem iterativen Prozess unterschiedliche Ansätze erarbeitet, welche breit aufgestellt diverse Zugänge berücksichtigten. Durch die Vielzahl an Aspekten, die in dem Hörversuch evaluiert wurden, konnten zwar einige Möglichkeiten nur sehr oberflächlich betrachtet werden, eröffneten allerdings viele Teilaspekte für zukünftige Untersuchungen. Es konnten neben den gestalterischen Erkenntnissen mit dem 3D Audioformat Ambisonics auch eindeutige Tendenzen für die Konzipierung einer umhüllenden akustischen Identität gefunden werden.

Der Hörversuch hat gezeigt, dass Bewegungsmuster, die einfacher nachvollziehbar waren, auch besser bewertet wurden. Daher erscheint es in diesem Zusammenhang von Bedeutung,

¹⁷⁰ Sharma, „Komponieren mit skulpturalen Klangphänomenen in der Computermusik.“

die Balance zwischen einer auditiven Imitation des visuellen Logos und einer Erweiterung eines statischen Sound Logos mit effektvollen dynamischen Bewegungen zu finden.

Um eine möglichst präzise Darstellung des Klangobjektes erreichen zu können, sollten die Klangbewegungen auf der horizontalen Ebene stattfinden. Vertikale Informationen können durch die ungenaue Lokalisierung leicht verloren gehen.

Weiters konnte festgestellt werden, dass sich der Raum zwischen dem Lautsprechersystem und der Abhörposition kaum mit Klangobjekten bespielen lässt. Die Dominanz des Direktschalls führt auch bei vorbeifahrenden Trajektorien zu Verzerrungen, welche sich kaum kompensieren lassen. Daher sollte sich die Klangzone eher hinter dem Lautsprecher befinden, zumal sich eine Distanz zu den Klangobjekten durch Lautstärkenunterschiede und durch die Veränderung der ambionischen Ordnung gut realisieren lässt.

Visuelle Anker haben die Wahrnehmung der Klangbewegungen deutlich unterstützt. So konnten auch mit Hilfestellungen komplexere Klangobjekte in unterschiedlichen Dimensionen erkannt werden. Da alle Sound Logos neu geschaffen wurden, hatten die Proband*innen keine Berührungspunkte und abgespeicherte Informationen zu den fiktiven Marken. Es wurde im Rahmen dieser Hörversuche nicht untersucht, in wie weit ein abgespeichertes Bild, vor allem in Bezug auf das visuelle Logo einer Marke, als Hilfestellung für die Lokalisierung dient. Die Untersuchungen lassen jedoch darauf schließen, dass jegliche Form einer Konditionierung, wie sie in einem holistischen Konzept der Markenkommunikation verfolgt wird, für die Wahrnehmung der Klangobjekte zuträglich ist.

Da im Rahmen dieser Hörversuche nur mit dem IKO gearbeitet wurde, sollten weiterführende Untersuchungen mit unterschiedlichen Endgeräten und Wiedergaberäumen das Anforderungsprofil für die Konzipierung von 3D Audio Sound Logos weiter schärfen.

Zusätzlich kann die Ausarbeitung eines 3D Audio Konzepts für eine konkrete Marke weitere Einblicke in die Zusammenhänge zwischen der konditionierten Vorerfahrung und dem Erkennen der Bewegungsmuster liefern.

„Hearing does not offer a meta-position; there is no place where I am not simultaneous with the heard. [...] I cannot hear it if I am not immersed in its auditory object[...].“¹⁷¹

¹⁷¹ Voegelin, Listening to noise and silence.

Man kann Klänge nicht von außen betrachten, sondern sie umhüllen einen und man wird natürlicherweise Teil dieser Klangumgebung. Mit immer besserer 3D Audio Technologie kann genau diese Klangumgebung aktiv mitgestaltet werden. Für die akustische Markenkommunikation öffnet sich somit ein Raum an Möglichkeiten, der Platz für neue Innovationen bringt und maßgeblich mitbeeinflussen kann, wie wir eine Marke erleben und wie uns eine Marke in Erinnerung bleibt. Vor allem durch 3D Audio-fähige Lautsprecher-systeme und zunehmend audiophilere Interaktionen durch Sprachassistenten, wird 3D Audio zukünftig ein wesentlicher Bestandteil der auditiven Markenkommunikation sein. Dabei gilt es, die klangliche Identität facettenreich und zukunftsorientiert zu gestalten, um nicht Passagier, sondern ein Pionier im Übergang in die neue Dimension des Sound Brandings zu sein.

