

Masterarbeit

Vergleich der immersiven Qualität
der Wiedergabeverfahren 2.0 und 5.1
unter Berücksichtigung
der Gestalttheorie

Kim Schicklang, Matrikelnummer 27866

31. Januar 2016

Erstprüfung: Prof. Ursula Drees (HdM Stuttgart)
Zweitprüfung: Prof. Oliver Curdt (HdM Stuttgart)

Hochschule der Medien, Stuttgart
Electronic Media Master, AM

Eidesstattliche Erklärung

Hiermit erkläre ich, Kim Schicklang, an Eides statt, dass ich die vorliegende Masterarbeit selbstständig verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel benutzt habe. Die Stellen meiner Arbeit, die dem Wortlaut oder dem Sinn nach anderen Werken entnommen wurden, sind in jedem Fall unter Angabe der Quelle kenntlich gemacht. Dasselbe gilt sinngemäß für Tabellen und Abbildungen. Die Arbeit ist noch nicht veröffentlicht oder in anderer Form als Prüfungsleistung vorgelegt worden.

Stuttgart, den 30. Januar 2016

Kim Schicklang

Abstract

The goals of this study are related to the effects of immersion depending on the choice of the playback systems 2.0 and 5.1 and if there is an measurable influence on immersion depending on the authenticity and consistency of perceived audio-things in an audio-image.

If immersion means, that we leave one space of perception and enter the other we should be strongly interested in the factors that make this transition in audio-environments easier or more difficult.

In the last decades audio reproduction systems had been expanded with more and more loudspeaker-systems. The increase of the amount of loudspeakers always had been justified with an improvement of hearing experience. It's argued that more loudspeakers arranged in a circle or sphere around the listerners position would intensify a feeling of sensual inclusiveness.

On the other hand there are different forms of media whithout audiovisual spheres in a technical sense, like books or smart phones, that can be highly immersive.

This thesis examines if the perception of „Gestalt“ has an influence on immersion amongst the physical surrounding with sound and how deep the consistency of a perceived audio-image affects the feeling of immersion.

On three days a hearing experiment with 35 subjects had been carried out to detect if there are differences in perception between the playback systems 2.0 and 5.1. depending on the listeners position, in the middle or outside (to the left or right of) the sweet-spot and to draw conclusions which factors do have a greater influence: The physical circumstances or the consistency of Gestalt.

Kurzfassung

Die Fragestellung dieser Arbeit ist, ob die immersive Wirkung der Wiedergabesysteme 2.0 und 5.1. von der Authentizität und Stimmigkeit wahrgenommener Hördinge eines Hörbildes abhängen oder nicht. Wenn Immersion bedeutet, dass ein Übertritt von einem Wahrnehmungsraum in den anderen stattfindet, sind die Faktoren, die diesen Übergang leichter oder schwerer gestalten, auch in auditiven Umgebungen von besonderem Interesse.

In den letzten Jahrzehnten wurden Wiedergabesysteme um immer mehr Lautsprecher erweitert. Die Zunahme der Lautsprecheranzahl wurde in allen Fällen mit einem besseren Hörerlebnis begründet. Zuhörer sollen durch ein Mehr an Lautsprechern, die kreis- oder kugelförmig um den Zuhörpunkt angeordnet werden, das Gefühl haben, noch stärker in die Szene eingebunden zu sein.

Demgegenüber steht, dass auch Medien, die nicht technisch umhüllend sind, wie beispielsweise Bücher oder Smartphones, immersiv sein können.

Die vorliegende Arbeit untersucht, ob - neben der physikalischen Umhüllung mit Schall - die Gestaltwahrnehmung einen Einfluss auf die Immersion hat und wie stark die Konsistenz des wahrgenommenen Hörbildes das Gefühl des „Eintauchens“ beeinflusst.

In einer dreitägigen Studie mit 35 Probanden wurde ein Hörversuch durchgeführt um Wahrnehmungsunterschiede zwischen den Wiedergabesystemen 2.0 und 5.1. auf drei unterschiedlichen Sitzpositionen, inmitten des Sweet-Spots und rechts, sowie links davon, zu erfassen und daraus Rückschlüsse zu ziehen, von welchen Faktoren die Immersion mehr beeinflusst wird: von den physikalischen Gegebenheiten oder der Stimmigkeit der Gestalt.

Inhaltsverzeichnis

Eidesstattliche Erklärung.....	1
Abstract.....	2
Kurzfassung.....	3
Einleitung.....	6
Theorieteil.....	8
Verschiedene Räume.....	8
Gestaltgrundlagen und Wahrnehmung.....	14
Phänomenologie.....	14
Gestalttheorie.....	15
Semiotik vs. Phänomene.....	21
Übertrag Bild auf Ton.....	23
Räumliches Hören.....	25
Zweiohriges Richtungshören.....	27
Einohriges Richtungshören.....	28
Monaurale Tiefenreize.....	29
Physikalische Grundlagen, Wiedergabeverfahren.....	30
Mono.....	32
2.0.....	32
5.1.....	35
Wo ist hinten? Wo ist vorne?.....	36
Untersuchung.....	40
Konzept.....	40
Beschreibung der Untersuchung, Untersuchungsmethodik.....	41
Vorstudie.....	41
Beschreibung.....	41
Versuchsanordnung.....	41
Das Szenendesign.....	43
Fragebogen.....	47
Ablauf der Studie.....	49
Verfahren der Datenauswertung.....	52
Ergebnisse.....	56
Zusammenfassung der Ergebnisse.....	56
Die Ergebnisse im Detail.....	57
Kurvenverlauf/Wegstrecke des wahrgenommenen Hördings.....	57
Lokalisation der Zahlen auf der Stereobasis.....	59

2.0 ausserhalb des Sweet-Spots, Vorschlag für eine Visualisierung.....	66
Genauigkeit der Lokalisation im Vergleich zum Ideal.....	68
Sicherheit und Unsicherheit in Abhängig von der Panningrichtung.....	71
Qualitative Beurteilung.....	74
Einfluss der Position der Untersuchungsleitung im Raum.....	77
Demografie.....	78
Zusammenfassung.....	79
Literaturverzeichnis.....	81

Einleitung

Inwiefern unterscheidet sich die Authentizität der Wiedergabe von Tonaufnahmen über 2.0-Wiedergabesysteme von der Raumwahrnehmung bei dem, aus mehreren Lautsprechern bestehenden, Wiedergabeverfahren 5.1? Die Arbeit untersucht, wie sich eine Hörposition ausserhalb des theoretischen Optimums, dem sogenannten „Sweet-Spot“, auf die wahrgenommene Gestalt der Audiodarbietung auswirkt und diskutiert, ob entstehende Unstimmigkeiten innerhalb des Hörbildes Auswirkungen auf die Immersion haben können. Gesetzmässigkeiten und Zusammenhänge aus der Gestalttheorie werden auf „Hördinge“ bezogen, die mit „Sehdingen“ vergleichbar sind.

Ausgangspunkt der Arbeit ist die Fragestellung, ob aus dem Trend zu immer mehr Lautsprechern, der mit der Einführung von 5.1 Ende der 80er-Jahre begonnen hat, auch ein stetig zunehmender Mehrwert an Erlebnisqualität abzuleiten ist, oder ob Einflüsse auf die Stimmigkeit der Wahrnehmung, auch unabhängig der technischen Voraussetzungen, vor allem in der Konsistenz der wahrgenommenen Gestalt zu finden sind.

Ob zwischen der theoretischen Steigerung der Qualität des wahrgenommenen Raumes durch Hinzufügen weiterer Lautsprecher und der tatsächlichen Wahrnehmung ein Unterschied besteht, ist Inhalt dieser Arbeit. Es wurde eine empirische Untersuchung durchgeführt.

Im Tonstudio der HdM wurden 35 Probanden in einer Versuchsanordnung abwechselnd 2.0 und 5.1-Darbietungen vorgespielt, ohne dass diese vorher darüber in Kenntnis gesetzt wurden, welches Wiedergabeverfahren sie jeweils hörten. Bei der Darbietung handelte es sich um ein Panning einer Sprachaufnahme von links nach rechts bzw. von rechts nach links. Die Probanden wurden aufgefordert, sowohl die Form des Hördings zu bestimmen, als auch Angaben zur Qualität der Immersion zu machen. Die Probanden wiederholten die Beurteilung an verschiedenen Sitzpositionen, sowohl innerhalb, als auch ausserhalb des „Sweet-Spots“.

Im ersten Teil der vorliegenden Arbeit werden verschiedene wissenschaftliche Raumbegriffe und Raumkonzepte vorgestellt und der Zusammenhang zur Phänomenologie, der Grundlage der Gestalttheorie, hergestellt. Da es in dieser Arbeit um eine Untersuchung geht, welche sowohl die Gestaltwahrnehmung, als auch die physikalischen Grundlagen berührt, besteht die zweite Hälfte des Theorieteils aus einer Vorstellung der grundsätzlichen Unterschiede zwischen den beiden Wiedergabesystemen 2.0 und 5.1.

Der zweite Teil der Arbeit beschreibt die Untersuchung, deren Ergebnisse dann im dritten Teil besprochen werden.

Hinweis: In der Arbeit wird von 5.1 gesprochen, obwohl bei der Untersuchung nur 5 von 6 Lautsprechern zum Einsatz kamen und auf Grund der Nichtverwendung des LFE-Kanals die Bezeichnung 5.0 richtiger wäre. Da der Begriff 5.1 ein verbreiteter Begriff ist und die Untersuchung auch auf 5.1 übertragen werden kann – über den LFE-Kanal werden „nur“ nicht-ortbare tieffrequente Signale wiedergegeben – wird in der Arbeit von 5.1 gesprochen.

Theorieteil

Verschiedene Räume

Medien erschaffen Raum. Egal ob wir ein Buch lesen, einen Film anschauen, uns in einer Filterbubble in Social Media befinden, einem Hörspiel lauschen oder Musik hören – Medien geben uns die Möglichkeit das, was wir für Realität halten, zu verlassen und in eine andere Welt, einen anderen Raum einzutauchen. Damit stellen sich bereits zwei Fragen: Was ist Raum und wann empfinden wir einen medialen Raum als so überzeugend, dass er uns plausibel genug erscheint um ihn zu betreten?

Die technische Weiterentwicklung von medialen Wiedergabesystemen führen zu einer weiteren Überlegung: Wie wirkt sich die Weiterentwicklung der Technik und die damit verbundene Änderung technischer Parameter auf die Raumwahrnehmung aus? Die Voraussetzung um hierzu Aussagen treffen zu können, bedingt, sich vorab darüber zu verständigen, was unter „Raum“ verstanden werden kann. Genauso wichtig ist es, sich darüber Gedanken zu machen, was es benötigt, um eine Raumerfahrung machen zu können.

Raubegriffe

Länge und Zeit sind physikalische Grundgrößen, Raum und Zeit Dimensionen unserer Wahrnehmung. Es existieren aber unterschiedliche Raum- und Zeitdefinitionen in der Wissenschaft und Philosophie. Isaac Newton formulierte einen „absoluten Raum“ (Günzel 24) und eine „absolute Zeit“, die ohne „Beziehung auf irgendeinen äußern Gegenstand“ verflösse (Schmied 32). Albert Einsteins Relativitätstheorie besagt, dass Raum und Zeit relativ sind – also sowohl abhängig des Bewegungszustandes sind, den ein Betrachter innehat – als auch von der Masse beeinflusst werden.

