

MASTERARBEIT

EVALUIERUNG DES PRODUKTIONSTECHNISCHEN POTENZIALS UND KONSUMENTENORIENTIERTEN MEHRWERTS VON BINAURAL-3D-AUDIO IM KONTEXT VON POPMUSIK



HOCHSCHULE DER MEDIEN
FAKULTÄT ELECTRONIC MEDIA
AUDIOVISUELLE MEDIEN (M.ENG)



Hochschule der Medien Stuttgart
Fakultät Electronic Media
Studiengang Audiovisuelle Medien (M.Eng.)

Evaluierung des produktionstechnischen Potenzials und konsumentenorientierten Mehrwerts von Binaural-3D-Audio im Kontext von Popmusik

Masterarbeit

zur Erlangung des akademischen Grades
Master of Engineering

vorgelegt von

Marvin Pfeifer

Matrikelnummer 33758

Wunnensteinstr. 33, 70186 Stuttgart

marvinpfeifer.hello@gmail.com

Erstprüfer: Prof. Oliver Curdt
Zweitprüfer: Daniel Deboy
Bearbeitungszeitraum: 19. Juni 2018 – 18. Dezember 2018

Eidesstattliche Erklärung

Hiermit versichere ich, Marvin Pfeifer, ehrenwörtlich, dass ich die vorliegende Masterarbeit mit dem Titel: „Evaluierung des produktionstechnischen Potenzials und konsumentenorientierten Mehrwerts von Binaural-3D-Audio im Kontext von Popmusik“ selbstständig und ohne fremde Hilfe verfasst und keine anderen als die angegebenen Hilfsmittel benutzt habe. Die Stellen der Arbeit, die dem Wortlaut oder dem Sinn nach anderen Werken entnommen wurden, sind in jedem Fall unter Angabe der Quelle kenntlich gemacht. Die Arbeit ist noch nicht veröffentlicht oder in anderer Form als Prüfungsleistung vorgelegt worden.

Ich habe die Bedeutung der ehrenwörtlichen Versicherung und die prüfungsrechtlichen Folgen (§26 Abs. 2 Bachelor-SPO (6 Semester), § 24 Abs. 2 Bachelor-SPO (7 Semester), § 23 Abs. 2 Master-SPO (3 Semester) bzw. § 19 Abs. 2 Master-SPO (4 Semester und berufsbegleitend) der HdM) einer unrichtigen oder unvollständigen ehrenwörtlichen Versicherung zur Kenntnis genommen.

Stuttgart, den 18. Dezember 2018

Marvin Pfeifer

Danksagung

Ich möchte mich herzlich bei allen Personen bedanken, die zu dieser Arbeit beigetragen haben. Ich danke *Daniel Deboy* für die fachkundige Betreuung sowie Unterstützung mit Ressourcen und Kontaktpersonen. Auch diesen danke ich für ihren hilfreichen Input. Dazu gehören *Matthias Frank*, *Brian F. G. Katz* sowie *Magdalena Piotrowska*. Letzterer sei für ihre professionelle, qualitativ hochwertige Arbeit beim Mastering der Hörproben ein besonderer Dank ausgesprochen. Weiterhin danke ich den Sponsoren *Dear Reality* und *New Audio Technologies* für ihre Bereitstellung der beiden Softwarelösungen *DearVR Pro* und *Spatial Audio Designer*. Namentlich möchte ich hier *Tom Ammermann* und *Achim Fell* sowie *Simon Zimmermann* und *Armin Ghajar Jazi* vom *DearVR*-Entwicklungsteam nennen und mich für ihren konstruktiven Support bedanken. Ebenfalls bedanke ich mich bei der Musikgruppe *Django Django* für die Übermittlung von Einzelspuren sowie bei der Akustik-Abteilung des *Fraunhofer-Institut für Bauphysik* für die Bereitstellung ihres 3D-Audiolabors. Außerdem danke ich *Prof. Oliver Curdt* für zahlreiche lehrreiche Vorlesungen, Workshops und Exkursionen. Schließlich will ich mich bei allen bedanken, die mich bei meinem Studium und darüber hinaus unterstützt haben. Dazu zählen meine Eltern und meine Familie sowie meine Freunde und WG-Mitbewohner.

Glossar

AD/DA	Analog Digital / Digital Analog
ADM	Audio Definition Model
BRIR	Binaural Room Impulse Response (binaurale Raumimpulsantwort)
DAW	Digital Audio Workstation
dB	logarithmische Verhältnisgröße zur Angabe von Pegeln und Pegeldifferenzen
EQ	Equalizer
FFT	Fast Fourier Transformation
FOA	First Order Ambisonics
HOA	Higher Order Ambisonics
HpTF	Headphone Transfer Function
HRIR	Head Related Impulse Response
HRTF	Head Related Transfer Function
IKL	Im-Kopf-Lokalisation
ILD	Interaural Level Differences
ITD	Interaural Time Differences
MPEG-H	Übertragungsstandard der Moving Picture Experts Group
OLE	Overall Listening Experience
Panning	Positionierung von Audioobjekten im räumlichen Panorama
Sweet Spot	Optimale Abhörposition bei Beschallung mit Lautsprechern
Vpn	Versuchsperson(en)
VR	Virtual Reality

Kurzfassung

Die technischen Fortschritte von 3D-Audio entwickeln sich kontinuierlich weiter und manifestieren sich in einer wachsenden Relevanz in Bereichen wie der Filmindustrie, Virtual Reality und 360°-Video. Gleichzeitig gibt es eine steigende Tendenz Musik über Kopfhörer an Stelle von Lautsprechersystemen zu konsumieren. Anlässlich dieser beiden Trends geht die vorliegende Arbeit der Frage nach, welche Chancen sich hieraus für die Musikindustrie ergeben, dem Nutzer ein dreidimensionales, klangästhetisch ansprechendes Hörerlebnis zu bieten. Im Mittelpunkt steht dabei die Frage des potentiellen Mehrwerts, den das Binaural-3D-Format dem Rezipienten gegenüber dem Stereoformat im Kontext von Popmusik ermöglichen kann.

Um diese Frage zu ergründen, werden Veränderungen des Musik-Nutzerverhaltens und der Musikbranche sowie aktuelle technische Entwicklungen auf dem Gebiet von 3D-Audio erörtert. Aus den Einzelspuren hochwertiger Stereo-Mischungen wird im Praxisteil Binaural-3D-Musik produziert. Dabei werden die derzeit besten Softwarelösungen recherchiert, verwendet und die technischen sowie gestalterischen Potentiale neuartiger Produktionstechniken evaluiert. Mit dem produzierten Binaural-3D-Material wird anschließend in einer Nutzerstudie ermittelt, ob die Binauralsynthese für den Musikrezipienten im Vergleich zum konventionellen Stereoformat einen Mehrwert generieren kann.

Schließlich wird aus den gewonnenen Erkenntnissen ein Ausblick auf die Chancen für die sich wandelnde Musikindustrie und auf die Marktrelevanz des Binaural-3D-Formats vorgenommen.

Schlüsselworte: 3D Audio, Binaural, Kopfhörer

Abstract

The development of 3D-sound technology is subject to constant progress and expresses itself in an increasing relevance of the format in areas such as the film industry, virtual reality applications and 360°-video. Simultaneously there is a growing tendency to use headphones instead of loudspeaker setups for music consumption. Following these trends this thesis deals with the resulting potential for the music industry to provide the user with a three-dimensional, aesthetic and immersive listening experience. Furthermore, this investigation puts a special emphasis on the potential benefit of the binaural format for the recipient in contrast with the conventional stereo format in the context of contemporary pop music.

To examine this question further, current developments of consumer listening habits as well as the music industry and technological developments in the discipline of 3D-audio are considered. The individual stems of high-quality stereo music productions are remixed to create appealing binaural-3D-music with the best software tools currently available on the market. In this process, technical as well as artistic potentials are evaluated. Furthermore, the benefits of the binaural-3D-format are investigated in a user study.

Finally, the results of the study are discussed with an outlook on new opportunities and the relevance of the binaural format for the music industry.

Keywords: 3D Audio, Binaural, Headphones

Inhaltsverzeichnis

Eidesstattliche Erklärung.....	ii
Danksagung	iii
Glossar	iv
Kurzfassung.....	v
Abstract.....	vi
Inhaltsverzeichnis	vii
1 Einführung.....	1
1.1 Ausgangssituation	2
1.2 Motivation und Zielsetzung	3
1.3 Vorgehen	4
2 Grundlagen	6
2.1 Historie des Binauralformats.....	6
2.2 Räumliches Hören.....	11
2.3 HRTF Grundlagen.....	12
2.4 HRTF Messung	14
2.5 Binauralsynthese.....	16
2.6 Wiedergabe.....	17
2.7 Faktoren für Lokalisationsschärfe	21
3 Musik-Nutzerverhalten	24
3.1 Entwicklung der Musikbranche und Veränderungen des Nutzerverhaltens 24	
3.2 Veränderungen der Hörgewohnheiten über die Zeit.....	27
4 Aktuelle Entwicklungen.....	30
4.1 Aktuelle Anwendungen von Binauraltechnik.....	30

4.2	Distribution und Wiedergabeformate	33
4.3	Aktuelle Untersuchungen zum Thema Binauraltechnik	37
5	Produktion der Hörproben	40
5.1	Kriterien der Musikauswahl.....	40
5.2	Musikauswahl.....	42
5.3	Signalkette.....	45
5.4	Gestalterische Umsetzung.....	49
5.5	Gestalterische Stilmittel.....	53
5.6	Prozessierung der Binauralsignale.....	55
5.7	Anwendung psychoakustischer Erkenntnisse	57
6	Hörversuch	59
6.1	Ansatz	59
6.2	Versuchsaufbau	60
6.3	Versuchsablauf.....	61
6.4	Qualitative Untersuchung der Transferierbarkeit zweier 3D-Audio-Mischungen	64
7	Auswertung	65
7.1	Analysemethode.....	65
7.2	Versuchspersonen	65
7.3	Ergebnisse	66
7.3.1	Vergleich Probandengruppen Labor / Online.....	67
7.3.2	Einzelbewertung.....	69
7.3.3	Relation zwischen Kopfhörermodell und Präferenz.....	72
7.3.4	Kommentarfelder	72
8	Schlussbetrachtung	74
8.1	Diskussion der Ergebnisse.....	74
8.2	Fazit und Ausblick.....	76
9	Literaturverzeichnis.....	80

Abbildungsverzeichnis	90
10 Anhang	92
A.1 Hörversuchsergebnisse dargestellt als gestapelte Säulendiagramme	92
A.2 Package (deutsch)	98
A.2 Package (englisch)	100
A.3 Baustein für Hörversuch (Max 7-Patch)	103
A.4 Daten-DVD	104

1 Einführung

Ich glaube nicht daran, dass das jetzt ein für alle Mal begraben ist. Ich denke, man wird's wieder irgendwann ausgraben und dann vielleicht unter verbesserten technischen Bedingungen erleben können. Ich bin ganz sicher, dass das nochmal wiederkommt.¹

Diese Worte sprach der als „Kunstkopfguru“ bekannte Hörspielregisseur Ulrich Gerhardt, welcher das erste im Rundfunk veröffentlichte Kunstkopfstereofonie-Hörspiel produzierte, in einem Interview mit dem *Bayerischen Rundfunk* (ausgestrahlt 2013). Während die Technik damals noch einigen Limitationen unterworfen war und diejenigen Hörer ausschloss, die keine Kopfhörer zur Verfügung hatten, lässt sich in den letzten Jahren wieder eine verstärkte Relevanz des kopfhörerbezogenen 3D-Audioformats beobachten. Zum einen ist dies bedingt durch eine Vielzahl an technischen Entwicklungen wie *Virtual Reality*-Applikationen oder 360°-Film, die in den letzten Jahren ihre Marktreife erlangen konnten und einer dreidimensionalen Kopfhörerwiedergabe bedürfen, um ihr Immersionspotential vollends zu entfalten. Auch der steigende Einfluss von Lautsprecher-basierten 3D-Wiedergabesystemen wie *Dolby Atmos*, *Auro 3D* oder *Auro Max* dürfte die Möglichkeiten der Binauraltechnik bezüglich der Entwicklung nutzerorientierter Medieninhalte wieder stärker in den Fokus rücken lassen. Begünstigt wird dies auch durch die Veränderung der Abhörpräferenzen der Nutzer, mitunter bedingt durch technische Entwicklungen wie die Möglichkeit des mobilen Streamings von Audioinhalten. So setzen Stand 2013 93 % der 14- bis 29-jährigen einen Kopfhörer für die Musikwiedergabe ein.² Vor dem Hintergrund

¹ vgl. Arnold 2013

² vgl. Veltkamp 2013

dieser Entwicklungen drängt sich die Frage der Relevanz von 3D-Audio für den Bereich des Kopfhörerentertainments sowie der potentielle Mehrwert des Binaural-3D-Formats in den Fokus.

1.1 Ausgangssituation

Aktuell wird die Binauraltechnik bereits in einigen Bereichen eingesetzt. Dazu gehören an prominentester Stelle die Unterhaltungsmedien Virtual Reality, 360°-Video, Augmented Reality und wie eingangs erwähnt das Hörspiel. Vor allem der Bereich Virtual Reality ist dabei aktuell ein großer Wachstumsmarkt, dessen Umsatz sich laut Prognosen bis 2020 jährlich verdoppeln wird.³ Auch in der Musikproduktion lassen sich einige Beispiele von Kunstkopfaufnahmen und binauraler Synthese finden. In den letzten Jahren wurden zudem einige Kopfhörerkonzerte durchgeführt, bei denen die Musik im 3D-Binauralformat live gemischt wurde. Darüber hinaus existieren Systeme, welche die binaurale Wiedergabe von InEar-Monitoring für Livekonzerte ermöglichen. Für den Produktionsprozess von Toninhalten bieten verschiedene Hersteller Raumsimulationssoftware an, welche die Auralisierung von Abhörräumen und spezifischen Lautsprecheranordnungen mittels Binauralsynthese erlauben.

In Kapitel 4.1 werden einige aktuelle binaurale Anwendungen im Segment der Unterhaltungsindustrie präsentiert. Außerhalb dieser findet die Binauraltechnik auch in der Forschung, Messtechnik, Planung von Raumakustik und Beschallungsanlagen sowie der akustischen Rekonstruktion historischer Klangereignisse Einsatz.⁴

Auf Produktionsseite bieten mittlerweile eine Vielzahl an Herstellern hochqualitative Softwarelösungen zur Binauralisierung von Schallquellen an. Einige dieser Werkzeuge integrieren darüber hinaus realistisch anmutende 3D-

³ vgl. Statista 2017

⁴ vgl. Dickreiter, et al. 2014

Hallräume. In Kapitel 5.3 wird eine Auswahl der aktuellen Binauralrenderer präsentiert.

Mit Übertragungsstandards wie dem *Audio Definition Model* oder dem Kodierungsformat *MPEG-H* wurden in den letzten Jahren weiterhin Strukturen entwickelt, welche die Bereitstellung von immersiven, dreidimensionalen Audioinhalten unabhängig vom Wiedergabeformat ermöglichen. Parallel dazu wurde mit dem Kopfhörermodell *Audeze Mobius Blue* im Jahr 2018 der weltweit erste Kopfhörer mit integrierter Binaural-Prozessierung veröffentlicht. Auch wenn dieses in seiner ersten Version lediglich *First Order Ambisonics* integriert, sind für die Zukunft laut Hersteller die Implementation von *MPEG-H* und *Higher Order Ambisonics* geplant.⁵

1.2 Motivation und Zielsetzung

Die Produktion von Musik für 3D-Lautsprechersysteme ermöglicht neue Möglichkeiten, dem Rezipienten ein räumlicheres, immersiveres sowie emotional intensiveres Hörerlebnis zu bieten, wie in einer Studie belegt werden konnte.⁶ Besonders im Direktvergleich zum herkömmlichen Stereoformat machen sich diese Unterschiede bemerkbar. Die Möglichkeiten ein derartiges Klangerlebnis im Kontext von Musik bezeugen zu können sind jedoch recht überschaubar. So werden 3D-Lautsprechersysteme bislang vor allem im Kinobereich eingesetzt, wobei die Verbreitung selbst hier recht schmal ausfällt. Auch im Heimbereich dürfte die aufwändige Installation der Vielzahl an für ein umhüllendes Klangbild bislang nötigen Lautsprechern eher abschreckend wirken. An dieser Stelle ist das Prinzip der Binauralsynthese dank seiner Niederschwelligkeit sehr attraktiv. So können binaurale Produktionen mit konventionellen Wiedergabegeräten und Kopfhörern abgespielt werden und gewinnen durch den aktuell steigenden Trend der mobilen Kopfhörer-Musikwiedergabe an Relevanz (s. Kapitel 3.2). Die

⁵ vgl. Audeze 2018

⁶ vgl. Hahn 2018

vorliegende Arbeit ist somit in erster Linie durch das intensivere Klangerlebnis motiviert, welches 3D-Musik bei Lautsprecherwiedergabe heute bereits erzeugen kann. Dabei steht die Frage im Fokus, ob und inwieweit mit den aktuellen technischen Möglichkeiten auf Produktions- sowie Wiedergabeseite dem Musikrezipienten auf niederschwellige Art und Weise ein immersiveres, emotional ansprechenderes Hörerlebnis geboten werden kann, als dies mit dem konventionellen Stereoformat bislang möglich ist.

1.3 Vorgehen

Zur Klärung dieser zentralen Fragestellung bezieht sich die weitere Vorgehensweise auf verschiedene Studien, laut welchen der Musikhörer einen größeren Wert auf die Verstärkung oder Veränderung der emotionalen Stimmung legt als auf die Klangqualität.^{7 8 9} Während bereits einige Untersuchungen zu technischen Aspekten des Binauralformats durchgeführt wurden, fokussiert sich diese Arbeit auf die subjektive Empfindung des Rezipienten.

Dafür werden zunächst die Grundlagen der Materie betrachtet und zusammenfassend erläutert. Anschließend wird ein Überblick über aktuelle Tendenzen der Musikindustrie sowie Veränderungen der Hörgewohnheiten über die Zeit gegeben. In einem weiteren Schritt werden der derzeitige Stand der technischen Rahmenbedingungen für 3D-Audio, konkrete Anwendungen des Binauralformats sowie Ergebnisse aktueller Untersuchungen zur Materie präsentiert. Der zweite Teil der Arbeit befasst sich mit der praktischen Umsetzung der Produktion von Binauralmusik. Dabei werden die einzelnen Schritte der Sondierung aktueller Softwarelösungen, Auswahl und Akquirierung des Musik-Quellmaterials sowie der Produktion der Versuchssystems erläutert. Der letzte Teil dieser Arbeit widmet sich der Beschreibung und Durchführung des Hörversuchs

⁷ vgl. Krumhansl 2002

⁸ vgl. Roe 1985

⁹ vgl. Kamalzadeh, Baur und Möller 2012

sowie dessen Auswertung. Zuletzt werden die Ergebnisse präsentiert, diskutiert und in den Kontext eines Ausblicks auf zukünftige Studien und Entwicklungen gestellt.

2 Grundlagen

Der Mensch ist in der Lage, Schallquellen aus jeglichen Richtungen zu verorten. Grundlegend dafür sind verschiedene physische Faktoren, welche sich das Übertragungsverfahren der Kunstkopftechnik zu Nutze macht, um alle Richtungen aus dem oberen Hallfeld sowie Entfernungen von Schallereignissen korrekt abzubilden und somit einen realistischen Raumeindruck zu erzeugen. Dieser äußert sich beim Rezipienten bei Wiedergabe in einer positiv wahrgenommenen Umhüllung. Die Kunstkopftechnik oder kopfbezogene binaurale (lat. für „beidohrig“) Übertragung lässt sich bereits in ihren Anfängen im 19. Jahrhundert finden¹⁰ und ist in ihrem Funktionsprinzip grundlegend für die Binauraltechnik, die für einige aktuelle Anwendungen in der Audiobranche wieder eine relevante Rolle spielt. Sowohl die geschichtlichen Ursprünge als auch die physischen und technischen Grundlagen der Binauraltechnik sollen in diesem Kapitel näher betrachtet werden.

2.1 Historie des Binauralformats

Im 19. Jahrhundert galt es noch als umstritten, ob der Mensch überhaupt in der Lage sei räumlich zu hören. So konnte der englische Physiker Lord Rayleigh (1842 – 1919) in einem Experiment anhand einer um den Kopf der Testperson kreisenden Stimmgabel zeigen, dass zwar der Winkel korrekt lokalisiert, die Verortung nach vorne oder hinten jedoch leicht miteinander verwechselt wurde.¹¹ 1881 wurde in Paris das von Clement Ader entwickelte Théatrophone vorgestellt (s. Abbildung 1). Mit diesem konnten per Mikrofonpaar aufgenommene Theater- und Operaufführungen über das Telefonnetz übertragen und über zwei

¹⁰ vgl. Görne 2014

¹¹ vgl. Görne 2014

Ohrmuscheln wiedergegeben werden. Somit könnte man das Théatrophone als das erste binaurale (im Sinne von „beidohrige“) System bezeichnen.¹²

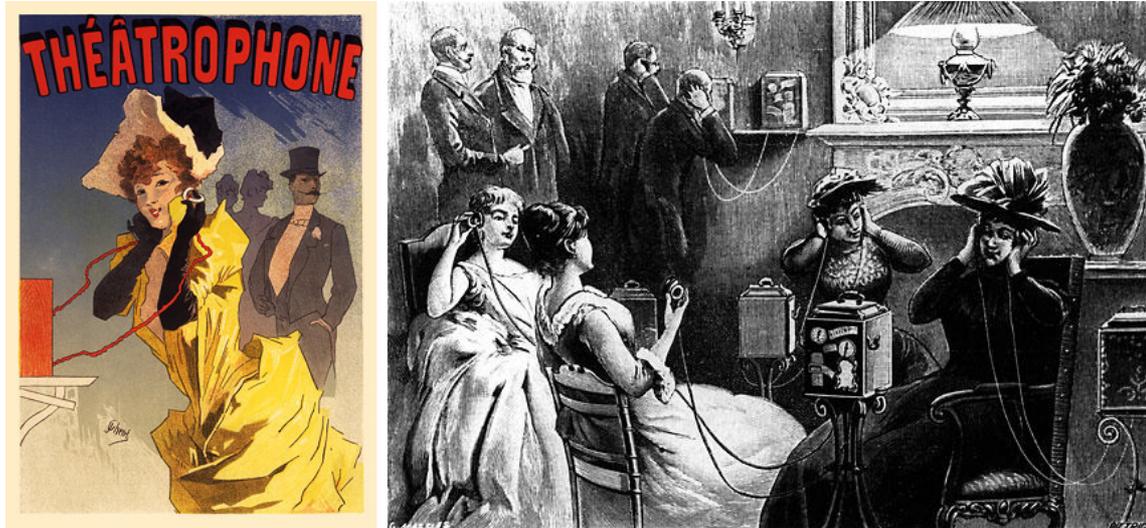


Abbildung 1: links: Werbung für das Théatrophone, rechts: Münzautomaten¹³

Die Geburtsstunde der eigentlichen Kunstkopftechnik fiel schließlich in die 1920er Jahre. Zwischen 1924 und 1927 wurden vier verschiedene Systeme aus den USA und Frankreich patentiert, jedoch nie umgesetzt. Das erste tatsächlich realisierte Exemplar wurde 1930 von F.A. Firestone an der Universität Michigan konstruiert, welches er aus einer Schaufenster-Wachsbüste baute und für Experimente, Messungen und Recherchearbeiten verwendete.¹⁴ Anfang der 1970er wurde die Kunstkopftechnik dann zum ersten Mal im Rundfunkbereich eingesetzt. Das erste binaurale Hörspiel „Demolition“ von Ulrich Gerhardt feierte auf der Funkausstellung 1973 seine Premiere und erntete begeisterte Kritiken. So wurde es als die „Sensation der Berliner Funkausstellung 1973“ betitelt und Gerhardts „Spitzenleistung“ der „unvorstellbar realistischen“ Produktion hervorgehoben. Der Regisseur hatte sein Ziel, sämtliche Wirkungsmöglichkeiten der Kunstkopf-Stereofonie optimal zur Geltung zu bringen, eindrucksvoll

¹² vgl. Brech und Paland 2015

¹³ Daniels 2002

¹⁴ vgl. Brech und Paland 2015

erreicht.¹⁵ Dies motivierte ca. 60 weitere derartige Produktionen bis zum Ende der 1970er Jahre.¹⁶ Bei diesen wurde vor allem der 1971 von der Firma *Neumann* veröffentlichte Kunstkopf *KU80* (s. Abbildung 2) verwendet, welcher für seinen guten Raumeindruck und seine realistische Entfernungsabbildung gelobt wurde. Als problematisch erwiesen sich die schlechte vorne/hinten-Lokalisation, Klangverfärbungen bei der Wiedergabe über Lautsprecher, eine mangelhafte Hochtonwiedergabe sowie ein recht hohes Grundrauschen.¹⁷



Abbildung 2: Neumann KU 80¹⁸

Neben diesen klanglichen Defiziten bewirkte jedoch auch der im Vergleich zur herkömmlichen Hörspielproduktion sehr hohe Kostenmehraufwand, dass nach der Hochphase der 1970er kaum weitere Kunstkopfproduktionen realisiert wurden. Während für eine Hörspielproduktion im Monoformat fünf bzw. im Stereoformat zehn Tage aufgebracht werden mussten, benötigte eine Kunstkopfproduktion 20 – 30 Tage und blieb somit ein Nischenprodukt. Erst in den 1990ern wurden für die öffentlichen Radiosender wieder einige neue Kunstkopf-Hörspiele produziert.¹⁹

¹⁵ vgl. Bayerischer Rundfunk 2018

¹⁶ vgl. Arnold 2013

¹⁷ vgl. Leonhard und de Gruyter 2001

¹⁸ Georg Neumann GmbH 2018

¹⁹ vgl. Arnold 2013

Mit dem Nachfolger *KU80* wurde versucht die oben erwähnten Probleme zu korrigieren. Der *KU81* war in seiner Form detaillierter an durchschnittlichen Kopf- und Ohrmaßen angepasst. Durch seine Diffusfeldentzerrung konnte eine höhere Kompatibilität mit der Lautsprecherwiedergabe erreicht werden.²⁰ 2002 brachte *Neumann* schließlich mit dem Modell *KU100* die dritte Generation auf den Markt, welche im Vergleich zum Vorgänger eher in funktionaler Hinsicht weiterentwickelt wurde. Die Gesichtsform wurde etwas abstrahiert und als Mikrofonpaar zwei Druckempfänger des Typs *Neumann KM83* eingesetzt. Während die *Neumann* Kunstköpfe eher für Sprach- und Tonaufnahmen verwendet wurden, existieren noch eine Vielzahl weiterer Modelle, welche in der Forschung und Entwicklung Verwendung finden (s. Abbildung 3).²¹

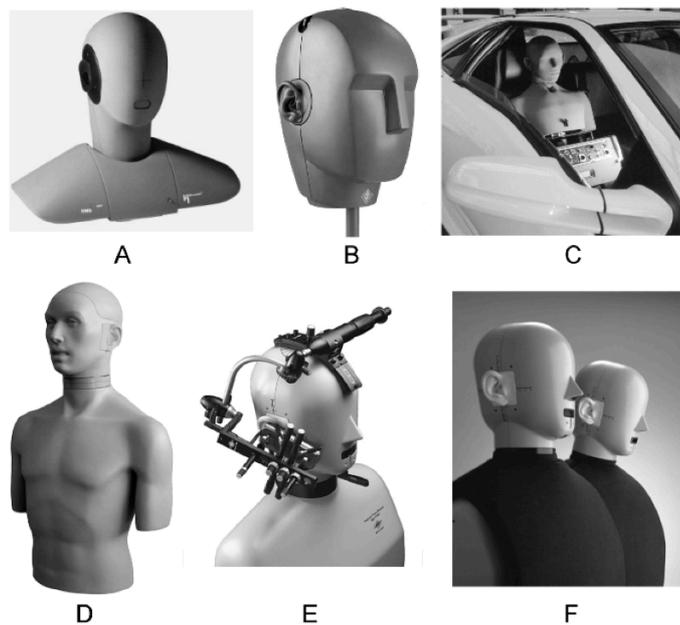


Abbildung 3: verschiedene Kunstkopfsysteme A: Head Acoustics HMS III, B: Neumann KU 100, C: Cortex MK1, D: KEMAR KB 4004, E: Brüel & Kjaer HATS 4128, F: Brüel & Kjaer HATS 4100²²

In der Musikproduktion spielte die Kunstkopftechnik bislang eine tendenziell untergeordnete Rolle. Zwar wurde auch hier bereits früh experimentiert. Als

²⁰ vgl. Dickreiter, et al. 2014

²¹ vgl. Weinzierl 2009

²² Weinzierl 2009

Beispiel ist das 1974 erschienene Album „Bourbon Skiffle Company - Kunstkopf-Stereofonie“ der WAM-Schallplatten Prod. GmbH zu nennen (s. Abbildung 4).



Abbildung 4: Albumcover: Bourbon Skiffle Company - Kunstkopf-Stereofonie (Photo)

Auch die Musikgruppe *Pink Floyd* experimentierte bei der Produktion ihres Albums „The Final Cut“ mit binauraler Aufnahmetechnik.²³ Jedoch konnte sich das Aufnahmeverfahren auch hier nicht als wahre Alternative zur Stereoproduktion positionieren. Dies ist unter anderem durch die deutliche Klangfärbung bei Lautsprecherwiedergabe sowie die eingeschränkten Gestaltungsmöglichkeiten in der Postproduktion, wie beispielsweise die Verwendung weiterer Mikrofone, bedingt. Infolge der sogenannten „Cone of Confusion“ (s. Kapitel 2.2) kommt es zudem nach wie vor zu Verwechslungen der vorne/hinten-Lokalisation.²⁴ Schallquellen werden auf Grund der fehlenden visuellen Hinweise dabei nur selten vorne verortet, da außerhalb des Sichtfeld befindliche Signale erwartungsgemäß hinten positioniert sind.²⁵

Trotz der anfänglich großen Begeisterung über die neuen Möglichkeiten der Kunstkopftechnik, kam das neue Übertragungsverfahren nie über den Status einer interessanten Effektspielerei hinaus. Erst mit dem Aufkommen von Virtual Reality-Anwendungen und 360°-Video besteht seit einigen Jahren ein wachsender Bedarf an dreidimensionalem Ton zur Intensivierung der Immersion. Grundlegend

²³ vgl. Mabbett 2010

²⁴ vgl. Weinzierl 2009

²⁵ vgl. Arnold 2013

für dementsprechend synthetisch generierte Binauralsignale ist die sogenannte Binauraltechnik, welche in den folgenden Abschnitten näher beschrieben wird.

2.2 Räumliches Hören

Für die Lokalisation von Schallereignissen sind einige verschiedene Faktoren ausschlaggebend. Neben offensichtlichen visuellen Indizien sind interaurale Laufzeitdifferenzen (engl. „Interaural Time Differences“ oder abgekürzt ITD) und interaurale Intensitätsdifferenzen (engl. „Interaural Intensity Differences“ oder abgekürzt IID) maßgeblich (s. Abbildung 5).

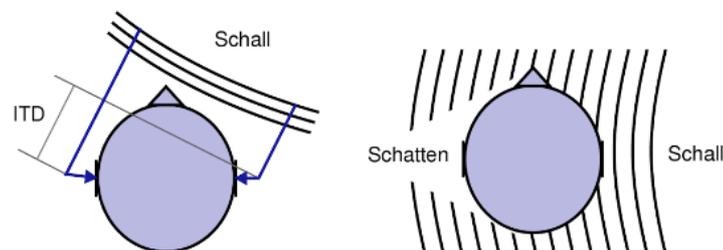


Abbildung 5: links: ITD, rechts: IID²⁶

Bei ersteren handelt es sich um Unterschiede der beiden Ohrsignale in Laufzeit und Hüllkurve, letztere bedingt durch Beugung der Frequenzen am Kopf. Bei einem durchschnittlichen Kopfdurchmesser von 17 cm funktioniert diese Art der Ortung jedoch nur für Frequenzen unterhalb von 1,5 kHz. Für diese gilt, entsprechend der kleinsten wahrnehmbaren Laufzeitdifferenz von 0,03 ms, eine Lokalisationsschärfe von ca. 3 - 5°. Für alle Frequenzen oberhalb 1,5 kHz gelten interaurale Intensitätsdifferenzen: Die Schallwellen werden vom Kopf abgeschattet und resultieren bedingt durch den Abstand der Ohren und deren entgegengesetzte Richtwirkung in Pegeldifferenzen des linken und rechten Ohrsignals.²⁷ Sowohl IDT als auch IID sind jedoch wirkungslos für Schallereignisse mit gleichem Abstand zu beiden Ohren. Schallsignale, die diese Bedingung

²⁶ Butz und Krüger, Mensch-Maschine-Interaktion 2014

²⁷ vgl. Dickreiter, et al. 2014

erfüllen, liegen auf der sogenannten „Cone of Confusion“ (dt. Verwirrungskegel, s. Abbildung 6).