Literaturverzeichnis

Agrawal, Sarvesh, Adèle Simon, und Søren Bech. „Defining Immersion: Literature Review and Implications for Research on Immersive Audiovisual Experiences.“ Paper presented at the AES 147th Convention, New York, USA, Oktober 16-19,2019.

Altman, Marcus, Kurt Krauss, Joel Susal, Nicolas Tsingos. „Immersive Audio for VR.“ Paper presented at the Conference on Audio for Virtual and Augmented Reality, Los Angeles, USA, September 30 - Oktober 1, 2016.

Anzenbacher, Christoph. Audiologos: integrative Gestaltungsmaßnahmen vor dem Hintergrund der Musikpsychologie. 1. Auflage. Praxisforum Medienmanagement, Band 17. Baden-Baden: Nomos, 2012.

Begault, Durand R. 3-D Sound for Virtual Reality and Multimedia. Cambridge, Mass [u.a.]: AP Professional, 1994.

Berg, Jan. „The Contrasting and Conflicting Definitions of Envelopment.“ Paper presented at the 126th AES Convention, Munich, Germany, Mai 7-10, 2009.

Bishop, Christopher W., Sam London, und Lee M. Miller. „Neural Time Course of Visually Enhanced Echo Suppression.“ *Journal of Neurophysiology* 108, Nr. 7 (Oktober 2012): 1869–83. <https://doi.org/10.1152/jn.00175.2012>.

Blauert, Jens. „Localization and the Law of the First Wavefront in the Median Plane.“ *The Journal of the Acoustical Society of America* 50, Nr. 2B (August 1971): 466–70. <https://doi.org/10.1121/1.1912663>.

Blauert, Jens. *Spatial Hearing: The Psychophysics of Human Sound Localization*. Rev. ed. Cambridge, Mass: MIT Press, 1997.

Boren, Braxton, und Agnieszka Roginska. „The Effects of Headphones on Listener HRTF Preference.“ Paper presented at the AES 131st Convention, New York, USA Oktober 20-23, 2011.

Bregman, Albert S. Auditory Scene Analysis: The Perceptual Organization of Sound. Cambridge, Mass. [u.a.]: MIT Press, 1994.

Brown, Andrew D, G Christopher Stecker, und Daniel J Tollin. „The Precedence Effect in Sound Localization.“ Journal of the Association for Research in Otolaryngology 16, (2015): 1-28. doi: 10.1007/s10162-014-0496-2.

Brughera, Andrew, Larisa Dunai, und William M. Hartmann. „Human Interaural Time Difference Thresholds for Sine Tones: The High-Frequency Limit.“ The Journal of the Acoustical Society of America 133, Nr. 5 (Mai 2013): 2839–55. <https://doi.org/10.1121/1.4795778>.

Chion, Michel. Audio-Vision: Ton und Bild im Kino. Herausgegeben von Jörg U. Lensing. 1. Auflage. Edition Mediabook. Berlin: Schiele & Schön, 2012.

Clapham, Christopher, und James Nicholson. trajectory. Oxford University Press, 2014. <https://doi.org/10.1093/acref/9780199679591.013.2840>.

Donovan, Jeff M., Brian S. Nelson, und Terry T. Takahashi. „The Contributions of Onset and Offset Echo Delays to Auditory Spatial Perception in Human Listeners.“ The Journal of the Acoustical Society of America 132, Nr. 6 (Dezember 2012): 3912–24. <https://doi.org/10.1121/1.4764877>.

Flückiger, Barbara. Sound design: die virtuelle Klangwelt des Films. 6. Auflage. Zürcher Filmstudien 6. Marburg: Schüren, 2017.

Frank, Matthias, Franz Zotter, und Alois Sontacchi. „Producing 3D Audio in Ambisonics“, Paper presented at the 57th AES International Conference, Hollywood, USA, März 6-8, 2015.

Friesecke, Andreas. Die Audio-Enzyklopädie: ein Nachschlagewerk für Tontechniker ; mit 715 Abbildungen und 145 Tabellen. München: Saur, 2007.

Füg, Simone, Stephan Werner, und Karlheinz Brandenburg. „Controlled Auditory Distance Perception Using Binaural Headphone Reproduction – Algorithms and Evaluation.“ Paper presented at the 27th Tonmeistertagung - VDT International Convention, Köln, Deutschland, November 22-25, 2012.

Getzmann, Stephan, und Jörg Lewald. „Localization of Moving Sound“. Perception & Psychophysics 69, Nr. 6 (August 2007): 1022–34. <https://doi.org/10.3758/BF03193940>.