Die Phänomenologie, eine besondere Form der Philosophie die Anfang des 20. Jahrhunderts von Edmund Husserl geprägt wurde (Günzel 105), beschreibt den Versuch, Raum bzw. Räumlichkeit über Wahrnehmung zu erklären. Ein Raum ist aus phänomenologischer Sicht nicht nur im Einsteinschen Sinn nach physikalisch relativ, sondern davon abhängig, welchen Sinnzusammenhang dieser für einen Wahrnehmenden enthält. Was wir wahrnehmen, ist zu den physikalischen Gegebenheiten nicht identisch: Raum konstituiert sich so, wie wir ihn begreifen.

Eine immer grösser werdende Verfügbarkeit von Medien in der Gesellschaft hat auch Auswirkungen auf die Vorstellungsräume, von denen wir umgeben werden. Der französische

Philosoph Michel Foucault, bezeichnet die Epoche, in der wir uns befinden, als „Epoche des Raumes“.

*Wir sind in der Epoche des Simultanen, wir sind in der Epoche der Juxtaposition, in der Epoche des Nahen und des Fernen, des Nebeneinander, des Auseinander.
(Foucault 34)*

Laut Foucault hätten Menschen sich bis zum 19. Jahrhundert vor allem damit beschäftigt, Geschichten zu erzählen und Mythen weiterzutragen. Räume seien hierarchisch strukturiert gewesen (Foucault 36).

Heute haben wir es mit einer dynamischen Welt zu tun, in der unterschiedliche Räume verschiedener Qualitäten aufeinander treffen, je nachdem, welche Rolle wir als Mensch gerade einnehmen. In welchem Raum wir uns befinden, hängt nicht nur von den physikalischen Gegebenheiten ab, wie beispielsweise einer architektonischen Umgebung, sondern auch von der Sphäre, der wir uns in einem Moment, oder für eine gewisse Zeit, unsere Aufmerksamkeit widmen. Eine Sphäre ist der Bereich, der uns jeweils umgibt. Ein „Raum“ entsteht aus dem Netz der Beziehungen der Elemente, die in einer Sphäre enthalten sind und darin ihren Platz finden.

Seit der Erfindung und flächendeckenden Nutzung von Medien, hat das Angebot, unterschiedliche Räume betreten zu können, zugenommen. Wir hören einen Radiosender, der Musik spielt, die in einem Studioraum aufgenommen wurde und eine Geschichte erzählt, die selbst einen Raum beschreibt, sitzen in einem Zimmer vor unserem Laptop, und unterhalten uns über soziale Medien mit einem Bekannten, der auf der anderen Seite des Atlantiks wohnt. Mediale Live Räume werden genauso zahlreich angeboten, wie aufgezeichnete Räume. Das Angebot umfasst Bild- wie Tonmedien.

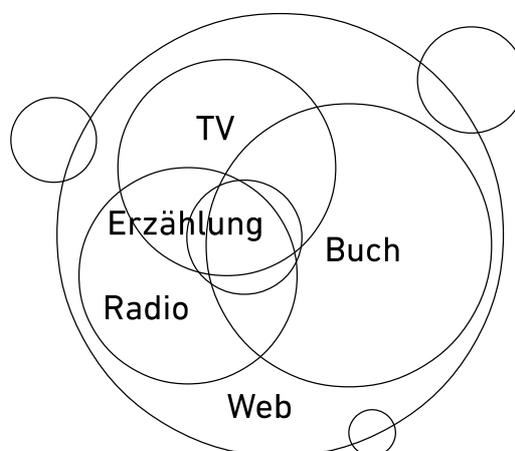


Abbildung 1: Simultanität der Räume (eigene Grafik)

Ein Orchesteraufnahme, die mit Mikrofonen aufgezeichnet und auf schallspeichernden Medien konserviert wird, kann zu einer anderen Zeit, an einem anderen Ort erklingen. In einem architektonischen Wiedergaberaum erzeugen Lautsprecher ein Schallfeld, das sich über die Zeit verändert. Dieses physikalische Raumphänomen enthält die Information über den aufgenommenen Raum und wird dann selbst wieder als räumliches „Bild“ wahrgenommen. Ähnlich wie bei russischen Matroschka-Puppen haben wir es mit einer Verschachtelung verschiedener Räume zu tun.

Bei dem Vergleich zwischen 2.0 und 5.1 ist darauf zu achten, von welchem Raum gesprochen wird. Für die Wahrnehmung entscheidend ist nicht etwa der Ton, der messbar im Wiedergaberaum verbleibt, als physikalischer Schall zwischen den Lautsprechern, sondern welches Hörbild wahrgenommen wird, wie wir Räumlichkeit in unserem Kopf organisieren. Ein Hörbild gestaltet Raum.

Um die Unterscheidung der in sich verschachtelten Räume zu verdeutlichen: Ein Sweet-Spot beispielsweise ist ein physikalischer, meist zentraler Punkt zwischen den Lautsprechern, an welchem, der Theorie nach, die Abhörposition optimal ist. Der Sweet-Spot bezieht sich auf den physikalischen Raum.

Die Mitte des Hörbildes kann sich an einer anderen Position befinden: In der Mitte dessen, was wir als Mitte des wahrgenommenen Hörbildes erachten. Hörposition und die Mitte des „Hörbildes“ hängen zwar voneinander ab, doch bei 2.0 und 5.1. sind hier unterschiedliche Effekte feststellbar, wenn sich beispielsweise die Abhörposition verändert. Unterschiedliche physikalische Gegebenheiten, wie das Vorhandensein oder Fehlen eines zentralen Lautsprechers – dem sogenannten Center, der sich in der physikalischen Mitte zwischen dem linken und rechten Lautsprecher vor dem Zuhörer befindet – beeinflussen, wie sich der vorgestellte Raum (den das Hörbild beschreibt) beim Wahrnehmenden konstituiert.

Unabhängig physikalischer Voraussetzungen, ist in allen Fällen medialer Raumdarstellung entscheidend, wie leicht es uns fällt, uns auf das Angebot, einen vorgestellten, mentalen Raum betreten zu können, einzulassen.

Immersion

Das Betreten eines mentalen Raumes mittels technischer Apparate wird spätestens seit der Entwicklung von sogenannter „Virtual Reality“ (computergenerierten 3D-Welten) mit dem Begriff „Immersion“ in Verbindung gebracht.

Mathias Bauer beschreibt 2011 im „Jahrbuch immersiver Medien“ Immersion als Übergang in einen Anschauungsraum, der durch die wechselseitige Durchdringung technisch-apparativer und psychisch-imaginativer Projektion entstehe (Bauer 20). Der Zuschauer erlebe sich u.a. auf Grund der Herabsetzung der Schwelle zwischen Innen und Aussen, Umwelt und Bewusstsein, von Alter und Ego als Teil der Szene. Bauers Definition zu Folge ist ein Medium dann immersiv, wenn die Grenze zwischen hier und da möglichst klein, oder – im besten Fall – gar nicht mehr vorhanden ist.

Immersion sei die Ästhetik des Eintauchens meint Laura Bieger (Bieger 9).

Sie ist eine Ästhetik des emphatischen körperlichen Erlebens und keine der kühlen Interpretation. Und: sie ist eine Ästhetik des Raumes, da sich das Eintaucherleben in einer Verwischung der Grenze zwischen Bildraum und Realraum vollzieht. Immersive Räume sind ein markanter Teil der Ästhetisierung von Lebenswelten, die unsere heutige Kultur so nachhaltig prägt. Es sind Räume, in denen Welt und Bild sich überblenden und wir buchstäblich dazu eingeladen sind, uns in die Welt des Bildes zu begeben und in ihr zu bewegen. (Bieger 9)

Es besteht weitgehende Einigkeit darin, dass Immersion den Übergang von einem Raum in den anderen meint, während uns der eine als real erscheint und der andere erst betreten werden will. Immersiv ist ein Wahrnehmungserlebnis, wenn, nachdem der Übergang vollzogen ist, uns der betretene Raum ebenso real erscheint, wie der Raum, den wir mental verlassen haben.

Der Begriff „Immersion“ kann auch auf klassische Medien wie Film, TV oder auch Tonmedien angewandt werden. Jede audiovisuelle Darstellung ist immersiv, wenn durch die Darstellung eine neue räumliche Sphäre initiiert wird, in die ich bewusst eintauchen kann. Wird dieses „Eintauchen“ als Übergang verstanden, so ist ein Medium, in dem eine Darstellung zum Raum-Bild wird, dann besonders immersiv, wenn dieser Übergang von einem mentalen Raum in den anderen möglichst einfach stattfinden kann. Eine bildliche Darstellung erscheint dann immersiv, wenn sie nicht mehr als Bild aufgefasst wird, sondern als „neuer“ schlüssiger Wahrnehmungsraum, in dem ich mich nach dem Übergang befinde.

In den letzten Jahren hat sich sowohl bei Bildmedien, als auch bei den Tonmedien die Annahme verbreitet, dass Immersion durch Hinzufügen von Rauminformation, wie beispielsweise durch 3D-Kinoprojektion oder durch mehrkanalige Wiedergabeverfahren,

vergrößert werde. Die Idee dahinter ist eine theoretische und kaum belegt. Insbesondere die Auswirkungen der Inkonsistenzen zwischen räumlicher Darstellung und den Beschränkungen des Wiedergaberaums auf den Effekt der Immersion sind noch kaum wissenschaftlich untersucht worden.

Theoretisch wäre es denkbar, dass eine 3D-Kinoprojektion am Übergang zwischen Bild und Off - dem Bildkader - bei bestimmten Bildkompositionen Inkonsistenzen der wahrgenommenen Dinge entstehen, welche die Raumwahrnehmung negativ beeinflussen (u.a. durch angeschnittene Körper). Das Fehlen einer sich durch die Bewegung des Zuschauers verändernden Parallaxe innerhalb des Kinobildes bei konventionellen 3D-Wiedergabeverfahren wäre ein weiterer physikalischer Grund, der immersions-verhindernde Auswirkungen haben könnte.

Auch bei der Tonwiedergabe wären solche Unstimmigkeiten innerhalb des dargestellten Raumbildes vorstellbar. Durch Hinzufügen weiterer Lautsprecher könnten unter gewissen Umständen neue sich auf Immersion negativ auswirkende Effekte entstehen. Dazu zählen beispielsweise Verschiebungen der Balance zwischen umhüllendem Schallfeld, das über die hinteren Lautsprecher wiedergegeben wird und der Wiedergabe von vorne (Dickreiter 194). Es ist möglich, dass durch ein Mehr an Lautsprechern auch die geometrischen Zusammenhänge innerhalb eines Hörbildes negativ beeinflusst werden könnten.