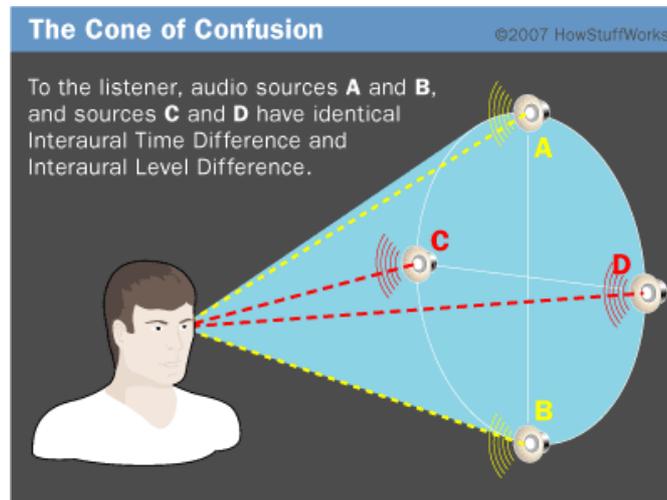


Abbildung 6: „Cone of Confusion“, A und B sowie C und D haben jeweils identische ILD und IID²⁸

Da diese identische IID und IDT besitzen, ist das Gehör bei der vorne/hinten-Lokalisierung nicht in der Lage präzise zu differenzieren. Weiterhin lassen sich tiefe Frequenzen auf Grund ihrer großen Wellenlängen und Beugungseigenschaften durch die zwei erläuterten Mechanismen nicht orten.

Dass die präzise Lokalisierung dennoch im Alltag zumeist gelingt, liegt zum Teil an kleinen Mikrobewegungen, die der Kopf unterbewusst stets durchführt und somit die Schallereignisse entlang der Cone of Confusion voneinander trennen kann. Diese Lokalisierungsmethode wird Peilung genannt. Damit diese erfolgreich funktioniert, bedarf es einer minimalen Schallereignislänge von 200 ms.²⁹

2.3 HRTF Grundlagen

Ein weiteres maßgebliches Prinzip zur Lokalisierung ist durch die Form von Ohr, Kopf und Torso gegeben. Dieses Prinzip wird kopfbezogene oder Außenohr-Übertragungsfunktion, zumeist jedoch kurz HRTF („Head Related Transfer Function“) bezeichnet. Es beschreibt die Übertragungsfunktion von einem

²⁸ How Stuff Works Tech 2016

²⁹ vgl. Weinzierl 2009

beliebigen Punkt im ungestörten Schallfeld bis zum Gehöreingang und beinhaltet die Veränderung von Ankunftszeit, Amplitude, Frequenzgang sowie Phase eines Schallereignisses durch Beugung, Reflexion oder Abschattung am Körper. Maßgeblich sind dabei die Körperteile Torso, Schulter, Kopf, Ohrmuscheln, Ohrkanaleingang und der Ohrkanal selbst, wobei der größte Einfluss dabei durch Kopf, Ohrmuscheln und Schultern (± 5 dB) sowie Torso (± 3 dB) gegeben ist.^{30 31} Dabei ist zu beachten, dass sich die Ohrform von Mensch zu Mensch unterscheidet (s. Abbildung 7).

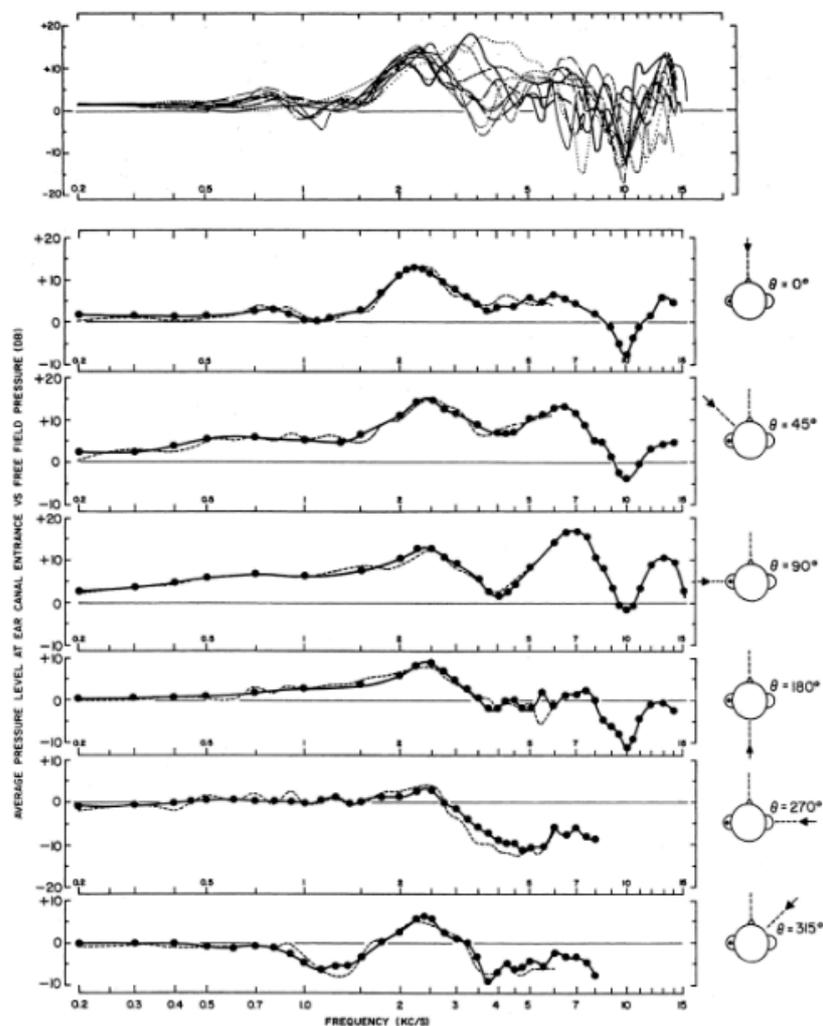


Abbildung 7: winkelabhängige HRTF. Oben: Streuung für frontalen Schalleinfall. Unten: winkelabhängige gemittelte HRTF-Kurven³²

³⁰ vgl. Lerch, Sessler und Wolf 2009

³¹ vgl. Weinzierl 2009

³² Shaw 1966

So konnte die Hamburger Polizei im Jahr 2012 einen 33-jährigen Mann auf Grund eben jener überführen. Er hatte vor der Tat an den Türen seiner Opfer gelauscht und seinen individuellen Ohrabdruck hinterlassen.³³ Diese Verhältnismäßigkeit stellt für die Kunstkopf- und Binauraltechnik das Problem dar, dass im Normalfall ein nicht-individueller, gemittelter HRTF-Satz Verwendung findet und somit nicht allen Nutzern die gleiche, realistische Raumabbildung geboten werden kann. Diese hängt von der Deckungsgleichheit der individuellen mit der gemittelten HRTF ab.³⁴

2.4 HRTF Messung

Im Idealfall sollten bei der Wiedergabe von Binauralsignalen die individuellen HRTF des Rezipienten verwendet werden. Problematisch ist dabei jedoch der sehr aufwändige Prozess der Messung, welcher entweder per direkter oder reziproker Methode durchgeführt wird. Bei ersterer sind die Messmikrofone im Hörgang der zu vermessenden Person positioniert, während bei der reziproken Methode kleine Lautsprecher im Ohr und viele Messmikrofone in kugelförmiger Anordnung um das Subjekt herum installiert sind (s. Abbildung 8). In beiden Fällen wird als Testsignal wie auch bei der Vermessung von Raumimpulsantworten üblich ein exponentieller „Sweep“ abgespielt. Der Vorteil der reziproken Methode ist hierbei, dass das Impulssignal nur einmal abgespielt werden muss. Nachteilig ist hierbei jedoch die verminderte Qualität der Messungen, da zum einen die kleinen Lautsprecher keine tiefen Frequenzen wiedergeben können und nicht den Pegel erreichen, der für einen ausreichenden Signal-Rausch-Abstand nötig wäre. Im Normalfall findet daher die direkte Messmethode Anwendung. Die Person sitzt hierbei auf einem drehbaren Stuhl und ist von einer kreisförmigen Lautsprecher-Anordnung umgeben (siehe Abbildung 8).

³³ vgl. Günther 2012

³⁴ vgl. Leonhard und de Gruyter 2001



Abbildung 8: links: direkte Messung, rechts: reziproke Messung³⁵

Die Messmikrofone sind dabei am Eingang des geblockten Ohrkanals positioniert, da die interindividuellen Differenzen hier am geringsten ins Gewicht fallen. Der Einfluss des hinter dem Messmikrofon liegenden Hörkanals ist richtungsunabhängig und somit für die HRTF unbedeutend. Er wird durch eine richtungsunabhängige Entzerrung ausgeglichen. Per „Multiple Exponential Sweep“-Methode (MESM), bei der das Impulssignal parallel überlappend abgespielt wird, kann für möglichst jeden Winkel des Lautsprecher-Arrays zum Messsubjekts dessen HRTF gemessen werden. Der Winkel sollte hierbei möglichst klein gewählt sein, um eine hohe Auflösung der Übertragungsfunktion zu gewährleisten. Weiterhin ist für ein qualitatives Ergebnis wichtig, dass der Kopf stets zentral im Mittelpunkt der Anordnung positioniert ist. Weiterhin sollte die Messung in einem reflexionsarmen Raum durchgeführt werden, sodass nur der akustische Einfluss von Torso, Kopf und Ohr erfasst werden. Die Dauer der Einmessung kann auf diese Weise bis zu 45 Minuten dauern.³⁶

³⁵ Zaar, Vermessung von Außenohrübertragungsfunktionen mit reziproker Messmethode 2010

³⁶ vgl. Weinzierl 2009

Es wird also recht deutlich, dass der eingangs erwähnte Idealfall der Implementation der jeweils individuellen HRTF bei der Wiedergabe nach aktuellem Stand für den Großteil der Nutzer noch nicht praktisch realisierbar ist.

Wird die Messung statt in einem reflexionsarmen in einem akustisch wirksamen Raum durchgeführt, erhält man eine binaurale Raumimpulsantwort. In dieser sind neben der HRTF auch die akustischen Eigenschaften des Raums codiert. Anhand dieser sogenannten BRIR („Binaural Room Impulse Response“) bzw. BRTF („Binaural Room Transfer Function“) lässt sich der entsprechende Raum virtuell synthetisieren.³⁷

Ein recht simpler Weg, um dennoch individuelle binaurale Signale generieren zu können, besteht darin einen Kopfhörer mit in den Ohrstöpseln installierten Mikrofonkapseln zu tragen und damit das zu binauralisierende Schallereignis aufzuzeichnen.³⁸

2.5 Binauralsynthese

Die Binauralsynthese oder Binauraltechnik beschreibt im engeren Sinne nicht die direkte Aufzeichnung von Ohrsignalen, beispielsweise mit einem Kunstkopf, sondern die Synthetisierung dieser durch digitale Signalverarbeitung. Bei dieser sogenannten Auralisierung wird das Audiosignal mit der entsprechenden, richtungsbestimmenden HRTF gefaltet.³⁹ Diese Faltung ist ein simpler Rechenalgorithmus, welcher das Signal mit einem Impulssignal multipliziert. Da diese Multiplikation extrem rechenaufwändig ist, wird zuerst eine „Fast Fourier Transformation“ (FFT) durchgeführt, welche das Signal aus dem Zeit- in den Frequenzbereich wandelt. Anschließend wird das Nutzsignal mit dem Übertragungsfrequenzgang, welcher aus der gemessenen Impulsantwort

³⁷ vgl. Dickreiter, et al. 2014

³⁸ vgl. Roginska und Geluso 2017

³⁹ vgl. Weinzierl 2009

resultiert, gefaltet und schließlich wieder per inverser Fast Fourier Transformation zurück in den Zeitbereich gebracht (siehe Abb. 9).

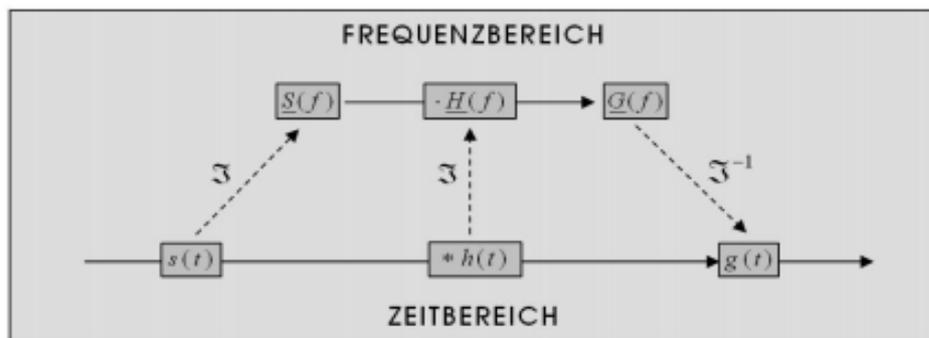


Abbildung 9: Schnelle Faltung⁴⁰

Dieses Verfahren der schnellen Faltung ist besonders geeignet für Echtzeitverarbeitung. Auch mehrkanalige Formate wie 5.1 Surround lassen sich durch diese Methode binauralisieren. Für eine derartige Surround-Simulation über Kopfhörer würden jeder Kanal mit der zweikanaligen kopfbezogenen Impulsantwort (engl. Head Related Impulse Response, HRIR), der jeweiligen Richtung gefaltet und die resultierenden Signale schließlich zu einem Stereo-Signal aufsummiert werden.⁴¹

Wie im vorherigen Abschnitt erwähnt, lassen sich auf diese Weise auch Räume virtuell darstellen, indem das Signal nicht mit der HRIR, sondern mit der BRIR gefaltet wird. Dieses Verfahren ist jedoch deutlich rechenintensiver, da abhängig von der Nachhallzeit des darzustellenden Raums die zu multiplizierenden Impulsantworten um ein Vielfaches länger ausfallen und somit mehr Samples berechnet werden müssen.⁴²

2.6 Wiedergabe

Die Wiedergabe von Binauralsignalen sollten über Kopfhörer erfolgen, da bei Lautsprecherwiedergabe die Raumakustik, das Übersprechen beider Kanäle und

⁴⁰ Keinath und Tarnow 2009

⁴¹ vgl. Görne 2014

⁴² vgl. Weinzierl 2009

die erneute Filterung durch die Außenohrübertragungsfunktion die korrekte Abbildung der Räumlichkeit zunichte machen würden. Der Kopfhörer sollte dabei nach der ITU-Empfehlung [ITU-R BS.708] diffusfeldentzerrt sein. Nach dieser sollte die Kopfhörer-Übertragungsfunktion mit der menschlichen Außenohrübertragungsfunktion im Diffusfeld übereinstimmen, um Klangfärbungen zu minimieren.⁴³ Weiterhin kommt es durch die Interaktion von Kopfhörer und Ohr zu Resonanzen sowie Verzerrungen bedingt durch alle weiteren Elemente der Signalkette wie D/A-Wandler, Verstärker etc. Die Entzerrung dieser sogenannten Kopfhörer-Transferfunktion (abgekürzt HpTF) kann die Lokalisierung verbessern.⁴⁴ Für die Lautsprecherwiedergabe existieren zwar Ansätze der klangtreuen Übertragung binauraler Signale. Zu diesen gehört beispielsweise die Implementation eines sogenannten „Crosstalk Cancellation Filters“ auf das Ausgangssignal, welcher dem Lautsprecher vorgeschaltet ist und die Übersprechanteile der beiden Kanäle minimiert (siehe Abbildung 10).

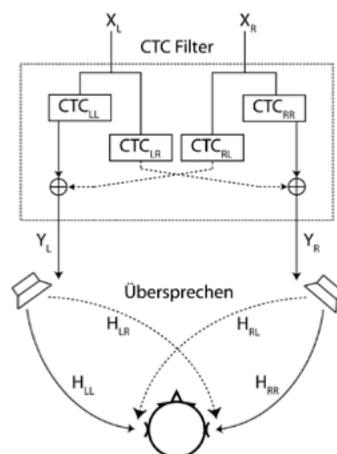


Abbildung 10: Crosstalk Cancellation⁴⁵

Da dieses Transaural-Filter jedoch nicht präzise genug arbeitet und zudem auf die jeweilige Kopfausrichtung des Hörers angepasst werden müsste, bleibt dieses System eher eine theoretische Lösungsmethode. Auf dieser aufbauend wurde am

⁴³ vgl. Dickreiter, et al. 2014

⁴⁴ vgl. Roginska und Geluso 2017

⁴⁵ Weinzierl 2009

Institut für Rundfunktechnik (IRT) der sogenannte „Binaural Sky“ entwickelt, welcher aus einem über der Hörerin installierten Lautsprecheresetup besteht und per Wellenfeldsynthese und Verfolgung der Kopfposition (engl. Headtracking) in der Lage ist, Schallquellen wie zum Beispiel Binauralsignale direkt neben den Ohren zu simulieren (s. Abbildung 11).⁴⁶

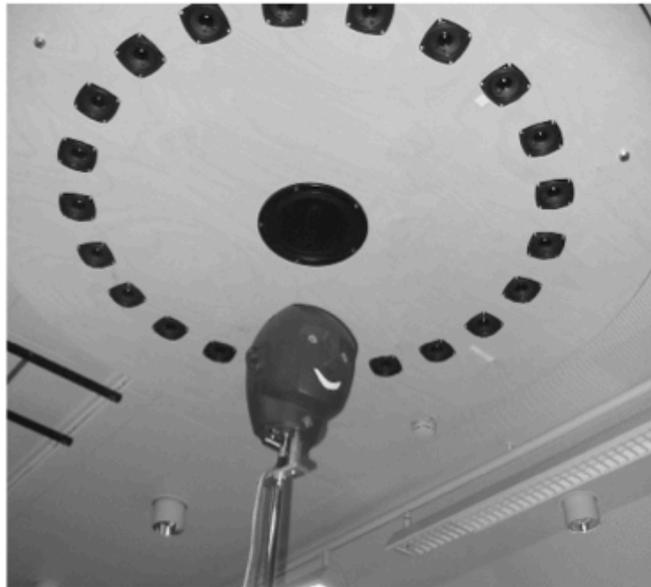


Abbildung 11: Binaural Sky⁴⁷

Werden wiederum nicht-binaurale Signale über Kopfhörer wiedergegeben, kommt es auf Grund des fehlenden Einflusses des Außenohrs auf das Schallfeld zur „Im-Kopf-Lokalisation“. Die Phantomschallquellen werden als zwischen den Ohren im Kopf positioniert wahrgenommen. Dieser Höreindruck wird nach einer Weile als ermüdend wahrgenommen. Dem kann entgegengewirkt werden, wenn das Schallsignal vorher mit einer HRTF gefaltet wird.⁴⁸ Es ist nicht gänzlich geklärt, wodurch das Phänomen der Im-Kopf-Lokalisation bedingt ist. Eine mögliche Theorie besagt, dass die eigene Stimme als im Kopf positioniert wahrgenommen wird und das einzige Objekt darstellt, welches bei Kopfbewegungen mitversetzt

⁴⁶ vgl. Dickreiter, et al. 2014

⁴⁷ Weinzierl 2009

⁴⁸ vgl. Görne 2014

wird. Im Umkehrschluss verortet das Gehör Schallquellen, welche sich nicht synchron zur Kopfbewegung verschieben, zwischen den Ohren.⁴⁹

Untersuchungen am Institut für Rundfunktechnik zeigten weiterhin, dass das Gehör Änderungen der interauralen Ohrsignalbeziehungen erwartet.⁵⁰ Dies kann bei der Wiedergabe von Binauralsignalen über Kopfhörer berücksichtigt werden, indem die Kopfposition des Hörers verfolgt und die Faltung entsprechend der Änderung des Winkels zwischen Ohr und Schallereignis angepasst wird (s. Abbildung 12). Dies erhöht den Lerneffekt, sich auf nicht-individuelle HRTF einzustellen und verbessert die wie oben erläutert problematische vorne/hinten-Lokalisation in signifikantem Ausmaß. Nötig ist dabei, dass die Nachführung der jeweiligen Impulsantwort auf die Kopfbewegung und der Faltungsprozess möglichst verzögerungsfrei erfolgen. Experimente von Sandvad 1996 und Karamustafoglu 1999 wiesen eine Latenz von 60 ms als maximale Grenze aus. Weitere Untersuchungen zeigten zudem, dass nur der Direktschall und die frühen Reflexionen der binauralen Impulsantworten nachgeführt werden müssen. Der diffuse Nachhall besitzt keinen relevanten richtungsbestimmenden Einfluss. Das Headtracking erfolgt in der Regel per an den Kopfhörern angebrachten oder in einer Virtual-Reality-Brille integrierten Positionssensoren.⁵¹

⁴⁹ vgl. Roginska und Geluso 2017

⁵⁰ vgl. Dickreiter, et al. 2014

⁵¹ vgl. Weinzierl 2009

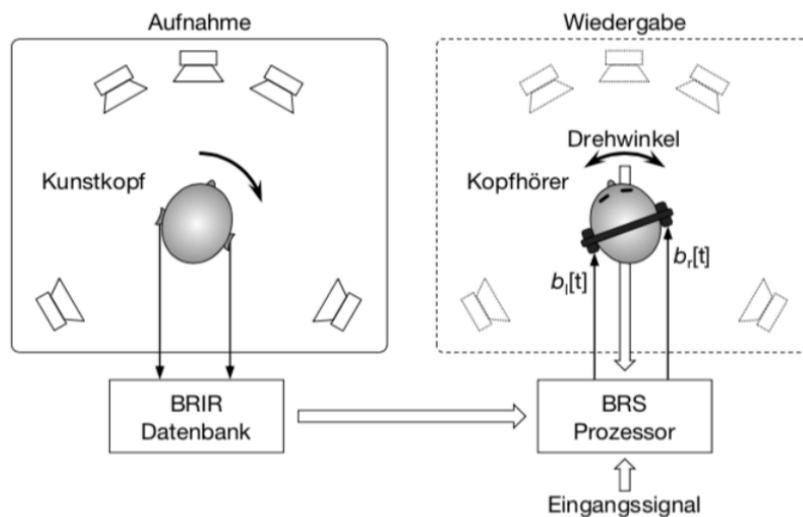


Abbildung 12: Prinzip der Binauralsynthese: Links Messung der BRIR mit Kunstkopf, Rechts Wiedergabe mittels Headtracking⁵²

2.7 Faktoren für Lokalisationsschärfe

Neben der Methode des Headtrackings und der Verwendung individueller HRTF-Sätze existieren weitere Faktoren, welche die Lokalisationsschärfe positiv beeinflussen. Der Psychoakustiker Jens Peter Blauert belegte in einem Experiment (Blauert 1696/70) die Existenz von richtungsbestimmenden Frequenzbändern, welche folglich auch unter dem Begriff blauertsche Bänder bekannt sind. Den Probanden wurde hierbei als schmalbandiges Signal Terzrauschen aus verschiedenen Richtungen der Medianebene, also der vertikal zwischen den Ohren verlaufenden Ebene, vorgespielt. Dabei zeigte sich, dass nicht die tatsächliche Einfallsrichtung der Schallquelle für die Lokalisierung entscheidend war. Stattdessen konnten für die Richtungen vorne, oben und hinten richtungsbestimmende Frequenzbänder identifiziert werden. So konnte das oben positionierte Terzrauschen nur bei einer Terzmittenfrequenz von ca. 8 kHz von den meisten Probanden korrekt lokalisiert werden (s. Abbildung 13).

⁵² Laumann 2016

Wurden als Stimuli statt des schmalbandigen Terzrauschens ein Breitbandsignal oder ein der Hörerin bekanntes Signal wie zum Beispiel Sprachaufnahmen verwendet, konnte die tatsächliche Einfallsrichtung zielsicher identifiziert werden.⁵³

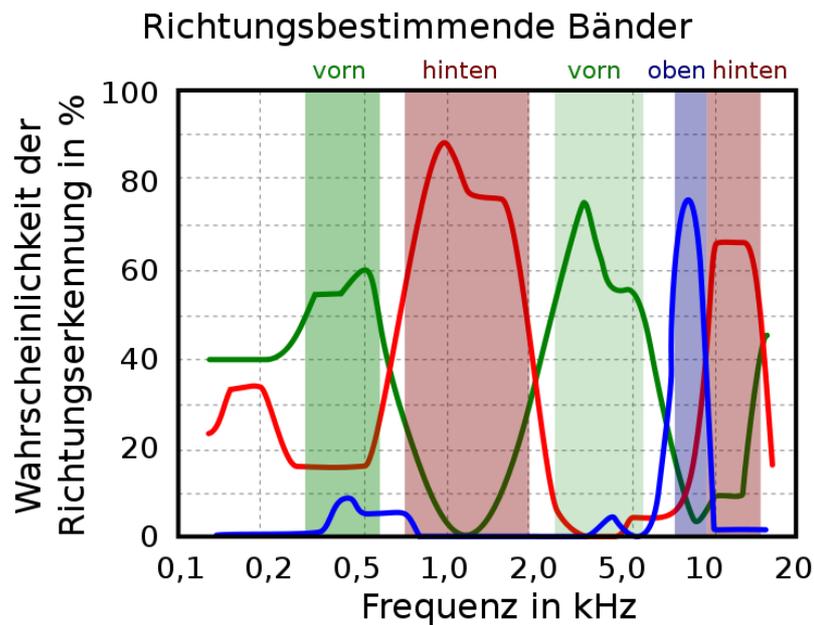


Abbildung 13: Richtungsbestimmende Bänder nach Jens Blauert⁵⁴

Auch visuelle Eindrücke können die Ortung von Schallereignissen positiv beeinflussen. Das Bellen eines Hundes kann beispielsweise zielsicherer einer Richtung zugeordnet werden, sofern in dieser dazu passend das Tier abgebildet ist. Dieses Prinzip ist in Anwendungen wie 360°-Video oder Virtual Reality relevant. Die korrekte Lokalisierung eines Schallereignisses ist zudem wahrscheinlicher, insofern es sich um ein vertrautes Geräusch handelt und dieses zudem mit einer erwartungsgemäßen Richtung verknüpft ist. Dies könnte zum Beispiel ein über dem Kopf kreisender Hubschrauber sein.⁵⁵

Die Lokalisierungspräzision ist durch die physischen Eigenschaften des Menschen beschränkt. Die kleinste Änderung eines Schallereignisses in seinem Winkel zum

⁵³ vgl. Weinzierl 2009

⁵⁴ Blauert, Räumliches Hören 1974

⁵⁵ vgl. Butz und Krüger, Mensch-Maschine-Interaktion 2014

Hörer, die zu einer veränderten wahrgenommenen Position der Quelle führt, liegt für die seitliche Abweichung in der Horizontalebene bei 4 Grad. In der Medianebene hört der Mensch mit einer weitaus niedrigeren Auflösung. Hier liegt die sogenannte Lokalisationsunschärfe bei 10 Grad.⁵⁶

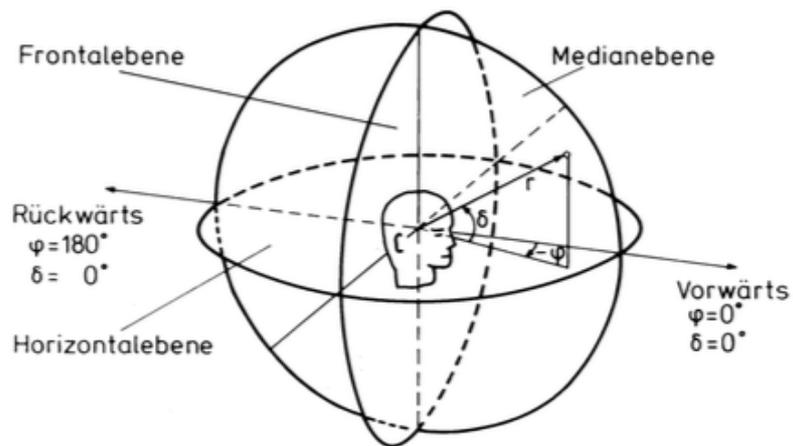


Abbildung 14: Kopfbezogenes Koordinatensystem nach Blauert⁵⁷

⁵⁶ vgl. Weinzierl 2009

⁵⁷ Blauert, Räumliches Hören 1974

3 Musik-Nutzerverhalten

3.1 Entwicklung der Musikbranche und Veränderungen des Nutzerverhaltens

Technischer Fortschritt und Nutzerverhalten sind stets in einer Wechselwirkung, welche auch die Musiklandschaft in den letzten Jahrzehnten gravierend geprägt haben. Ein entscheidender Faktor war dabei die Erfindung des MP3-Formats durch das *Fraunhofer Institut für integrierte Schaltungen*. Es wurde 1992 als Teil des MPEG-1-Standards festgeschrieben und dient der Speicherung und Übertragung von Audioinhalten. Der Schlüssel ist dabei eine Komprimierung um circa 85 % gegenüber dem verlustfreien CD-Signal bei weitgehendem Qualitätserhalt, was die Übertragung von Audioinhalten über das Internet (engl. Streaming) ermöglicht. Die Komprimierungsalgorithmen haben sich seitdem stetig weiterentwickelt und die Qualitätseinbußen werden immer geringer. Gleichzeitig wurde mit der Erfindung des Smartphones die Infrastruktur geschaffen, um das mobile Streaming von Musik zu ermöglichen. Derzeit sind weltweit 2,32 Milliarden Smartphones im Umlauf (Stand 2018).⁵⁸ Diese technologischen Entwicklungen bilden die Basis für aktuelle Veränderungen des Nutzerverhaltens. So nutzen mittlerweile 41 % der deutschen Konsumenten Musik-Streaming-Dienste täglich.⁵⁹ Der Anteil der Musikstreaming-Nutzer wuchs dabei von 9 % im Jahr 2013 auf aktuell 50 % (Stand 2018).⁶⁰ 86 % aller Internetnutzer weltweit hören Musik per Streaming und 75 % über ihr Smartphone.⁶¹ Als Wiedergabegerät werden dabei Kopfhörer immer beliebter: Der Absatz von Kopfhörern wuchs von 236 Mio. (2013) auf 368 Mio. (2017) an.⁶²

⁵⁸ Statista 2018

⁵⁹ Statista 2018

⁶⁰ Statista 2018

⁶¹ Statista 2018

⁶² Statista 2017

Mittlerweile geben 58 % der Konsumenten an Audioinhalte über Kopfhörer zu konsumieren, welche sich mit 63 % knapp hinter der Nennung „Radio, Stereoanlage“ auf Platz zwei in dieser Befragung einreihen.⁶³ Bei den musikkaffinen Internetnutzern bilden In-Ear-Kopfhörer mit 40 % sogar das beliebteste Wiedergabegerät (s. Abbildung 15).⁶⁴

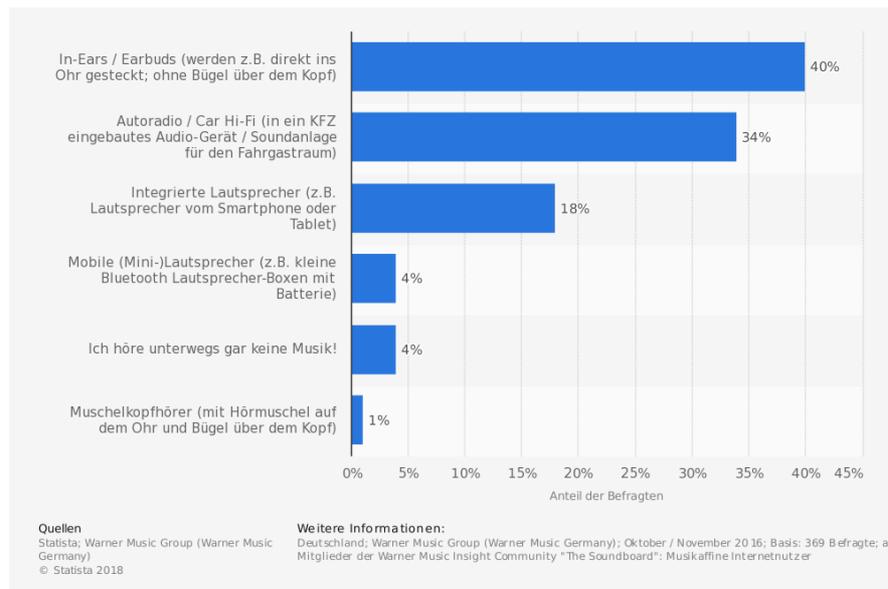


Abbildung 15: Anteil der befragten musikkaffinen Internetnutzer, die am liebsten über die folgenden Geräte unterwegs Musik hören

Der Trend geht dabei zu teureren, hochwertigen Modellen; Kopfhörer generieren mittlerweile mehr Geld als Lautsprecher.⁶⁵ Neben der Mobilität und der Unabhängigkeit von der Abhörumgebung bieten sie mehr Dynamik, kein Übersprechen zwischen den Kanälen und ein intensiver wahrgenommenes Hörerlebnis.⁶⁶

Insgesamt lässt sich also ein Trend weg von Lautsprecherwiedergabe und gekaufter Musik, ob in Form von Tonträgern oder digitalen Downloads, hin zu

⁶³ Statista 2015

⁶⁴ Statista 2017

⁶⁵ vgl. Günther 2012

⁶⁶ vgl. Köhler 2017

mobilem Streaming und Kopfhörerwiedergabe beobachten. Dabei werden mittlerweile nicht mehr ganze Alben, sondern eher Singles und Playlists konsumiert. Der Anteil von Alben am Musikkonsum der Befragten lag lediglich bei 22 %, während er bei Playlists 31 % und bei Singles 46 % betrug.⁶⁷

Was hat all dies für Auswirkungen auf den Musikmarkt? Wie müssen Muskschaffende auf die schnelllebige Internetkultur reagieren, um relevant zu bleiben? Mit der Abnahme von Downloads und dem gleichzeitigen Zuwachs von Streams, welche im Durchschnitt 0,005 \$ generieren, wächst der Druck auf die Künstler präsent zu bleiben. Statt wie früher üblich alle zwei Jahre ein Album zu veröffentlichen, macht es besonders für kleinere Künstler Sinn regelmäßig Musik zu veröffentlichen, um bei den Fans nicht in Vergessenheit zu geraten. Diesem Prinzip folgte die Band *White Violet*, welche ein Jahr lang zwei Songs im Monat veröffentlichte und diese schließlich auf einem Album kompilierte. Im gleichen Stil ging das Produzenten-Duo *The Chainsmokers* vor, als es einen Sommer lang ein Musikstück im Monat zusammen mit Musikvideos und weiterem „Social-Content“ wie „Behind-The-Scenes“-Material herausbrachte und im Herbst 2015 gebündelt auf ihrer EP *Bouquet* veröffentlichte. Ihr Manager Adam Alpert äußerte sich hierzu wie folgt:

“To stay hot, you have to deliver to your fans, and they consume quickly. It’s not just about delivering music. It’s about delivering all types of content.”⁶⁸

Für die meisten Künstler dürfte es schwierig sein, derart viel Musik zu generieren. An dieser Stelle könnten Remixe ihrer Songs ein Weg sein, mit überschaubarem Aufwand regelmäßig neues Material zu veröffentlichen. Auch Binaural-3D-Versionen könnten an dieser Stelle eine attraktive Rolle spielen.