Gerzon, Michael A., „Periphony: With-Height Sound Reproduction.“ Journal of the Audio Engineering Society 21, 1 (Jänner/ Februar 1973): 2-10.

Groves, John. Commusication: From Pavlov’s Dog to Sound Branding. Cork: Oak Tree Press, 2011.

Hafsati, Mohammed, Nicolas Epain, Rémi Gribonval, und Nancy Bertin. „Sound Source Separation in the Higher Order Ambisonics Domain.“ Proceedings of the 22nd International Conference on Digital Audio Effects (DAFx-19), Birmingham, UK, September 2-6, 2019.

Hahn, Ephraim. „Musical Emotions Evoked by 3D Audio.“ Paper presented at the Conference on Spatial Reproduction, Tokyo, Japan, August 6-9, 2018.

Hammerschmidt, David, und Clemens Wöllner. „Audio-Visual Quality Perception in Musical Performance Videos.“ Musikpsychologie 27, (Dezember 2017): 112-127. <https://doi.org/10.23668/PSYCHARCHIVES.2808>.

Herre, Jürgen, Johannes Hilpert, Achim Kuntz, und Jan Plogsties. „MPEG-H Audio - The New Standard for Universal Spatial/ 3D Audio Coding.“ Paper presented at the 137th AES Convention, Los Angeles, USA, Oktober 9-12, 2014.

Kabel, Peter. Dialog zwischen Mensch und Maschine: Conversational User Interfaces, Intelligente Assistenten und Voice-Systeme. Wiesbaden: Springer Gabler. in Springer Fachmedien Wiesbaden GmbH, 2020.
<http://public.eblib.com/choice/PublicFullRecord.aspx?p=6221634>.

Kendall, Gary S. „Spatial Perception and Cognition in Multichannel Audio for Electroacoustic Music.“ *Organised Sound* 15, Nr. 03 (Dezember 2010): 228–38.
<https://doi.org/10.1017/S1355771810000336>.

Kleinjohann, Michael. Marketingkommunikation mit Acoustic Branding: Planung, Einsatz und Wirkung von Stimme, Ton und Klang für die Corporate Identity. Wiesbaden: Springer Gabler, 2020.

Kronlachner, Matthias. „Spatial Transformations for the Alteration of Ambisonic Recordings“ Masterarbeit, Universität für Musik und Darstellende Kunst Graz, 2014.

Laitinen, Mikko-Ville, Archontis Politis, Ilkka Huhtakallio, und Ville Pulkki. „Controlling the Perceived Distance of an Auditory Object by Manipulation of Loudspeaker Directivity.“ *The Journal of the Acoustical Society of America* 137, Nr. 6 (Juni 2015): EL462–68.
<https://doi.org/10.1121/1.4921678>.

Leclère, Thibaud, Mathieu Lavandier, und Fabien Perrin. „On the Externalization of Sound Sources with Headphones without Reference to a Real Source.“ *The Journal of the Acoustical Society of America* 146, Nr. 4 (Oktober 2019): 2309–20. <https://doi.org/10.1121/1.5128325>.

Linke, Julian, Florian Wendt, Franz Zotter, und Matthias Frank. „How Masking Affects Auditory Objects of Beamformed Sounds.“ Paper presented at the DAGA Conference, Munich, Germany, März 19-22, 2018.

Litovsky, Ruth Y, H Steven Colburn, William A Yost, und Sandra J Guzman. „The Precedence Effect“, *The Journal of the Acoustical Society of America* 106, 4 (Oktober 1999): 1633- 1654.

Middlebrooks, John C, und David M Green. „Sound Localization by Human Listeners.“ *Annu. Rev. Psychol.* 42, (1991): 135-159.

Miller, Carolyn Handler. *Digital Storytelling: A Creator's Guide to Interactive Entertainment*. 4. Aufl. Florida; London; New York: CRC Press, 2019. <https://doi.org/10.1201/9780429440045>.

Nipkow, Lasse, Zielinsky, Gregor, und Ammermann, Tom. „Die Bedeutung von 3D bei Immersive Audio.“ Paper presented at the 30th Tonmeistertagung - VDT International Convention, Köln, Deutschland, November 11-17, 2018.

Pulkki, Ville. „Virtual Sound Source Positioning Using Vector Base Amplitude Panning.“ *Journal of the Audio Engineering Society* 45, No. 6 (Juni 1997): 456-466.

Platz, Friedrich, und Reinhard Kopiez. „When the Eye Listens: A Meta-Analysis of How Audio-Visual Presentation Enhances the Appreciation of Music Performance.“ *Music Perception* 30, Nr. 1 (September 2012): 71–83. <https://doi.org/10.1525/mp.2012.30.1.71>.