Genauso ist unklar, ob das Verwischen des Übergangs zwischen einem Raum und dem anderen, also das Verkürzen des Weges von einem Hier nach dem Dort, die Immersion fördert, oder ob sie dadurch verloren geht. Der Gedanke, dass ein Kunstwerk als solches nicht mehr wahrgenommen würde, wenn die Distanz verloren ginge, wurde in der Literatur bereits mehrfach geäußert, beispielsweise von Rudolph Arnheim (Arnheim 222) oder Christian Metz (Balázs u. a. 192).

Rudolf Arnheim, der 1928 in Berlin bei den Gestalttheoretikern Max Wertheimer, Wolfgang Köhler und Kurt Lewin promovierte, kritisierte in „Film als Kunst“ die Wirkung räumlicher Bilder, wie z.B. stereoskopischer Fotografien und sah in Ihnen den Verlust von gestalterischer Qualität. Der Bildrahmen verlöre den Charakter der zwingenden Begrenzung einer bewussten Komposition (Arnheim 222).

Wird das Filmbild raumhaft, so liegt es nicht mehr in der Ebene des Rahmens sondern hinter ihm, hat gar keine optische Beziehung mehr zu ihm sondern wirkt als ein bloßes vorgeschaltetes Cachet - zufällig, hinderlich, nicht mehr positiv gestaltend! (Arnheim 222)

Hanno Loewy, Literatur- und Medienwissenschaftler, sowie Direktor des Jüdischen Museums Hohenems, beschreibt im Nachwort zu „Der Geist des Films“ von Béla Balázs, wie der französische Filmtheoretiker und Semiotiker Christian Metz den Willen zu distanzierter

Betrachtung eines Kunstwerks als „Lust des Voyers“ bezeichnet (Balázs u. a. 192).

Der Voyeur achtet peinlich darauf, einen Abstand zu wahren, einen leeren Raum zwischen seinem Objekt und seinem Auge. Sein Blick heftet das Objekt in der richtigen Entfernung fest, so wie der Zuschauer im Kino, der darauf achtet, nicht zu nah und nicht zu weit von der Leinwand entfernt zu sitzen. Diese Entfernung darf nicht aufgehoben werden. Die Gefahr, vom Objekt überwältigt zu werden, wäre gleichbedeutend mit der, das Objekt selbst zu verschlingen, es nicht mehr sehen zu können.(Balázs u. a. 192)

Der Gedanke lässt sich weiterführen, wenn wir an die technische Weiterentwicklung immersiver Medien denken. Nähmen wir einmal an, die Grenze zwischen dem einen Raum und dem anderen wäre überhaupt nicht mehr vorhanden, wie könnten wir dann überhaupt noch von einer Ästhetik des Eintauchens sprechen, wenn da nichts mehr wäre, in das wir eintauchen könnten? Und auch eine weitere Frage wurde noch nicht behandelt: Wenn die Schwelle vom Hier ins Da hoch wäre, wie sähe es dann auf dem Rückweg aus, wenn es um das Auftauchen ginge?

Ein Medium ist dann immersiv, wenn wir die spezielle immersive Wirkungsweise des jeweiligen Mediums nicht oder nicht mehr bewusst wahrnehmen und der mentale Raum dem Raum entspricht, den wir für die Realität halten. Die Schwelle zwischen Hier und Da ist der Bruch zwischen bewusst unterschiedlichen Wahrnehmungswelten. Diese Schwelle kann innerhalb eines Mediums angelegt sein, wie beispielsweise bei einem Szenenwechsel in einem Film oder einem Musikstück, oder sie zeigt sich an der Oberfläche des jeweiligen Mediums, in das wir eintauchen. Ein Hörbild wird z.B. durch Schallwellen transportiert. Die Schwelle zwischen den Wahrnehmungswelten konstituiert sich einerseits durch den physikalischen Schall, andererseits durch die Sinnhaftigkeit, sowie Plausibilität des Hörbildes.

Gestaltgrundlagen und Wahrnehmung

Phänomenologie

Die Phänomenologie beschäftigt sich mit der Sinnhaftigkeit des Wahrgenommenen. Nach Husserl ist die Welt der „Gesamtinbegriff von Gegenständen möglicher Erfahrung und Erfahrungserkenntnis“ (Husserl 11). Zentraler Gedanke in Husserls Lehre ist die Annahme eines reinen Bewusstseins, das Realität sinnhaft zusammensetzt (Husserl 120). Die phänomenologische Welt ist eine Welt der Gegebenheit der Dinge, unabhängig davon, ob diese Dinge als physikalisch aufgefasst werden könnten, oder es sich um Ideen handelt. Die Unterscheidung ergibt aus phänomenologischer Sicht keinen Sinn. Für die Wahrnehmung spielt es keine Rolle, ob das Wahrgenommene als physikalisch existierend definiert wird, oder nicht: Der wahrgenommene Gegenstand ist in beiden Fällen derselbe.

Im Gegenteil: Um sich der Wahrnehmung der Dinge bewusst nähern zu können, ist es im Sinne der Phänomenologie nötig, Erklärungen für das Wahrgenommene beiseite zu lassen und sich auf die Wahrnehmung an sich zu konzentrieren. Husserl nennt dies „Reduktion“ (Höfliger 4). Ziel der Reduktion sei es, sich klar zu werden, wahrnehmend zu sein und sich Subjektivität einzugestehen (Husserl und Luft 247). Am Ende der Reduktion stünde ein „reines Ich“ (Friedrich und Gleiter 22). Voraussetzung dafür sei, die natürliche Welt bei der Wahrnehmung auszuschalten (Husserl 122). Husserls „Reduktion“ formuliert einen grundlegenden Gedanken, der auch heute noch seine Gültigkeit besitzt, wenn wir an Gestaltung oder Design denken: All das wegzulassen, was den Blick auf das Wesentliche versperrt. Sich auf das Wesentliche zu konzentrieren bedeutet, sich des Ballastes kultur-abhängiger Erklärungen und zeitgenössischer Betrachtungsweisen zu entledigen.

Die Gestalttheorie knüpft an den Ansatz der Phänomenologie Husserls, die Wahrnehmung in den Mittelpunkt der Betrachtung zu stellen, um die Welt zu beschreiben, an.

Gestalttheorie

Während die Anwendung gestalttheoretischer Grundlagen im Grafikdesign nicht wegzudenken ist, sind in vielen Tonstudios heute immer noch Menschen anzutreffen, die sich über Gestaltgesetze noch nie bewusst Gedanken gemacht haben. Alleine der Vergleich von google-Suchergebnissen zeigt, dass das Wissen über Gestaltgesetze in der „Bildgestaltung“ zu den Grundlagen gehört, während bei „Tongestaltung“ immer noch davon ausgegangen wird, dass es sich bei „Tongestaltung“ lediglich um die Arbeit mit „Ton“ handelt, ohne auf gestalterische Zusammenhänge einzugehen. Der Wikipedia-Artikel „Bildgestaltung“ (Wikipedia „Bildgestaltung“) beinhaltet den Hinweis auf psychologische Grundlagen, der entsprechende Artikel „Tongestaltung“ (Wikipedia: „Tongestaltung“) verweist auf das Verwenden von Soundarchiven oder der Nachvertonung von Filmen.

Ein Grund mag sein, dass das Bildmedien häufig als wichtiger erachtet werden, als Tonmedien. Viele Toningenieure, die schon einmal am Set eines Filmes mitgearbeitet haben, sind sich dieser Hierarchie bewusst: Das Bild zählt mehr, als der Ton.

„Wir leben primär in einer Sehwelt“ meint Oliver Krämer in „Strukturbilder, Sinnbilder, Weltbilder“ und beschreibt, wie die tiefe Vertrautheit mit dem Sichtbaren über lange Zeit gewachsen sei, wir uns in der Klangwelt aber nicht gleichermassen heimisch fühlten.

Das wohl wichtigste Moment einer solchen Funktionsübertragung geht mit der Erfindung von Bild und Schrift einher. Die Weitergabe von Erfahrungswissen, die anfänglich vor allem auf mündlicher Überlieferung beruhte und an die gesprochene Sprache, an das Zuhören und das Klanggedächtnis gebunden war, wurde dadurch entpersonalisiert. Mit dem Bild und der Schrift als Vehikel konnten Wissensinhalte nun unabhängig von der Anwesenheit ihrer Urheber große Entfernungen und Zeiträume überbrücken. Die klangbasierte orale Weitergabe wurde von bildbasierter und literarischer Tradierung abgelöst. Der damit einhergehende Paradigmenwechsel vom Hören zum Sehen lässt sich auch anderenorts historisch nachweisen. (Krämer 11)

Es verwundert, dass die Grundlagen der Gestalttheorie und die Beschreibung von Wahrnehmungs-Phänomenen als Dinge auf Überlegungen von tonaffinen Wissenschaftlern zurückzuführen ist.

Der Begriff „Gestalt“ geht auf Christian Freiherr von Ehrenfels zurück, dessen Arbeit „Über Gestaltqualitäten“ im Jahr 1890 erschien, der sich darin selbst auf eine Schrift von Ernst Mach bezieht, die 1886 unter dem Titel „Beiträge zur Analyse der Empfindungen“ veröffentlicht wurde (von Ehrenfels). Ehrenfels beschreibt in dem Text unter anderem, dass Melodien mehr sind, als die Summe der Einzeltöne eines Liedes.

Weitere Untersuchungen sind auf den Philosophen, Psychologen und Musiktheoretiker Carl Stumpf zurückzuführen. Er führte Ende des 19. Jahrhunderts Untersuchungen über die Ordnungsstruktur der Töne durch. Sein Hauptwerk „Tonpsychologie“ erschien in zwei Bänden in den Jahren 1883 bis 1890. Stumpf führte den Gedanken aus, dass, wenn Töne jeweils in ihrer Höhe oder in ihrer Lautheit „skalierbar“ seien, man aus dieser Vergleichbarkeit die Existenz von Grundverhältnissen ableiten könne (Ebeling 14).

Ein solches Vergleichsurteil setzt aber schon voraus, dass verschiedene Einzelempfindungen festgestellt wurden, also im Fall von Tönen müssen mindestens zwei Töne unterschieden worden sein. Das Grundverhältnis der Mehrheit ist die Voraussetzung für einen Vergleich. Stumpf benennt Mehrheit, Steigerung, Ähnlichkeit und Verschmelzung als die wichtigsten Grundverhältnisse. Das einfache Tonhöhenurteil umfasst also bereits die drei Grundverhältnisse der Mehrheit, der Ähnlichkeit und der Steigerung. (Ebeling 16)

Dass heisst: Eine Melodie, die transponiert wird, oder leiser oder lauter erklingt, wird immer noch als solche erkennbar bleiben, da die Grundverhältnisse erhalten bleiben.