⁶⁷ vgl. Jakobsen 2016

⁶⁸ Alpert 2015

3.2 Veränderungen der Hörgewohnheiten über die Zeit

Neben dem Nutzerverhalten haben auch die Hörgewohnheiten über die Jahrzehnte hinweg einige Veränderungen erfahren. Sie bilden sich durch Hörerfahrungen aus, mit welchen Gedächtnisinhalte aufgebaut werden. Jede neue Hörerfahrung wird dabei subjektiv interpretiert und mit den bisherigen Erfahrungen abgeglichen, und prägt damit die weitere Entwicklung der Hörgewohnheiten.⁶⁹ Diese werden ebenso wie das Nutzerverhalten dabei durch technische Entwicklungen beeinflusst. Während bei Einführung des Stereoformats dieses anfangs von mancher Seite noch recht kritisch gesehen wurde,⁷⁰ ist das Monoformat heutzutage für die Musikproduktion weitgehend irrelevant. Selbst bei der Produktion des Albums „Sgt. Pepper’s Lonely Hearts Club Band“ der für ihre experimentelle Nutzung der technischen Möglichkeiten der damaligen Zeit bekannte Band *The Beatles* wurden für die Erstellung der Monomischung drei Wochen investiert, während die Stereomischung in drei Tagen in Abwesenheit der Musikgruppe erstellt wurde. Im Jahr 2018 wurde schließlich in einem aufwändigen Verfahren eine Stereoversion des Albums produziert und somit den Hörgewohnheiten der heutigen Zeit angepasst.⁷¹ Auch beim Panning, also der Verteilung von Schallquellen im Stereopanorama, lässt sich eine Entwicklung der Herausbildung von Konventionen beobachten. Während anfangs noch extreme Panningmethoden geläufig waren, bei denen der Bass und das Schlagzeug rechts und der Gesang sowie die restlichen Instrumente links positioniert wurden, entwickelte sich mit der Zeit der Brauch der Positionierung tragender sowie tieffrequenter Objekte wie Gesang, Schlagzeug und Bass in der Panorama-Mitte. Diese Entscheidung war anfangs bedingt durch technische Aspekte wie zum einen der Limitationen der Vinylschallplatte sowie zum anderen der

⁶⁹ vgl. Köhler 2017

⁷⁰ vgl. Modell 2010

⁷¹ vgl. Tingen 2018

Gewährleistung von Monokompatibilität. Obwohl beide Faktoren heutzutage für den Großteil der Nutzer nicht mehr derart relevant sind, blieben diese Panning-Konventionen bestehen und sind Teil heutiger Hörgewohnheiten.⁷² Deren Beeinflussung durch technische Innovationen lässt sich auch am Beispiel der Klangcharakteristik der Popmusik der 1980er Jahre erkennen, welche durch künstliche synthetische Klänge, digitale Halleffekte und maschinelle, quantisierte Rhythmusinstrumente geprägt wurde.⁷³ Konkret lässt sich diese Veränderung anhand der Produktion der Schlagzeugaufnahmen identifizieren. Während in den 1970ern deren Klang noch eher trocken und direkt ausfiel, änderte sich dies, als bei den Schlagzeugaufnahmen von Peter Gabriels Album „Melt“ das Talkback-Mikrofon der Studioregie versehentlich mit aufgenommen wurde. Dieses war mit einem Kompressor und einem Noisegate versehen, welches der Snare einen peitschenartigen Klang gab. Der Schlagzeuger dieser Produktion Phil Collins entschied sich bei den Aufnahmen seiner eigenen Komposition „In The Air Tonight“ 1981 diesen Klangcharakter zu replizieren, indem er die Aufzeichnung des Schlagzeugs in einem Studio mit Steinwänden durchführte und damit die Blaupause für den Schlagzeugklang der 1980er kreierte. Begünstigt durch die Entwicklung des digitalen Halleffekts konnte diese Ästhetik von zahlreichen Künstlern imitiert werden und prägte damit die Hörgewohnheiten der Rezipienten in diesem Jahrzehnt.⁷⁴ Die digitale Revolution und die Erfindung des MP3-Formats hat neben ihrem erwähnten Einfluss auf das Nutzerverhalten sicherlich auch die Hörgewohnheiten insofern beeinflusst, als dass der Faktor Klangqualität auf der Prioritätenliste dem der schnellen Verfügbarkeit weichen musste. Dies trifft insbesondere auf die 2000er Jahre zu, als die Kompressionsverfahren noch nicht derart fortgeschritten waren und hörbare Artefakte verursachten. Auch der

⁷² vgl. Izhaki 2013

⁷³ vgl. Inglis 2015

⁷⁴ vgl. Vox 2017

Einfluss des sogenannten Lautheitskriegs, begünstigt durch die technischen Möglichkeiten des digitalen Zeitalters, auf die dynamikbezogenen Hörgewohnheiten der Rezipienten sei hier als prägender Faktor zu nennen.⁷⁵ Für die Zukunft wird es in diesem Sinne interessant sein zu sehen, inwiefern aktuellen technische Entwicklungen wie Virtual Reality, 360°-Video und 3D-Audio sich am Markt durchsetzen können und die Hörgewohnheiten der Nutzer in Zukunft beeinflussen werden.

⁷⁵ dpa 2016

4 Aktuelle Entwicklungen

4.1 Aktuelle Anwendungen von Binauraltechnik

In den letzten Jahren lässt sich eine immense Ausweitung binauraler Produktionen feststellen. Neben den zu Beginn erwähnten Bereichen der Forschung und Entwicklung ist dies vor allem in Unterhaltungsmedien wie 360°-Film, Virtual Reality und Hörspiel der Fall. Einige konkrete Beispiele aus der Unterhaltungsindustrie werden im Folgenden präsentiert.

2016 veröffentlichte der Entwickler *Dowino* mit „A Blind Legend“ ein Spiel, welches auf die Bildebene verzichtet und per binauralem Ton auch Menschen mit Sehbehinderung ein immersives Erlebnis bietet.⁷⁶ Parallel zu der Premiere des Kurzfilms „Notes On Blindness: Into Darkness“ von Peter Middleton und James Spinney wurde die gleichnamige interaktive VR-Applikation veröffentlicht, welche mit Hilfe von binauralem Ton die biografische Geschichte auf immersive Art erlebbar macht und dafür mit sehr positiven Kritiken gewürdigt wurde.⁷⁷ Der Kurzfilm „Final Stop: A 3D Audio Thriller“, produziert vom Unternehmen *Sennheiser* zur Präsentation ihrer Ambisonics-Produkte, verfolgt hier einen etwas anderen Ansatz und experimentiert mit einer Kombination aus konventionellem 2D-Bildinhalt und binauralem Filmtone.⁷⁸ Auch die öffentlich-rechtliche Rundfunkanstalt des Vereinigten Königreichs *BBC* hat bereits einige Dokumentarfilme veröffentlicht, welche im 360°-Videoformat aufgenommen sind und dem Zuschauer mit interaktivem binauralen Ton die jeweiligen Themeninhalte auf immersiver Ebene nahe bringen. So begleitet der Kurzdokumentarfilm „Himalayas: a Trek to School in 360 video – BBC News“ zwei Schüler auf ihrem Schulweg im Himalaya Gebirge.⁷⁹

⁷⁶ Dowino 2016

⁷⁷ Notes On Blindness 2016

⁷⁸ Sennheiser 2018

⁷⁹ BBC News 2017

Im Musikbereich ist die Zahl aktueller Binauralproduktionen recht überschaubar. Die Recherche wird dabei erschwert durch die Wortüberlagerung des 3D-Binauralformats mit dem akustischen Phänomen der „binauralen Beats“, welches jedoch mit der dieser Arbeit zugrunde liegenden Materie nicht zu verwechseln ist. Es beschreibt das Phänomen einer aus zwei überlagerten Frequenzen resultierenden Schwebefrequenz, welche einigen Theorien zufolge eine stimulierende Wirkung auf das menschliche Gehirn haben soll.⁸⁰ Trotz diesen Umstands konnten einige relevante Produktionen binauraler Musik ausgemacht werden. Der Künstler Ottmar Liebert veröffentlichte 2008 das mit einem Kunstkopf aufgenommene Album „Up Close“, welches auf Grund seines überzeugenden Raumeindrucks nennenswert ist. Der Musiker Beck konzipierte 2013 ein mit einer 360°-Kamera aufgezeichnetes Konzert, bei dem der Ton mittels eines Kunstkopfs aufgenommen wurde, welcher mit vier Blickrichtungen und somit acht Ohren bzw. Mikrofonen ausgestattet war (s. Abbildung 16). Die Wiedergabe per VR-Brille ermöglichte die Verfolgung der Kopfposition und somit die interaktive Abbildung des korrespondierenden Kunstkopfsignals.⁸¹



Abbildung 16: Kunstkopf mit vier Blickrichtungen⁸²

Darüber hinaus wurden in den letzten Jahren einige weitere Musikproduktionen basierend auf dem Prinzip der Binauralsynthese veröffentlicht. Auf der

⁸⁰ vgl. Fessl 2017

⁸¹ MarDie1530HD 2018

⁸² Roginska und Geluso 2017

Internetplattform *Youtube* existieren zahlreiche Kanäle für binaurale 3D-Musik. Es finden sich populäre Stücke wie „Bohemian Rhapsody“ von *Queen* oder „Wake Me Up“ von *Avicii*, welche mit millionenstelligen Klickzahlen und einer Vielzahl positiver Bewertungen und Kommentaren gewürdigt werden.⁸³ Der *Youtube*-Nutzer Andrew Ward äußert sich zur binauralisierten Version von „Bohemian Rhapsody“ wie folgt:

Absolutely amazing, I loved this song already but this version is probably my favourite. I really wish more songs were done in this way. It really brings the performance to life.⁸⁴

Die Stücke haben dabei gemein, dass sie sehr drastische, effektvolle Bewegungen der Klangobjekte verwenden. So wird die Gesangsstimme in ostentativer Manier um den Kopf herumbewegt. Dies mag zur Verdeutlichung der Möglichkeiten des Binauralformats sinnvoll sein und beim ersten Hören imposant wirken. Es darf jedoch angenommen werden, dass sich dieser Effekt nach einiger Zeit abnutzt und als störend empfunden wird. Neben dieser eher amateurhaft produzierten Binauralmusik existieren jedoch auch Beispiele, welche von den Künstlern selbst oder erfahrenen Toningenieuren angefertigt wurden. Die Firma *Sennheiser* veröffentlichte zur Demonstration ihrer Ambisonics-Produkte in Zusammenarbeit mit den Musikern Robin Schulz und Nico Santos einen entsprechenden Song. Dieser wurde zwar mit positiven Bewertungen und Kommentaren rezipiert, indes bezogen sich diese eher auf die Komposition an sich statt auf den eher dezent ausgestalteten Raumeindruck.⁸⁵ Auch die deutsche Musikgruppe *Kraftwerk* veröffentlichte im Mai 2017 ihre Albumkollektion „3-D Der Katalog“, deren siebte CD neun Binaural-3D-Mischungen enthält. Diese

⁸³ Dawn of Music 2018

⁸⁴ A. Ward 2016

⁸⁵ Sennheiser 2016

entstanden in Zusammenarbeit der Künstler mit dem Entwickler der 3D-Audio-Software *Spatial Audio Designer*, Tom Ammermann. Das Gesamtwerk wurde sehr positiv rezipiert und erhielt eine *Grammy Award*-Auszeichnung als „Best Dance/Electronic Album“ sowie eine Nominierung als „Best Surround Sound Album“. ⁸⁶ Neben einzelnen Künstlern, die mit dem Binauralformat experimentieren, existieren mittlerweile auch eigene Labels, welche sich auf die Mischung von räumlicher, immersiver Musik fokussieren, wie etwa das 2018 gegründete Label *The Zero System*.⁸⁷ Auch im Genre der klassischen Musik wurden in den letzten Jahren einige binaurale Mischungen veröffentlicht. So zeichnet die *Research & Development*-Abteilung der *BBC* seit 2016 die Konzertreihe *Proms* mit 3D-Mikrofonssystemen auf und stellte die binauralisierten Versionen der Öffentlichkeit zur Verfügung.⁸⁸ Sowohl auf der Tonmeistertagung 2016 als auch 2018 fanden diese Produktionen viel Beachtung.

4.2 Distribution und Wiedergabeformate

Aktuell existieren eine Vielzahl verschiedener 3D-Formate mit jeweils unterschiedlichen Konfigurationen und Wiedergabesystemen. Vor allem kanalbasierte Systeme wie Auro-3D oder 22.2 verlangen eine Mischung, die direkt auf das jeweilige Lautsprecher-Setup angepasst ist. Da es jedoch derzeit noch kein Verfahren gibt, das sich vorherrschend durchgesetzt hätte, ist diese Arbeitsweise nicht sehr praktikabel und es müsste für jedes System eine eigene Version gemischt werden. Dieses Problem gehen objektbasierte Verfahren wie *Dolby Atmos* deutlich besser an. Hierbei besteht jedes Audio-Objekt aus jeweils seinem Inhalt und zusätzlichen Metadaten, welche unter anderem die Position, Ausbreitung und Bewegungsenergie definieren. So kann das Schallfeld flexibel auf eine beliebige Lautsprecher-Anordnung gerendert werden.

⁸⁶ Recording Academy Grammy Awards 2018

⁸⁷ The Zero System 2018

⁸⁸ BBC 2018

Auch bei *Ambisonics* handelt es sich um ein Aufnahme- und Wiedergabeverfahren mit variablen Wiedergabeformaten und zudem einer beliebig hohen Genauigkeit. Es wurde 1970 von dem britischen Mathematiker Michael Gerzon zusammen mit Peter Carven und Richard Lee an der Universität Oxford entwickelt.⁸⁹ In seiner einfachsten Form wird das Schallfeld mit vier Nierenmikrofonen aufgenommen, die in tetraedrischer Form angeordnet sind. Diese vier Kanäle werden als sogenanntes A-Format zusammengefasst und durch eine Rechnung in das B-Format gewandelt. Der Kanal W beschreibt dabei den omnidirektionalen Anteil beziehungsweise den Schalldruck, während die Kanäle X, Y, Z den Druckgradienten in die jeweilige Richtung beinhalten.⁹⁰ In dieser einfachen Form, auch bekannt als *Ambisonics* erster Ordnung, kann jedoch nur eine recht ungenaue Abbildung des Schallfelds erreicht werden. Um diese zu verbessern, muss eine höhere Ordnung gewählt werden (engl. *Higher Order Ambisonics*, s. Abbildung 17). So besitzt das *Ambisonics*-Format dritter Ordnung 16 Kanäle. Dies ermöglicht eine bessere Richtungsabbildung und eine Vergrößerung des „Sweet Spots“, welcher ab höheren Ordnungen zur „Sweet Area“ wächst.⁹¹ Es gibt jedoch stets eine obere Frequenzgrenze, bis zu der das Schallfeld perfekt produziert werden kann. Bei *First Order Ambisonics* (FOA) liegt diese bei bereits 600 Hz, bei *Third Order Ambisonics* bei 1.800 Hz. Um die für das menschliche Gehör relevante Auflösung von 20.000 Hz zu erreichen, wäre die Verwendung von *Ambisonics* 32ster Ordnung mit wiedergabeseitig 1.000 Lautsprechern notwendig.⁹²

Von diesem Umstand abgesehen bietet *Ambisonics* die Vorteile, dass es stets abwärtskompatibel ist, flexibel an das Wiedergabeformat angepasst werden

⁸⁹ vgl. Gerzon, *With Heigh Sound Reproduction* 1973

⁹⁰ vgl. Gerzon, *The Design of Precisely Coincident Microphone Arrays for Stereo and Surround Sound* 1975

⁹¹ vgl. Sontacchi 2003

⁹² vgl. Ward und Abhayapala 2001

kann, auch die Wiedergabe über Kopfhörer ermöglicht und aufnahmeseitig ein recht unkompliziertes System zur Aufzeichnung von Schallfeldern darstellt. Auf dem Musikmarkt konnte es sich zwar bislang nicht durchsetzen, ist aber dank Virtual Reality und 360°-Video in den letzten Jahren wieder vermehrt als Aufnahme- und Wiedergabeverfahren in Erscheinung getreten.⁹³

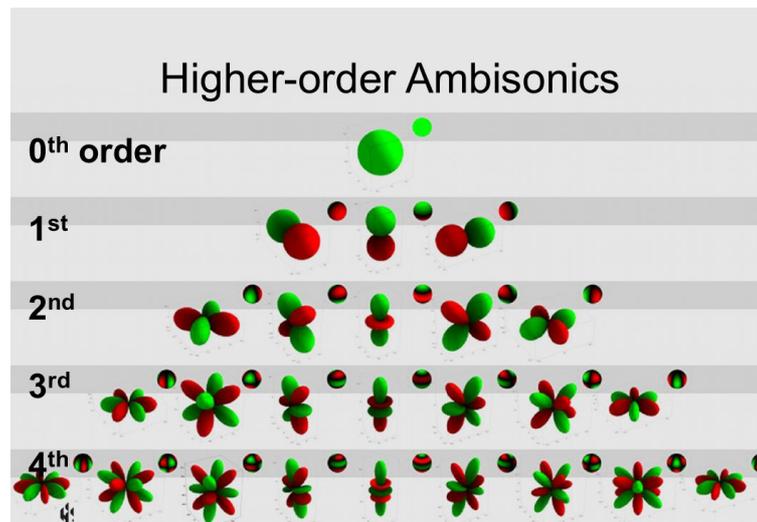


Abbildung 17: Higher Order Ambisonics, dargestellt: 0. Bis 4. Ordnung⁹⁴

Für die Zukunft dürften jedoch Formate interessanter sein, welche ähnlich wie *Dolby Atmos* den objektbasierten Ansatz verfolgen und sich an die jeweilige Wiedergabeplattform und den verwendeten Medientyp anpassen. Das Ziel ist dabei eine sogenannte ‚Universal Delivery‘ (dt. „universelle Bereitstellung“) zu ermöglichen. Auch hier ist jedes Objekt mit Metadaten versehen, welche unter anderem die Bewegung im Raum beschreiben, allerdings auch Informationen bezüglich der Sprache, Untertitel oder Kommentare beinhalten können. Die raumbezogenen Metadaten werden beim Wiedergabegerät durch eine Rendering-Einheit interpretiert und für die jeweilige Plattform, beispielsweise Lautsprechersystem (Stereo, Surround, 3D), Soundbar oder Kopfhörer, berechnet.⁹⁵ Die *Internationale Fernmeldeunion* (engl. *International*

⁹³ vgl. Elen 2001

⁹⁴ ResearchGate 2010

⁹⁵ vgl. Martins 2018

Telecommunication Union) hat mit dem *Audio Definition Model* bereits einen solchen offenen Standard entwickelt. Hierbei handelt es sich gemäß der Empfehlung ITU-R BS.2076-1 um ein Modell zur Beschreibung von Audioobjekten, welches die Kompatibilität von objektbasierter Übertragung mit Metadaten über verschiedene Plattformen ermöglicht.⁹⁶ Im Rahmen des EU-Projekts *Orpheus* wurde mit *MPEG-H 3D-Audio* weiterhin ein Audio-Standard entwickelt, welcher die Codierung von Audio-Objekten, Metadaten sowie Audiokanälen und Higher Order Ambisonics ermöglicht. Der MPEG-H Dekodierer rendert hierbei den Bitstream bei Wiedergabe auf die jeweilige Lautsprecheranordnung oder per Binauralsynthese auf Kopfhörer (s. Abbildung 18). Dabei werden bis zu 128 Kanäle oder 128 Audioobjekte und Ambisonics bis zur 29. Ordnung unterstützt.⁹⁷ Ein effizientes Codierungsverfahren ermöglicht die Übertragung per Streaming oder über Rundfunk bei Bitraten in der Größenordnung heutiger 5.1-Surround-Übertragung. MPEG-H wird Stand 2018 bereits in Korea und den Vereinigten Staaten sowie in Europa und China in den neuen TV-Standards implementiert.

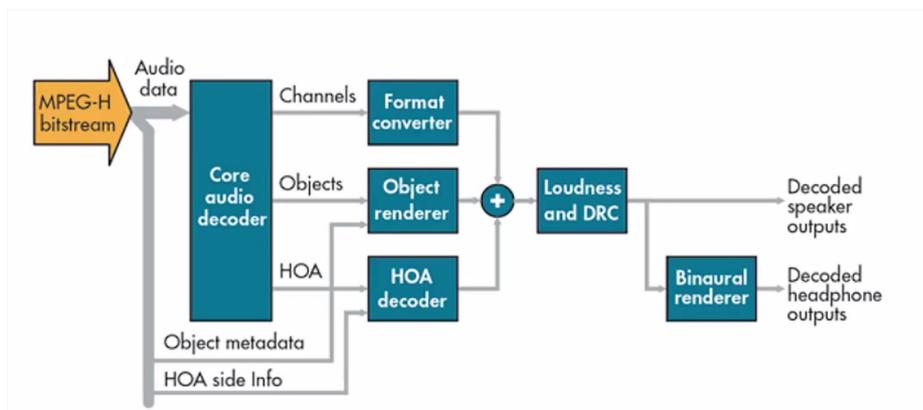


Abbildung 18: MPEG-H Audiocodierung, simplifiziertes Schema⁹⁸

Auf Produktionsseite existieren mittlerweile mit den Software-Lösungen *Pyramix 11.1* von *Merging Technologies* und dem *Spatial Audio Designer* von *New Audio*

⁹⁶ International Telecommunication Union 2017

⁹⁷ vgl. Fleischmann 2017

⁹⁸ mixingroom.de 2017

Technologies Softwarelösungen für den Export von ADM und MPEG-H.⁹⁹ Die gebotenen Möglichkeiten durch diese neuen Standards sind mannigfaltig und die Ausschöpfung des vollen Potentials von objektbasiertem Audio befindet sich erst im Anfangsstadium. Es lässt sich jedoch vermuten, dass sie die technische und infrastrukturelle Basis für die zukünftige Marktrelevanz von 3D-Audio legen.

Auch auf Wiedergabeseite wurden in den letzten Jahren vielversprechende Entwicklungen veröffentlicht. Hervorzuheben sind dabei zum einen der Binauralisierungs-Prozessor *Smyth Realiser*, welcher die Messung der individuellen HRTF ermöglicht und die Virtualisierung von bis zu 16 Lautsprechern sowie die Dekodierung von 3D-Formaten wie *Dolby Atmos*, *Auro-3D* und *DTS:X* ermöglicht. In Kombination mit einem Headtracking-System verspricht der Hersteller *Smyth Research* sehr realistische Ergebnisse.¹⁰⁰ Auch die Entwicklung der Soundbar-Technologie ist mittlerweile so weit fortgeschritten, dass die Erzeugung eines überzeugenden 3D-Klangfelds per Nutzung der Raumreflexionen ermöglicht wird. Dabei kann der Installierungsaufwand für die Konsumentin minimal gehalten werden bei gleichzeitiger Bereitstellung eines gemäß bisheriger Kritiken umhüllenden, realistischen 3D-Klangerlebnisses.¹⁰¹

4.3 Aktuelle Untersuchungen zum Thema Binauraltechnik

Die Renaissance des binauralen Formats, vorangetrieben durch technische Entwicklungen wie Virtual Reality und 360°-Film, äußert sich auch in einer intensivierten Auseinandersetzung der Forschung mit dem Thema Binauraltechnik. Einige dieser Erkenntnisse werden im Folgenden kurz zusammengefasst und präsentiert.

Bereits 2007 wurde eine Studie zur Evaluation des Binauralformats im Kontext von Popmusik durchgeführt. Als Versuchsstem wurde eine aus Gesang und Gitarre

⁹⁹ vgl. Martins 2018

¹⁰⁰ Smyth Research 2018

¹⁰¹ vgl. Jahn 2018

bestehende Musikdarstellung in Stereo und mit einem Kunstkopf aufgezeichnet sowie zusätzlich binaural synthetisiert. In einer Nutzerstudie wurden den Probanden diese Items präsentiert. Es konnte gezeigt werden, dass die Räumlichkeit der binauralen Versionen im Vergleich positiver bewertet wurde. Bezüglich der Klangfarbe wurde jedoch die Stereoversion bevorzugt. Insgesamt konnte keine Präferenz für das Binauralformat aufgezeigt werden, welches unter anderem mit den technischen Limitationen der Technik begründet wurde.¹⁰² Auch gilt es zu hinterfragen, ob die Möglichkeiten des Binauralformats mit nur zwei akustischen Schallquellen hinreichend ausgereizt werden können.

Eine der zentralen Limitationen des Binauralformats ist die Verwendung nicht-individueller HRTF-Sätze, welche sich in Klangfärbungen und einer suboptimalen Lokalisierung äußert. Daher ist die Individualisierung von Standard-HRTF ein dringliches Thema einiger aktueller Untersuchungen. Ein vielversprechendes Prinzip konnte in einer Studie der Hochschule Düsseldorf ausgemacht werden. Dabei wurden die ITD von der vorhandenen HRTF getrennt, auf Grundlage des Ohrabstandes der Nutzerin neu berechnet und anschließend zur Laufzeit auf die beiden Binauralsignale hinzuaddiert. Die Entfernung der IDT aus der HRTF gestaltet sich dabei auf konventionellem Weg als recht fehleranfällig. In der Studie konnte eine neuartige Methode identifiziert werden, bei der zuerst die Transformation in eine minimalphasige HRTF unternommen wird. Diese Methode erwies sich in Tests als weitgehend fehlerfrei.¹⁰³ In einer weiteren Studie von 2018 wurden verschiedene Ansätze der Individualisierung evaluiert. Zwar erreichte keine Methode das qualitativ hochwertige Ergebnis einer individuellen Messung, jedoch erwiesen sich einige Ansätze, wie beispielsweise die Erstellung eines geometrischen Scans, als vielversprechend.¹⁰⁴ Auch die Möglichkeit, aus einer

¹⁰² vgl. Fontana, Farina und Grenier 2007

¹⁰³ vgl. Bau, Epe und Leckschat 2018

¹⁰⁴ vgl. Guezenoc und Séguier 2018

Datenbank mit mehreren HRTF die bestpassendste auswählen zu können, lieferte positive Ergebnisse.¹⁰⁵

Erwähnenswert ist zudem eine aktuelle Studie, in der die Steigerung der Intensität sowie Qualität von musikinduzierten Emotionen in Abhängigkeit vom 3D-Audioformat untersucht wurde. Hierfür wurden verschiedene Ausschnitte klassischer Musik in den Formaten Stereo, Surround 5.1 sowie Auro 9.1 produziert und den Probanden präsentiert. Es konnte gezeigt werden, dass eine signifikante Korrelation zwischen der subjektiv wahrgenommenen emotionalen Intensität der Musik und dem Wiedergabeformat existiert. Dabei wirkte sich die Steigerung der emotionalen Intensität bei dem Format Auro-3D 9.1 am stärksten aus (s. Abbildung 19).¹⁰⁶

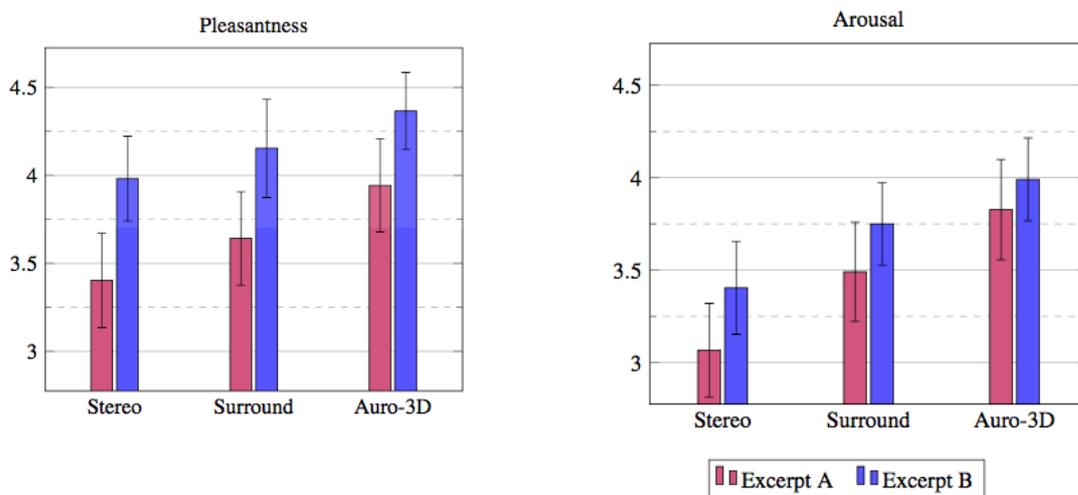


Abbildung 19: Gefälligkeit (links) und emotionale Erregung (rechts) in Abhängigkeit vom Wiedergabeformat¹⁰⁷

¹⁰⁵ vgl. Roginska und Geluso 2017

¹⁰⁶ vgl. Hahn 2018

¹⁰⁷ Hahn 2018

5 Produktion der Hörproben

Im Bereich der populären Musik ist die geringe Relevanz des Binauralformats wie beschrieben besonders ausgeprägt. Die meisten Binauralproduktionen sind im Genre der elektronischen oder klassischen Musik beheimatet, wobei im letzteren Fall in der Regel das Ziel verfolgt wird, den Klangeindruck eines Konzertbesuchs möglichst naturgetreu darzustellen. Um jedoch den Mehrwert des Binauralformats für die breite Masse der Hörerschaft feststellen zu können, sollten auch die Stimuli für die Studie möglichst massenkompatibel gewählt sein. Aus diesem Grundansatz resultierte die Notwendigkeit, den Anforderungen entsprechende Hörproben neu zu generieren. Da für die Produktion von binauralen Audioinhalten noch kein etablierter Arbeitsablauf existiert, mussten weiterhin für den vorliegenden Anwendungsfall geeignete Softwarelösungen sondiert und Produktionsmethoden entwickelt werden. Diese werden im Folgenden vorgestellt und beschrieben.