Rakerd, Brad, und Hartmann, William M. „Localization of Sound in Rooms, II: The Effects of a Single Reflecting Surface.“ *The Journal of the Acoustical Society of America* 78, Nr. 2 (August 1985): 524–33. <https://doi.org/10.1121/1.392474>.

Rakerd, Brad, und Hartmann, William M. „Localization of Sound in Rooms, III: Onset and Duration Effects.“ *The Journal of the Acoustical Society of America* 80, Nr. 6 (Dezember 1986): 1695–1706. <https://doi.org/10.1121/1.394282>.

Rakerd, Brad, und Hartmann William M. „Localization of Sound in Rooms. V. Binaural Coherence and Human Sensitivity to Interaural Time Differences in Noise.“ *The Journal of the Acoustical Society of America* 128, Nr. 5 (November 2010): 3052–63. <https://doi.org/10.1121/1.3493447>.

Rayleigh, Lord. „XII. On Our Perception of Sound Direction.“ The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science 13, Nr. 74 (Februar 1907): 214–32. <https://doi.org/10.1080/14786440709463595>.

Roginska, Agnieszka, und Paul Geluso, Hrsg. Immersive Sound: The Art and Science of Binaural and Multi-Channel Audio. New York ; London: Routledge, Taylor & Francis Group, 2018.

Rummukainen, Olli, Jing Wang, Zhitong Li, Thomas Robotham, Zhaoyu Yan, Zhuoran Li, Xiang Xie, Frederik Nagel, und Emanuël A P Habets. „Influence of Visual Content on the Perceived Audio Quality in Virtual Reality.“ Paper presented at the 145th AES Convention, New York, USA, Oktober 17-20, 2018.

Rumsey, Francis. „Spatial Audio: Channels, Objects, or Ambisonics?“ J. Audio Eng. Soc. 66, Nr. 11 (2018): 987-992.

Sakamoto, Hiroo, und Yoichi Haneda. „Sound Localization of Beamforming-Controlled Reflected Sound from Ceiling in Presence of Direct Sound.“ Paper presented at the 144th AES Convention, Milan, Italy, Mai 23-26, 2018.

Sharma, Gerriet K. „Komponieren mit skulpturalen Klangphänomenen in der Computermusik.“ PhD diss., Universität für Musik und Darstellende Kunst Graz, 2016.

Shivappa, Shankar, Martin Morrell, Deep Sen, Nils Peters, und S. M. Akramus Salehin. „Efficient, Compelling and Immersive VR Audio Experience Using Scene Based Audio / Higher Order Ambisonics.“ Paper presented at the Conference on Audio for Virtual and Augmented Reality, Los Angeles, USA, September 30 - Oktober 1, 2016.

Sontacchi, Alois. „Dreidimensionale Schallfeldreproduktion für Lautsprecher- und Kopfhöreranwendungen.“ PhD diss., Technische Universität Graz, 2003.

Steiner, Paul, und Erich Joachimsthaler. Sound Branding: Grundlagen akustischer Markenführung. 3., aktualisierte und erweiterte Auflage. Research. Wiesbaden: Springer Gabler, 2018.

Stevenson, Angus. immersive. Oxford University Press, 2010.
https://doi.org/10.1093/acref/9780199571123.013.m_en_gb0401380.

Tabry, Vanessa, Robert J. Zatorre, und Patrice Voss. „The Influence of Vision on Sound Localization Abilities in Both the Horizontal and Vertical Planes.“ *Frontiers in Psychology* 4 (Dezember 2013): 1-7. <https://doi.org/10.3389/fpsyg.2013.00932>.

Udesen, Jesper, Tobias Piechowiak, und Fredrik Gran. „Vision Affects Sound Externalization.“ Paper presented at the 55th AES International Conference, Helsinki, Finland, August 27-29, 2014.

Voegelin, Salomé. *Listening to noise and silence: towards a philosophy of sound art*. New York: Continuum, 2010.

Voss, Patrice, Franco Lepore, Frédéric Gougoux, und Robert J. Zatorre. „Relevance of Spectral Cues for Auditory Spatial Processing in the Occipital Cortex of the Blind.“ *Frontiers in Psychology* 2 (März 2011): 1-12. <https://doi.org/10.3389/fpsyg.2011.00048>.

Wallach, Hans, E. B. Newman, und M. R. Rosenzweig. „A Precedence Effect in Sound Localization.“ *The Journal of the Acoustical Society of America* 21, Nr. 4 (Juli 1949): 468–468. <https://doi.org/10.1121/1.1917119>.