Carl Stumpf hatte nicht nur einen grossen Einfluss auf Edmund Husserl, den Begründer der Phänomenologie, sondern war Professor von Max Wertheimer, Wolfgang Köhler und Kurt Koffka, welche die Gestalttheorie etablierten. Ein berühmtes Zitat, welches dem griechischen Denker und Philosophen Aristoteles zugeschrieben wird, fasst den Grundgedanken der Gestalttheorie zusammen:

*„Das Ganze ist mehr als die Summe seiner Teile.“
Zitat von Aristoteles (Staubmann 147)*

Die Gestalttheorie basiert auf Überlegungen zur Psychologie der Wahrnehmung und der Beobachtung von kultur-unabhängigen Gesetzmässigkeiten. Wahrgenommene Phänomene als Realität anzuerkennen, gehört zur Grundüberlegung der Gestalttheorie. Die Phänomene lassen sich zwischen dem Subjektiven, der Bewertung durch das Individuum, das verknüpft ist mit Gelerntem und Studiertem, und der physikalischen Materie verorten. Wenn wir wahrnehmen, fassen wir das Wahrgenommene sinnvoll zu Phänomenen zusammen. Wie uns ein gleichseitiges Dreieck nicht als Anordnung von drei gleich langen Strichen erscheint, so ist für uns die Aufnahme eines Musikstücks ein Lied und nicht die Abfolge von Einzeltönen unterschiedlicher Höhe und Lautheit.

Bei den aus der Gestalttheorie abgeleiteten Gestaltgesetzen, die Wertheimer 1922/1923 eingeführt hat (Stadler und Wildgen 2473), wird davon ausgegangen, dass die Wahrnehmung des Menschen auf sinnvollen Zusammenfassungen basiert.

Dazu zählen u.a. folgende Gesetze:

- Gesetz der Nähe (Was nah zusammen liegt, gehört zusammen)
- Gesetz der Ähnlichkeit (Was sich ähnelt, gehört zusammen)
- Gesetz der guten Gestalt (Die einfachste, sinnvollste Form)
- Gesetz der guten Fortsetzung (Der naheliegendste Weg)
- Gesetz der Geschlossenheit (Zusammengehörende, geschlossene Form)
- Gesetz des gemeinsamen Schicksals (Was sich gleich bewegt)

Der U.S.-amerikanische Wahrnehmungspsychologe James J. Gibson knüpft in den 60er-Jahren inhaltlich an den Gedanken der Skalierbarkeit, den Carl Stumpf in seinen tonpsychologischen Untersuchungen beschrieb, an und beschreibt, dass trotz sich ändernder Umweltbedingungen Parameter einer höheren Ordnung existierten, die zu stabilen und konstanten Wahrnehmungen führten (Gibson 20).

Der in seiner Umwelt aktiv tätige Beobachter erhält somit unveränderte Wahrnehmungen trotz ständig sich ändernder Sensationen. Was er wahrnimmt, ist ein gleichbleibendes Objekt anstelle ständig wechselnder Lichtsensationen, eine gleichbleibende Tastwahrnehmung trotz veränderter Drucksensationen, eine gleich laute Schallquelle trotz Änderung des Lautheitseindrucks in den Ohren. Ein Erklärungsversuch dafür kann sein, daß ein Individuum offenbar fähig ist, das Invariante in einer Reizsituation herauszuheben und zu benützen, wobei üblicherweise dem Wechsel von Sensationen keine Beachtung geschenkt wird. (Gibson 21)

Trotz sich ändernder raum-zeitlicher Verhältnisse, fassen Menschen Wahrgenommenes zu den gleichen Dingen zusammen. Dieses Zusammenfassen gehorcht der Regel der Sinnhaftigkeit. Nimmt man diese Dinge als existent an, ist ihr Raum, in dem sie sich befinden, relativ. Was stabil ist, sind die Verhältnisse zueinander.

Sinnvolle auditive Zusammenfassungen können analog der Begriffswelt der Gestalttheorie als „Hördinge“ beschrieben werden, als wahrgenommene Gestalten, die wir hören. Bei der Wahrnehmung gelten ähnliche Gesetzmässigkeiten, wie bei „Sehdingen“. Melodien werden von Menschen nicht als unzusammenhängende Abfolge von Einzeltönen wahrgenommen, sondern können transponiert werden und dennoch als gemeinsame Gestalt wiedererkannt werden.

Auch Rolf-Dieter Dominicus und Jens Blauert sprechen von „Hördingen“:

Dinge (Gegenstände): Dies sind die so genannten Sinneswahrnehmungen denn sie treten zumeist im Zusammenhang mit Aktivitäten der Sinnesorgane auf. Es gibt jedoch auch hier Ausnahmen, z.B. Augenflimmern bei Migräne. Das dinglich Wahrgenommene ist sinnesspezifisch, d. h. Sehding, Hörding, Riechding, Schmeckding oder Tastding. Hinzu kommt das ebenfalls gegenständliche, als kinästhetisch, statisch oder topisch Wahrgenommene - auch im Zusammenhang mit Propriozeption. (Blauert und Dominicus 29)

Blauert und Dominicus führen an, dass viele Toningenieure das Wahrgenommene, welches scheinbar der „eigentlichen“ Realität entbehre, aus ihrem Denken verbannten, da es ihnen zu „psychologisch“ oder „subjektiv“ vorkäme. Diese Sicht, so Blauert und Dominicus, die Welt in Realität und Wahrgenommenes zu trennen und zu versuchen, das Wahrgenommene als Abbild der Realität zu verstehen, habe zwar als objektivistischer Realismus in der Naturwissenschaft eine lange Tradition, sei aber nicht geeignet um die Prozesse der Wahrnehmung umfassend zu beschreiben. Der Schluss auf eine prinzipiell unzugängliche Welt ausserhalb der Wahrnehmung sei erkenntnistheoretisch problematisch und daher sei es wichtig, daran zu erinnern, dass eine alternative Weltsicht existiere, in welcher das Wahrgenommene als das eigentlich Reale im Vordergrund angesehen werde (Blauert und Dominicus 29).

Auch der Mitbegründer der kybernetischen Wissenschaft Heinz von Foerster bezeichnet die Idee, Wahrnehmung als reine Abbildung aufzufassen, als reduktionistisch.

Der von der Optik entlehnte Begriff „Abbildung“ setzt zwei Bereiche voraus, den „Gegenstandsraum“ und den „Bildraum“, wobei der Abbildungsprozeß es bewerkstelligt, jedem Gegenstandspunkt einen Bildpunkt zuzuordnen. Durch die Metapher, die Abbildung für Wahrnehmung setzt, wird die Verwunderung über das ursprüngliche Phänomen abgestumpft. (von Foerster 437)

Foerster ist Anhänger des radikalen Konstruktivismus, der philosophischen Richtung, in der davon ausgegangen wird, dass alle Inhalte des Erkennens ihren Ursprung in den Individuen haben (Haan und Rülcker 30). Wahrnehmung, im Sinne radikal-konstruktivistischer Annahmen, ist stets subjektiv. Da uns die Möglichkeit fehle, unsere Wahrnehmung mit einer objektiven Wahrheit zu vergleichen, könnten wir nicht davon ausgehen, dass diese Wahrheit existiere. Wahrnehmung als Abbildung aufzufassen, würde den Gedanken eines objektiven „Etwas“ voraussetzen, das – unabhängig unserer Wahrnehmung – messbar sei. Da wir immer Wahrnehmen - uns aber nie auf der Seite dessen befinden, was wir wahrnehmen - können wir das Objekt der Wahrnehmung niemals messen, sondern höchstens die Wahrnehmung an sich.

Dieser Gedanke radikaler Konstruktivisten ist mit der Phänomenologie verwandt. Der Ansatz, sich der Subjektivität der Wahrnehmung bewusst zu werden, wurde bereits von Husserl formuliert, der daher am Ende der „phänomenologischen Ausschaltung der Welt“ das „reine Ich“ sah (Husserl 124).

In der Gestalttheorie, die auf den phänomenologischen Grundannahmen basiert, steht Wahrnehmung im Mittelpunkt der Betrachtung. Es wird davon ausgegangen, dass ein Hauptmerkmal der Wahrnehmung ist, sinnvolle Zusammenfassungen zu machen. Das Ergebnis der Zusammenfassungen seien die Dinge, mit denen wir es zu tun haben.

Die Gestalttheorie betrachtet die Gruppierungsprinzipien als autarkes System mit der Aufgabe, Gegenstände zu organisieren. (Schmicking 89)

Nach Ansicht von Daniel Schmicking bestünden Rede, Musik und Klanggestalten der natürlichen und artefaktischen Umwelt des Menschen aus Folgen zahlreicher auditiver Elemente oder Teile, die in ihren spezifischen Ordnungen Gestalten bildeten. Er führt weiter aus, dass sich entgegen der traditionellen Auffassung von Hören als einem einfachen Sinn, auf der phänomenalen Ebene eine Vielzahl von Fragen ergäben, die auf Grund der Komplexität und Vielfalt der Gliederung des Hörfeldes, noch gar nicht alle beantwortet seien (Schmicking 90).

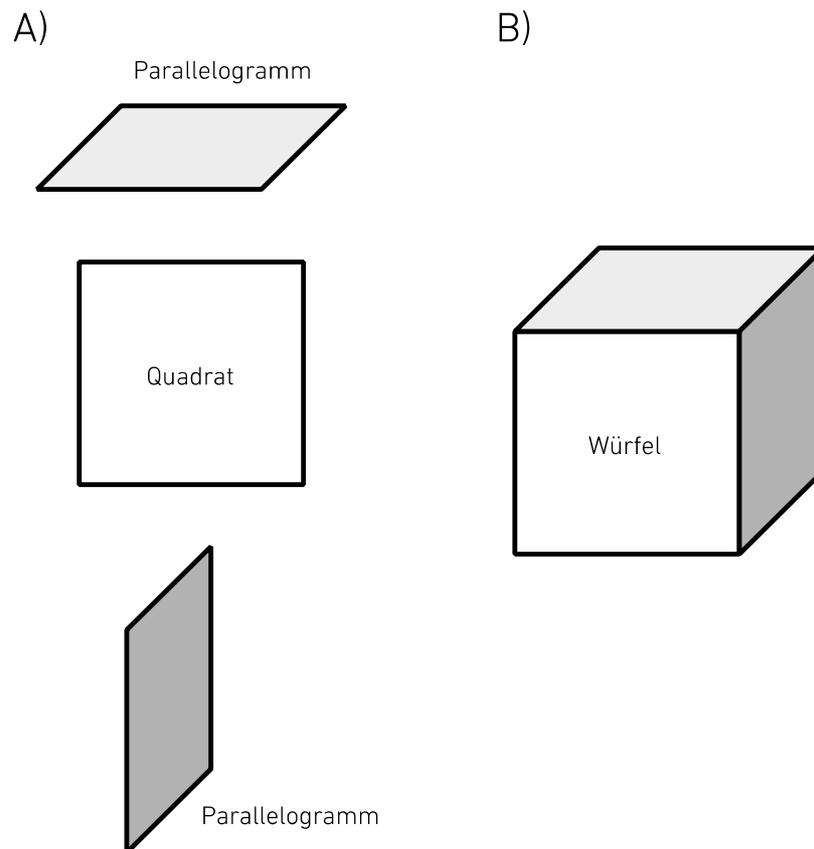


Abbildung 2: Sinnvolle Zusammenfassung von Parallelogramm und Quadrat zu einem Würfel (eigene Grafik)

Links: Drei unterschiedliche Dinge werden wahrgenommen (Senkrechtes Parallelogramm, Quadrat und waagerechtes Parallelogramm)

Rechts: Drei Einzelteile ergeben ein neues Ding, einen Würfel

Semiotik vs. Phänomene

Im Gegensatz zur Gestalttheorie wird in der Semiotik, der „Lehre“ der Zeichen, Wahrnehmung als auf Erfahrungen und kulturellen Übereinkommen referenzierend verstanden. Drei Hauptstränge der Semiotik, die Ende des 19. bis Anfang des 20. Jahrhunderts entwickelt wurden, können unterschieden werden: Der US-amerikanische Strang von Charles Sanders Peirce, die Tradition von Ferdinand de Saussure und die russisch-tschechische Tradition (Stadler und Wildgen 2474). Semiotik beschäftigt sich mit den Zeichen als solchen, mit der Art und Weise, wie sie in Systemen organisiert sind, und mit ihrem Kontext (Crow 14). Semiotik ist ein Teilgebiet der Sprachwissenschaft, der Linguistik. Wesentliches Merkmal der Semiotik ist, den Ursprung von Bedeutung zu erforschen.