5.1 Kriterien der Musikauswahl

Neben der erwähnten Massenkompatibilität und der damit einhergehenden Fokussierung auf das Genre der Popmusik wurden weitere Attribute ausgemacht, welche die Hörproben erfüllen sollten. Neben einer Basis an akustischen bzw. der Hörerin bekannten Klangquellen wie Gesang, Bass, Schlagzeug oder Gitarre sollten die Musikstücke weiterhin eine Vielzahl an synthetischen Instrumenten und Klängen beinhalten, welche frei im Raum verteilt und bewegt werden konnten, ohne Gefahr zu laufen, die Versuchspersonen (Vpn) in ihren Hörgewohnheiten zu irritieren. Während eine über der Hörerin positionierte Gesangsquelle leicht als störend wahrgenommen werden könnte, dürften frei im Raum positionierte synthetische Objekte nicht unnatürlich wirken. Dies gilt besonders für Klangquellen, die für die Struktur des Musikstücks keine tragenden Elemente darstellen, wie bspw. Dopplungen, Klangeffekte sowie Echo- und Hallwege. Weiterhin sollten die Musikstücke durch eine räumlich weite Grundästhetik

gekennzeichnet sein, sodass großzügige 3D-Halleffekte zur Verbesserung der Lokalisierung verwendet werden konnten, ohne das Klangbild der Binaural- zu drastisch von der Stereoverision abweichen zu lassen. Die Hörproben sollten zudem einem hohen qualitativen Anspruch genügen sowie im besten Fall bereits kommerziell veröffentlicht und positiv rezipiert worden sein, um das Risiko einer gleichgültig oder negativ ausfallenden Rezeption der Probanden gegenüber den Stimuli so gering wie möglich zu halten.

Auf Grundlage dieser Anforderungen wurde eine Liste entsprechender Musikstücke erstellt, deren Künstler bzw. Labels kontaktiert und um Freigabe der Einzelspuren gebeten wurden. Zu diesem Zweck wurde ein Infopaket erstellt, in welchem das Projekt vorgestellt wurde (s. Anhang). Die Resonanz auf die ca. 40 versendeten Anfragen erwies sich jedoch als sehr gering und es kam nur in einem konkreten Fall zu einem tatsächlichen Datentransfer. Um für den Hörversuch weitere Hörproben realisieren zu können, wurde auf frei erhältliche Einzelspuren kommerziell veröffentlichter Popmusik zurückgegriffen.

Popmusik beschreibt dabei laut dem Duden „massenhaft verbreitete populäre Musik bzw. Unterhaltungsmusik unterschiedlicher Stilrichtungen.“¹⁰⁸ Der Begriff wurde zwar zu der Zeit geprägt, als Elvis Presley mit seiner Adaption von Rhythm & Blues seinen Durchbruch erlangte und Jugendliche ihre eigene Kultur zur Abgrenzung von der Elterngeneration kreierten. Heute umfasst der Oberbegriff „Pop“ allerdings eher standardisierte, an Markt und Wirtschaftlichkeit orientierte Musik, welche simpel, leicht begreifbar und unterhaltsam sein sollte.¹⁰⁹ Für die vorliegende Untersuchung erschienen diese Attribute sinnvoll, da sie sicherstellen, dass der Großteil der Probanden die Hörproben nicht grundlegend ablehnt. Auch für das erklärte Ziel der Evaluierung der Marktrelevanz des Binauralformats erschien die Verwendung massenkompatibler Popmusik sinnvoll.

¹⁰⁸ Duden 2018

¹⁰⁹ vgl. Laut 2018

5.2 Musikauswahl

Auf Anfrage erklärte sich die britische, im Jahr 2009 gegründete Electronic Rock-Band *Django Django* bereit, Einzelspuren ihres 2018 auf ihrem dritten Studioalbum „Marble Skies“ veröffentlichten Songs „In Your Beat“ bereitzustellen. Dieser ist im Genre des Art Pops angesiedelt und ist mit einer Vielzahl an elektronischen Instrumenten und effektvollen Elementen bestückt. Insgesamt umfasste das transferierte Dateipaket 98 Einzelspuren (s. Abbildung 20).

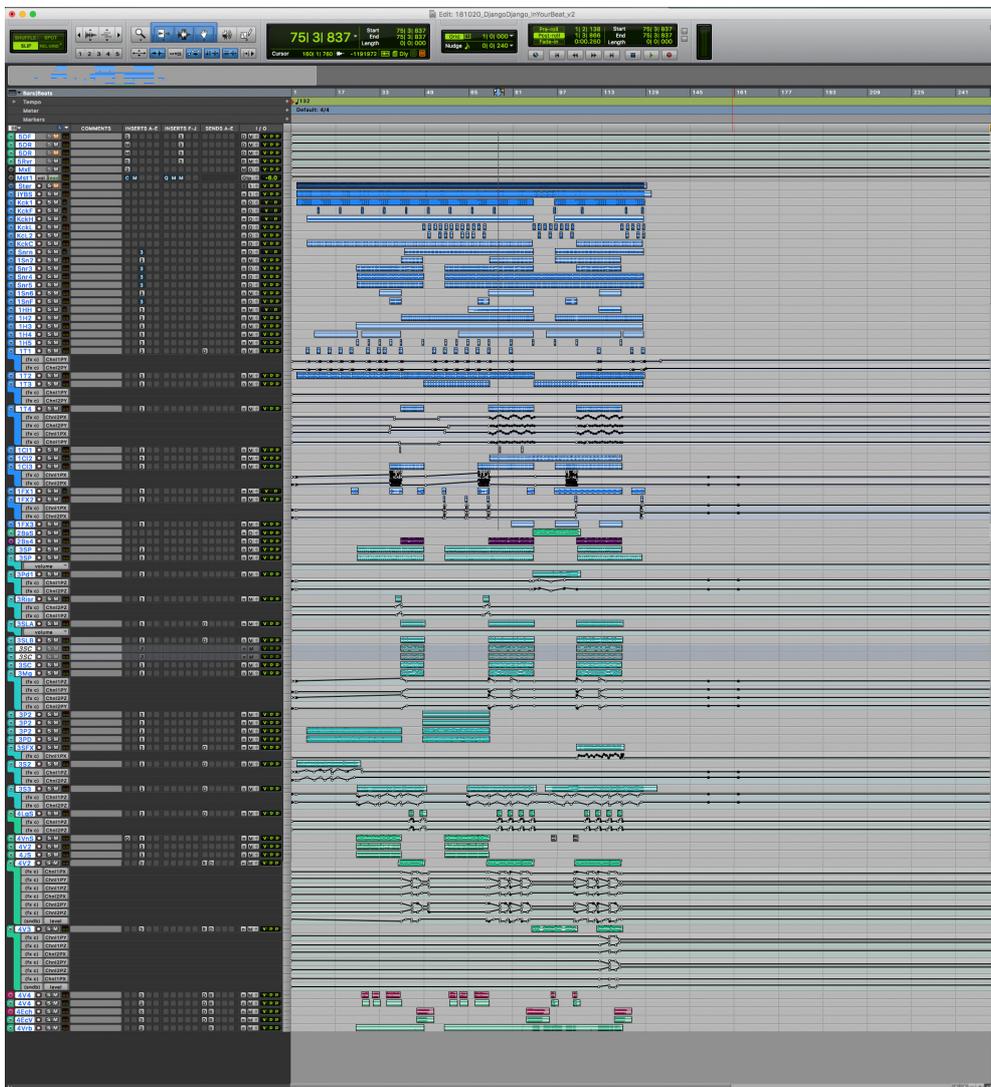


Abbildung 20: Pro Tools Projekt "In Your Beat" (Screenshot)

Für die Produktion weiterer Hörproben konnte auf zwei Demoprojekte zurückgegriffen werden, die mit der Musikbearbeitungssoftware *Logic X* des

Herstellers *Apple* mitgeliefert werden. Beide Projekte stammen von kommerziell veröffentlichten Musikproduktionen und gleichen diesen in ihren Elementen mit Ausnahme der verwendeten Mastering-Prozessierung. So wurde das Stück „Helena Beat“ der US-amerikanischen Indie-Pop-Band *Foster The People* von ihrem Frontmann Mark Forster mit der erwähnten Software produziert und 2010 auf ihrem Debütalbum „*Torches*“ sowie als zweite Single veröffentlicht.¹¹⁰ Bestehend aus insgesamt 97 Spuren und einer Vielzahl an vielschichtigen, elektronischen Effektspielereien stellt es den kantigen Synthesizer-Sound der Band in den Mittelpunkt. Bei Kritikern und Hörern wurde diese Klangästhetik sehr positiv aufgenommen und resultierte darin, dass „Helena Beat“ eines der meistgestreamten Songs 2011 wurde.¹¹¹

Im Jahr 2018 wurde die Projektdatei des Musikstücks „*Colors*“ des Künstlers Beck dem Lieferumfang von *Logic X* hinzugefügt. Dem US-Amerikaner Beck Hansen gelang 1994 mit der Single „*Loser*“ der weltweite Durchbruch. Seitdem ist der Komponist, Multi-Instrumentalist, Sänger und Produzent für seine eklektische Zusammenführung verschiedener Musikstile bekannt, welche sich unter dem Genre Alternative Rock bzw. Art Pop zusammenfassen lassen. Nach bisherigen fünf *Grammy Awards* veröffentlichte der Künstler im Jahr 2017 sein 13. Studioalbum „*Colors*“, dessen gleichnamige Singleauskopplung eine vielschichtige Modern-Pop-Komposition darstellt und auf insgesamt 101 Spuren konventionelle und synthetische Instrumente sowie eine Myriade an verspielten Klangeffekten und Samples zu einem kohärenten Gesamtwerk verflechtet (s. Abbildung 21).¹¹²

¹¹⁰ vgl. Lester 2010

¹¹¹ vgl. Murphy 2011

¹¹² Recording Academy Grammy Awards 2018

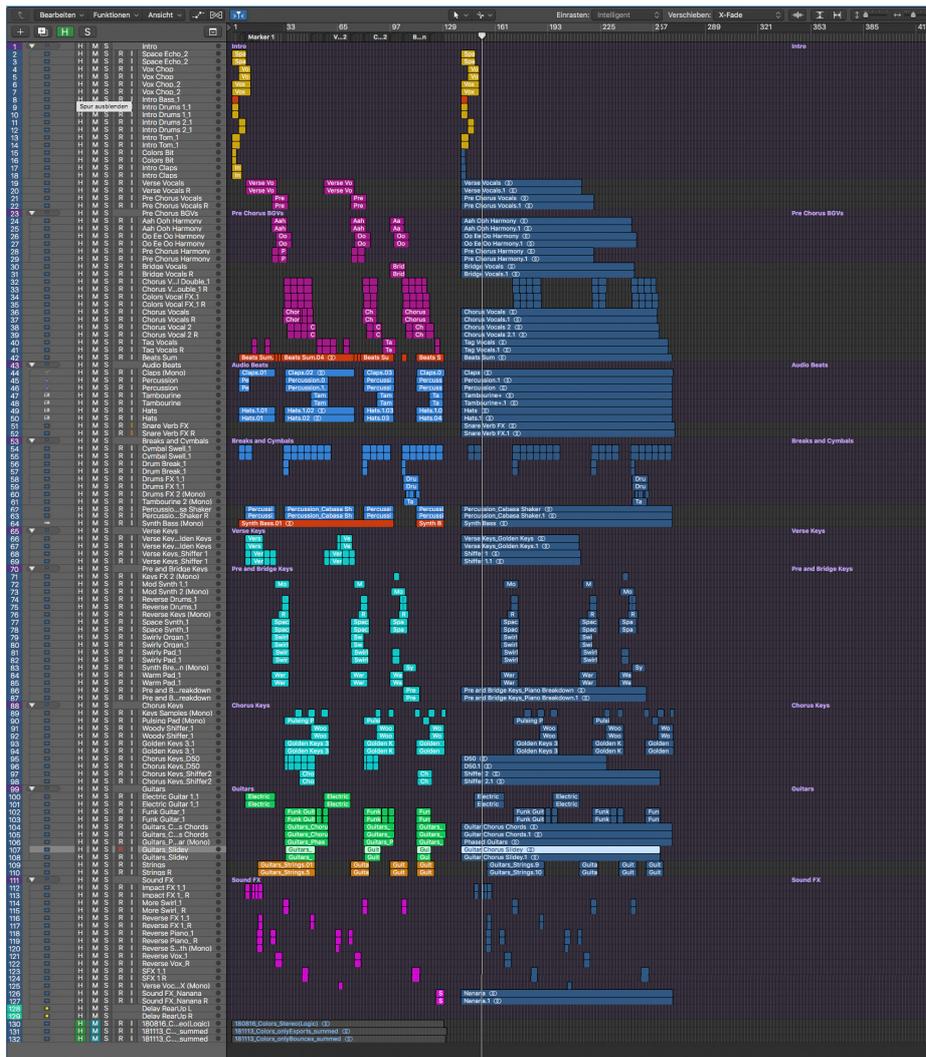


Abbildung 21: Logic Pro X Projekt "Colors" (Screenshot)

Da die drei erwähnten Produktionen allesamt eine sehr hohe Klangdichte aufweisen, wurde als vierte Hörprobe das eher minimalistische Stück „Sun“ des kanadischen Komponisten und promovierten Mathematikers Daniel Victor Snaithe ausgewählt, welcher unter seinem Künstlernamen *Caribou* sphärische Musik im Genre Electronica und Psychedelic Pop produziert. Das 2010 veröffentlichte Album „Swim“ wurde von Kritikern und Hörern positiv rezipiert und markierte Snaithe's kommerziellen Durchbruch.¹¹³ Auf diesem kombiniert er akustische Instrumente mit elektronischen Synthesizern und Samples zu einem organischen Klangbild, das auch das ausgewählte Stück „Sun“ kennzeichnet. Dieses bietet

¹¹³ vgl. All Music 2018

durch seine sphärische, schwebende Klangcharakteristik die Möglichkeit der großzügigen Verräumlichung der einzelnen Schallquellen und eine resultierende potentielle Steigerung der Einbettung des Rezipienten in die Klangwelt der Komposition. Hinsichtlich des Hörversuchs ist hierbei die Frage interessant, ob die Dichte des Klangbilds sich in einer unterschiedlichen Bewertung durch die Probanden bemerkbar macht.

5.3 Signalkette

Für das Mischen von Binauralinhalten wird im Gegensatz zur Produktion für lautsprecherbasierte Formate nur ein relativ überschaubarer Umfang an Equipment benötigt. Neben einer Musikbearbeitungssoftware und einem Kopfhörer ist vor allem die Softwarelösung zur Binauralisierung der Schallquellen für die schlussendliche Qualität der Produktion entscheidend. Von der Vielzahl an Anbietern konnte sich bislang noch keiner mit seinem Produkt als Standard durchsetzen. Daher wurde für die vorliegende Arbeit versucht, alle relevanten Softwarelösungen zu sondieren, um ein bestmögliches Ergebnis der Produktionen zu gewährleisten. Die maßgeblichen Gründe der Entscheidung werden im Folgenden kurz dargestellt.

Die Softwarefirma *facebook* bietet mit seiner *Spatial Workstation fb360* ein Werkzeug, welches eine weitgehend färbungsfreie Abbildung von Schallquellen im Raum ermöglicht. Jedoch macht sich der Fokus auf 360°-Video als Zielformat in der unübersichtlichen Darstellung der Elevation und der eingeschränkten Auswahl an Exportformaten deutlich.

Das Plugin *O3A* der Firma *Blue Ripple Sound* wurde auf Grund seiner umständlichen grafischen Oberfläche und mangelhaften Lokalisierbarkeit sowie Klangqualität nicht verwendet. Neben Artefakten machte sich auch ein Drücken auf den Ohren bei einem seitlichen Einfallswinkel von 90° bemerkbar.

Die Softwarelösung *Dolby Atmos* bietet ohne die kostspielige Rendering Unit keine Delaykompensation, was den Arbeitsprozess erschwert.

Das französische Institut *Ircam* offeriert mit der Werkzeugkollektion *Flux Spat Revolution* eine sehr realistische Räumlichkeit, eine Auswahl verschiedener HRTF sowie viele Optionen der Klangoptimierung und Ausspielung. Da es sich jedoch nicht direkt in eine Digital Audio Workstation einbinden lässt, sondern die Audiosignale per Drittsoftware wie *Dante* oder *JackPilot* übermittelt werden müssen, erwies sich der Arbeitsprozess als umständlich und instabil.

Auch der Pluginhersteller *Waves* bietet mit *WavesNx* eine Lösung zur Binauralisierung an. Auch wenn der Ansatz des mitgelieferten Headtrackers reizvoll ist, bietet es nur eine virtuelle Studioumgebung mit limitierten Einstellungsmöglichkeiten und als Exportformat lediglich First Order Ambisonics an.

Für die Produktion der Hörproben wurden schließlich zwei Systeme ausgewählt. Die Firma *New Audio Technologies* bietet mit dem *Spatial Audio Designer* neben einer sehr benutzerfreundlichen grafischen Oberfläche (s. Abbildung 22) und für den Arbeitsprozess nützlichen Funktionen wie gepaarter Bearbeitung von Stereokanälen eine überzeugend klingende Räumlichkeit sowie eine Vielzahl an Exportformaten wie bspw. MPEG-H. Durch die Zusammenführung und Prozessierung aller Spuren in der sogenannten „Mix-Engine“ kann eine recht CPU-sparsame Performance erreicht werden. Im Klang macht sich jedoch der Umstand negativ bemerkbar, dass nicht die Klangobjekte direkt sondern Lautsprecheranordnungen binauralisiert werden. Die Klangelemente werden über die virtualisierten Lautsprecher wiedergegeben, was besonders bei der Positionierung als Phantomschallquelle Mängel im Klang bewirkt. Auch die zur Auswahl stehenden Räume begünstigen zwar die Lokalisierbarkeit der Schallquellen, weisen jedoch eine färbende Klangcharakteristik auf, welche sich nicht unbedingt zum Vorteil der Klangfarbe ausprägt. Die Reflexionen können jedoch in der Software weitgehend eingedämmt werden.



Abbildung 22: Spatial Audio Designer, unten links: Ansicht von oben, unten rechts: Ansicht von hinten¹¹⁴

Als zweiter Binauralrenderer wurde das Plugin *DearVR Pro* des Softwareunternehmens *DearReality* verwendet. Dieses zeichnet eine sehr gute Lokalisation und große Auswahl an realistisch klingenden, auf algorithmischer Berechnung und Faltung basierenden 3D-Hallräumen aus, zu welchen zusätzlich Reflektionen hinzugeschaltet werden können. Die grafische Oberfläche ist dabei sehr übersichtlich aufbereitet und ermöglicht eine intuitive Bedienung (s. Abbildung 23).

¹¹⁴ New Audio Technologies 2018



Abbildung 23: Grafische Oberfläche DearVR, 1: Position, 2: Reverb, 3: Reflections, 4: Master¹¹⁵

Zu bemängeln ist der Umstand, dass in der aktuellen Version (v1.2) Stereoquellen nicht zusammen in einer Instanz bearbeitet werden können, sondern zunächst in zwei Monospuren gesplittet werden müssen, was bei größeren Projekten einen erheblichen Mehraufwand darstellt. Laut Hersteller soll die Funktion der Mehrkanalbearbeitung implementiert werden. Weiterhin sind neben den bereits existierenden Exportformaten (Binaural-)Stereo und Ambisonics bis zur dritten Ordnung weitere Optionen der Auspielung geplant.¹¹⁶ Ob in *DearVR* ein eigens gemessener HRTF-Satz erstellt oder auf bereits existierende Messungen zurückgriffen wurde, konnte auf Anfrage nicht in Erfahrung gebracht werden, da die Information unter das Firmengeheimnis fällt.

Also DAW wurde zum einen die Software *Pro Tools 11* von *Avid* verwendet. Zwar ermöglicht diese lediglich den Export von Stereo als Ausgabeformat, bietet jedoch die Option, das Projekt in der Version *Pro Tools 12 Ultimate* mit einem geringen Mehraufwand im Ambisonics-Format dritter Ordnung zu exportieren. Dieses Verfahren wurde für eine der Hörproben mit den Kodierungsplugins des Software-Herstellers *SSA aX* getestet. Da sich dabei jedoch in negativer Weise

¹¹⁵ DearVR 2018

¹¹⁶ vgl. Obst 2017

Klangverfärbungen bemerkbar machten, wurde dieser Pfad nicht intensiver verfolgt.

Da *Pro Tools 11* auf 128 simultan aktive Spuren limitiert ist und das Projekt „Colors“ diese überschritt, wurde für die Bearbeitung dieser Hörprobe das bereits erwähnte, auf 256 Spuren beschränkte *Logic Pro X* verwendet.

Die Produktionen wurden mit dem geschlossenen, diffusfeldentzerrten Studiokopfhörer *Beyerdynamic dt 770 Pro* durchgeführt, welcher einen ausgewogenen Frequenzgang sowie eine präzise Abbildung bietet. Als Zweitkopfhörer wurde mit dem für den Konsumentenmarkt konzipierten, etwas bass- und höhenbetonenen Modell *Sony WH-1000XM2* das Klangbild abgeglichen. Als Audiointerface diente dabei das mit hochwertigen *Burr Brown* AD/DA-Wandlern ausgestattete Gerät *Audient iD14*.

5.4 Gestalterische Umsetzung

Für den Bearbeitungsprozess der Mischung existieren in der Musikproduktion keine direkten, festgeschriebenen Vorgaben und Regeln, sondern eher Konventionen, welche das Ziel der Balance und Klarheit auf verschiedenen Wiedergabemedien verfolgen. Verschiedene Mittel der Klangbearbeitung finden hierbei Verwendung. So können die einzelnen Elemente hinsichtlich ihrer Lautstärke, Dynamik, Frequenz, Räumlichkeit sowie ihrer Platzierung im Panorama zu einem einheitlichen Klangbild geformt werden. Während die jahrzehntelange Tradition der Mischung von Stereoinhalten bereits einen großen Umfang an Richtlinien und Konventionen hervorgebracht hat, befindet sich 3D-Musik noch in ihrem Anfangsstadium und bietet viel Raum für die Exploration neuer Mischansätze. Vor allem bei klassischen Produktionen steht dabei oft das Ziel der möglichst naturgetreuen Abbildung einer Konzertsituation im Mittelpunkt, die dem Hörer einen Besuch dieser simulieren soll. Im vorliegenden Anwendungsfall der Popmusik liegt der Fokus unabhängig vom Endformat tendenziell eher auf einer intensiveren emotionalen Reaktion des Rezipienten. Das 3D-Format kann

diese wie erwähnt durch das Potential der höheren Immersion in die Klangwelt des Musikstücks begünstigen.¹¹⁷ Um im Hörversuch eine Vergleichbarkeit der beiden Formate zu gewährleisten, wurde auf weitere Werkzeuge der Bearbeitung, abgesehen von der 3D-Panorama-Veränderung und dem Einsatz von 3D-Hallräumen, verzichtet.

Grundsätzlich wurden bei der Bearbeitung der Mischungen einige Prinzipien ausgemacht, die ein qualitatives Endergebnis begünstigen. Ähnlich wie bei der Stereomischung empfiehlt es sich, eine feste Basis an Schallquellen zu haben, die der Mischung ein stabiles Grundgerüst verleihen. Dementsprechend wurden Gesang, Schlagzeug, Bass sowie tragende Instrumente eher mittig und vor der Hörerin positioniert. Wie erwähnt gilt für den vorliegenden Anwendungsfall die Vorgabe, die Hörgewohnheiten der Probanden nicht zu missachten und irritationsbedingte Verzerrungen in der Bewertung zu riskieren. Im direkten Vergleich wurde weiterhin deutlich, dass tieffrequente Klangquellen wie Bassdrum und Bass durch die Binauralisierung ihre Kompaktheit verlieren und für ein druckvolles Klangerlebnis nicht verräumlicht werden sollten. Die erwähnte Im-Kopf-Lokalisation kann an dieser Stelle einen klanglichen Vorteil darstellen. Aufbauend auf der dieserart konstruierten Basis konnten weitere Objekte frei im virtuellen Raum verteilt werden. Auch wenn dies einen weitgehend vom subjektiven Geschmack des Produzenten abhängigen Entscheidungsprozess darstellt, wurden auch hier einige Richtlinien ausgemacht. Tieffrequente Klänge wurden eher auf dem Elevationslevel des Rezipienten positioniert, da sie über dem Hörer eher bedrückend wirkten. Dort ließen sich hochfrequente Klänge hingegen mit positivem Ergebnis platzieren und lokalisieren. Weiterhin wirkte sich die HRTF-bedingte Klangfärbung der Objekte abhängig von ihrer Positionierung im Raum teilweise negativ auf die Klangfarbe aus. Somit musste hier ein Kompromiss zwischen künstlerischer Intention in der Positionierung der einzelnen

¹¹⁷ vgl. Hahn 2018

Elemente und einem ausgewogenen Klangbild gefunden werden. Ausgehend von einer in sich bereits stimmigen statischen Mischung konnten dann einzelne Objekte in ihrer Bewegung automatisiert werden. Analog zum beschriebenen Headtracking können sich bewegende Klangelemente der schärferen Lokalisation zuträglich sein. Bewegen sich jedoch zu viele Elemente gleichzeitig im Raum, verschlechtert sich in der Konsequenz die Lokalisationsschärfe, da einzelne Bewegungswege nicht mehr klar verfolgt werden können. Bei der Arbeit mit dem *Spatial Audio Designer* wurde zudem registriert, dass statische Schallquellen tendenziell bei den virtuellen Lautsprechern positioniert werden sollten, um Unschärfefeffekte der interpolierten Phantomschallquellen zu vermeiden (s. Abbildung 24).

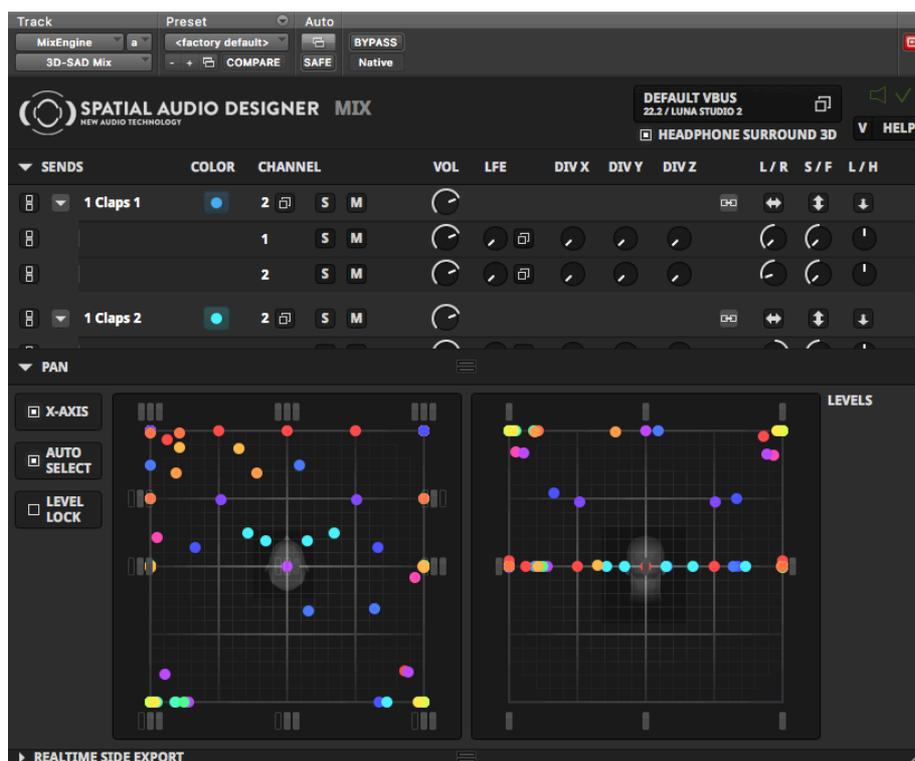


Abbildung 24: Anordnung der Schallquellen im Spatial Audio Designer (Screenshot)

Bei der Verwendung der in *DearVR Pro* integrierten „Bass Boost“-Funktion, welche die tiefen Frequenzen des Eingangssignals anhebt, galt es zu bedenken, dass dies die Schärfe der räumlichen Ortung etwas beeinträchtigt. Sie wurde daher lediglich eingesetzt, falls sich daraus ein klanglicher Vorteil für die Gesamtmischung ergab. Die Signalkette des Arbeitsprozesses mit *DearVR Pro* ist

in Abbildung 25 dargestellt. Dem gegenüber demonstriert Abbildung 26 eine Übersicht des Signalflusses der Binauralproduktion mit dem SAD.

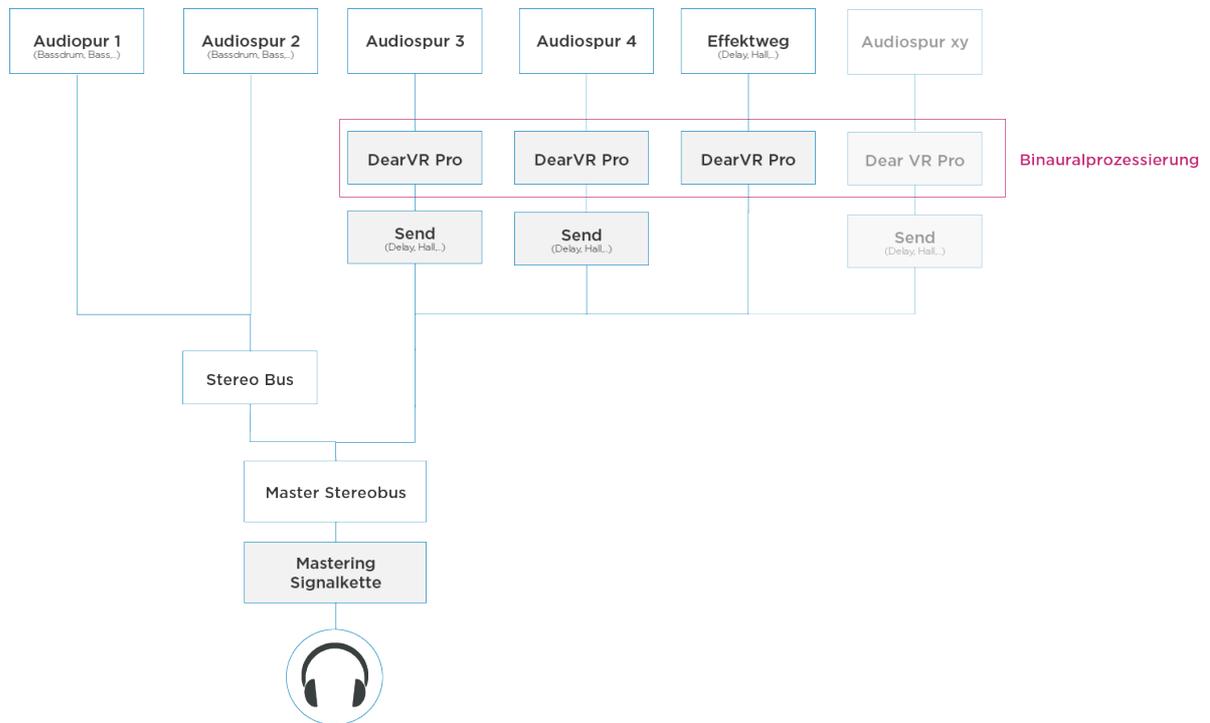


Abbildung 25: Signalkette der Binauralproduktion mit DearVR Pro (eigene Darstellung)

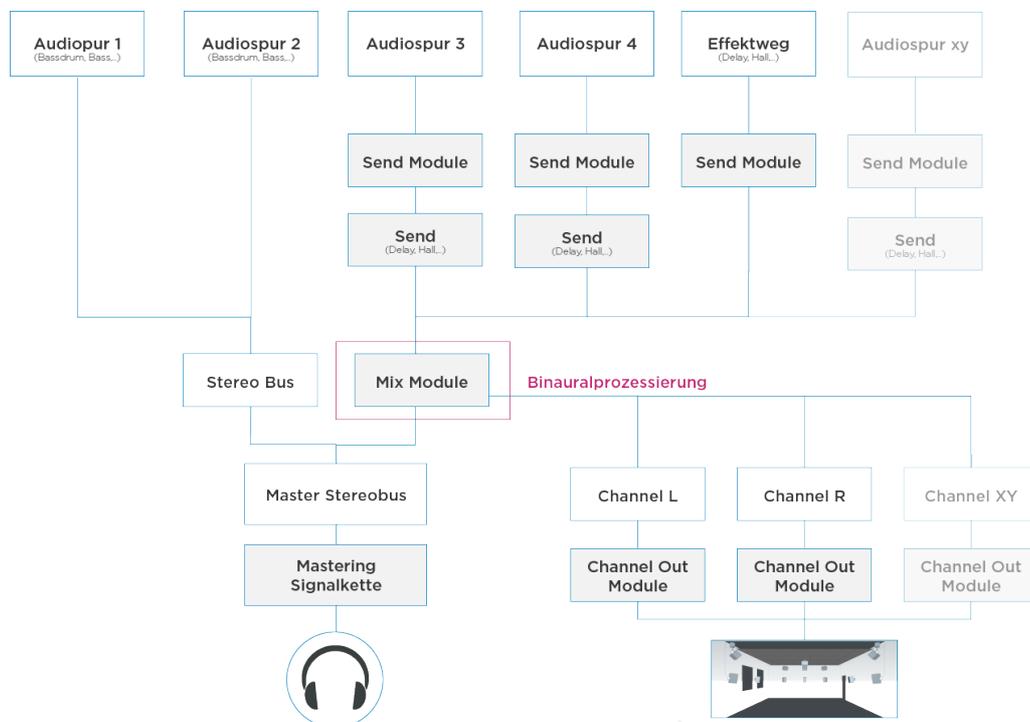


Abbildung 26: Signalkette der Binauralproduktion mit dem SAD (eigene Darstellung)

Weiterhin sollte ähnlich wie bei der Mischung für Surroundformate eine ausgewogene Balance zwischen dem Signal der vorne und auf Elevationsebene des Hörers platzierten Objekte sowie dem Signal der im hinteren und oberen Schallfeld angelegten Elemente angestrebt werden, da sonst das Klangbild in eine Richtung zu kippen droht und in der Folge ein irritierender Höreindruck entsteht. Bevor im folgenden Abschnitt einzelne Produktionstechniken etwas näher beschrieben werden, sei noch erwähnt, dass bei der Verwendung der in der Software *DearVR Pro* integrierten 3D-Hallräume von drastischen klangverändernden Maßnahmen abgesehen wurde, um die Vergleichbarkeit der Hörproben zu gewährleisten. Gesetzt den Fall, dass das zu Grunde liegende Klangobjekt bereits durch eine sehr räumliche Klangcharakteristik gekennzeichnet war, wurde eine dazu passende Einstellung des Effekts ausgewählt und in einem adäquaten Verhältnis hinzu gemischt.