Weinzierl, Stefan, und Verband Deutscher Tonmeister, Hrsg. *Handbuch der Audiotechnik*. Berlin: Springer, 2008.

Wendt, Florian, Gerriet K. Sharma, Matthias Frank, Franz Zotter, und Robert Höldrich. „Perception of Spatial Sound Phenomena Created by the Icosahedral Loudspeaker.“

Computer Music Journal 41, Nr. 1 (März 2017): 76–88.
https://doi.org/10.1162/COMJ_a_00396.

Woodcock, James, William Davies, und Trevor Cox. „Influence of Visual Stimuli on Perceptual Attributes of Spatial Audio.“ *Journal of the Audio Engineering Society* 67, Nr. 7/8 (14. August 2019): 557–67. <https://doi.org/10.17743/jaes.2019.0019>.

Zaunschirm, Markus. „The (Virtual IKOsahedral Loudspeaker Array) VIKO.“ Letzter Zugriff August 25, 2020. <https://iaem.at/projekte/osil/vikocall.zip/view>

Zhang, Chenyan, Andrew Perkis, und Sebastian Arndt. „Spatial immersion versus emotional immersion, which is more immersive?“ Paper presented at the 9th International Conference on Quality of Multimedia Experience (QoMEX 2017), Erfurt, Germany, Mai 31 - Juni 2, 2017. <https://doi.org/10.1109/QoMEX.2017.7965655>.

Zotter, Franz, und Frank, Matthias. „Investigation of Auditory Objects Caused by Directional Sound Sources in Rooms“. *Acta Physica Polonica A* 128, Nr. 1A (Juli 2015): A-5-A-10. <https://doi.org/10.12693/APhysPolA.128.A-5>.

Zotter, Franz. „Analysis and Synthesis of Sound-Radiation with Spherical Arrays.“ PhD diss., Universität für Musik und Darstellende Kunst Graz, 2009.

Zotter, Franz, und Matthias Frank. „All-Round Ambisonic Panning and Decoding.“ *J. Audio Eng. Soc.* 60, Nr. 10 (Oktober 2012): 807-820.

Zotter, Franz, und Frank, Matthias. *Ambisonics: A Practical 3D Audio Theory for Recording, Studio Production, Sound Reinforcement, and Virtual Reality*. Bd. 19. Springer Topics in Signal Processing. Cham: Springer International Publishing, 2019. <https://doi.org/10.1007/978-3-030-17207-7>.

Zotter, Franz, Matthias Frank, und Alois Sontacchi. „The Virtual T-Design Ambisonics-Rig Using VBAP“, Paper presented at the 1st European Congress on Sound and Vibration, Ljubljana, Slowenien, September 15-18, 2010.

Zotter, Franz, Markus Zaunschirm, Matthias Frank, und Matthias Kronlachner. „A Beamformer to Play with Wall Reflections: The Icosahedral Loudspeaker.“ *Computer Music Journal* 41, Nr. 3 (September 2017): 50–68. https://doi.org/10.1162/comj_a_00429.

Online Quellen

Amp. „Best Audio Brands 2020 Ranking.“ Letzter Zugriff August 25, 2020.

<http://ampsoundbranding.com/best-audio-brands-ranking/>

Apple. „Logic Pro X.“ Letzter Zugriff August 25, 2020.

<https://www.apple.com/de/logic-pro/>

DearVR. „Lufthansa 3D Audio Podcast by ZDF Digital.“ Letzter Zugriff August 25, 2020.

<https://blog.dearvr.com/overview/lufthansa-3d-audio-podcast-by-zdf-digital>

Fraunhofer IIS. „MPEG-H Audio.“ Letzter Zugriff August 25, 2020.

<https://www.iis.fraunhofer.de/en/ff/amm/broadcast-streaming/mpegh.html>

International Telecommunication Union (ITU), “Multichannel sound technology in home and broadcasting applications.” Letzter Zugriff August 25, 2020.

https://www.itu.int/dms_pub/itu-r/opb/rep/R-REP-BS.2159-4-2012-PDF-E.pdf

Institut für Elektronische Musik und Akustik. „IEM Plugin Suite.“ Letzter Zugriff August 25, 2020.

<https://plugins.iem.at>

Institut für Elektronische Musik und Akustik. „Plug-in Descriptions.“ Letzter Zugriff August 25, 2020.

<https://plugins.iem.at/docs/pluginDescriptions/>

Institut für Elektronische Musik und Akustik. „The IEM Icosahedral Loudspeaker.“ Letzter Zugriff August 25, 2020.