Während in der Phänomenologie versucht wird, durch Weglassen von Interpretationen oder Bedeutungszusammenhängen zu einem reinen Wesen, dem „Eidos“ zu gelangen (Husserl 14), verhält sich der Grundgedanke der Semiotik dazu gegenteilig. Semiotiker begreifen Wahrgenommenes als kulturell übermitteltes oder festgelegtes Zeichen, in phänomenologischen Ansätzen wie z.B. der Gestalttheorie geht es um die Betrachtung der Dinge und ihrer Verhältnisse zueinander. Im Gegensatz zur Semiotik geht es bei der Gestalttheorie um die Wahrnehmung von Dingen im hier und jetzt, während es in der Semiotik einen weiteren Schritt benötigt: Die Decodierung von verschlüsselter Information.

Dass sich Semiotik und Gestalttheorie polar gegenüberstehen, führen Stadler und Wildgen an. Sie verweisen auf den US-amerikanischen Semiotiker Charles W. Morris, dessen Ansatz sich an Charles Sanders Peirce orientiert (Stadler und Wildgen 2474). In Morris Arbeiten sehen sie eine „konträre Beziehung von Semiotik und Gestalttheorie“. Stadler und Wildgen sehen, zumindest bei Peirce, keine geistesgeschichtliche Nähe zwischen Semiotik und Gestalttheorie (Stadler und Wildgen 2475).

Die Existenz beider wissenschaftlicher Ansätze zeigt, dass die Wahrnehmung des Menschen sowohl auf Erfahrung und Codes, als auch auf der Wahrnehmung von Grundformen basierend aufgefasst werden kann. Es kommt darauf an, was im Einzelnen betrachtet oder untersucht werden soll: die Gestalt des Wahrgenommenen oder seine Geschichte.

Übertragen auf die Wahrnehmung von Tonaufnahmen heisst das, dass beispielsweise das, was ein OFF-Sprecher in einem Radiobeitrag zu uns sagt, einerseits Sprache, andererseits als Tongestalt aufgefasst werden kann, die sich aus Phonemen zusammensetzt.

Der ehemalige Dirigent und Musiklehrer Sergiu Celibidache bezog auf einem Vortrag, den er 1985 in München hielt, klar für die phänomenologische Betrachtung von Musik Stellung. Er brachte seine Ablehnung gegenüber der Idee zum Ausdruck, Musik als Sprache zu verstehen.

Ist Musik eine Sprache? Nein. Musik ist alles andere als eine Sprache. Die Sprache bedient sich einer in der Diskursivität erscheinenden konventionellen Symbolik und polyvalenter semantischer Bedeutungen. Sie kann durch prädikative Randbewegungen auf verschiedenen Wegen zu einem zentralen sinnhaften Kern kommen. Der Ton spricht den Menschen direkt, unentrinnbar an, unabhängig von jeder spezifischen individuellen Determination wie Rasse, Geschlecht, Zustand, Alter, und ruft freie, nicht konditionierte Reflexe hervor. (Lang u. a. 40)

Für Celibidache findet sich die Wirkmächtigkeit von Musik nicht primär in der Geschichte, die ein Musikstück vermittelt, sondern im Ton an sich. Ton-Grundformen seien kulturunabhängig.

Rolf-Dieter Dominicus und Jens Blauert nehmen an, dass die Objektbildung und die Formation von Audioszenen überwiegend nach relativ festen Regeln ablaufen (Blauert und Dominicus 38). Sie führen als wesentliche Gestaltgesetze das Gesetz der guten Fortsetzung, der Geschlossenheit, der Einfachheit, der Gewöhnung und Trägheit an.

Der Musikpsychologe Herbert Bruhn verweist darauf, dass sinnhafte Zusammenfassungen von Gehörtem auch einen praktischen Nutzen habe: Die Reduktion der zu verarbeitenden Informationen (Bruhn 445). Der Mensch nehme kaum jemals alle verfügbaren Informationen auf.

Gerade bei der Musikverarbeitung kann gezeigt werden, wie weit die Anzahl der aufgenommenen Informationen und damit auch die individuellen Interpretationen von Musik streuen. Die Menschen nutzen in der Wahrnehmung nie alle Informationen aus, die die physikalische Umwelt ihnen bietet. Der Wahrnehmende bricht die Informationsverarbeitung ab, sobald das Ergebnis ausreichend präzise scheint, um eine Handlungsentscheidung zu treffen. (Bruhn 445)

Laut Bruhn müssten Menschen, die Musik hören, auch nicht alle Aspekte eines erklingenden Musikstücks aufnehmen, um dieses zu genießen. Einzelparameter müssten nicht aufgenommen werden, um Musik als ganzheitliches Erlebnis zu empfinden (Bruhn 446).

Übertrag Bild auf Ton

Obwohl die Ansicht, dass Bildmedien Medien des Raumes sind und Tonmedien mit der Zeit zu tun haben in der Wissenschaft weit verbreitet war, kann diese Sicht – insbesondere dann, wenn eine phänomenologische Perspektive eingenommen wird – als verkürzt bezeichnet werden. Eine zeitliche Abfolge von Hördingen kann einen Raum beschreiben, wie auch ein Bild einen längeren Zeitabschnitt oder einen Moment darstellen kann.

Die Gegenüberstellung von Bildender Kunst als Raumkunst und Musik als Zeitkunst, gehe laut Constanze Peres, Kunstphilosophin an der HfBK Dresden, auf den Dichter und Theologen Johann Gottfried Herder zurück (Peres 10–11). Malerei solle keine „Zeitfolge in uns erwecken“ und Musik keine Gegenstände im Raum schildern.

Dabei bezieht sich die objektive Verfaßtheit des Werkes nach Herder sowohl auf die sukzessive oder simultane Existenzweise der Werke als auch auf das, was sie darstellen: ein simultan-räumlich existierendes Werk kann bzw. soll eben nur nebeneinander Seiendes im Raum, ein sukzessiv-zeitlich existierendes Werk nur nacheinander Seiendes in und durch Zeit vorführen. (Peres 12)

Herder war der Ansicht, dass das, was sich über die „Zeitfolge erstreckt“ nur aus der Zeitfolge erkannt werden könne (Herder u. a. 36). Dieser Vorstellung nach seien Raum und Zeit zwei unterschiedliche Entitäten, die auch nur getrennt zu Kunst verarbeitet werden könnten. Diese Anschauung berücksichtigt nicht, dass Raum und Zeit zusammen gehören und auch zusammen wahrgenommen werden. Ein „Später“ kann in einer Musikmischung beispielsweise Auskunft über die Raumgröße geben. In einem zentralperspektivischen Bild endet das Weiter-Weg in der räumlichen wie zeitlichen Unendlichkeit.

Peres führt an, dass beide Künste, sowohl bildende Kunst, als auch Musik, in einem Raum-Zeit-Kontinuum zur Existenz kämen, wenn Raum und Zeit nicht als Behälter für Dinge, sondern als relativ zu den Dingen verstanden würden (Peres 15).

Dieser Gedanke greift die Überlegungen Einsteins auf, in dessen Relativitätstheorie die Vorstellung eines absoluten Raumes oder einer absoluten Zeit, wie sie Newton formuliert hatte, abgelehnt wird. Das Verhältnis der Dinge zueinander und seine Beziehung zum Wahrnehmenden schaffen das jeweilige Raum-Zeit-Kontinuum.

Ausgerechnet in der Sprache, dem Werkstoff der Semiotiker, findet sich ein Hinweis auf die Räumlichkeit der Zeit: Mit dem Begriff „Zeitraum“ wird die räumliche Komponente von Zeit deutlich. Die Zeit ist es, welchen den Raum möglich macht.

Die Überführung von Zeiterleben in Raumvorstellung betreffe die Grunddimension musikalischer Wirklichkeit, meint Oliver Krämer.

Bei der Überführung von Zeiterleben in Raumvorstellung formen wir Nachzeitigkeit in Gleichzeitigkeit und Nacheinander in Nebeneinander um. Diese Umformung der einen in die andere Erlebnisgrundlage ist Voraussetzung unseres Musikdenkens. Auf dieser Grundlage transformieren wir das Erleben von Dauern in die Vorstellung von Strecken. Wir sprechen von Zeitabschnitten und Zeitspannen. (Krämer 96)

Krämer geht davon aus, dass es sich bei der Überführung von Ereignishaft-Dynamischem in Gestalthaft-Figürliches um einen Transformationsakt handelt (Krämer 96), steht damit aber im Widerspruch zur phänomenologischen Sicht der Gestalttheorie, dernach eine solche Transformation nicht von Nöten ist um ein „Hörding“ als solches wahrzunehmen. Eine Melodie beispielsweise ist bereits ein „Hörding“ und muss nicht erst zu einem solchen transformiert werden.

Die Verbindungen zwischen Ton und Bild wurden auch von der Synästhesieforschung aufgegriffen und Analogien zwischen Lautstärke und Grösse oder zwischen Tonhöhe und wahrgenommener „Höhe“ der Hördinge aufgeführt.

Der „Tonraum“ ist auch durch Begriffe aus dem visuellen Bereich beschreibbar, Gestaltgesetze, wie beispielsweise Figur-Grund-Beziehungen, Gesetze der Ähnlichkeit, der gestaltgerechten Fortsetzung, der Nähe, des gemeinsamen Schicksals und weiteren aus der Auseinandersetzung mit Sehdingen vertrauten Gesetzmässigkeiten sind auch auf Hördinge übertragbar.

Das Eintauchen in Audiodarbietungen wird immer auch ein Raumerlebnis hervorrufen. Selbst in einer abstrakten Szene werden phänomenologische Zusammenfassungen zu „Hördingen“ stattfinden. Es genügt bereits die Bewegung eines Hördings von einer Seite des architektonischen Raumes auf die andere Seite, also beispielsweise ein Panning von einem Lautsprecher zum anderen, wenn es sich um ein 2.0-Wiedergabeverfahren handelt. Aber auch Lautheitsunterschiede zwischen gleichen Hördingen werden räumlich verstanden.