5.5 Gestalterische Stilmittel

Wie bereits erwähnt wurden bei der Mischung der Binauralversionen Konventionen der Stereoproduktion berücksichtigt. Ein beliebtes Stilmittel, um einzelne Abschnitte eines Songs besser zur Geltung kommen zu lassen, ist hierbei die Kontrastierung. Der Refrain eines Musikstücks kann eine größere Wirkung erzielen, wenn ihm eine etwas dezentere und im Panorama eher schmal angelegte Passage vorausgeht. Dieses Stilmittel wurde bei der Gestaltung der Hörprobe „Colors“ eingesetzt, indem die Objekte der Strophen eher kompakt und mittig vor der Hörerin platziert sind. Im Pre-Chorus öffnet sich dann das Panorama durch die weitläufige Anordnung einiger sphärischer Objekte im Raum. Dieses Prinzip der räumlichen Dynamik wurde weiterhin im Stück „Helena Beat“ eingesetzt. Das Intro ist hier zunächst sehr schmal gestaltet. Mit Einsetzen des Lacheffektes (00:15 min) wird das Klangbild kurz sehr weitläufig, verkleinert sich jedoch zu einer kleinen Schallquelle zwischen den Ohren und lässt somit den Eindruck der einsetzenden Instrumente ab 00:22 min umso eindrucksvoller

erscheinen. Dieser Effekt wird auch in der gleichen Hörprobe beim recht reduzierten, räumlich kompakt gehaltenen dritten Refrain deutlich (03:24 min), welchem ein sehr großräumig angelegter C-Teil vorausgeht und der dann mit seiner Wiederholung seine volle räumliche Größe zurückerlangt (03:38 min).

Abhängig vom Klangcharakter eines Objekts lassen sich unterschiedliche Effekte erzielen. Sphärische, flächige Klänge können durch Hinzufügen eines 3D-Halleffekts mit langer Hallfahne und einer weiten Verteilung um den Hörer herum dessen Gefühl der Immersion potentiell steigern. Als Beispiel ist hier der C-Teil des Stücks „Colors“ zu nennen (03:12 min). Perkussive Klänge hingegen lassen sich durch ihre definierten Transienten etwas leichter lokalisieren und eignen sich, um interessante Akzente zu setzen und die Räumlichkeit zu verdeutlichen. So wechseln die Klatsch-Sounds bei „Helena Beat“ während der Strophe links und rechts im Wechsel hin und her (00:59 min). Auch die Unterstreichung der semantischen Ebene der Komposition bietet sich teilweise an. Das Stück „In Your Beat“ setzt direkt mit einem perkussiven Synthesizer-Klang ein, der in seiner Tonhöhe im langsamen Wechsel hoch- und runtermoduliert wird. Diese Auf- und Ab-Bewegung wird durch die Automation der Elevation in kreisender Manier um die Hörerin herum entsprechend gespiegelt. Als weiteres Beispiel ist der Refrain von „Colors“ zu nennen, dessen Gesang ein Konglomerat aus klanglich differenziert ausgestalteten Einwüfen darstellt. Um diesen Eindruck der Extravaganz zu unterstreichen, sind die einzelnen Vokal-Einwürfe in verschiedenen Positionen angeordnet (s. Abbildung 27).

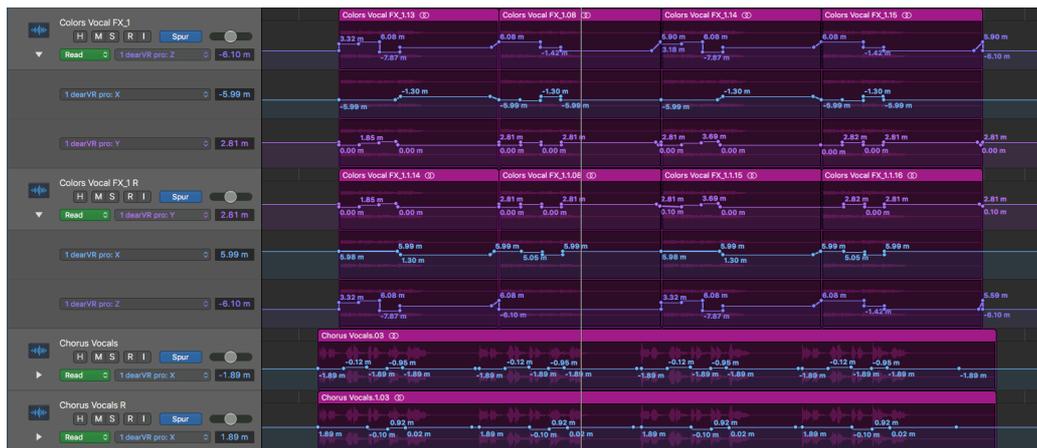


Abbildung 27: Automation der Vokal-Einwürfe im Refrain von "Colors" (Screenshot)

Bei der Produktion von „Sun“ wurde auf Grund dessen minimalistischer Zusammensetzung auf schnelle, effektvolle Automationen weitgehend verzichtet und eher der Ansatz einer stimmigen, ansprechenden Räumlichkeit verfolgt, welche die Hörer leichte in das sphärische Klangbild eintauchen lässt. Dafür wurde unter anderem eine eigene Aux-Spur mit einem Echoeffekt angelegt, räumlich hinten oben positioniert und einigen Signalen wie dem Gesangssample in homöopathischer Dosis hinzugemischt. Obgleich der Effekt nur sehr dezent zu vernehmen ist, trägt er zu einer weiteren, immersiveren Räumlichkeit bei.

5.6 Prozessierung der Binauralsignale

Da das Grundprinzip binauraler Signale auf der frequentiellen Filterung der Eingangssignale aufbaut, sind jegliche Eingriffe wie Veränderung des Frequenzgangs oder der Phasen einzelner Frequenzbänder sowie Dynamikeinengungen mit Vorsicht durchzuführen, da derartige Prozessierungen den dreidimensionalen Höreindruck zerstören können.¹¹⁸ Dieser Umstand steht in Konflikt mit dem in der Musikproduktion elementaren Arbeitsschritt des Masterings, bei welchem die Audioproduktion für die Veröffentlichung aufbereitet und klanglich optimiert wird.¹¹⁹ In einer privaten, unveröffentlichten

¹¹⁸ vgl. Theile, Wittek und Reisinger 2002

¹¹⁹ vgl. Dickreiter, et al. 2014

Quelle beschreibt der Leiter des Bereichs Akustik an der Universität Sorbonne Brian F.G. Katz einen Fall, bei dem eine kommerzielle Musikproduktion zusätzlich zur Stereoverision auch im Binaural-3D-Format veröffentlicht werden sollte. Da das Mastering von einem bezüglich Binauralsynthese unerfahrenen Ingenieur durchgeführt wurde und das räumliche Klangbild zerstörte, wurde schließlich nur die Stereoverision veröffentlicht.

Neben der Masteringprozessierung ist in der Musikproduktion die summierte Bearbeitung mehrerer zusammengefasster Spuren ein beliebtes Stilmittel. Vor allem bei der Bearbeitung einer Schlagzeugaufnahme durch einen Summenkompressor engt diese nicht nur die Dynamik ein, sondern kann auch eine erhebliche Klangveränderung zur Folge haben. Da diese Bearbeitung aus genannten Gründen nicht nach der Binauralisierung geschehen sollte, muss beim vorherigen Export der Einzelspuren aus dem Stereoprojekt darauf geachtet werden, dass diese die entsprechenden klanglichen Einflüsse der Summenprozessierung beinhalten. Dabei reicht es nicht, jede Spur einzeln zu exportieren, da in diesem Fall der Einfluss aller anderen in den Summenkompressor sendenden Signale nicht im zu exportierenden Einzelsignal implementiert ist. Hierfür müssen die entsprechenden Spuren den Summenkompressor per „pre-fade“-geschalteten Keychain-Sendewegen steuern. Dies ermöglicht den Export jeder einzelnen Spur inklusive der klanglichen Einflussnahme der Summenprozessierung bei Stummschaltung der restlichen Spuren. Auf diesem Weg lässt sich das Problem der Prozessierung binauraler Signale zumindest auf dieser Ebene umgehen. Dies gilt allerdings nicht für das Mastering, welches der finale Arbeitsschritt der Klangoptimierung bleiben muss. Dieser Bearbeitungsschritt wurde für alle vier Hörproben von der erfahrenen Mastering-Ingenieurin Magdalena Piotrowska durchgeführt. Dabei wurden sowohl die Stereo- als auch die Binauralversionen durch die gleiche Signalkette prozessiert. Diese bestand dabei aus den folgenden Bestandteilen:

- digitaler EQ (Linearer Breitband EQ und Linearer Tiefband EQ der Firma Waves)
- D/A-Konvertierung
- Analoge Kette: Kompressor zur Gain-Einstellung, EQ (*Pultec Design*), Kompressor
- A/D-Konvertierung
- Digitales Limiting

Es wurde bei der Auswahl der Signalketten-Bausteine darauf geachtet, keine Werkzeuge zu benutzen, welche die Phasenverhältnisse des Signals verzerren würden. Somit wurde auf Multiband-Kompressoren sowie non-lineare Phasen-EQs verzichtet. Bei der Bearbeitung wurde das Ziel einer klanglichen Kohärenz zwischen den Stereo- und Binauralversionen verfolgt, sodass der Hauptunterschied in der räumlichen Dimension bestand. Dieses Ziel konnte zum Großteil erreicht werden, auch wenn die binaurale Prozessierung zum Teil hörbare Klangverfärbungen vor allem in den hohen Mitten- und den hohen Frequenzbändern verursachte.

5.7 Anwendung psychoakustischer Erkenntnisse

In Abschnitt 2.7 wurde bereits der Einfluss der Blauertschen Bänder auf die Lokalisation eines Schallereignisses erläutert. Die Positionierung der einzelnen Objekte bei der Produktion der Hörversuchsstimuli berücksichtigte diese und weitere Faktoren, welche nach Blauert die Lokalisationsschärfe beeinflussen. So ist für die Differenzierung zwischen vorne und hinten laut Blauert eine Breitbandigkeit des Signals in den Höhen bis zu einer unteren Grenzfrequenz zwischen 2 und 8 kHz nötig. Sobald das zu lokalisierende Signal Frequenzanteile oberhalb von 2 kHz aufweist, verbessert sich die Genauigkeit der Ortung.¹²⁰ Für

¹²⁰ vgl. Blauert, Untersuchungen zum Richtungshören in der Medianebene bei fixiertem Kopf 1969

die Lokalisationsunschärfe in der seitlichen Veränderung der Position nennt Blauert einen Winkel von 4° . In der Medianebene fällt die Unschärfe jedoch mit ca. 10° deutlich höher aus.¹²¹ Untersuchungen von Roffler und Butler haben gezeigt, dass dieser Wert verbessert werden kann, wenn im Signal Frequenzen oberhalb von 7 kHz vorhanden sind.¹²² Diese Feststellung deckt sich mit dem von Blauert für die Höhenwahrnehmung als notwendig identifizierten Frequenzband von 8 kHz.¹²³ Dies wurde bei der Produktion der Stimuli berücksichtigt, insofern vorrangig Objekte mit einem entsprechenden Frequenzanteil in der Höherebene positioniert wurden.

¹²¹ vgl. Dickreiter, et al. 2014

¹²² vgl. Roffler und Butler 1968

¹²³ vgl. Blauert, Räumliches Hören 2014

6 Hörversuch

Während sich ein Großteil bisheriger Studien eher auf technische Aspekte von 3D-Audio bzw. des Binauralformats bezieht wie beispielsweise Umhüllung, Transparenz oder Einbettung,¹²⁴ bedarf es bedingt durch die steigende Marktrelevanz von binauralen Inhalten einer intensiveren Evaluation der Qualität des Hörerlebnisses für den Nutzer. Dieses ist definiert durch das Maß an Gefälligkeit, das der Rezipient bei der Nutzung einer Applikation oder eines Mediums empfindet. Es bildet sich aus der Erwartungserfüllung sowie der persönlichen Präferenzen und momentanen Befindlichkeit der Nutzerin heraus.¹²⁵ Der Fokus der zugrunde liegenden Studie liegt somit auf der Evaluation der Qualität des subjektiven Hörerlebnisses für die Rezipienten.

6.1 Ansatz

Das Hören von Musik kann erwiesenermaßen Emotionen beim Rezipienten induzieren.¹²⁶ Die Verstärkung oder Veränderung des emotionalen Zustands wird dabei vom Hörer als wichtiger eingestuft als die Klangqualität, wie in einigen Studien gezeigt werden konnte.^{127 128 129} Dabei stellt die Klangfarbe eine der zentralen Kriterien zur Bewertung von Klangqualität dar und der Nutzer ist eher bereit einen Kompromiss bei der räumlichen Auflösung als bei der Klangfarbe einzugehen, wie in einer Studie belegt wurde.¹³⁰ Die Binauralsynthese beinhaltet eine Überschneidung dieser Faktoren. In erster Linie ermöglicht sie eine dreidimensionale Räumlichkeit und Verortung der Schallquellen, was laut der in

¹²⁴ vgl. Theile und Wittek 2011

¹²⁵ vgl. Le Callet, Möller und Perkis 2013

¹²⁶ vgl. Schmidt-Atzert 1982

¹²⁷ vgl. Krumhansl 2002

¹²⁸ vgl. Roe 1985

¹²⁹ vgl. Kamalzadeh, Baur und Möller 2012

¹³⁰ vgl. Rumsey, et al. 2005

Kapitel 4.3 vorgestellten Studie einen positiven Einfluss auf die emotionale Reaktion des Rezipienten bewirkt. Das Grundprinzip der Binauraltechnik beruht dabei jedoch auf frequentieller Filterung der Signale und hat somit inhärent Klangfärbungen zur Folge. Daher stellt diese Studie die Relation der Parameter „allgemeines subjektives Klangerlebnis“, „Klangfarbe“ und „Räumlichkeit“ in den Mittelpunkt ihrer Untersuchung mit einem besonderen Fokus auf erstgenanntem Merkmal. Dieses ist unter seinem englischen Äquivalent „Overall Listening Experience“ ein erprobtes gefühlsbezogenes Messinstrument für die Evaluierung der wahrgenommenen Qualität von Audioinhalten und wird in dieser Arbeit zu Gunsten der Übersichtlichkeit mit „OLE“ abgekürzt.

6.2 Versuchsaufbau

Für den Versuch wurden wie in Kapitel 5 beschrieben vier verschiedene Items in jeweils zwei Versionen (Stereo, Binaural-3D) produziert und von einer erfahrenen Mastering-Ingenieurin aufeinander abgestimmt sowie in ihrer Lautstärke aneinander angepasst. Es wurde dann für jedes Item eine 60-sekündige Sequenz ausgewählt, welche das Stück in aussagekräftiger Weise repräsentiert und zumindest eine Strophe und Refrain beinhaltet. Alle Stimuli wurden im Format 48 kHz / 16 Bit WAV exportiert und mit einem 1s-Fade versehen.

Die Studie wurde vordergründig als webbasierter Hörversuch konzipiert. Dabei wurde die Plattform *unipark.com* verwendet. Der webbasierte Ansatz erschien sinnvoll, um die tatsächliche Breite der realistischen Nutzer-Abhörbedingungen darzustellen und ein möglich aussagekräftiges Ergebnis zu erhalten, das den alltäglichen Musikkonsum widerspiegelt und eine höhere externe Validität innehat. Zur Klärung des Einflusses der Abhörbedingungen auf das Testergebnis wurde der Hörversuch zudem unter kontrollierten Bedingungen durchgeführt. Zu diesen gehörte die Verwendung eines akustisch abgeschirmten Raums und die gleiche Signalkette, die auch zur Produktion der Hörproben verwendet wurde (s. Kapitel 5.3). Weiterhin wurden die Versuchsitens in unkomprimierter Fassung

verwendet, während den Probanden der Webstudie AAC-komprimierte Dateien mit einer Übertragungsrate von 320 kB/s präsentiert wurden. Da der Anbieter *Unipark* keine Implementation unkomprimierter Dateien erlaubt, wurde für die Laborstudie der Hörversuch zunächst in der Software *Max MSP* nachgebaut (s. Anhang). Um eine einheitliche Formatierung der Daten zu erhalten, wurde diese Entscheidung jedoch verworfen und stattdessen die Oberfläche des webbasierten Versuchs in Kombination mit einer Pro Tools-Session, in der die unkomprimierten Fassungen angelegt waren, verwendet.

Die produzierten Musikstücke wurden den Versuchssitemen wie folgt zugeordnet:

- Familiarisierung: „In Your Beat“ (Django Django)
- Hörprobe 1: „Helena Beat“ (Foster The People)
- Hörprobe 2: „Colors“ (Beck)
- Hörprobe 3: „Sun“ (Caribou)

Zur Bewertung der Merkmale „Allgemeines Klangerlebnis“, „Klangfarbe“ und „Räumlichkeit“ wurde eine fünfstufige Likert-Skala verwendet.

6.3 Versuchsablauf

Auf der Startseite des Hörversuchs findet sich zunächst eine kurze Erklärung des Hörversuchs und der Hinweis, dass neben einem intakten Hörvermögen das Tragen von Kopfhörern Voraussetzung für die Teilnahme ist. Weiterhin werden die Vpn mit folgenden Anweisungen instruiert:

In diesem Hörversuch werden Sie verschiedene Musikausschnitte präsentiert bekommen. Für jeden Ausschnitt werden Sie unter anderem Ihr subjektives Hörerlebnis auf einer simplen Skala bewerten. Berücksichtigen Sie dabei bitte alle Faktoren, die auch Ihr Musikhörverhalten im Alltag beeinflussen (persönlicher Geschmack, Audioqualität etc.).

Nachdem auf Seite 2 Alter und Kopfhörermodell abgefragt werden, wird den Vpn auf Seite 3 eine Hörprobe zur Familiarisierung präsentiert. Diese erlaubt der Probandin, sich mit der Materie und dem Versuchsschema vertraut zu machen sowie einen angemessenen Abhörpegel einzustellen, welcher für den Rest des Hörversuchs laut Hinweis nicht geändert werden soll. Auf den folgenden drei Seiten werden jeweils in einem 60-sekündigen Videoclip beide Versionen der Hörprobe im AB-Verfahren abwechselnd in einem Turnus von zehn Sekunden mit der jeweiligen Titeleinblendung „Version A“ bzw. „Version B“ präsentiert. Dabei ist die Zuordnung des verwendeten Formats der Hörprobe zur Versionsbezeichnung A oder B nicht konstant. Auf einer fünfstufigen Likert-Skala wird anschließend markiert, welche Version bezüglich des subjektiven allgemeinen Hörerlebnisses bevorzugt wird. Die möglichen Antwortmöglichkeiten lauten dabei „Ganz klar A“, „Eher A“, „Neutral“, „Eher B“, „Ganz klar B“. Es wird hier der Direktvergleich gewählt, da die Unterschiede zwischen den Versionen teilweise recht gering ausfallen. In einem Kommentarfeld kann die Entscheidung jeweils optional begründet werden (s. Abbildung 28). Wurden die drei Items auf diese Weise bewertet, werden sie in einer zweiten Runde jeweils einzeln bezüglich der Merkmale „Allgemeines Klangerlebnis“, „Klangfarbe / Timbre“ und „Räumlichkeit“ bewertet. Die Likert-Skala besteht dabei aus den Merkmalsausprägungen „Sehr schlecht“, „Schlecht“, „Neutral“, „Gut“, „Sehr gut“. Auf diese Weise können beide Versionen ohne automatisierten A/B-Wechsel angehört werden. Zudem ermöglicht dieses Verfahren den Abgleich der Ergebnisse des Merkmals „Allgemeines Klangerlebnis“ aus beiden Runden sowie den Ausschluss unstetiger Eingaben (s. Abbildung 29). In beiden Runden verhindert weiterhin die unstetige Zuordnung des Formats zu Version A oder B die Bewertung des Klangerlebnisses auf Grund erfahrungsbedingter Annahmen.

Im folgenden Videoclip werden zwei Versionen desselben Musikstücks im Wechsel abgespielt. Bitte tragen Sie auf dem darunter positionierten Schieberegler ein, welche Version Sie bevorzugen.



Welche der beiden Versionen bevorzugen Sie bezüglich Ihres allgemeinen Hörerlebnisses?



Kommentarfeld (Optional)

Welche Unterschiede konnten Sie zwischen den beiden Versionen feststellen? Weshalb haben Sie sich für die jeweilige Version entschieden?

An empty rectangular text input field with a thin gray border and a small cursor icon at the bottom right corner.

Abbildung 28: A/B-Direktvergleich des Items (Screenshot)

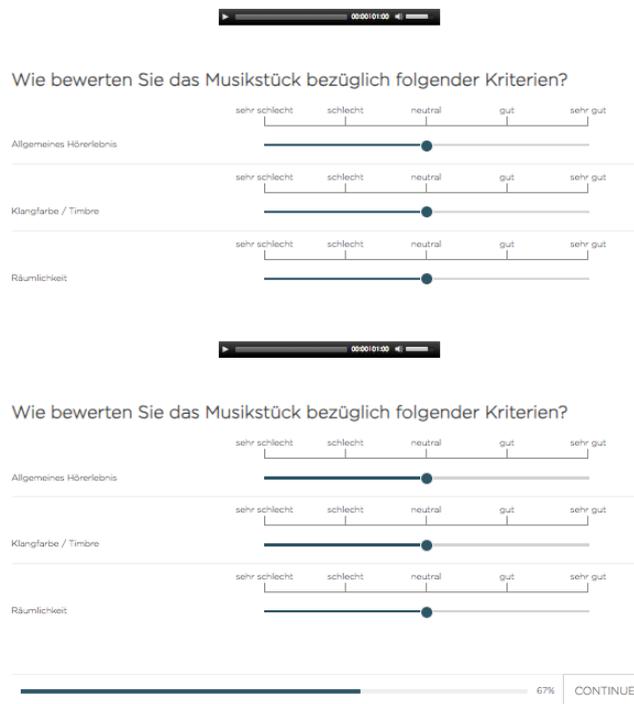


Abbildung 29: Einzelbewertung des Items (Screenshot)

6.4 Qualitative Untersuchung der Transferierbarkeit zweier 3D-Audio-Mischungen

Der Erfolg von 3D-Audio hängt, unter anderem aus Gründen der Wirtschaftlichkeit, maßgeblich von der Übertragbarkeit der 3D-Mischung auf verschiedene Wiedergabesysteme ab. Da der *Spatial Audio Designer* die Auswahl eines flexiblen Ausgabeformats ermöglicht, konnten die zwei mit dieser Software produzierten Hörproben auf diesen Aspekt hin evaluiert werden (s. Anhang 4 Daten-DVD). Dafür wurden die 3D-Versionen auf dem Auro-3D 13.1-System der Abteilung Akustik des *Fraunhofer-Institut für Bauphysik* (s. Abbildung 30) von einer kleinen Gruppe Tonexperten verglichen und auf Unterschiede hin evaluiert. Die Ergebnisse dieser qualitativen Untersuchung und der in den vorherigen Abschnitten erläuterten quantitativen Studie werden im folgenden Kapitel präsentiert.

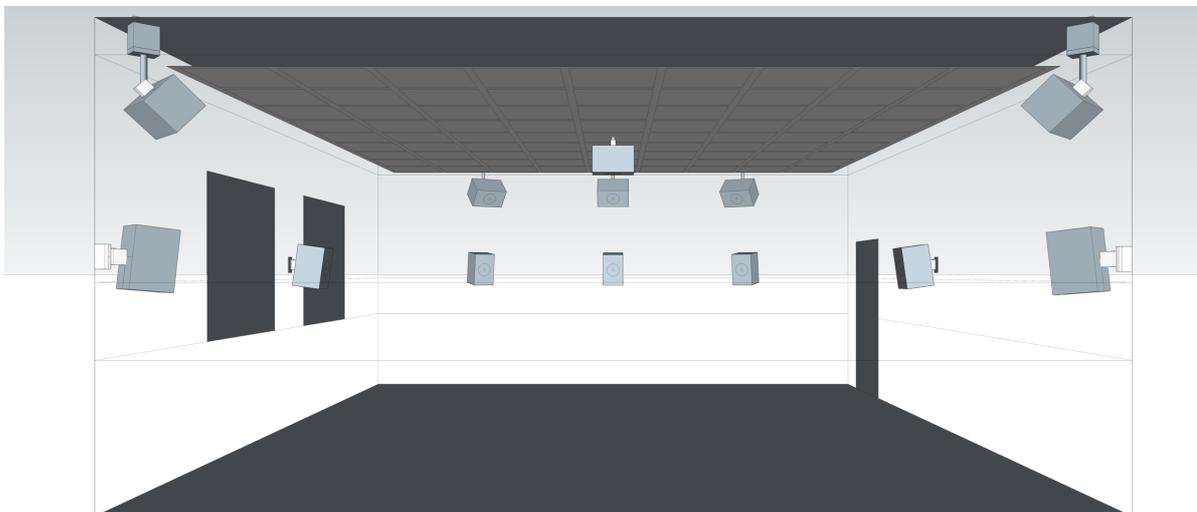


Abbildung 30: 3D-Audio-Labor des *Fraunhofer-Institut für Bauphysik* (CAD-Skizze)¹³¹

¹³¹ Fraunhofer-Institut für Bauphysik 2018

7 Auswertung

7.1 Analysemethode

Zur Auswertung der Daten werden in erster Linie gestapelte Säulendiagramme verwendet, welche eine übersichtliche Darstellung der relativen Verteilung aller Bewertungen bieten. Zusätzlich wird für jeden Parameter das jeweilige arithmetische Mittel auf der Säule markiert, sodass eine schnelle Einordnung der durchschnittlichen Bewertung ermöglicht wird. Ausgehend von den gestapelten Säulendiagrammen werden Verhältnismäßigkeiten abgeleitet, welche anhand des Wilcoxon-Vorzeichen-Rang-Tests auf Signifikanz untersucht werden ($p < .05$). Bei diesem handelt es sich um einen nichtparametrischen statistischen Test, welcher die Prüfung der Gleichheit der zentralen Tendenzen der Grundgesamtheiten anhand jeweils zwei gepaarter Stichproben ermöglicht. Die Ergebnisse dieser Untersuchung werden, wo sinnvoll, in Klammern hinter der jeweiligen Hypothese vermerkt.

7.2 Versuchspersonen

Insgesamt nahmen 217 Personen am Hörversuch teil, wovon 44,24 % diesen vollendeten. Dies entspricht einer Stichprobe von $n=96$ Vpn. Davon nahmen 23 Probanden unter Laborbedingungen und 73 Personen an der online-basierten Studie teil.

Laut einer Untersuchung des Einflusses des Alters auf die Fähigkeit der Lokalisierung wird diese mit zunehmendem Alter unpräziser.¹³² Daher wurde das Alter der Probanden abgefragt, dessen arithmetisches Mittel bei $\bar{x}_{\text{Alter}}=26,98$ Jahren liegt. Da lediglich 14 der 96 Probanden ein Alter von >30 Jahren angaben, wurde die Untersuchung des Einflusses des Alters auf die Bewertung der Hörproben aus Gründen der ungenügend großen Stichprobe nicht weiterverfolgt.

¹³² vgl. Dobрева, O'Neill und Paige 2011

Die Daten wurden vor der Analyse von inkonstanten Eingaben bereinigt. Für den Fall, dass sich die OLE-Bewertung eines Items in der ersten Runde als widersprüchlich zu der OLE-Bewertung des gleichen Items in der zweiten Runde zeigte, wurden diese Daten von der weiteren Auswertung ausgeschlossen. Bedingung war hierbei, dass der Unterschied zwischen den beiden Bewertungen mindestens 3 Abstufungen betrug. Auf dieser Grundlage wurden für die erste Hörprobe zehn, für die zweite Hörprobe zwei und für die dritte Hörprobe vier Bewertungen von der weiteren Untersuchung ausgeschlossen. Die niedrige Konstanz der OLE-Bewertungen der ersten Hörprobe könnte zum einen mit dem Lerneffekt, sich auf kleine Unterschiede im Laufe des Versuchs besser fokussieren zu können, zusammenhängen. Möglich ist auch, dass der Unterschied durch die Wiedergabe in der zweiten Runde ohne A/B-Wechsel bedingt ist, da den Vpn hier dreißig weitere Sekunden der Hörprobe präsentiert wurden.

7.3 Ergebnisse

Das Diagramm „Präferenz OLE“ (s. Abbildung 31) zeigt die Bewertung der OLE der drei Hörproben im A/B-Direktvergleich. Je höher hierbei das arithmetische Mittel (grüner Punkt), desto stärker wird die Stereoversion präferiert. Über die drei Hörproben gemittelt liegt das arithmetische Mittel bei $\bar{x}_{\text{Hörprobe1-3,gemittelt}}=3,16$ und damit leicht über dem Wert „neutral“. Diese leichte Tendenz zur Präferenz der Stereoversion zeigt sich auch in den relativen Häufigkeiten. So gaben $f_{4,5(\text{Hörprobe1-3gemittelt})}=47,96$ % an, die Stereoversion „eher“ oder „ganz klar“ zu bevorzugen. Umgekehrt liegt dieser Wert für die Binauralversion bei $f_{1,2(\text{Hörprobe1-3gemittelt})}=37,92$ %. Dementsprechend gaben $f_{3(\text{Hörprobe1-3gemittelt})}=14,06$ % der Probanden an, keine Präferenz für die binaurale oder Stereo-Version zu empfinden.