<https://iem.kug.ac.at/projects/osil/about-the-iko.html>

iZotope. „The Basics of Convolution in Audio Production.“ Letzter Zugriff 25, 2020.

<https://www.izotope.com/en/learn/the-basics-of-convolution-in-audio-production.html>

Kronlachner, Matthias. „Ambisonics plug-in suite.“ Letzter Zugriff August 25, 2020.

<http://www.matthiaskronlachner.com/?p=2015>

Kronlachner, Matthias. „multichannel audio plug-in suite.“ Letzter Zugriff August 25, 2020.

<http://www.matthiaskronlachner.com/?p=1910>

Ringe, Cornelius. „Audio Branding Guide.“ Letzter Zugriff August 25, 2020.

<http://www.radiozentrale.de/kreation/whitepaper/>

Sonible. „d:24.“ Letzter Zugriff August 25, 2020.

<https://www.sonible.com/d24/>

Sonible. „IKO.“ Letzter Zugriff August 25, 2020.

<https://iko.sonible.com/en.html>

Reaper. „Technical.“ Letzter Zugriff August 25, 2020.

<https://www.reaper.fm/about.php#technical>

Wikipedia. „Transient (acoustics).“ Letzter Zugriff August 25, 2020.

[https://en.wikipedia.org/wiki/Transient_\(acoustics\)](https://en.wikipedia.org/wiki/Transient_(acoustics))

Abbildungsverzeichnis

ABBILDUNG 1: KOORDINATENSYSTEM FÜR DIE BESCHREIBUNG DER RICHTUNGSWAHRNEHMUNG DURCH DEN AZIMUT AUF DER HORIZONTALLEBENE UND DIE ELEVATION AUF DER MEDIANEBENE.	11
ABBILDUNG 2: VERANSCHAULICHUNG DER RICHTUNGSWAHRNEHMUNG AUF DER HORIZONTALLEN EBENE DURCH PEGEL- UND LAUFZEITUNTERSCHIEDE.....	12
ABBILDUNG 3: HEAD-SHADOW EFFECT UNTERHALB UND OBERHALB DES GRENZBEREICHES UM 1500HZ.	13
ABBILDUNG 4: A: VERSUCHSANORDNUNG MIT ZWEI LAUTSPRECHERN UND EINEM LEAD- UND LAG-SIGNAL	20
ABBILDUNG 5: GRUNDPRINZIP DES BEAMFORMINGS, DURCH DIE GEZIELTE VERWENDUNG VON WANDREFLEXIONEN MIT HILFE VON GEBÜNDELTEN SCHALLSTRAHLEN	21
ABBILDUNG 6: APPROXIMATIVE NACHBILDUNG EINES RÄUMLICHEN SCHALLFELDES DURCH KUGELFLÄCHENFUNKTIONEN.	25
ABBILDUNG 7: DARSTELLUNG DER KUGELFLÄCHENFUNKTIONEN IN EINEM POLAREN KOORDINATENSYSTEM. DIE REIHEN ZEIGEN DIE ANZAHL DER KANÄLE VON DER 0 BIS ZUR 3. ORDNUNG, SOWIE DIE ANORDNUNG DER KUGELFLÄCHENFUNKTIONEN FÜR DIE MATHEMATISCHE REPRODUKTION DES SCHALLFELDES.	26
ABBILDUNG 8: ERGEBNISSE DER UNTERSUCHUNG VON HAHN IN BEZUG AUF DIE EMOTIONALE ERREGUNG UND GEFÄHREN.	30
ABBILDUNG 9: IKO, EIN 20-FLÄCHIGES IKOSAEDERFÖRMIGES LAUTSPRECHERSYSTEM DER FIRMA SONIBLE UND DES INSTITUTS FÜR ELEKTRONISCHE MUSIK UND AKUSTIK GRAZ.	32
ABBILDUNG 10: SCREENSHOT DES MXFC_CONVOLVER MIT DER FILTERVORLAGE FÜR DEN IKOs	34
ABBILDUNG 11: SCREENSHOT DES ENERGYVISUALIZER.	34
ABBILDUNG 12: SCREENSHOT DES MCFX_ENCODER_03.	35
ABBILDUNG 13: SCREENSHOT DES STEREO ENCODER.	36
ABBILDUNG 14: ROUTING DER SPUREN IN REAPER SOWOHL FÜR DEN IKO, ALS AUCH FÜR DIE VIRTUELLE UMGEBUNG ÜBER EINE BINAURALE WIEDERGABE.	36
ABBILDUNG 15: GRUNDRISS DES HÖRSAALS DES INSTITUTS FÜR ELEKTRONISCHE MUSIK UND AKUSTIK INKL. DER RAUMMAßE UND WANDERMATERIALIEN. DER HÖRSAAL HAT EINEN PARKETTBODEN UND AKUSTISCHE FLIESEN AN DER DECKE.	38
ABBILDUNG 16: BEWEGUNGSMUSTER DES GESUCHTEN KLANGES DES SOUND LOGOS 1. DAS BEWEGUNGSMUSTER A BEFINDET SICH AUF DER HORIZONTALLEN EBENE, DAS BEWEGUNGSMUSTER B AUF DER VERTIKALEN EBENE UND DAS BEWEGUNGSMUSTER C BEFINDET SICH AUF DER SAGITTALLEN EBENE.	40
ABBILDUNG 17: BEWEGUNGSMUSTER DES GESUCHTEN KLANGES DES SOUND LOGOS 2. DIE BEWEGUNGSMUSTER A UND B BEFINDEN SICH AUF DER HORIZONTALLEN EBENE UND BEWEGUNGSMUSTER C BEFINDET SICH AUF DER SAGITTALLEN EBENE.	41
ABBILDUNG 18: BEWEGUNGSMUSTER DES GESUCHTEN KLANGES DES SOUND LOGOS 3. ALLE DREI BEWEGUNGSMUSTER FINDEN AUF DER HORIZONTALLEN EBENE STATT.	42
ABBILDUNG 19: BEWEGUNGSMUSTER DES GESUCHTEN KLANGES DES SOUND LOGOS 4. WÄHREND DAS BEWEGUNGSMUSTER A AUF DER HORIZONTALLEN EBENE UND DIE BEWEGUNGSMUSTER B UND C AUF DER VERTIKALEN EBENE STATTFINDEN, BEWEGT SICH DAS SOUND LOGO D GLEICHZEITIG AUF DER HORIZONTALLEN, ALS AUCH AUF DER VERTIKALEN EBENE.	44
ABBILDUNG 20: LINKS: GRAPHISCHE DARSTELLUNG DER RÄUMLICHEN DIMENSIONEN.	49
ABBILDUNG 21: AUSWAHLMÖGLICHKEITEN FÜR DAS BEWEGUNGSMUSTER DES SOUND LOGOS 4.	50