Räumliches Hören

Je nach Profession existieren unterschiedliche Definitionen von „Räumlichem Hören“: Menschen, die im weitesten Sinne der Physik verhaftet sind, gehen – im Gegensatz zu Phänomenologen – meist von einem Newtonschen Raumverständnis aus. Nach Newton befinden wir uns in einer Art Raum-Zeit-Schachtel, in der Raum und Zeit allgemeingültig sind. Räumliches Hören wäre aus Newtonscher Sicht beschreibbar durch die Ausbreitung von Schallwellen im absoluten Raum, die sich zu einem messbaren Zeitpunkt an einer bestimmten Stelle im Raum befinden und die Ohren erreichen.

Volker Schoff von der Arbeitsgemeinschaft für Forschung des Landes NRW beispielsweise versteht unter „Räumlichem Hören“ die Wahrnehmung, die auf besondere geometrische Konfiguration von Schallfeldern zurückzuführen ist (Aschoff 9). Dazu gehöre, dass Schallquellen oder Phantomschallquellen mehr oder minder genau lokalisiert werden könnten und eine Aussage über die Grösse und Beschaffenheit eines geschlossenen Raumes getroffen werden könnte, in welchem das Schallereignis stattfindet.

Diese Beschreibung kann deswegen nicht vollständig sein, da Wahrnehmung kein reiner Abbildungsprozess ist. Spätestens seit dem wir von verschiedensten medialen Hördingen umgeben sind, müssten wir uns erst darauf einigen, auf welchen Raum sich diese Angaben beziehen. Eine Tonaufnahme von einem umfallenden Mammutbaum wird, abgespielt auf einer Stereoanlage in einem kleinen Zimmer, nicht zu einer Zimmerpflanze umgedeutet werden.

Nur bei vertrauten Schällen in vertrauten Räumen sind absolute Entfernungen bestimmbar (Bruhn und Michel 652). Fehlen diese Informationen, nehmen wir Hördinge nur in Relation zueinander wahr. Auch die Beschaffenheit eines Raumes abzuschätzen, ist eine Frage des Kontextes, in welchem wir das Schallereignis verorten.

Jens Blauert definiert Räumliches Hören als Beziehung zwischen den Orten und den räumlichen Ausdehnungen der Hörereignisse, die mit anderen Ereignissen, vorwiegend Schallereignissen, aber auch physiologischen Vorgängen und Ereignissen anderer Sinnesgebiete, korrelieren (Blauert und Braasch 1). Blauerts Definition des „Räumlichen Hörens“ berücksichtigt, dass „Hören“ mehr ist, als ein Abbildungsvorgang einer – als real angenommenen – physikalischen Realität.

Ausgehend von Blauerts Ausführungen über „räumliches Hören“ haben wir es vereinfacht mit zwei Welten zu tun: Der physikalischen Welt und der wahrgenommenen Welt – dem physikalischen Raum und dem wahrgenommenen Raum, den aus phänomenologischer Sicht entscheidenden. Wird die physikalische Welt als Teil der wahrgenommenen Welt angesehen, dann kann das, was wahrgenommen wird, nie mittels physikalischer Erklärungen in Gänze beschrieben werden.

Wird von „räumlichem Hören“ gesprochen macht es demnach Sinn, den wahrnehmenden Menschen in den Mittelpunkt zu stellen.

Um die Verständigung einfacher zu machen, werden bei „Räumlichem Hören“ drei unterschiedliche Ebenen unterschieden:

- Die Frontalebene
- Die Horizontalebene
- Die Medianebene

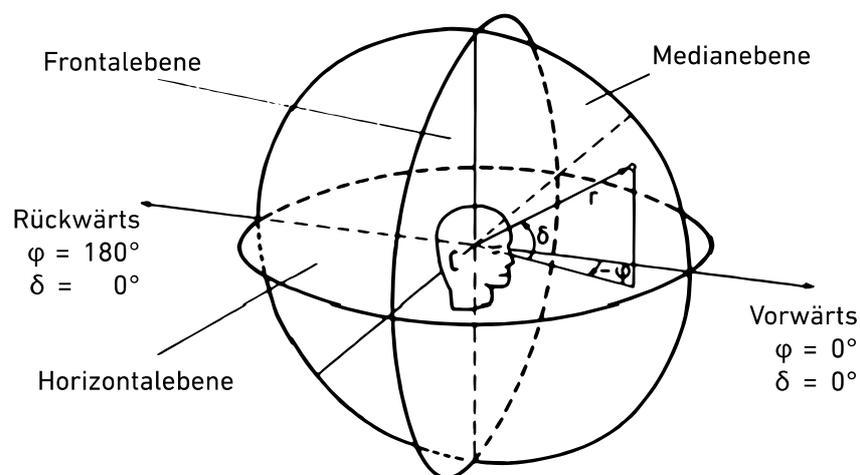


Abbildung 3: Ebenen räumlichen Hörens (Blauert, Jens, und Jonas Braasch, Seite 88)

Die Frontalebene trennt den Kopf in vorne und hinten, die Horizontalebene in oben und unten und die Medianebene in rechts und links. Ist von „Räumlichkeit“ die Rede, bezieht sich dies oft auf die Horizontalebene: Sowohl 2.0 und 5.1 sind Wiedergabeverfahren in denen sich die Schallquellen und Phantomschallquellen in der Horizontalebene verteilen. Auf die Studie von Blauert, der die Lokalisation auf der Medianebene genauer untersucht hat (Blauert 1969), wird an dieser Stelle nicht näher eingegangen. Es soll der Hinweis genügen, dass schmalbandige Schallereignisse mit Hauptfrequenzen bei 300Hz oder 3kHz von vorne, mit 8kHz von oben und mit 1kHz oder 10kHz eher von hinten lokalisiert werden (Bruhn und Michel 651).

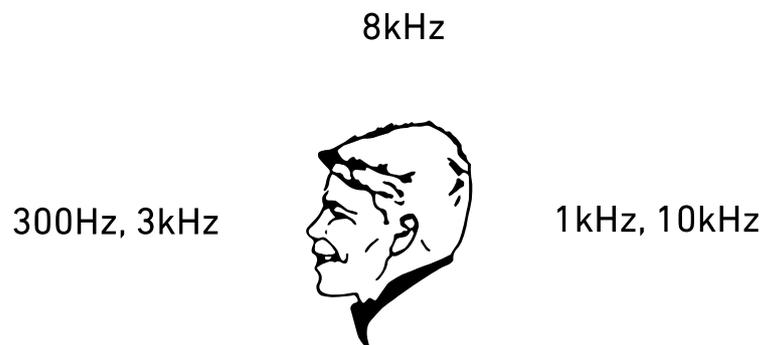


Abbildung 4: Schematic illustration of the localization of narrow-band sounds in the median plane, irrespective of the position of the sound source (Zwicker, Eberhard und Fastl, Hugo, Seite 319)

Blauert hat 1969 herausgefunden, dass schmalbandige Schallquellen je nach Frequenz entweder von vorne, von oben oder von hinten lokalisiert werden. Die von ihm entdeckten richtungsbestimmenden Bänder werden auch „Blauertsche Bänder“ genannt.

Zweiohriges Richtungshören

Unter „zweiohrigem Richtungshören“ wird das Hören mit zwei Ohren verstanden. Lokalisation einer Schallquelle erfolgt auf Grund von Laufzeit-, Intensitäts- und Frequenzgangunterschieden zwischen den am Ohr ankommenden Schallwellen.

Befindet sich eine zu lokalisierende Schallquelle nicht frontal vor dem Kopf, also in Position der Sichtachse, sondern beispielsweise auf der Horizontalebene schräg links oder schräg rechts, trifft das Signal an einem der beiden Ohren sowohl später, als auch leiser, mit geringerer Intensität, ein.

An das abgewendete Ohr wird die Wellenfront etwas später gelangen als an das zugewandte Ohr (einfach weil der Schallweg zum abgewendeten Ohr größer ist). Außerdem wird der Schall das abgewendete Ohr etwas schwächer treffen als das andere (es liegt sozusagen im Schallschatten des Kopfes). In der üblichen Terminologie spricht man in einem solchen Fall von einer Zeitdifferenz und einer Intensitätsdifferenz der Empfindungen in Bezug auf die beiden Ohren. (Gibson 112)

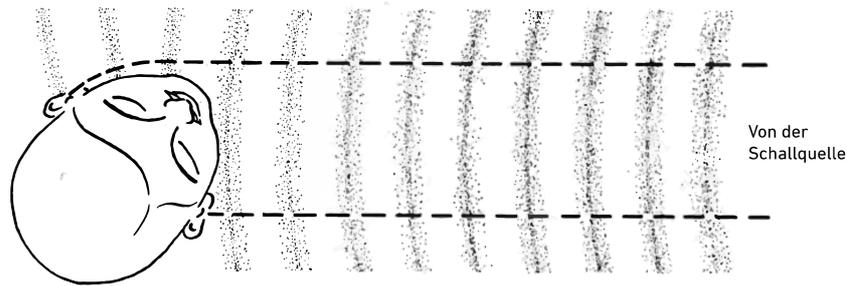


Abbildung 5: Unterschiede der Schallwellen, die die beiden Ohren erreichen
(Gibson, James J., Seite 112)

Ist die Wellenlänge gross im Vergleich zu den Abmessungen des Kopfes, ergeben sich nur geringe Unterschiede des Schalldrucks zwischen beiden Ohren (Aschoff 13).

Erst bei höheren Frequenzen, bei denen die Wellenlänge in die Größenordnung der Kopfabmessungen fällt, wird dieser Unterschied deutlich und zeigt dann insbesondere im Bereich bis etwa 50° einen gleichmäßigen Anstieg mit der Schalleinfallrichtung. (Aschoff 13)

Der Pegelunterschied spielt demnach nur bei hohen Frequenzen eine Rolle.

Oberhalb von 1000 Hz entspricht die Wellenlänge des eintreffenden Schalls der Grösse des Wegunterschiedes. Diese Uneindeutigkeit wirkt sich insbesondere auf die Ortbarkeit periodisch gleichförmig verlaufender Signale höherer Frequenzen negativ aus.

Sind die Flanken der Signale steil, wie bei Explosivlauten der Sprache, ist die Laufzeit das entscheidende Kriterium für Richtungshören. Es zeigt sich eine gleichförmige Abhängigkeit des Einfallswinkels in Bezug zur Laufzeit über den gesamten Frequenzbereich (Aschoff 13).

Einohriges Richtungshören

Auch mit einem einzigen Ohr ist Richtungshören möglich. Ab Frequenzen, die grösser sind als 1,5 kHz, werden schmalbandige Reflexionen an der Ohrmuschel, die den Frequenzgang am Trommelfell verändern, als Richtungsinformation gedeutet (Jecklin 10). Da jeder Mensch unterschiedliche Ohrmuscheln besitzt, ist das entstehende Frequenzmuster individuell verschieden. Diese ohr- und richtungsabhängigen Überhöhungen oder Absenkungen in der Frequenzkurve werden als Head-Related Transfer Function bezeichnet.