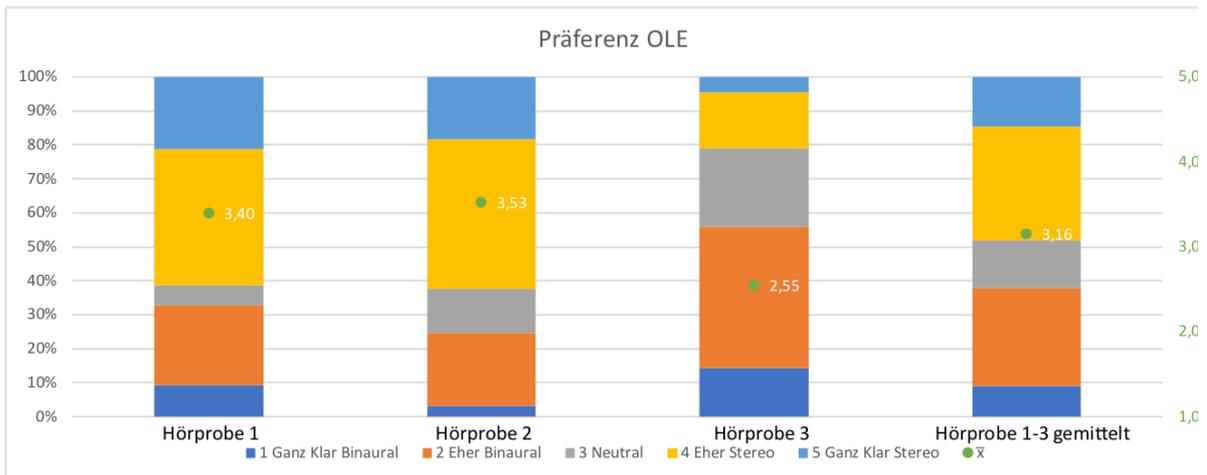


Abbildung 31: Präferenz OLE

Während der gemittelte Unterschied hier noch recht dezent ausfällt, zeigen sich bei individueller Betrachtung der Versuchssitems deutlichere Tendenzen. Für Hörprobe 2 fällt die Präferenz der Stereoersion mit den Werten $f_{4,5(\text{Hörprobe}2)}=62,37\%$, $f_{1,2(\text{Hörprobe}2)}=24,73\%$ und einem arithmetischen Mittel von $\bar{x}_{\text{Hörprobe}2}=3,53$ am deutlichsten aus. Während diese Tendenz bei Hörprobe 1 ähnlich angesiedelt ist, zeigt sich bei Hörprobe 3 eine stärkere Präferenz für die Binauralversion. Diese wurde von $f_{1,2(\text{Hörprobe}3)}=56,04\%$ der Probanden als leicht bzw. deutlich besser eingestuft. Lediglich $f_{4,5(\text{Hörprobe}3)}=20,88\%$ präferierten die Stereoersion, $f_{3(\text{Hörprobe}3)}=23,08\%$ bewerteten beide Versionen als ebenbürtig. Das arithmetische Mittel liegt hier bei $\bar{x}_{\text{Hörprobe}3}=2,55$. Es lässt sich feststellen, dass bei Hörprobe 3 die Präferenz für die Binauralversion signifikant höher ausfällt als bei Hörprobe 1 ($Z=-6,2095$, $p=,00001$) und Hörprobe 2 ($Z=-7,1623$, $p=,00001$).

7.3.1 Vergleich Probandengruppen Labor / Online

Das Diagramm „Präferenz OLE (Hörprobe 1-3 gemittelt)“ (s. Abbildung 32) zeigt die OLE-Präferenzen der beiden Probandengruppen „Labor“ (kontrollierte Abhörumgebung) und „Online“ (Onlinestudie). Es zeigt sich insgesamt, dass die Versuchsgruppe „Labor“ stärker zum Stereoformat tendiert, wenn auch nur in leichtem Ausmaß ($\bar{x}_{\text{Labor}}=3,31$; $\bar{x}_{\text{Online}}=3,11$). Am deutlichsten zeigt sich diese

Tendenz bei Hörprobe 1, für welche $\bar{x}_{\text{Labor}}=3,65$ und $\bar{x}_{\text{Online}}=3,32$ gilt (s. Abbildung 33).

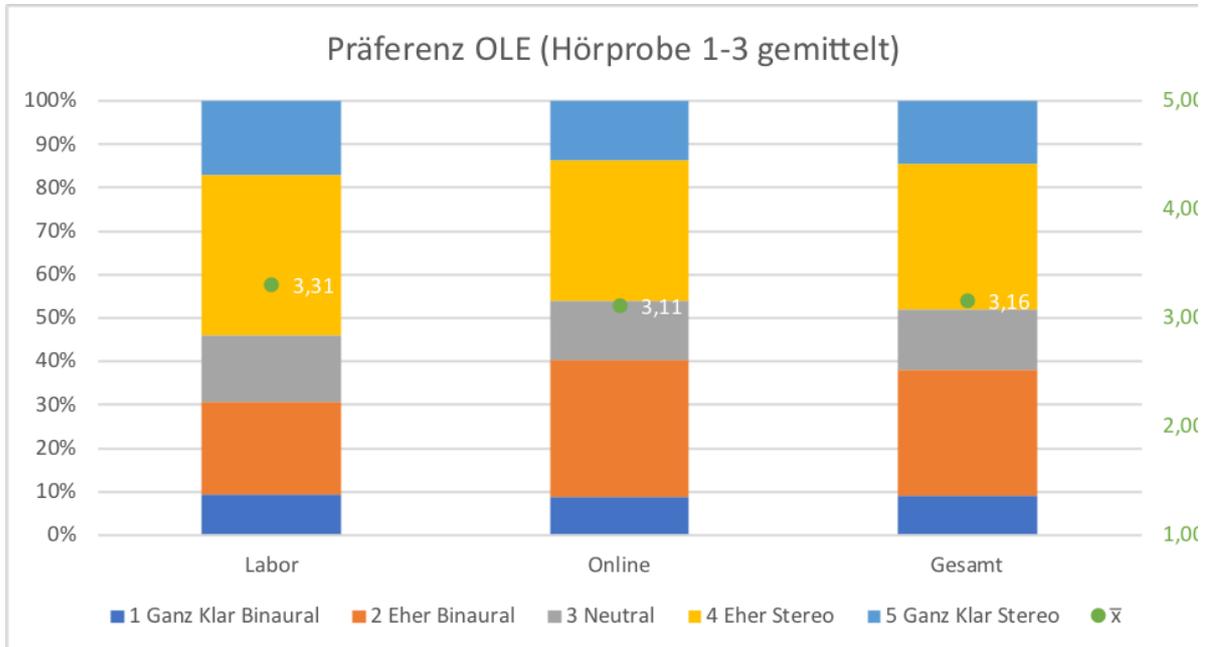


Abbildung 32: Präferenz OLE (Hörprobe 1-3 gemittelt)

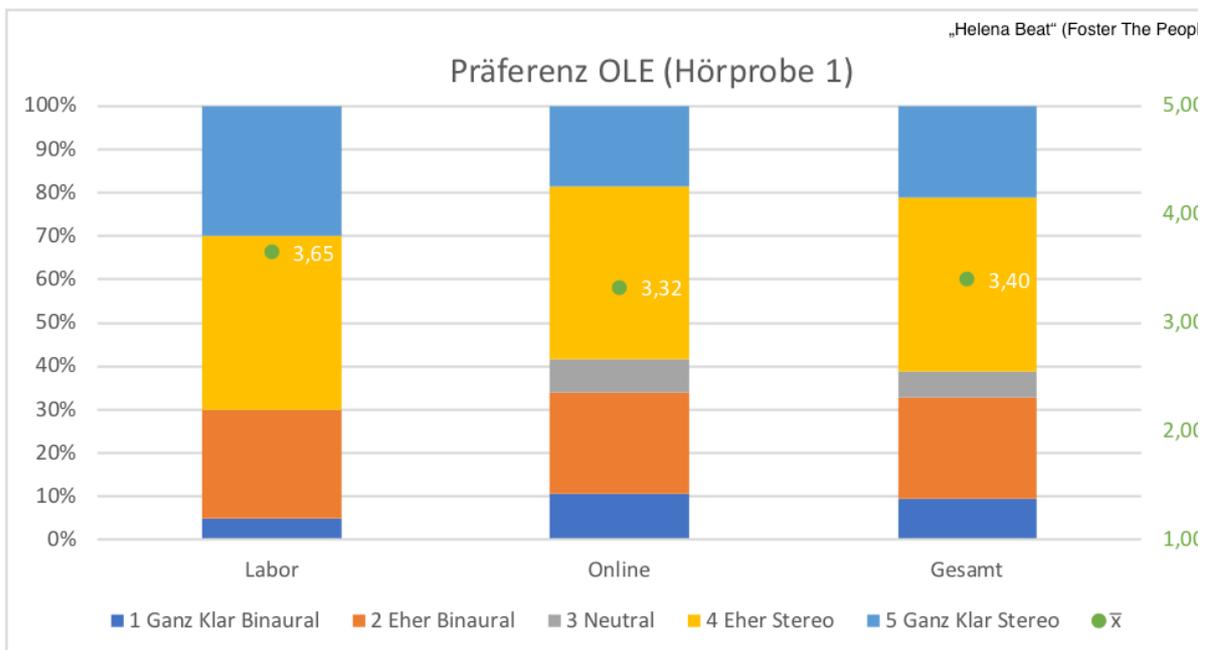


Abbildung 33: Präferenz OLE (Hörprobe 1)

7.3.2 Einzelbewertung

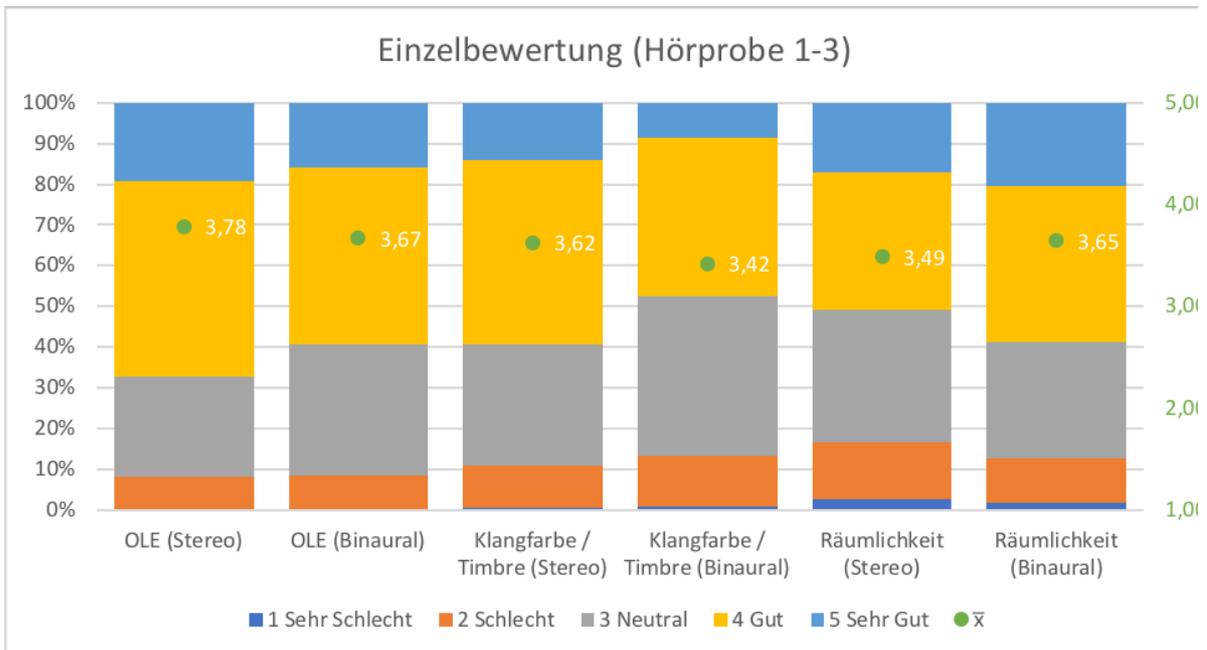


Abbildung 34: Einzelbewertung (Hörprobe 1-3)

Im Diagramm „Einzelbewertung (Hörprobe 1-3)“ (s. Abbildung 34) sind die Bewertungen der Merkmale OLE, Klangfarbe/Timbre und Räumlichkeit der Stereo- und Binauralversionen gemittelt über alle drei Hörproben dargestellt. Zuerst lässt sich feststellen, dass ein Großteil der Bewertungen im Bereich „neutral“ bis „gut“ angesiedelt ist: Für das Stereoformat gilt hierbei $f_{3,4(\text{Hörprobe1-3,OLE}(\text{Stereo}))} = 72,5\%$ und für das Binauralformat $f_{3,4(\text{Hörprobe1-3,OLE}(\text{Binaural}))} = 75,46\%$. Die der Studie zugrunde liegenden Versuchsitems können also für die vorliegende Untersuchung als geeignet betrachtet werden, da der Großteil der Probanden die Musik nicht prinzipiell ablehnt oder unabhängig vom Einfluss des Formats als „sehr gut“ einstuft.

Das allgemeine Klangerlebnis wurde auch hier, kongruent mit dem Ergebnis des A/B-Direktvergleichs, für das Stereoformat leicht, aber signifikant besser bewertet ($Z = -5,568$, $p = ,000$). Dabei bewerteten $f_{4,5(\text{Hörprobe1-3,OLE}(\text{Stereo}))} = 67,23\%$ das Stereoformat als „gut“ bis „sehr gut“. Für die Binauralversionen lag dieser Wert bei $f_{4,5(\text{Hörprobe1-3,OLE}(\text{Binaural}))} = 59,48\%$.

Die Bewertung der Klangfarbe spiegelt diese Tendenz wieder. Die Stereoersionen wurden bezüglich dieses Merkmals in $f_{4,5(\text{Hörprobe1-3,Klangfarbe(Stereo)})}=59,48$ % der Fälle als „gut“ bis „sehr gut“ bewertet. Die Klangfarbe der Stereoersionen wurde lediglich von $f_{4,5(\text{Hörprobe1-3,OLE(Binaural)})}=59,48$ % der Probanden derart positiv empfunden. Die Klangfarbe der Stereoersionen wurde somit als signifikant besser bewertet, als die der Binauralversionen ($Z=-7,416$, $p=,000$).

Bei Betrachtung der Bewertung der Räumlichkeit lässt sich hingegen die Tendenz beobachten, dass hinsichtlich dieses Merkmals die Binauralversionen leicht, aber signifikant besser bewertet wurden ($Z=-6,557$, $p=,000$). Dabei bewerteten etwas mehr als die Hälfte der Probanden die Räumlichkeit der Stereoersionen als „gut“ bis „sehr gut“ ($f_{4,5(\text{Hörprobe1-3,Räumlichkeit(Stereo)})}=50,93$ %), während dieser Wert für die Binauralversionen bei $f_{4,5(\text{Hörprobe1-3,Räumlichkeit(Binaural)})}=58,74$ % lag.

Am deutlichsten zeigen sich diese Tendenzen bei Betrachtung der Ergebnisse von Hörprobe 2 und Hörprobe 3. Für erstere bewerteten $f_{4,5(\text{Hörprobe2,Klangfarbe(Stereo)})}=76,34$ % die Klangfarbe der Stereoersion und $f_{4,5(\text{Hörprobe2,Klangfarbe(Binaural)})}=32,26$ % die der Binauralversion als „gut“ bzw. „sehr gut“ (s. Abbildung 35). Für die Stereoersion von Hörprobe 2 wurden die Merkmale „Klangfarbe“ ($Z=-4,5614$, $p=,00001$), „OLE“ ($Z=-3,8888$, $p=,0001$) und selbst „Räumlichkeit“ ($Z=-3,3095$, $p=,00094$) signifikant besser bewertet.

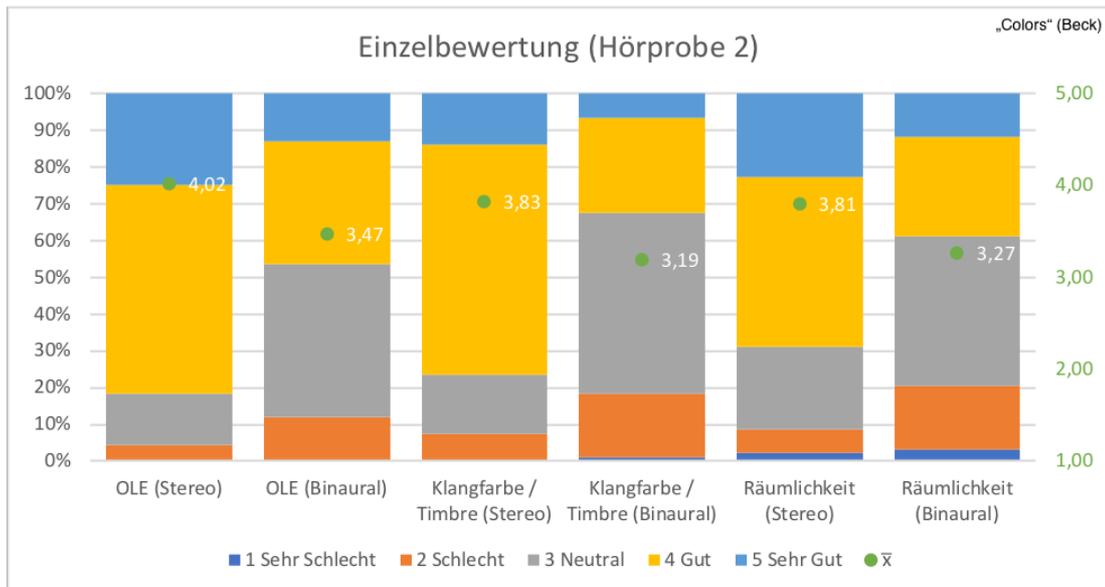


Abbildung 35: Einzelbewertung Hörprobe 2

Letztere wurde im Gegensatz dazu bei Hörprobe 3 für die Binauralversion von $f_{4,5}(\text{Hörprobe3}, \text{Räumlichkeit}(\text{Binaural})) = 79,12\%$ der Vpn als „gut“ bis „sehr gut“ eingestuft (s. Abbildung 36).

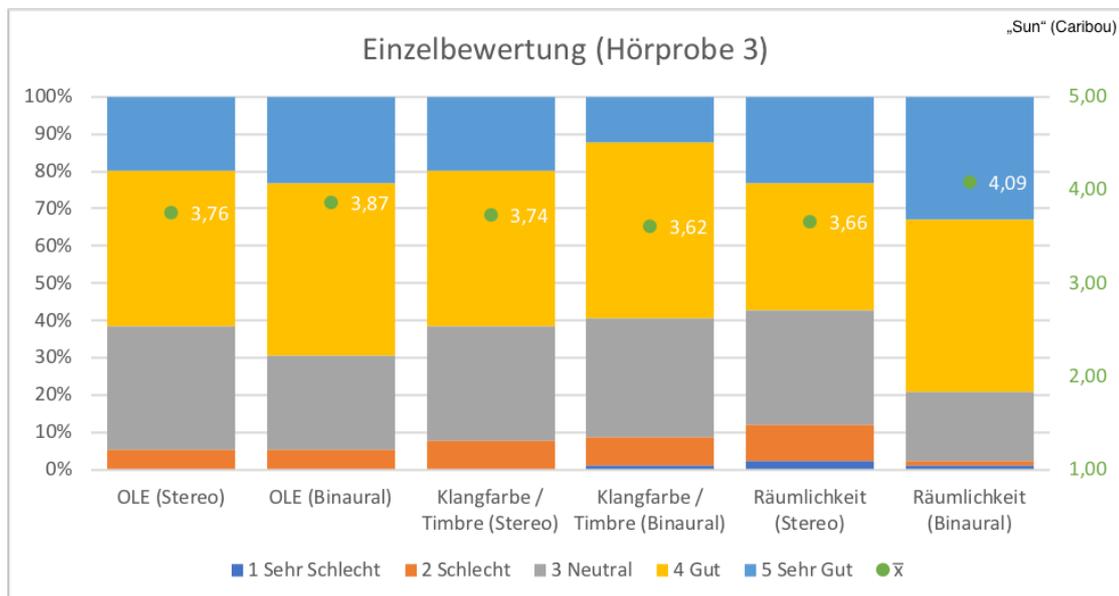


Abbildung 36: Einzelbewertung Hörprobe 3

Der arithmetische Mittelwert liegt hier bei $\bar{x}_{\text{Hörprobe3}, \text{Räumlichkeit}(\text{Binaural})} = 4,09$, welches zugleich die stärkste Tendenz der vorliegenden Studie darstellt. Für die Stereoersion gilt im Gegensatz dazu $f_{4,5}(\text{Hörprobe3}, \text{Räumlichkeit}(\text{Stereo})) = 57,14\%$ bzw. $\bar{x}_{\text{Hörprobe3}, \text{Räumlichkeit}(\text{Stereo})} = 3,62$. Die Räumlichkeit der Binauralversion wurde für Hörprobe 3 gegenüber der Stereoersion als signifikant besser bewertet ($Z = -$

2,7025, $p=,00694$). Auch das allgemeine Klangerlebnis der Binauralversion wurde leicht, jedoch nicht signifikant besser bewertet.

7.3.3 Relation zwischen Kopfhörermodell und Präferenz

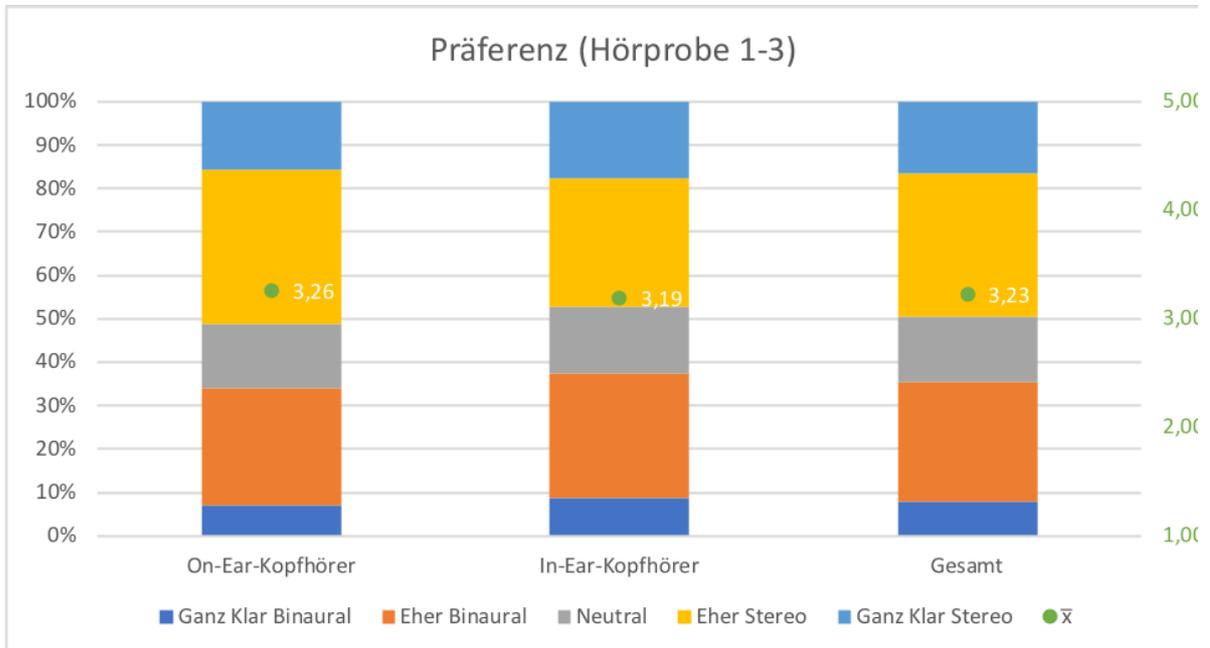


Abbildung 37: Präferenz (Hörprobe 1-3) in Abhängigkeit der verwendeten Kopfhörer

Das Diagramm „Präferenz (Hörprobe 1-3)“ (s. Abbildung 37) zeigt die über alle drei Hörproben gemittelte OLE-Bewertung in Abhängigkeit der verwendeten Kopfhörer. Es lassen sich jedoch keine signifikanten Unterschiede zwischen der Verwendung von On-Ear- oder In-Ear-Kopfhörern feststellen. Die Verteilung der Bewertungen über die Skala ist bei beiden Probandengruppen nahezu identisch. Es konnte also kein signifikanter Einfluss der Verwendung von On-Ear- oder In-Ear-Kopfhörern auf die Empfindung des allgemeinen Klangerlebnisses beobachtet werden.

7.3.4 Kommentarfelder

In der ersten Runde des Hörversuchs wurde den Probanden bezüglich des A/B-Direktvergleichs die Möglichkeit gegeben, ihre Entscheidung zu begründen. Die Sichtung der Kommentarfelder ergab, dass die Stereoversion eher als klar, direkt, intim, kompakt und sauber, dafür jedoch auch als flacher beschrieben wurde. Mehrmals wurde auf die klarere Abbildung der Stimme hingewiesen. Im

8 Schlussbetrachtung

Im Rahmen der vorliegenden Arbeit wurde ein Überblick über den aktuellen Stand binaural-basierter Anwendungen und Produktionen, technischer Entwicklungen sowie Veränderungen der Musikindustrie und des Nutzerverhaltens gegeben. Es wurden vier Musikstücke im Binaural-3D-Format produziert. Dabei wurden gestalterische Möglichkeiten sowie Limitationen im Misch- sowie Masteringprozess evaluiert. Aufbauend auf den produzierten Items wurde ein Hörversuch zur Klärung des Einflusses des Binauralformats auf OLE, Klangfarbe und Räumlichkeit entworfen, durchgeführt und bezüglich verschiedener Aspekte ausgewertet.

8.1 Diskussion der Ergebnisse

Diese Arbeit wurde durch die Frage motiviert, ob das Binauralformat im Kontext moderner Popmusik einen Mehrwert für den Rezipienten darstellen kann. Auf Grundlage der durchgeführten Studie konnte festgestellt werden, dass das allgemeine Klangerlebnis und die Klangfarbe durch die Binauralsynthese in einem kleinen, aber signifikanten Maß negativ beeinträchtigt werden, die Räumlichkeit jedoch von dieser Prozessierung signifikant profitieren kann. Besonders deutlich äußerte sich dies bei der binauralen Version des minimalistischen Elektropop-Stücks „Sun“ von Caribou, deren Räumlichkeit durchschnittlich als „gut“ ($\bar{x}_{\text{Hörprobe3,Räumlichkeit(Binaural)}}=4,09$) bewertet wurde und der damit die insgesamt höchste Bewertung dieser Studie zuteil wurde. Es lässt sich vermuten, dass das Ausmaß, zu dem die Klangfarbe negativ sowie OLE und Räumlichkeit positiv durch die Binauralisierung beeinflusst werden, von der Klangcharakteristik der einzelnen Elemente des Musikstücks abhängig ist. Von der kleinen Auswahl an drei Hörproben lässt sich die Tendenz erkennen, dass sehr dichte Produktionen mit Gesang im Mittelpunkt nach wie vor im Stereoformat besser zur Geltung kommen. Dies könnte durch die Fähigkeit des Gehörs bedingt sein, bei

vertrauten Klangquellen wie der menschlichen Stimme eher eine Klangverfärbung feststellen zu können. Da die interindividuellen HRTF-Differenzen vor allem im hochfrequenten Bereich angesiedelt sind, prägen sich Klangverfärbungen hier besonders aus und bewirken die in den Kommentarfeldern vermehrt monierte blecherne Klangfarbe. Weiterhin darf vermutet werden, dass eher minimalistische, luftige Kompositionen mit synthetischen Klängen einen Mehrwert durch den Binauralisierungsprozess erfahren können, da hier einzelne Klangobjekte besser im Raum wahrgenommen werden können und nicht durch eine Vielzahl weiterer Schallelemente überlagert werden.

Die Ergebnisse spiegeln jedoch auch wieder, dass es den meisten Vpn schwer fiel eindeutige Unterschiede zwischen den Versionen auszumachen. Es lässt sich somit, basierend auf dem aktuellen Stand der Technik, kein nennenswerter Mehrwert des Binauralformats für den Popmusikmarkt feststellen. Auch ein positiver Einfluss einer kontrollierten Abhörumgebung mit einer hochwertigen Signalkette auf die Wahrnehmung der Binauralversion konnte nicht festgestellt werden. Die Streuung der Bewertungen sowie die positive Rezeption von Hörprobe 3 lassen zumindest das Potential des Formats erahnen, wenn neben eines geeigneten zugrunde liegenden Musikstücks auch die technischen Bedingungen, wie die Verwendung individueller(er) HRTF, vorliegen und eine bessere Lokalisierung bei geringerer Klangverfärbung ermöglichen.

Neben diesen Bedingungen ist es für die Zukunftsfähigkeit des Binauralformats in Bezug auf die Musikproduktion entscheidend, dass nicht für jedes Wiedergabemedium eigene individuelle Mischungen generiert werden müssen. Die qualitative Untersuchung der Transferierbarkeit zweier 3D-Audio-Mischungen (s. Kapitel 6.4) konnte hier vielversprechende Ergebnisse aufzeigen. Die binaural, auf Kopfhörern gemischten Stücke „In Your Beat“ (Django Django) und „Sun“ (Caribou) zeigten zwar eine Überbetonung der Lautsprecherkanäle „Left-Side“ und „Right-Side“, jedoch lässt sich dies auf die binaurale Platzierung der Schallquellen an den Punkten der kleinsten negativen Klangverfärbung

zurückführen, welche eher seitlich als vor der Hörerin angesiedelt sind. Anhand einiger kleiner Korrekturingriffe konnte mit recht überschaubarem Arbeitsaufwand ein stimmiges 3D-Klangbild erreicht werden. Wenn die 3D-Mischung bereits bei der Produktion auf verschiedenen 3D-Abhörssystemen kontrolliert und abgestimmt wird, lässt sich vermuten, dass durch technologische Entwicklungen wie MPEG-H, ADM und dem Spatial Audio Designer ein auf verschiedenen Wiedergabemedien gleichmäßig überzeugendes 3D-Audio-Produkt ermöglicht werden kann.

Der in dieser Studie eingebettete Zusammenhang zwischen der subjektiven Präferenz und der wahrgenommenen Qualität der Klangfarbe reiht sich in die Erkenntnisse der in Kapitel 4.3 vorgestellten Untersuchungen ein. Nach aktuellem Stand bestätigt sich nach wie vor eine der Räumlichkeit übergeordnete Priorität des Timbres bei der Beurteilung des allgemeinen Hörerlebnisses.

8.2 Fazit und Ausblick

Diese Arbeit wurde durch das Ziel motiviert, den potentiellen Mehrwert des Binauralformats für moderne Popmusik im Rahmen aktueller technischer Möglichkeiten zu evaluieren und Bedingungen zu identifizieren, die ein gesteigertes immersives Hörerlebnis ermöglichen. Ein bedeutender Mehrwert konnte dabei angesichts der Ergebnisse der Nutzerstudie nicht festgestellt werden. Im Kontext der aktuellen Produktionsbedingungen lässt sich dem Binauralformat keine übermäßige Marktrelevanz attestieren und auch eine „Binauralrevolution“ der Musikindustrie scheint in naher Zukunft eher unwahrscheinlich. Gleichzeitig lässt die positive Resonanz der Räumlichkeit und OLE der dritten Hörprobe erahnen, dass binaurale Prozessierung seine eigene Daseinsberechtigung als Stilmittel im Kontext klangspezifischer Musik einnehmen könnte. Auch die Herausbildung eines Genres eigens für das Binauralformat zugeschnittener Kompositionen könnte ein zukünftiges Szenario darstellen. Angemerkt sei an dieser Stelle jedoch auch, dass die Herausbildung eines

geeigneten Arbeitsprozesseses noch der Herausforderung gegenübersteht, die binaural bedingten Klangfarbenveränderungen der Schallquellen in den Produktionsprozess miteinzubeziehen und zu berücksichtigen. Diese beschränken nach aktuellem Stand noch die gestalterische Freiheit der Positionierung der Schallquellen in wesentlichem Ausmaß. Damit das Format für Künstler, Produzenten und Konsumenten interessant wird, müssen weitere technische Fortschritte erzielt werden. Eine zentrale Rolle spielt dabei die flexible Individualisierung der HRTF und bessere Prozessierungsalgorithmen der Softwarehersteller, um Klangfärbungen zu minimieren und die Lokalisierungsschärfe zu verbessern. Eine Möglichkeit könnte an dieser Stelle auch die Einführung eines Headtracking-Systems darstellen, welches nur Kopfbewegungen bis zu einem bestimmten, möglichst kleinen Winkel berücksichtigt, sodass die Lokalisierung durch die in Kapitel 2.2 beschriebene Peilung verfeinert werden kann, ohne dass sich das 3D-Klangfeld komplett in eine Richtung verschiebt. Flexible, am Anwender orientierte Produktions- und Distributionstechnologien wie in Kapitel 4.2 beschrieben müssen weiterentwickelt werden, um den Arbeitsaufwand bei der Generierung von 3D-Mischungen möglichst gering zu halten. Reizvoll erscheinen auch Ansätze, die Klangobjekte bereits bei der Stereoproduktion mit für die 3D-Positionierung relevanten Tags zu versehen, auf deren Basis automatisiert ein in sich stimmiger 3D-Upmix generiert werden kann. Weiterhin sollten objektbasierte Formate wie ADM oder MPEG-H direkt in den Arbeitsprozess der DAW integriert werden. War bislang die Installation von 3D-Lautsprecher-Systemen noch für den Nutzer mit einem unverhältnismäßig hohen Arbeitsaufwand verbunden, machen Entwicklungen wie die Soundbar das 3D-Format auch für den Heimbereich attraktiv. Für die Relevanz des Binauralformats für die zukünftige Medienlandschaft gilt somit die Voraussetzung, dass anstelle einer auf 2 Kanäle heruntergerechneten binauralen Stereodatei ein flexibles, das Klangfeld beschreibendes Dateiformat wie MPEG-H oder ADM Verwendung findet.