ABBILDUNG 22: MITTELWERT INKLUSIVE FEHLERINDIKATION DER BEWERTUNG DER SOUND LOGOS IN BEZUG AUF EINBEZOGEN-SEIN, ANZAHL DER EREIGNISSE UND NACHVOLLZIEHBARKEIT.	52
ABBILDUNG 23: LINKS: AUSWAHLMÖGLICHKEIT FÜR DAS BEWEGUNGSMUSTER DES SOUND LOGOS 4B, C.	53

Anhang – Sound Logos

Die Sound Logos können unter folgendem Link heruntergeladen werden.

<https://www.dropbox.com/sh/lztvs43s471g6tg/AABhGEJowxOvOVA7E2g3yLLAa?dl=0>

Fragebogen Hörversuch - 3D Audio Sound Branding

Die nachfolgenden Hörversuche beschäftigen sich mit 3D-Audio im Sound Branding. Es sollten im Rahmen einer Masterarbeit Gestaltungskriterien für 3D-Audio fähige Wiedergabegeräte abgeleitet werden. Die Hörversuche werden auf dem 20 flächigen Ikosaederlautsprecher (IKO) durchgeführt.

Die Hörversuche beinhalten vier Entwürfe von Sound Logos, für die jeweils drei unterschiedliche Bewegungsmuster angefertigt wurden. Bei jedem Sound Logo werden die Aspekte Nachvollziehbarkeit der Bewegung und die Wirkung einer 3D-Audio Wiedergabe beurteilt.

Mit der Teilnahme an diesem Hörversuch stimmen Sie der anonymen Auswertung Ihrer hier angegebenen Daten zu Forschungszwecken zu.

Vielen Dank für Ihre Zeit!

Geschlecht: w m d

Alter: unter 20
 20 - 30
 30 - 40
 über 40

Welcher Gruppe gehören Sie an?

Professionelle(r) Musiker/in
Hobymusiker/in
Ungeübt

Haben Sie bereits Erfahrung mit 3D-Audio?