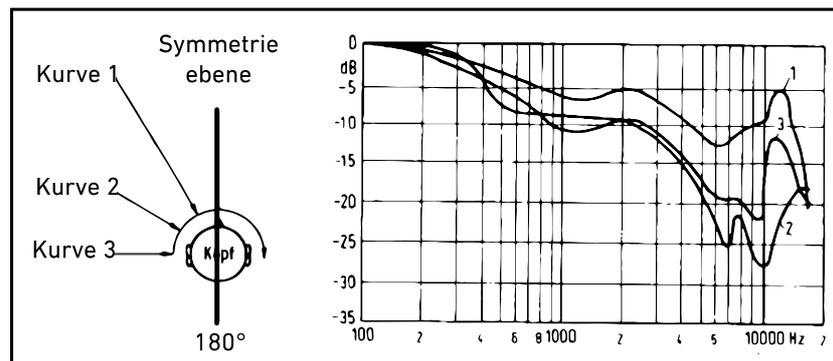


Abbildung 6: Einohriges Richtungshören (Jecklin, Jürg, Seite 12)

Blauert führt aus, dass leichte Kopfbewegungen zu einer Veränderung der am Ohr ankommenden Signale führen und daraus zusätzliche Richtungsinformationen resultieren (Blauert und Braasch 2). Laut Gibson ermöglicht das Ungleichgewicht an den Reizeingängen sich auf eine Schallquelle auszurichten (Gibson 113). Der Kopf wird in die Richtung gedreht, in der wir bereits eine Schallquelle lokalisiert haben.

Monaurale Tiefenreize

Auch beim einohrigen Hören ist es möglich, unterschiedliche räumliche Tiefen wahrzunehmen. Ein Beispiel für einen „monauralen Tiefenreiz“ ist das Leiserwerden eines Tons mit zunehmender Entfernung der Schallquelle (Herkner 78).

Daraus lässt sich schlussfolgern, dass Räumlichkeit, die durch die Richtungsinformationen links-rechts, vorne-hinten und oben-unten beschreiben werden kann, nicht allein vom Vorhandensein zweier Ohren abhängt. Unabhängig des Raumes, der sich in unserer Vorstellung bildet, wenn wir Medien wahrnehmen, kann bereits auf der Seite physikalischer Schallübertragung die Notwendigkeit einer „mehrkanaligen Schallübertragung“ für die Übermittlung von Rauminformation verneint werden.

Ungeachtet dessen ist in der Geschichte der Tonproduktions- und Wiedergabetechnik ein Hin zu noch mehr Wiedergabekanälen feststellbar.

Physikalische Grundlagen, Wiedergabeverfahren

Mittlerweile gibt es die unterschiedlichsten Wiedergabeverfahren: Mono, Stereo, Quadrophonie, 5.1, 7.1, Auro 3D, etc. Unterschieden wird in der Regel zwischen einkanaligen-, zweikanaligen und mehrkanaligen Wiedergabeverfahren.

Der Schweizer Tonmeister Jürg Jecklin vergleicht die Wirkung der unterschiedlichen Wiedergabeverfahren und schreibt, Mono gleiche einem Loch in der Wand, und Stereo der Wiedergabe von einer Loge aus (Jecklin 12–13). Er vertritt die Ansicht, dass ein echter Raumeindruck nur mit einer mehrkanaligen Wiedergabe erreicht werden könne. Ähnlich äussert sich Hubert Henle: mit der Stereotechnik sei das zweidimensionale, räumliche Hören in die Studiotechnik eingeführt worden (Henle 112). Sowohl Jecklin, als auch Henle klammern aus, dass Raum nicht allein durch das Vorhandensein von Signalen in einem Wiedergaberaum entsteht, sondern für die Wahrnehmung entscheidend ist, welches Raumes sich ein Mensch bewusst wird.

Menschen, die einen Roman lesen, werden trotz der Flächigkeit jeder einzelnen Seite dennoch der Ansicht sein, dass die Geschichte dreidimensional erlebt wird. Menschen werden bei der Betrachtung von Architektur Fotografien davon ausgehen, dass sie das, was sie sehen, dreidimensional ist. Räumlichkeit ist das Ergebnis sowohl semiotischer Verknüpfungen - als codierte Rauminformation, die Mittels Zeichen übertragen wird - als auch der Gestaltwahrnehmung.

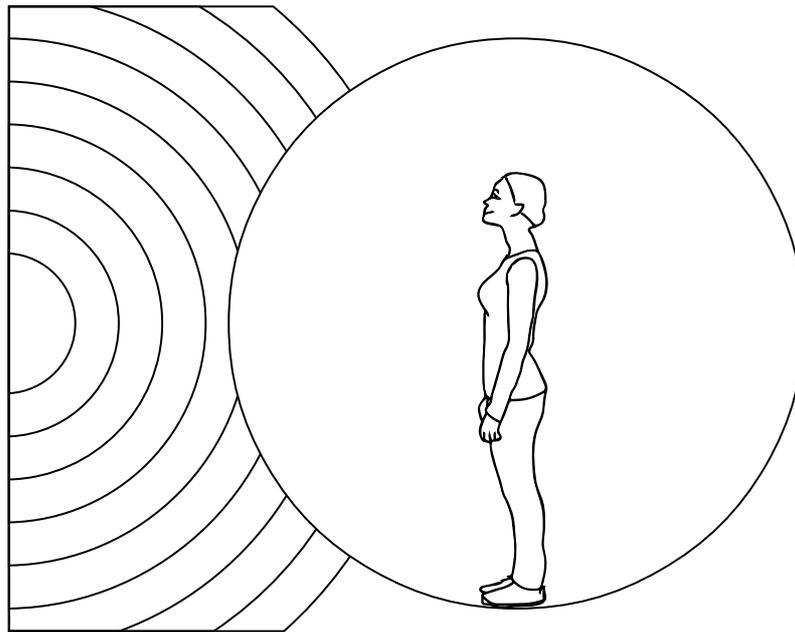


Abbildung 7: Physikalischer und wahrgenommener Raum sind zweierlei
(eigene Grafik)

Veranschaulichung des mentalen Raumes eines Menschen.

Die rechte „Kugel“ soll verdeutlichen, dass der wahrgenommene Raum eines Menschen ein anderer ist, als der physikalische Raum den ein Bild- oder Tonmedium skizziert.

Die Trennung zwischen Wiedergabesystemen, also der Konzentration auf die „Physik“ einerseits und der Wahrnehmung, dem „Hören“ andererseits, und das Anerkennen, dass zwischen beiden Betrachtungen Unterschiede bestehen, sind die Voraussetzungen für eine phänomenologische Auseinandersetzung mit Audioereignissen. Wenn von Wiedergabesystemen gesprochen wird, sollte dies nicht mit „Hören“ gleichgesetzt werden. Mono, Stereo und mehrkanalige Systeme wie 5.1 beschreiben Wiedergabeverfahren. „Hören“ ist etwas anderes: Hördinge zu einem sinnvollen Wahrnehmungsraum zusammenzufassen.

Zu Unterscheiden ist auch zwischen Aufnahmeverfahren und Wiedergabeverfahren. Aussagen über Aufnahmeverfahren beziehen sich beispielsweise auf die Art und Weise der Mikrofonierung (Stereofon oder Surround-Sound, Koinzidenzstereophonie oder Laufzeitstereophonie, etc.) oder ob – bei Popproduktionen üblich – ein Mehrspurverfahren zum Einsatz kommt, bei dem das Stereobild oder ein Surround-Mix erst im Studio entsteht, indem die einzelnen Klänge per panpot im Hörbild verteilt werden.

Wiedergabeverfahren beziehen sich auf die Anordnung und Anzahl der Lautsprecher im Raum und auf das, was mit ihnen wiedergegeben wird.

Aufnahmeraum, Wiedergaberaum und wahrgenommener Gestaltraum sind drei unterschiedliche Räume. Der semiotische Raum, der auf Assoziationen zu kulturell Übermitteltem basiert, ist ein weiterer.

Mono

Wird bei Wiedergabeverfahren von Mono gesprochen, ist damit grundsätzlich die Wiedergabe über eine einzige Lautsprecherbox gemeint. Alle aufgezeichneten Schallinformationen ertönen aus einer gemeinsamen Schallquelle.

Während bei mehrkanaligen Wiedergabeverfahren eine Konkurrenz zwischen der Räumlichkeit des Schallfeldes und der wahrgenommenen Räumlichkeit besteht, entsteht bei einkanaligen Wiedergabeverfahren Raum einzig auf der Wahrnehmungsseite. Die Möglichkeit Raum zu gestalten, ist beschränkt. Raumbedeutung resultiert verstärkt aus dem Sinnzusammenhang des Gehörten. Bereits vor der Einführung mehrkanaliger Wiedergabeverfahren wurde auf die gestalterische Wirkung von Hall, sowie durch Änderungen des Frequenzspektrums hingewiesen (Arnheim und Diederichs 38).

Frequenz	125Hz	250Hz	500Hz	1000Hz	2000Hz	4000Hz	8000Hz	16000Hz
Absorption in db/100m	0,035	0,075	0,18	0,42	1,1	2,7	10	35

Abbildung 8: Luftabsorption bei 50% Luftfeuchtigkeit und 20° Umgebungstemperatur (Andreas Friesecke, Seite 144)

Höhere Frequenzen werden in der Luft im Vergleich zu niedrigeren Frequenzen mit zunehmender Entfernung stärker gedämpft.

Entfernt sich ein Hörereignis, dann ändert sich – neben der wahrgenommenen Lautstärke – das Frequenzspektrum. Zuerst verliert sich die Tiefenanhebung im sogenannten Nahfeld, danach nimmt der Anteil hoher Frequenzen stetig ab (Friesecke 144). In einem Raum ändert sich das Verhältnis von Direktschall zu Diffusschall: Der Diffusschall, der „Hall-Anteil“, nimmt zu. Unter „Hallradius“ wird die Position im Raum verstanden, an dem der Pegel des Diffusschalls und des Direktschalls gleich gross sind.

Eine Hörbild, das Mono wiedergegeben wird, enthält Rauminformation.

2.0

2.0 bzw. Stereo bedeutet, dass für die Wiedergabe zwei Lautsprecherboxen oder Kopfhörer verwendet werden. Auf die kopfbezogene Stereophonie, die sich von der raumbezogenen Stereophonie dadurch unterscheidet, dass der Schall der beiden Kopfhörermuscheln nicht das jeweils gegenüberliegende Ohr erreichen, wie das bei der raumbezogenen Stereophonie der

Fall ist, soll hier nicht näher eingegangen werden.

Beim Stereo-Wiedergabeverfahren wird ausgenutzt, dass Menschen Phantomschallquellen wahrnehmen. Dies ist dann der Fall, wenn die Schallwellen, die das rechte und linke Ohr erreichen, gleich sind. Voraussetzung ist, dass die Schallwellen beider Lautsprecher korrelieren (Bruhn und Michel 652). Wenn an beiden Ohren Schall eintrifft, der sich weder in Klang, noch in Phasenlage unterscheidet, nehmen Menschen eine virtuelle Schallquelle wahr, die sich mittig zwischen beiden Lautsprechern befindet. Dies ist der Fall, wenn die Abhörposition so gewählt ist, dass beide Lautsprecher gleich weit entfernt sind. Bewegt sich der Zuhörer nach links oder rechts, wandert eine Phantomschallquelle, die in der Mitte wahrgenommen wird, mit. Man spricht auch von einer instabilen Mitte.