Es bleibt abzuwarten, wie dieses noch recht junge 3D-Format in verschiedenen Medien und Genres eingesetzt und rezipiert werden wird, aber die Veränderung der Hörgewohnheiten hin zu einer Präferenz räumlicherer, umhüllenderer Klangbilder scheint nicht unwahrscheinlich angesichts einer steigenden Akzeptanz und Verbreitung von Anwendungen wie Virtual Reality oder 360°-Video. Für die Musikindustrie könnte zudem die Möglichkeit attraktiv sein, bereits existierende Produktionen im 3D-Format zu mischen und neu zu vermarkten. Hierbei ist die Verwendung eines einheitlichen Sprachgebrauchs bei der Bewerbung des Formats entscheidend. Kritisch sind an dieser Stelle Praktiken wie die Vermarktung von Surround-Formaten unter dem Oberbegriff 3D-Audio, wie es beispielsweise der Kopfhörerhersteller Audeze in der Beschreibung seines Modells „Mobius“ formuliert.¹³³ Eine Richtlinie aus der Werbebranche besagt „If you confuse, you’ll lose.“ Auf dem komplexen Feld verschiedenster 3D-Audioformate muss ein Weg gefunden werden, den Konsumenten auf klare, verständliche und niederschwellige Art ein immersives, emotional ansprechendes Klangerlebnis nahe zu bringen.

Für die weitere Untersuchung der Möglichkeiten des Binauralformats wäre die Identifizierung von Genres und Klangquellen, die tendenziell von der Binauralprozessierung profitieren, interessant. Dabei könnte auch ein technischerer Ansatz, wie beispielsweise die frequentielle Analyse von gegenüber Stereo präferierter Binauralmusik, aufschlussreiche Erkenntnisse liefern. Auch eine detailliertere Analyse eines geeigneten Arbeitsprozesses für das Mastering scheint angemessen.

Um das eingangs erwähnte Zitat Ulrich Gerhardts aufzugreifen, sei abschließend formuliert, dass binaural-3D prozessierte Audioinhalte zur intensiveren Einbindung der Nutzer Stand 2018 wieder für viele Mediensparten von relevanter Bedeutung sind. Zu diesen Bereichen gehört die Musiklandschaft zwar derzeit nur

¹³³ Audeze 2018

in sehr bedingtem Ausmaß. Es darf jedoch mit ebenso spannungsgeladener, erwartungsfreudiger Haltung auf die Immersion und emotionale Attraktion der Musik der Zukunft geblickt werden, wie die eingangs zitierten Zeilen des Mitbegründers der binauralen Medienproduktion in sich tragen.

9 Literaturverzeichnis

- Ackermann, Phillip. *"Computer und Musik: Eine Einführung in die digitale Klang- und Musikverarbeitung"*. Springer-Verlag, 2013, 6 ff.
- All Music. *Caribou*. 2018. <https://www.allmusic.com/artist/caribou-mn0000609590/biography> Zugriff am 23. November 2018.
- Alpert, Adam. *Unless you're Adele, you have no business releasing album tracks all at once*. 3. November 2015. <https://qz.com/536000/unless-youre-adele-you-have-no-business-releasing-album-tracks-all-at-once/> Zugriff am 18. Oktober 2018.
- Arnold, Annegret. *40 Jahre Kunstkopf. Ein noch nicht abgeschlossenes Kapitel der Hörspielgeschichte*. 13. Dezember 2013.
- Audeze. *Audeze Mobius*. 2018. <https://www.audeze.com/products/mobius-series/mobius-headphone> Zugriff am 22. Oktober 2018.
- Auro Technologies. *auro-3d.com*. 2018. <http://www.auro-3d.com/buy/> Zugriff am 29. 11 2018.
- Bau, David, Christian Epe, und Dieter Leckschat. „SOFAPan - Ein Binauralrenderer mit austauschbaren und individualisierbaren HRTF.“ *30. Tonmeistertagung*. Köln: Hochschule Düsseldorf, 2018.
- Bayerischer Rundfunk. *Hörspiel: Demolition*. 24. September 2018. <https://www.br.de/radio/bayern2/programmkalender/sendung-2111916.html> Zugriff am 15. Oktober 2018.
- BBC. *Immerse yourself in music with spatial headphone mixes from the Proms*. 2018. <https://www.bbc.co.uk/programmes/p06drb3s> Zugriff am 26. November 2018.
- BBC News. *Himalayas: a Trek to School in 360 video - BBC News*. 10. Oktober 2017. <https://www.youtube.com/watch?v=Xpq5kBcSYCI> Zugriff am 26. November 2018.
- Beyerdynamic. *beyerdynamic.de*. 2018. <http://www.beyerdynamic.de/kopfhoeerer-headsets/kopfhoeerer-headsets/headzone/headzone-systemueberblick-und-anwendungsgebiete.html> Zugriff am 29. 11 2018.
- Binaulab Audio 3D. *Queen - Bohemian Rhapsody - AUDIO 3D [EN]*. 23. 03 2016. <https://www.youtube.com/watch?v=VnzIIhLNHqg> Zugriff am 2018. 11 29.

- Blauert, Jens. *Räumliches Hören*. Stuttgart: S. Hirzel Verlag, 1974.
- . *Räumliches Hören*. Stuttgart: Hirzel Verlag, 2014.
- . *Untersuchungen zum Richtungshören in der Medianebene bei fixiertem Kopf*. Aachen: na, 1969.
- Brech, Martha, und Ralph Paland. *Kompositionen für hörbaren Raum: Die frühe elektroakustische Musik und ihre Kontexte*. transcript Verlag, 2015, 152 f.
- Butz, Andreas, und Antonio Krüger. *Mensch-Maschine-Interaktion*. Walter de Gruyter GmbH & Co KG, 2014, 23 f.
- . *Mensch-Maschine-Interaktion*. De Gruyter Oldenbourg, 2014.
- Chris Pike, Zillah Watson. *Virtual Reality Sound in The Turning Forest*. 10. 05 2016.
<http://www.bbc.co.uk/rd/blog/2016-05-virtual-reality-sound-in-the-turning-forest>
 Zugriff am 29. 11 2018.
- Daniels, Dieter. *Kunst als Sendung*. C. M. Beck, 2002.
- Dawn of Music. *Avicii - Wake Me Up | 8D Audio [Use Headphones*. 5. Mai 2018.
<https://www.youtube.com/watch?v=EfMnWbr7uhk> Zugriff am 26. November 2018.
- DearVR. „dearVR pro Manual.“ 2018.
- Dickreiter, Michael, Volker Dittel, Hoeg Wolfgang, und Martin Wöhr. *Handbuch der Tonstudiotchnik*. Walter de Gruyter GmbH & Co KG, 2014, 362, 375, 1328 ff.
- Dobrev, Marina S., William E. O'Neill, und Gary D. Paige. „Influence of aging on human sound localization.“ In *Journal of Neurophysiology*, 2471–2486. American Physiological Society, 2011.
- Dowino. *A Blind Legend*. 07. 04 2016.
https://store.steampowered.com/app/437530/A_Blind_Legend/ Zugriff am 25. 11 2018.
- dpa. *Hoffnungsschimmer auf neue Hörgewohnheiten*. 16. Mai 2016.
<https://www.noz.de/deutschland-welt/kultur/artikel/714162/hoffnungsschimmer-auf-neue-horgewohnheiten> Zugriff am 1. Dezember 2018.

- . *Hoffnungsschimmer auf neue Hörgewohnheiten*. 16. Mai 2016.
<https://www.noz.de/deutschland-welt/kultur/artikel/714162/hoffnungsschimmer-auf-neue-horgewohnheiten> Zugriff am 1. Dezember 2018.
- Duden. *Popmusik, die*. 2018. <https://www.duden.de/rechtschreibung/Popmusik> Zugriff am 22. November 2018.
- Elen, Richard. „Ambisonics: The Surround Alternative.“ 2001.
<http://ambisonic.net/pdf/ambidvd2001.pdf> Zugriff am 16. 10 2018.
- Fessl, Sophie. *How Binaural Beats Affect Your Brain – and How They Don't*. 21. Dezember 2017. <https://knowingneurons.com/2017/12/21/binaural-beats/> Zugriff am 26. November 2018.
- Fleischmann, Jan. *MPEG-H – ein Audioformat der nächsten Generation (NGA)*. 14. Mai 2017.
<http://tech-magazin.de/2017/05/mpeg-h-ein-audioformat-der-naechsten-generation-nga/> Zugriff am 5. November 2018.
- Fontana, Simone, Angelo Farina, und Yves Grenier. *BINAURAL FOR POPULAR MUSIC: A CASE OF STUDY*. 2007. <http://biblio.telecom-paristech.fr/cgi-bin/download.cgi?id=7458> Zugriff am 11. 10 2018.
- Fraunhofer-Institut für Bauphysik. Stuttgart, 2018.
- Görne, Thomas. *Tontechnik: Hören, Schallwandler, Impulsantwort und Faltung, digitale Signale, Mehrkanaltechnik, tontechnische Praxis*. Carl Hanser Verlag GmbH Co KG, 2014, 112 f.
- Günther, Andreas. *Kopfhörer - Trend-Gerät und High-End-Produkt*. 2012.
<https://www.connect.de/news/kopfhoerer-trend-geraet-und-high-end-produkt-1457844.html> Zugriff am 11. 10 2018.
- Georg Neumann GmbH. *KU 80*. 2018. <https://de-de.neumann.com/ku-80> Zugriff am 14. Dezember 2018.
- Gerzon, Michael. „The Design of Precisely Coincident Microphone Arrays for Stereo and Surround Sound.“ *AES Convention 50*, 1975.
- Gerzon, Michael. „With Height Sound Reproduction.“ *AES Journal, Vol. 21, No. 1*, 1973.
- Guezenc, Corentin, und Renaud Séguier. „HRTF Individualization: A Survey.“ *145th Convention 2018 October 17 – 20*. New York: Audio Engineering Society, 2018.

- Hahn, Ephraim. „Musical Emotions Evoked by 3D Audio.“ Toyko: Audio Engineering Society, 2018.
- How Stuff Works Tech. *How Virtual Surround Sound Works*. 2016.
<http://electronics.howstuffworks.com/virtual-surround-sound2.htm> Zugriff am 30. 11 2018.
- ifun.de. *Nach 22 Jahren: Fraunhofer beendet MP3-Lizensierung*. 15. 05 2017.
<https://www.ifun.de/nach-22-jahren-fraunhofer-beendet-mp3-lizensierung-107452/>
 Zugriff am 13. 10 2018.
- Inglis, Sam. *Recreating The '80s Home Studio Experience*. November 2015.
<https://www.soundonsound.com/techniques/recreating-80s-home-studio-experience>
 Zugriff am 1. Dezember 2018.
- International Telecommunication Union. *Recommendation ITU-R BS.2076-1 Audio Definition Model*. ITU, 2017.
- Izhaki, Roey. *Mixing Audio: Concepts, Practices and Tools*. Taylor & Francis, 2013.
- Jahn, Dominic. *Gewagtes Debüt: Sennheiser AMBEO Soundbar mit Dolby Atmos & DTS:X*. 19. September 2018. <https://www.4kfilme.de/sennheiser-ambeo-soundbar-dolby-atmos-dts-x/> Zugriff am 10. Dezember 2018.
- Jakobsen, Laurie. *Playlists overtake albums in listenership, says Loop Study*. 2016.
<https://musicbiz.org/news/playlists-overtake-albums-listenership-says-loop-study/>
 Zugriff am 5. November 2018.
- Köhler, Michael. *Wie Hörgewohnheiten die Musikproduktion prägen. Entwicklungen und Veränderungen unter kulturellen, technischen und psychoakustischen Aspekten*. Diplomica Verlag, 2017, 6 ff.
- Kamalzadeh, M., D. Baur, und T. Möller. „A survey on music listening and management behaviours.“ *International Society for Music Information Retrieval Conference*. 2012. 373–378.
- Keinath, Daniel, und Christopher Tarnow. *Entwicklung und Erprobung einer Methode zur Integration binauraler Raumimpulsantworten bei der Mischung von Mehrspurproduktionen*. Diplomarbeit, Detmold: Hochschule für Musik Detmold Erich-Thienhaus Institut, 2009.

- Krumhansl, C.L. *Music: A link between cognition and emotion*. 2002.
- Laumann, Klaus Josef. „Über Methoden zur Qualitätsbeurteilung von virtuellen Kopfhörern. Lehrstuhl für Mensch-Maschine Kommunikation.“ 04. April 2016.
<https://mediatum.ub.tum.de/doc/1295008/document.pdf> Zugriff am 13. Dezember 2018.
- Laut. *Was ist Popmusik?* 2018. <https://www.laut.de/Genres/Pop-89> Zugriff am 22. November 2018.
- Le Callet, P., S. Möller, und A. Perkis. „Qualinet White Paper on Definitions of Quality of Experience, Version 1.2.“ *European Network on Quality of Experience in Multimedia Systems and Services (COST Action IC 1003)*, 2013.
- Leonhard, Joachim-Felix, und Walter de Gruyter. *Medienwissenschaft: ein Handbuch zur Entwicklung der Medien und Kommunikationsformen, Teil 2*. Walter de Gruyter, 2001, 1382 f.
- Lerch, Reinhard, Gerhard M. Sessler, und Dietrich Wolf. *Technische Akustik: Grundlagen und Anwendungen*. Springer-Verlag, 2009, 425-427.
- Lester, Paul. *Foster The People*. 11. Mai 2010.
<https://www.theguardian.com/music/2010/may/11/new-band-foster-people> Zugriff am 22. Oktober 2018.
- Mabbett, Andy. *Pink Floyd: The Music and the Mystery: The Music and the Mystery*. Omnibus Press, 2010.
- Macpherson, Ewan A., und John C. Middlebrooks. „Vertical-plane sound localization probed with ripple-spectrum noise.“ *The Journal of the Acoustical Society of America* 114 (1), DOI: 10.1121/1.1582174, 2003: 430–445.
- MarDie1530HD. *Beck - Sound And Vision (360 Audio Version)*. 2. Juni 2018.
<https://www.youtube.com/watch?v=vMsWB4OQvnw> Zugriff am 26. November 2018.
- Martins, Joao. *Object-Based Audio and Sound Reproduction* . 26. April 2018.
<https://www.audioxpress.com/article/object-based-audio-and-sound-reproduction>
 Zugriff am 13. Oktober 2018.

- mixingroom.de. *Hinter den Kulissen: MPEG-H, das Audioformat der Zukunft*. 16. Mai 2017.
<http://mixingroom.de/hinter-den-kulissen-mpeg-h-das-audioformat-der-zukunft/>
 Zugriff am 14. Dezember 2018.
- Modell, Josh. *Exclusive video: Bob Dylan wants you to embrace mono*. 18. 10 2010.
<https://news.avclub.com/exclusive-video-bob-dylan-wants-you-to-embrace-mono-1798222089> Zugriff am 23. November 2018.
- Moldrzyk, C, T Lentz, und S Weinzierl. „Perzeptive Evaluation binauraler Auralisationen.“
Fortschritte der Akustik, DAGA München, 2005.
- Murphy, Samantha. *Spotify Names Top 10 Songs Streamed in 2011*. 29. Dezember 2011.
<https://mashable.com/2011/12/29/spotify-names-top-10-songs-streamed-in-2011/?europe=true#wGAORjdaxZqV> Zugriff am 22. Oktober 2018.
- New Audio Technologies. „User Manual.“ 2018.
- New Audio Technology. *newaudiotechnology.com*. 29. 11 2016.
<http://www.newaudiotechnology.com/de/spatial-audio-designer/> Zugriff am 29. 11 2018.
- Notes On Blindness. *Notes On Blindness*. 2016. <http://www.notesonblindness.co.uk/vr/>
 Zugriff am 26. November 2018.
- Obst, Holger. *Test: Dear Reality DearVR Music*. 28. Oktober 2017.
<https://www.releasetime.de/test-dear-reality-dearvr-music/> Zugriff am 23. November 2018.
- Parnell, Tom. *Binaural Audio at the BBC Proms*. 02. 09 2016.
<http://www.bbc.co.uk/rd/blog/2016-09-binaural-proms> Zugriff am 29. 11 2018.
- Recording Academy Grammy Awards. *Beck*. 2018.
<https://www.grammy.com/grammys/artists/beck> Zugriff am 22. Oktober 2018.
- . *KRAFTWERK*. 2018. <https://www.grammy.com/grammys/artists/kraftwerk> Zugriff am 26. November 2018.
- ResearchGate. *3-D shape of Spherical Harmonic basis functions*. Juli 2010.
https://www.researchgate.net/figure/3-D-shape-of-Spherical-Harmonic-basis-functions_fig2_224165074 Zugriff am 12. Dezember 2018.
- Roe, K. *Swedish youth and music: Listening patterns and motivations*. 1985.

- Roffler, Suzanne K., und Robert A. Butler. „Factors That Influence the Localization of Sound in the Vertical Plane.“ *The Journal of the Acoustical Society of America* 43 (6), 1968: 1255-1259.
- Roginska, Agnieszka, und Paul Geluso. *Immersive Sound: The Art and Science of Binaural and Multi-Channel Audio*. Taylor & Francis, 2017.
- Rumsey, F., S. Zielinski, R. Kassier, und S. Bech. „On the Relative Importance of Spatial and Timbral Fidelities in Judgments of Degraded Multichannel Audio Quality.“ *Journal of the Acoustical Society of America*, 2005: 968–976.
- Salmon, Francois, Matthieu Aussal, Étienne Hendrickx, Jean-Christophe Messonnier, und Laurent Millot. „Optimization of interactive binaural processing.“ *Presented at the 143rd Convention 2017 October 18–21, New York, NY, USA*. New York: Audio Engineering Society, 2018.
- Schmidt-Atzert, L. „Emotionspsychologie und Musik.“ In *Gefühl als Erlebnis, Ausdruck als Sinn*, 26–46. 1982.
- Sennheiser. *Final Stop: A 3D Audio Thriller (headphones essential)* | Sennheiser. 21. August 2018. <https://www.youtube.com/watch?v=cQhqvVqxlEE&t=1s> Zugriff am 25. November 2018.
- . *ROBIN SCHULZ feat. NICO SANTOS – MORE THAN A FRIEND [Exclusive Song]* | Sennheiser. 15. Juni 2016. <https://www.youtube.com/watch?v=gRcjgCL2O18> Zugriff am 26. Dezember 2018.
- Shaw, E. A. G. *Earcanal Pressure Generated by a Free Sound Field*. The Journal of the Acoustical Society of America, 1966.
- Smyth Research. *A16 Realiser*. 2018. <https://smyth-research.com/> Zugriff am 10. 12 2018.
- Sontacchi, A. „Dreidimensionale Schallfeldreproduktion für Lautsprecher- und Kopfhöreranwendungen.“ *Dissertation Technische Universität Graz*. 2003.
- Sony Music. *Foster The People*. 2018. <https://www.sonymusic.de/kuenstler/foster-the-people> Zugriff am 22. Oktober 2018.
- Statista. *Absatz von Kopfhörern/ Headsets weltweit von 2013 bis 2017 (in Millionen Stück)*. August 2017.

- <https://de.statista.com/statistik/daten/studie/326416/umfrage/entwicklung-der-verkaufszahlen-von-kopfhoerern-weltweit/> Zugriff am 13. Oktober 2018.
- . *Anteil der befragten Internetnutzer, die Musikstreaming-Dienste nutzen, in Deutschland in den Jahren 2013 bis 2018.* Juni 2018.
- <https://de.statista.com/statistik/daten/studie/685445/umfrage/nutzung-von-musikstreaming-in-deutschland/> Zugriff am 13. Oktober 2018.
- . *Anteil der befragten musikkaffinen Internetnutzer, die am liebsten über die folgenden Geräte unterwegs Musik hören.* Juni 2017.
- <https://de.statista.com/statistik/daten/studie/723224/umfrage/die-beliebtesten-geraete-bei-der-mobilen-musiknutzung-in-deutschland/> Zugriff am 13. Oktober 2018.
- . *Anteil der Internetnutzer weltweit, die folgende Formen des Musikkonsums genutzt haben im Jahr 2018.* September 2018.
- <https://de.statista.com/statistik/daten/studie/788726/umfrage/formen-der-musiknutzung-von-internetnutzern-weltweit/> Zugriff am 13. Oktober 2018.
- . *Anteil der Musikstreaming-Nutzer in Deutschland in den Jahren 2015 bis 2017.* April 2018.
- <https://de.statista.com/statistik/daten/studie/880887/umfrage/anteil-der-musikstreaming-nutzer-in-deutschland/> Zugriff am 13. Oktober 2018.
- . *Prognose zum Umsatz mit Virtual-Reality-Software und -Services weltweit in den Jahren 2016 bis 2020 (in Milliarden US-Dollar).* 2017.
- <https://de.statista.com/statistik/daten/studie/426244/umfrage/prognose-zum-umsatz-mit-virtual-reality-software-weltweit/> Zugriff am 04. Dezember 2018.
- . *Prognose zur Anzahl der Smartphone-Nutzer weltweit von 2012 bis 2021 (in Milliarden).* 2018. <https://de.statista.com/statistik/daten/studie/309656/umfrage/prognose-zur-anzahl-der-smartphone-nutzer-weltweit/> Zugriff am 13. 10 2018.
- . *Welche Geräte nutzen Sie für den Konsum von Audioinhalten?* November 2015.
- <https://de.statista.com/statistik/daten/studie/481742/umfrage/von-digitalen-audiokonsumenten-genutzte-geraete/> Zugriff am 13. Oktober 2018.
- . *Wie häufig nutzen Sie Musik-Streaming-Dienste im Internet?* Juni 2018.
- <https://de.statista.com/statistik/daten/studie/189751/umfrage/nutzung-von-musik-streaming/> Zugriff am 13. Oktober 2018.

- Terpitz, Katrin. *Rekordumsatz für Sennheiser*. 08. 07 2015.
<http://www.handelsblatt.com/unternehmen/mittelstand/rekordumsatz-fuer-sennheiser-der-markt-fuer-kopfhoeerer-brummt/12027144-2.html> Zugriff am 30. 11 2018.
- The Zero System. *The Zero System*. 2018. <https://www.thezerosystem.com/welcome> Zugriff am 26. November 2018.
- Theile, G., H. Wittek, und M. Reisinger. „Wellenfeldsynthese-Verfahren: Ein Weg für neue Möglichkeiten der räumlichen Tongestaltung.“ *Proceedings 21st Tonmeistertagung 2002*. Hannover, 2002.
- Theile, Günther, und Helmut Wittek. „Principles in surround recordings with height.“ *AES Convention Paper 8403*. 2011.
- Tingen, Paul. *Neu gemischt: Beatles-Meisterwerk Sgt. Pepper's Lonely Hearts Club Band*. 4. Oktober 2018. <https://www.soundandrecording.de/stories/neu-gemischt-beatles-meisterwerk-sgt-peppers-lonely-hearts-club-band/> Zugriff am 1. Dezember 2018.
- Veltkamp, Niklas. *Kopfhörer-Trend durch Smartphone-Boom*. 04. 11 2013.
<https://www.bitkom.org/Presse/Presseinformation/Kopfhoeerer-Trend-durch-Smartphone-Boom.html> Zugriff am 11. 10 2018.
- Vox. *How a recording-studio mishap shaped '80s music*. 18. August 2017.
<https://www.youtube.com/watch?v=Bxz6jShW-3E> Zugriff am 29. November 2018.
- Ward, Andrew. *Queen - Bohemian Rhapsody - AUDIO 3D [EN]*. 23. 03 2016.
<https://www.youtube.com/watch?v=VnzIIhLNHqg> Zugriff am 30. 11 2018.
- Ward, Darren B, und Thushara D Abhayapala. „Reproduction of a Plane-Wave Sound Field Using an Array of Loudspeakers.“ *IEEE Transactions on Speech and Audio Processing Vol.9 No.6*, 2001.
- Wavearts. *wavearts.com*. 2018. <http://wavearts.com/products/plugins/panorama/> Zugriff am 29. November 2018.
- Waves. *waves.com*. 2018. <http://www.waves.com/plugins/nx#introducing-nx-virtual-mix-room> Zugriff am 29. November 2018.
- Weinzierl, Stefan. *Handbuch der Audiotechnik*. Springer-Verlag, 2009., 589, 671-680

Wikipedia. *Django Django*. 2018. https://en.wikipedia.org/wiki/Django_Django#cite_note-:0-1 Zugriff am 22. Oktober 2018.

Zaar, Johannes. *Vermessung von Außenübertragungsfunktionen mit reziproker Messmethode*. Projektarbeit, Graz: Institut für elektronische Musik und Akustik, 2010.

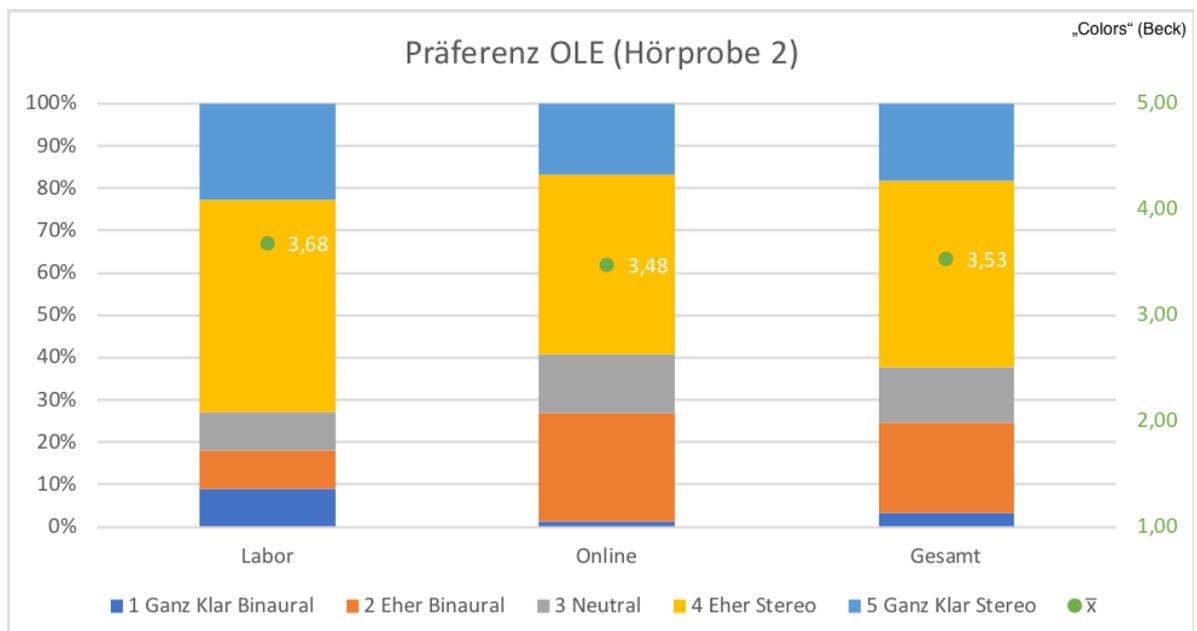
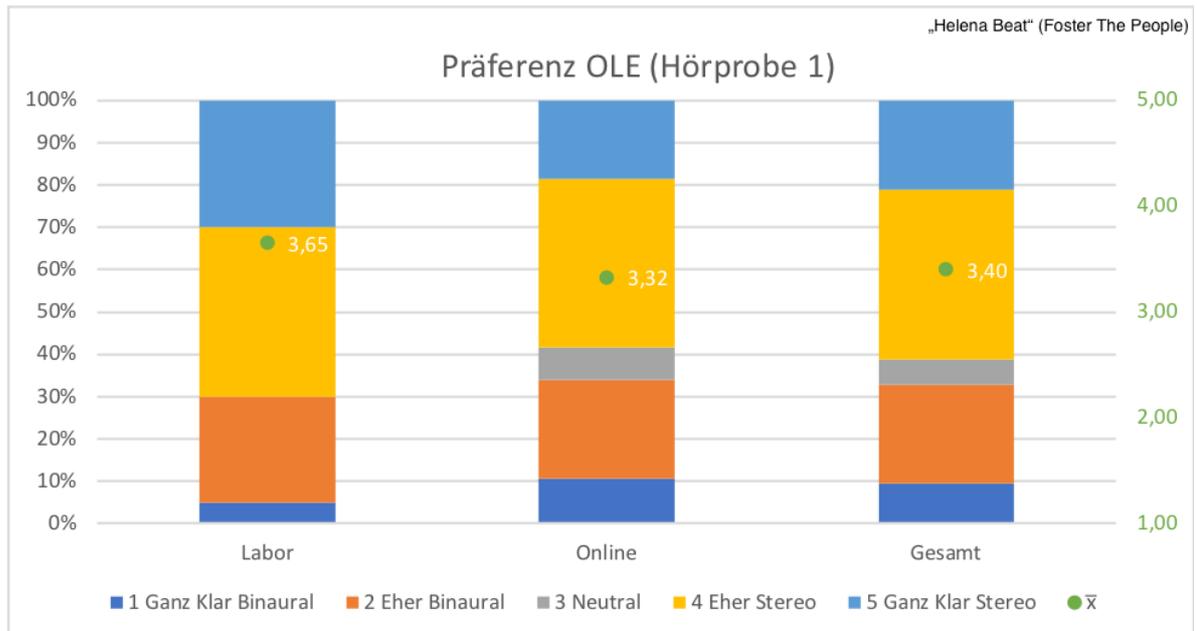
Abbildungsverzeichnis