Ja Nein

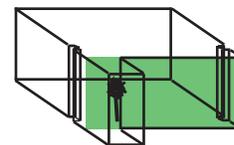
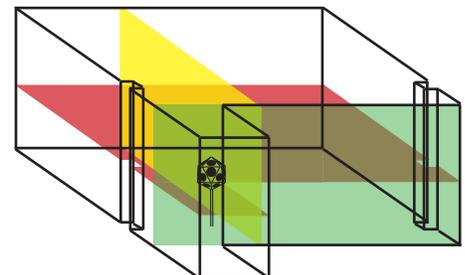
Haben Sie bereits Erfahrung mit dem IKO?

Ja Nein

Haben Sie Vorkenntnisse im Bereich des Sound Brandings?

Ja Nein

Räumliche Dimensionen



Vertikale



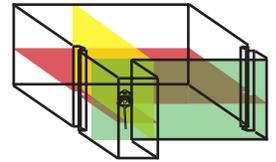
Horizontale



Sagittale

Sound Logo 1

Bitte zeichnen Sie das Bewegungsmuster des Klanges ein, der zuerst einzeln vorgespielt wird. Die Bewegung findet entweder horizontal, vertikal oder sagittal statt. Das Sound Logo kann beliebig oft abgespielt werden.



	A)	B)	C)
horizontal			
vertikal			
sagittal			

Welches Bewegungsmuster gefällt Ihnen am besten?

1. Platz _____ 2. Platz _____ 3. Platz _____

Wie sehr haben Sie sich vom Sound Logo miteinbezogen gefühlt?

sehr gar nicht

Wie empfinden Sie die Anzahl der Ereignisse (einzelne Klänge, Bewegungen)?

gerne mehr zu viel

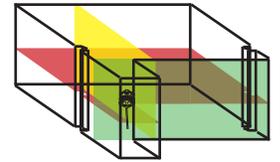
Wie gut waren die Bewegungen nachvollziehbar?

leicht schwer

Anmerkungen: _____

Sound Logo 2

Bitte zeichnen Sie das Bewegungsmuster des Klanges ein, der zuerst einzeln vorgespielt wird. Die Bewegung findet entweder horizontal, vertikal oder sagittal statt. Das Sound Logo kann beliebig oft abgespielt werden.



	A)	B)	C)
horizontal			
vertikal			
sagittal			

Wie oft wiederholt sich das Bewegungsmuster?

A _____

B _____

C _____

Welche Bewegungsmuster gefällt Ihnen am besten?

1. Platz _____

2. Platz _____

3. Platz _____

Wie sehr haben Sie sich vom Sound Logo miteinbezogen gefühlt?

sehr

gar nicht

Wie empfinden Sie die Anzahl der Ereignisse (einzelne Klänge, Bewegungen)?

gerne mehr

zu viel

Wie gut waren die Bewegungen nachvollziehbar?

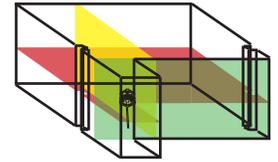
leicht

schwer

Anmerkungen:

Sound Logo 3

Bitte zeichnen Sie das Bewegungsmuster des Klanges ein, der zuerst einzeln vorgespielt wird. Die Bewegung findet entweder horizontal, vertikal oder sagittal statt. Das Bewegungsmuster ist nicht immer zusammenhängend und kann auch gleichzeitig stattfinden. Das Sound Logo kann beliebig oft abgespielt werden.



	A)	B)	C)
horizontal			
vertikal			
sagittal			

Welches Bewegungsmuster gefällt Ihnen am besten?

1. Platz _____ 2. Platz _____ 3. Platz _____

Wie sehr haben Sie sich vom Sound Logo miteinbezogen gefühlt?

sehr gar nicht

Wie empfinden Sie die Anzahl der Ereignisse (einzelne Klänge, Bewegungen)?

gerne mehr zu viel

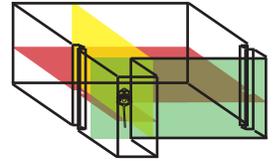
Wie gut waren die Bewegungen nachvollziehbar?

leicht schwer

Anmerkungen:

Sound Logo 4

Bitte wählen Sie das Bewegungsmuster des Klanges aus, der zuerst einzeln vorgespielt wird. Die Dimension der Bewegung ist bereits vorgegeben. Das Sound Logo kann beliebig oft abgespielt werden.



A)

horizontal

B)

vertikal

Für das Hörbeispiel C und D ändert sich der gesuchte Klang.
 Im Hörbeispiel D findet das Bewegungsmuster in beiden Dimensionen gleichzeitig statt.
 (Tipp: Stellen Sie sich zuerst die Bewegungsmuster räumlich vor)

C)

vertikal

D)

horizontal

vertikal

Anmerkungen: _____

Geschafft!!! Vielen Dank!