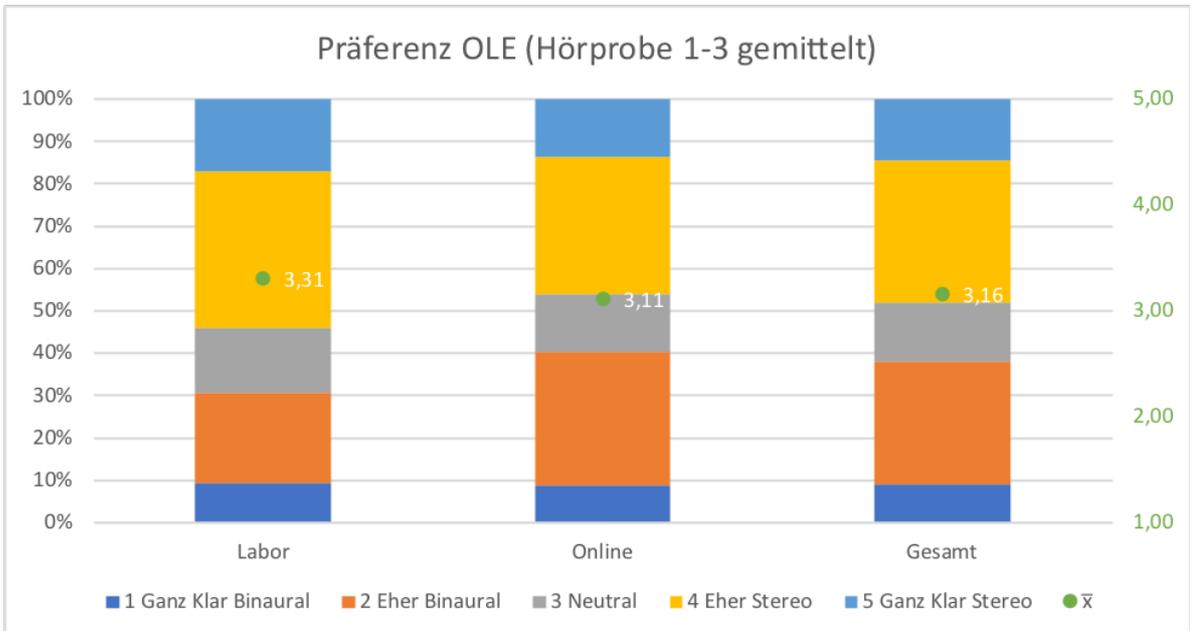
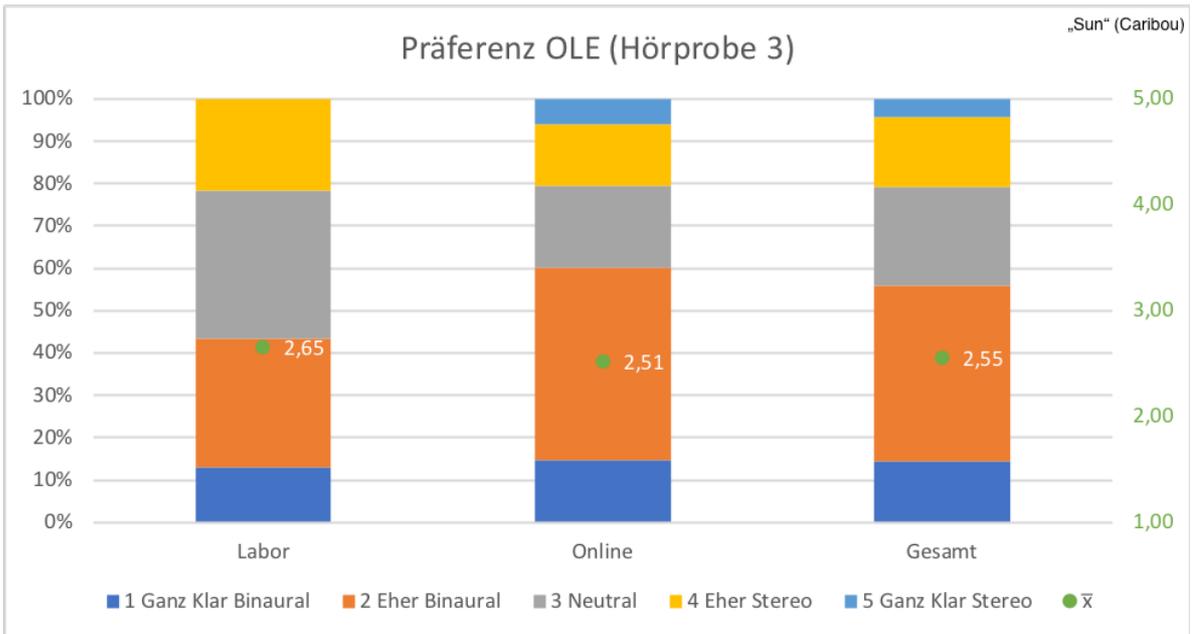
Abbildung 1: links: Werbung für das Théatrophone, rechts: Münzautomaten	7
Abbildung 2: Neumann KU 80	8
Abbildung 3: verschiedene Kunstkopfsysteme A: Head Acoustics HMS III, B: Neumann KU 100, C: Cortex MK1, D: KEMAR KB 4004, E: Brüel & Kjaer HATS 4128, F: Bruel & Kjaer HATS 4100.....	9
Abbildung 4: Albumcover: Bourbon Skiffle Company - Kunstkopf-Stereofonie (Photo).....	10
Abbildung 5: links: ITD, rechts: IID	11
Abbildung 6: „Cone of Confusion“, A und B sowie C und D haben jeweils identische ILD und IID	12
Abbildung 7: winkelabhängige HRTF. Oben: Streuung für frontalen Schalleinfall. Unten: winkelabhängige gemittelte HRTF-Kurven.....	13
Abbildung 8: links: direkte Messung, rechts: reziproke Messung.....	15
Abbildung 9: Schnelle Faltung.....	17
Abbildung 10: Crosstalk Cancellation.....	18
Abbildung 11: Binaural Sky	19
Abbildung 12: Prinzip der Binauralsynthese: Links Messung der BRIR mit Kunstkopf, Rechts Wiedergabe mittels Headtracking	21
Abbildung 13: Richtungsbestimmende Bänder nach Jens Blauert	22
Abbildung 14: Kopfbezogenes Koordinatensystem nach Blauert.....	23
Abbildung 15: Anteil der befragten musikkaffinen Internetnutzer, die am liebsten über die folgenden Geräte unterwegs Musik hören	25
Abbildung 16: Kunstkopf mit vier Blickrichtungen	31
Abbildung 17: Higher Order Ambisonics, dargestellt: 0. Bis 4. Ordnung	35
Abbildung 18: MPEG-H Audiocodierung, simplifiziertes Schema.....	36

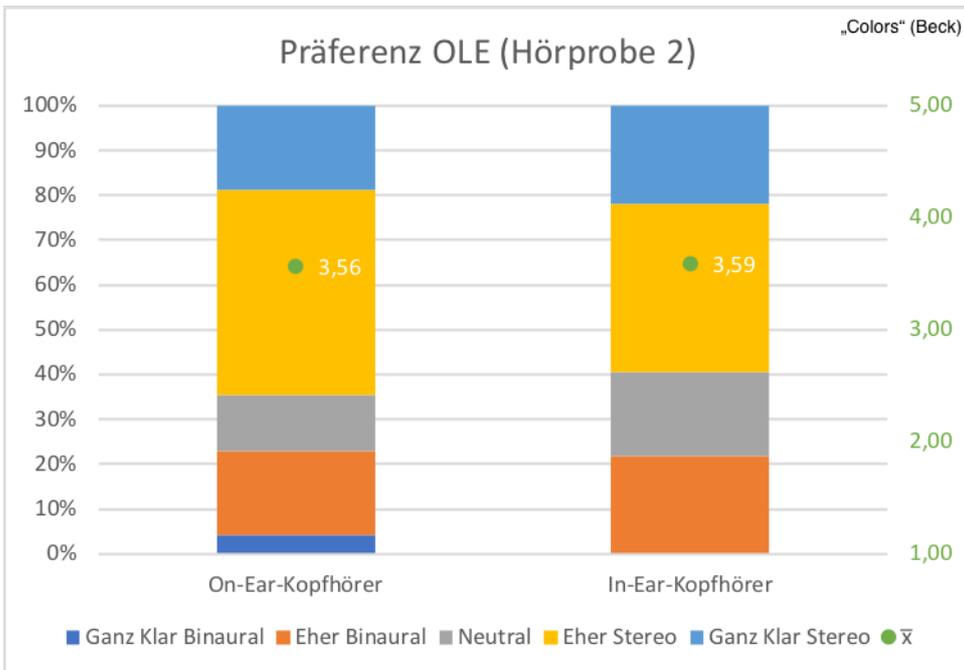
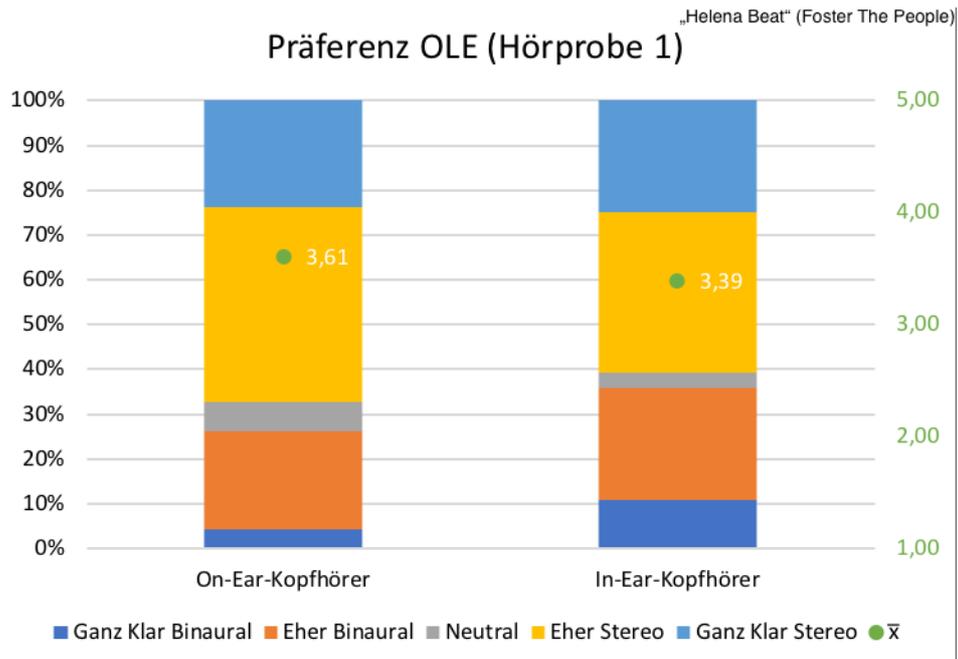
Abbildung 19: Gefälligkeit (links) und emotionale Erregung (rechts) in Abhängigkeit vom Wiedergabeformat	39
Abbildung 20: Pro Tools Projekt "In Your Beat" (Screenshot)	42
Abbildung 21: Logic Pro X Projekt "Colors" (Screenshot)	44
Abbildung 22: Spatial Audio Designer, unten links: Ansicht von oben, unten rechts: Ansicht von hinten	47
Abbildung 23: Grafische Oberfläche DearVR, 1: Position, 2: Reverb, 3: Reflections, 4: Master	48
Abbildung 24: Anordnung der Schallquellen im Spatial Audio Designer (Screenshot)	51
Abbildung 25: Signalkette der Binauralproduktion mit DearVR Pro (eigene Darstellung)	52
Abbildung 26: Signalkette der Binauralproduktion mit dem SAD (eigene Darstellung)	52
Abbildung 27: Automation der Vokal-Einwürfe im Refrain von "Colors" (Screenshot)	55
Abbildung 28: A/B-Direktvergleich des Items (Screenshot).....	63
Abbildung 29: Einzelbewertung des Items (Screenshot)	63
Abbildung 30: 3D-Audio-Labor des <i>Fraunhofer-Institut für Bauphysik</i> (CAD-Skizze)	64
Abbildung 31: Präferenz OLE.....	67
Abbildung 32: Präferenz OLE (Hörprobe 1-3 gemittelt).....	68
Abbildung 33: Präferenz OLE (Hörprobe 1).....	68
Abbildung 34: Einzelbewertung (Hörprobe 1-3).....	69
Abbildung 35: Einzelbewertung Hörprobe 2	71
Abbildung 36: Einzelbewertung Hörprobe 3	71
Abbildung 37: Präferenz (Hörprobe 1-3) in Abhängigkeit der verwendeten Kopfhörer.....	72
Abbildung 38: Wortwolke (eigene Darstellung)	73

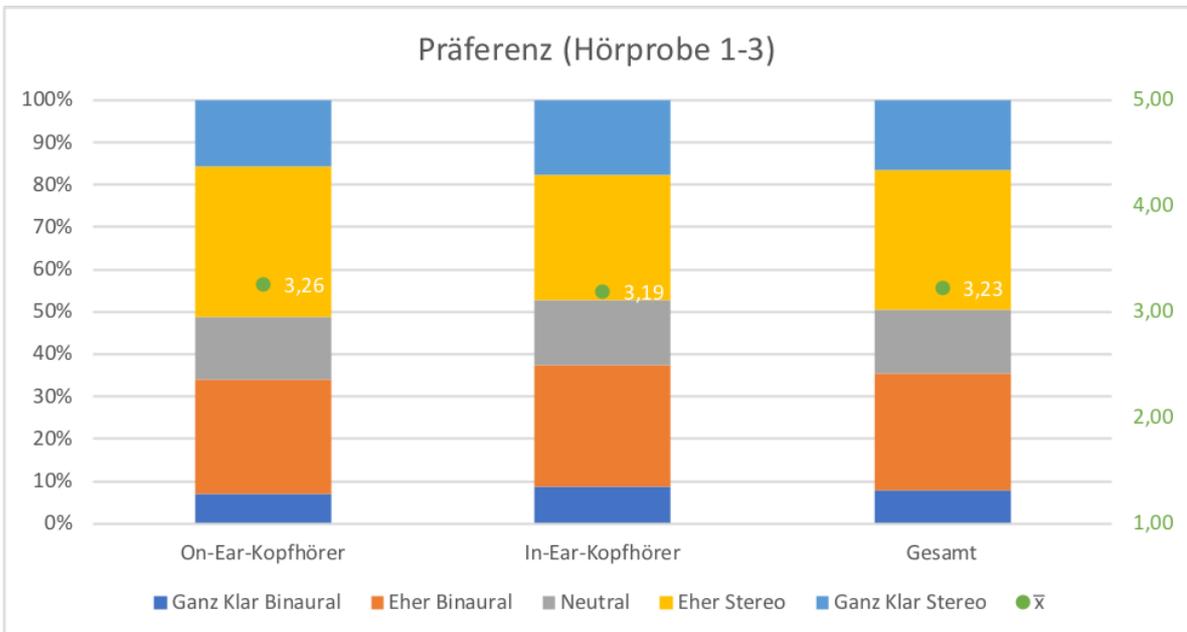
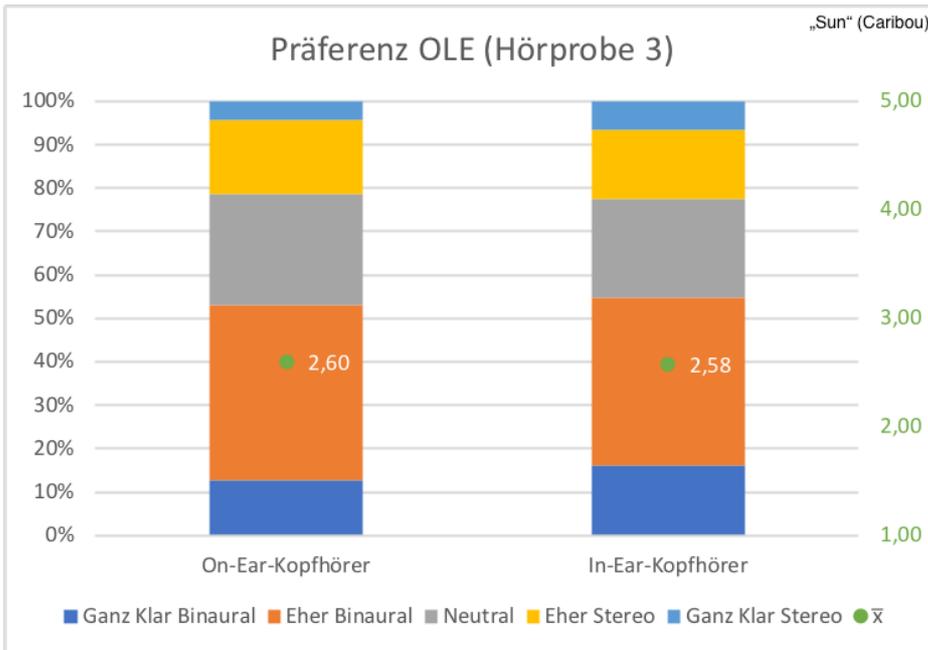
10 Anhang

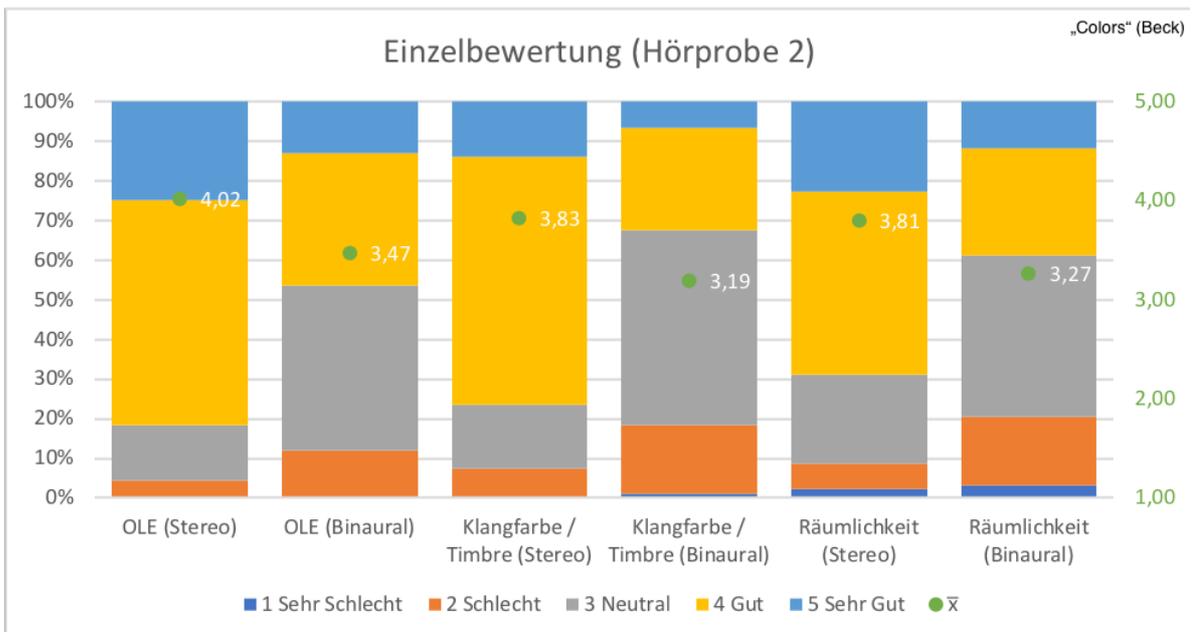
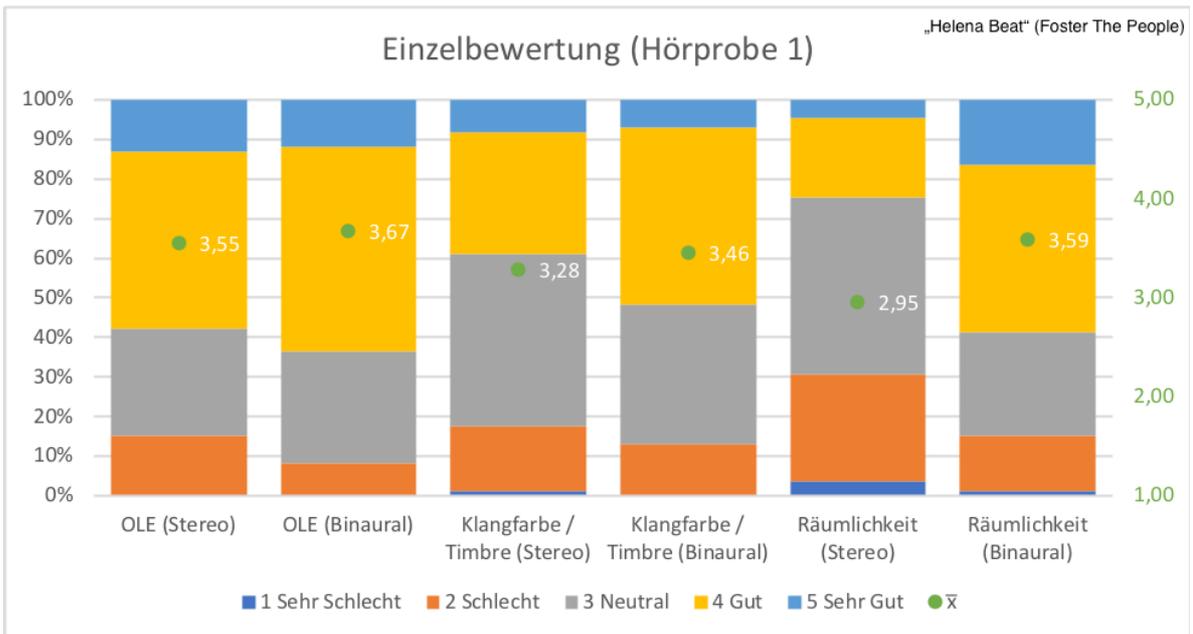
A.1 Hörversuchsergebnisse dargestellt als gestapelte Säulendiagramme

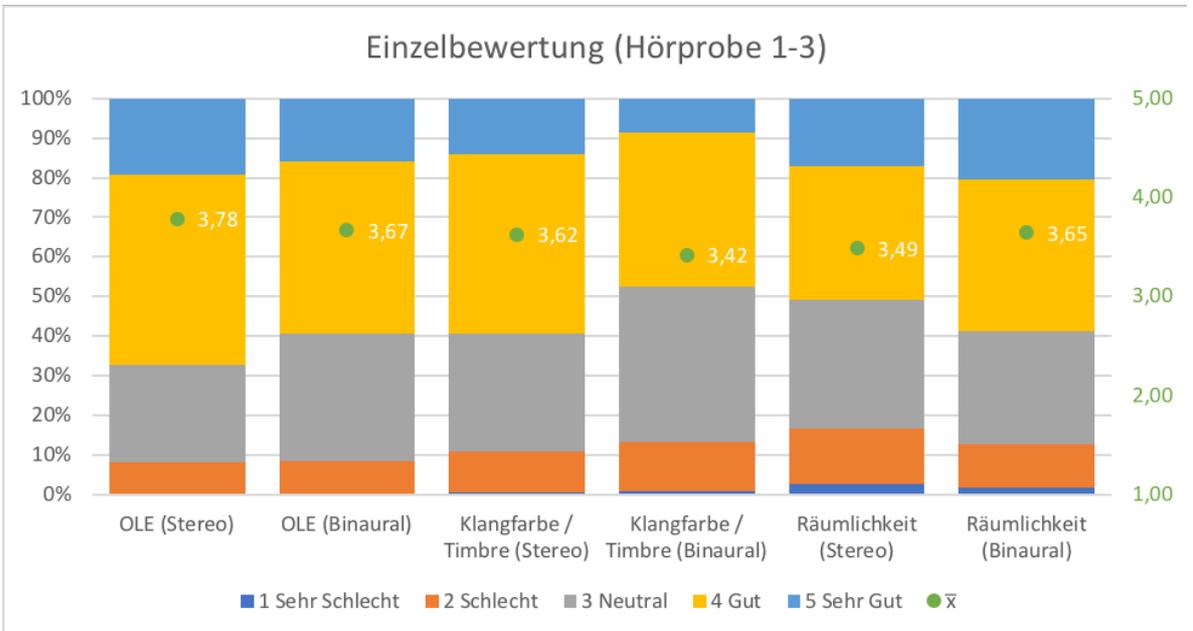
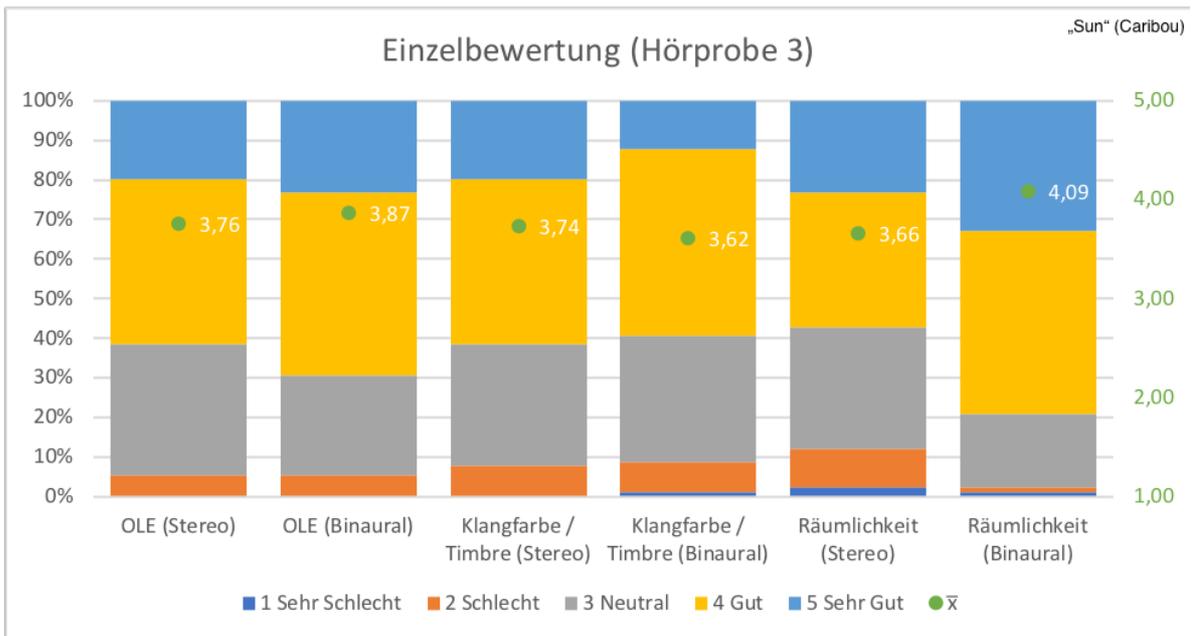












A.2 Package (deutsch)

BINAURAL POP 3D

DER MEHRWERT VON BINAURAL-3D-SOUND FÜR DEN MUSIKHÖRER,
POTENZIALE DES FORMATS UND CHANCEN FÜR DIE MUSIKBRANCHE



OB IM VR-GAME ODER 360°-FILM - WOHIN MAN AUCH HÖRT, 3D-SOUND LIEGT IM TREND. GLEICHZEITIG GREIFEN IMMER MEHR MUSIKLIEBHABER ZUM KOPFHÖRER STATT ZU HIFI-LAUTSPRECHERN. WELCHE CHANCEN ERGEBEN SICH HIERAUS FÜR KÜNSTLER, STUDIOS UND LABELS, DEM NUTZER AUCH BEIM MUSIKGENUSS EIN DREIDIMENSIONALES, KLANGÄSTHETISCH BEGEISTERNDEN HÖRERLEBNIS ZU BIETEN? MIT DEM INNOVATIVEN PROJEKT **BINAURAL POP 3D** SOLL GENAU DIES ERSTMALS ERFORSCHT WERDEN. AUS DEN EINZELSPUREN HOCHWERTIGER STEREO-MISCHUNGEN WIRD MIT OHR FÜRS DETAIL ANSPRECHENDE BINAURAL-3D-MUSIK PRODUZIERT. DABEI WERDEN DIE DERZEIT BESTEN SOFTWARELÖSUNGEN VERWENDET UND DIE POTENZIALE NEUARTIGER PRODUKTIONSTECHNIKEN EVALUIERT. MIT DEM BINAURAL-3D-MATERIAL WIRD DANN IN EINER NUTZERSTUDIE ERMITTELT, WELCHER MEHRGEWINN SICH DEM HÖRER IM VERGLEICH ZUM REINEN STEREOFORMAT BIETET UND WELCHE CHANCEN SICH HIERAUS FÜR DIE MUSIKLANDSCHAFT ERGEBEN.



NACH MEHRJÄHRIGER ERFAHRUNG IN DER PRODUKTION VON MUSIK IM STEREOFORMAT BESCHÄFTIGTE ICH MICH IM TONTECHNIK-MASTER AN DER HOCHSCHULE DER MEDIEN STUTTGART EINGEHEND MIT DER MISCHUNG VON MUSIK FÜR 3D-LAUTSPRECHER-SYSTEME. BEGEISTERT VON DEN ERWEITERTEN MÖGLICHKEITEN IM MISCHPROZESS UND DEM FASZINIERENDEN, UMHÜLLENDE HÖRERLEBNIS ENTWICKELTE ICH SCHNELL EINE LEIDENSCHAFT FÜR DIESES JUNGE FORMAT. AUS DER INTENSIVEN AUSEINANDERSETZUNG MIT DER MATERIE HERAUS ENTSPRANG DIE IDEE FÜR DAS MASTER-PROJEKT **BINAURAL POP 3D**. TREIBENDE MOTIVATION IST DABEI DIE VISION, DEN KLANGÄSTHETISCHEN HÖREINDRUCK VON 3D-SOUND ALLEN MUSIKLIEBHABERN OHNE UMSTÄNDLICHE INSTALLATION VIELER LAUTSPRECHER AUF EINFACHEM WEGE ZUGÄNGLICH ZU MACHEN.



[HIER](#) EIN ENTWURF, WIE BINAURAL-3D IM VERGLEICH ZU STEREO KLING KANN





MARVIN PFEIFER
+49 (0) 171 5382709
MPI30@HDM-STUTTGART.DE



A.2 Package (englisch)

**BINAURAL
POP 3D**

THE BENEFITS OF BINAURAL-3D-SOUND FOR THE MUSIC LISTENER,
POSSIBILITIES OF THE 3D-FORMAT AND CHANCES FOR THE MUSIC INDUSTRY



AS A CRUCIAL ELEMENT OF 360°-FILM OR VR, 3D-SOUND HAS BECOME INCREASINGLY IMPORTANT OVER THE PAST FEW YEARS. AT THE SAME TIME, MORE AND MORE USERS CHOOSE HEADPHONES OVER LOUDSPEAKERS TO LISTEN TO THEIR FAVOURITE MUSIC. AS A RESULT OF THESE TRENDS, WHAT ARE THE NEW OPPORTUNITIES FOR ARTISTS, STUDIOS AND LABELS TO PROVIDE A THREE-DIMENSIONAL, AESTHETICALLY THRILLING AND IMMERSIVE LISTENING EXPERIENCE FOR THE USER?

THE INNOVATIVE PROJECT **BINAURAL POP 3D** AIMS TO EXPLORE THIS QUESTION. THE STEMS OF HIGH QUALITY STEREO-PRODUCTIONS ARE REMIXED TO CREATE APPEALING BINAURAL-3D-MUSIC WITH AN EAR FOR DETAIL AND THE BEST SOFTWARE TOOLS CURRENTLY ON THE MARKET. IN THE PROCESS, THE POSSIBILITIES OF INNOVATIVE PRODUCTION METHODS WILL BE EXPLORED AND EVALUATED. FINALLY, THE BENEFITS OF THE 3D-FORMAT WILL BE INVESTIGATED IN A USER STUDY WITH AN OUTLOOK ON NEW OPPORTUNITIES FOR THE MUSIC INDUSTRY.



AFTER MULTIPLE YEARS OF EXPERIENCE IN THE PRODUCTION OF STEREO-MUSIC I FOCUSED ON MIXING MUSIC FOR 3D-SPEAKER-SETUPS IN MY MASTER STUDIES OF SOUND ENGINEERING AT THE UNIVERSITY OF THE MEDIA STUTTGART. THRILLED BY THE ADVANCED POSSIBILITIES IN THE MIXING STAGE AND BY THE FASCINATING, IMMERSIVE LISTENING EXPERIENCE, I QUICKLY GREW A PASSION FOR THIS YOUNG FORMAT WHICH FINALLY SPARKED THE IDEA FOR **BINAURAL POP 3D**. THIS PROJECT IS DRIVEN BY THE VISION TO PROVIDE THE MUSIC LOVERS WITH THE AESTHETICALLY ENGAGING LISTENING EXPERIENCE OF 3D-SOUND IN A SIMPLE AND STRAIGHT FORWARD WAY VIA CONVENTIONAL HEADPHONES.



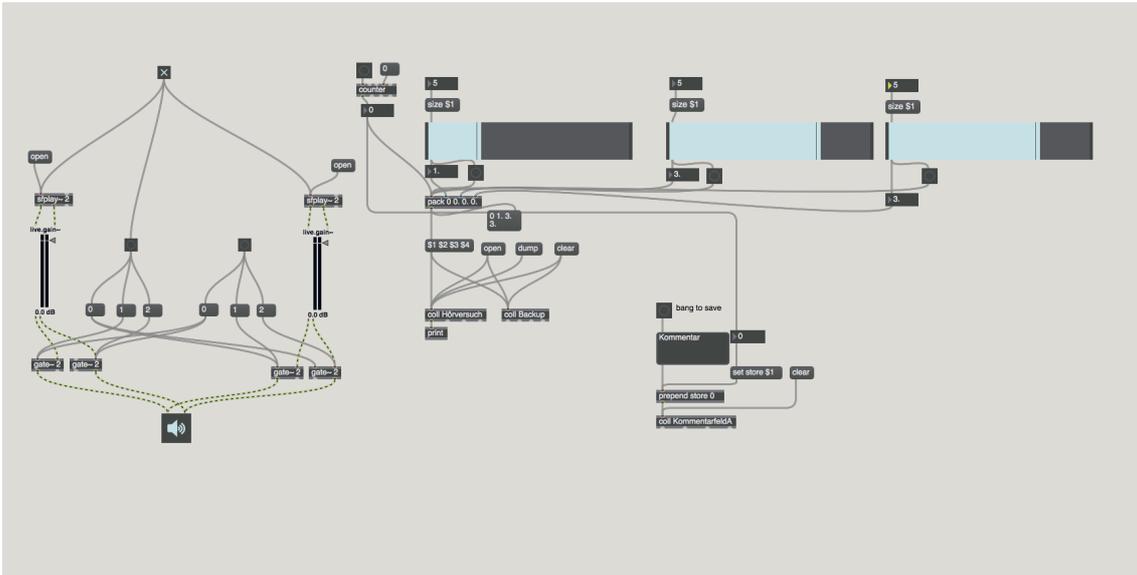
LISTEN TO A DRAFT OF WHAT BINAURAL 3D COULD SOUND LIKE COMPARED TO STEREO [HERE](#)



MARVIN PFEIFER
+49 (0) 171 5382709
MP130@HDM-STUTTGART.DE



A.3 Baustein für Hörversuch (Max 7-Patch)



A.4 Daten-DVD

Die Daten-DVD beinhaltet:

- digitale Version der schriftlichen Ausarbeitung im pdf-Format
- Hörproben in ungekürzter Fassung
- Videoclip: Automation der Schallquellen des Stücks „In Your Beat“ mit dem SAD
- Package zur Anwerbung von Kooperationspartnern (deutsch / englisch)
- Auro-3D-Versionen der Stücke „Helena Beat“, „In Your Beat“, „Sun“