Die Phantomschallquelle wird auch dann nicht mehr mittig gehört, wenn entweder die Wiedergabe auf einer Seite verzögert wird (Laufzeit/Phase) oder die Lautstärke einseitig verändert wird (Intensität/Pegel). Dieser Effekt wird dafür genutzt, um virtuelle Schallquellen im Stereopanorama verteilen zu können.

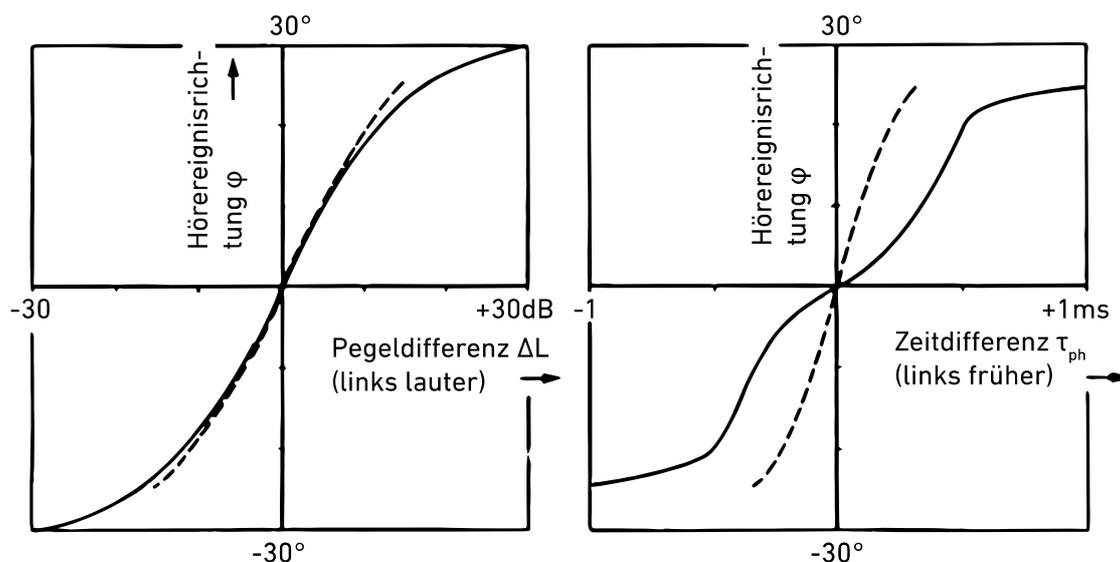


Abbildung 9: Summenlokalisationskurven für Pegel- bzw. Laufzeitdifferenzen der beiden Lautsprechersignale. Durchgezogene Linie: Zuhörerkopf fixiert; gestrichelte Linie: Kopf frei beweglich (Blauert, Jens, und Jonas Braasch, Seite 103, nach Wendt 1961)

Links: Hörereignisrichtung abhängig von Pegeldifferenz

Rechts: Hörereignisrichtung in Abhängigkeit zur Laufzeit

1. Pegel

Um eine Phantomschallquelle beim 2.0-Wiedergabeverfahren einmal von rechts nach links wandern zu lassen muss ein Pegelunterschied von ungefähr 18db zwischen beiden Lautsprechern bestehen. Ist beispielsweise der linke Lautsprecher 18db lauter als der rechte, wird die Phantomschallquelle ganz links auf der Stereobasis wahrgenommen. Einer Winkeländerung von 1° entspricht bei einer optimalen Abhörposition im Stereodreieck einer Änderung des Pegels um 0,6 dB.

2. Laufzeit

Auch bei einer Laufzeitdifferenz von 1,5 ms zwischen linkem und rechtem Lautsprecher ist eine Phantomschallquelle vollständig auf einer Seite der Stereobasis zu hören (Sengpiel, „Praktische Daten zur Lokalisation von Phantomschallquellen bei Intensitäts- und Laufzeit-Stereofonie - PraktischeDatenZurStereo-Lokalisation.pdf“ 1). Die wahrgenommene Schallquelle wandert in die Richtung des nicht-verzögerten Signals. Eine Winkeländerung von 1° entspricht einer Verzögerung von 0,05 ms. Steigt der Wert des Laufzeitunterschiedes zweier ansonsten identischer Lautsprechersignale über 30-40 ms, werden die Signale getrennt – und damit nicht mehr als Phantomschallquelle – wahrgenommen. Man spricht von der „Echoschwelle“.

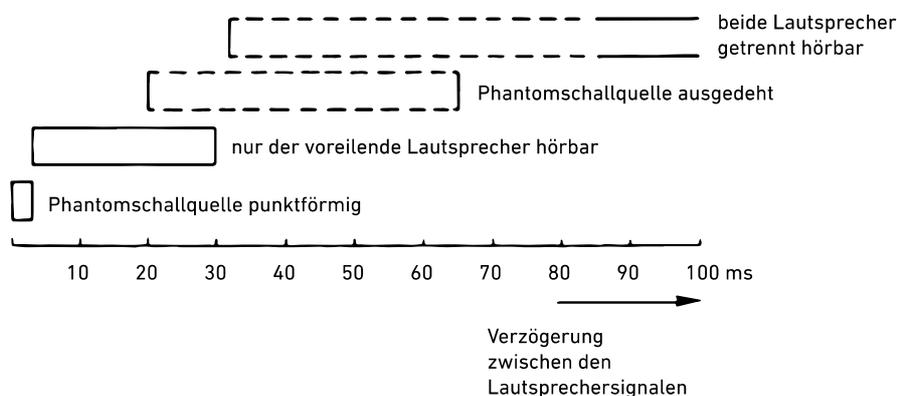


Abbildung 10: Wirkung der Zeitverzögerung zwischen identischen Signalen bei Stereowiedergabe (Dickreiter, Michael, Seite 129)

5.1

Die Ursprünge des 5.1-Wiedergabeverfahrens gehen auf die Weiterentwicklung des Tonfilms zurück. Die Mono-Wiedergabe durch einen einzigen Lautsprecher ist, auf Grund der Stabilität der Lokalisation diskreter Schallquellen und der relativ guten Übereinstimmung der Richtungsinformation von Filmtone und Filmbild, dafür geeignet, Dialoge wiederzugeben. Schon bei den ersten Überlegungen zur Einführung mehrkanaligen Tons im Kino, wurde daran gedacht, zwei Lautsprecher - links und rechts der Mitte - als Ergänzung hinzuzufügen (Henle 113). Im Jahr 1974 wurde der Kinostandard „Dolby Stereo“ eingeführt, bei dem auf zwei Aufnahmespuren vier Kanäle matriziert wurden: Links, Rechts, Center und Hinten (Heer 41).

Die „Society of Motion Picture and Television Engineers“ entwickelte 1987 die Grundlage für den 5.1.-Standard. Seitdem gilt, dass bei der Tonwiedergabe mit Bildbezug (Kino/TV), eine Mindestanzahl von Lautsprechern Verwendung finden soll, um ein ausreichendes Klangerlebnis zu gewährleisten (Kerins 330).

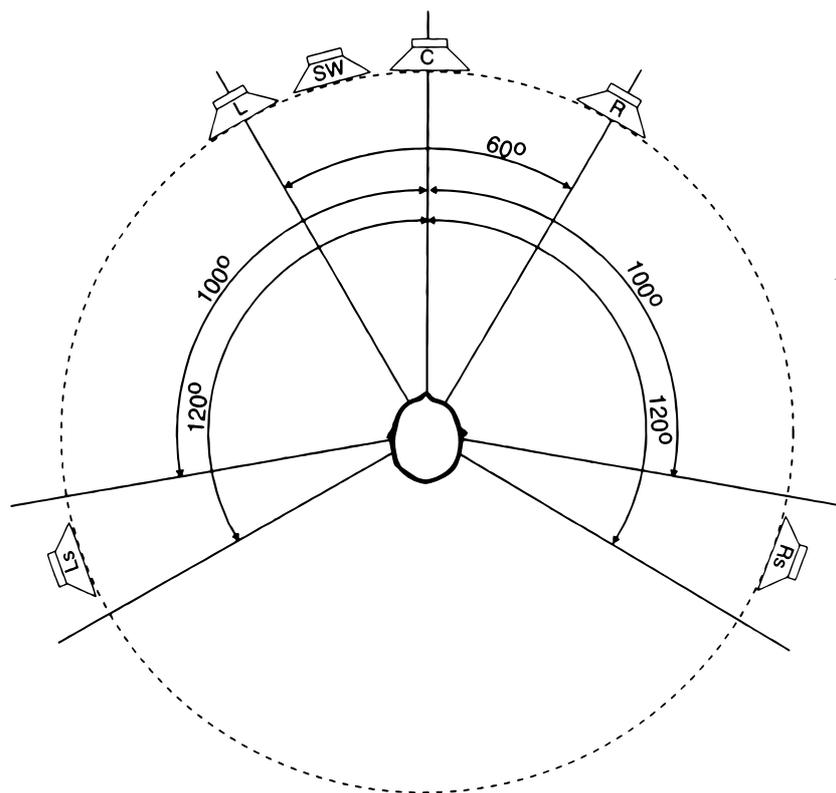


Abbildung 11: Lautsprecheranordnung nach ITU-R BS.775-1 (Henle, Hubert, Seite 17)

Das 5.1.-Wiedergabeverfahren besteht aus 6 Lautsprechern, von denen 5 im Raum verteilt sind, drei in Sichtrichtung (links, center, rechts), zwei als zusätzliche Lautsprecher entgegen

der Sichtrichtung (links hinten und rechts hinten) und ein weiterer Lautsprecher, der als LFE (Low Frequency Effects bzw. Low Frequency Enhancement) oder, je nach Einsatzzweck, als SW (Subwoofer) bezeichnet wird. Dieser unterhalb des Centers platzierte Lautsprecher ist für die Wiedergabe tiefer Frequenzen ausgelegt und wird bei Kinomischungen vorwiegend als Effekt-Lautsprecher eingesetzt (LFE), bei Musikmischungen als Subwoofer.

Der Center-Lautsprecher soll der Stabilisierung der Mitte dienen, die hinteren beiden Lautsprecher werden genutzt, den Raumeindruck zu verbessern.

Die Bildung von Phantomschallquellen zwischen seitlichen angeordneten Lautsprechern ist nicht, oder nur schwer möglich (Blauert und Braasch 17).

Wo ist hinten? Wo ist vorne?

Dass eine Summenlokalisation von seitlich angeordneten Lautsprechern nur schwer möglich ist, lässt sich auch phänomenologisch erklären.

Ausgehend von der Gestalttheorie sind „Sehdinge“ nicht gleichzusetzen mit der physikalisch vorhandenen Materie, die sich beispielsweise auf einem Blatt Papier befindet. Bereits die Tatsache, dass wir einen gezeichneten Würfel nicht als plattes Viereck, das von ebenso platten Parallelogrammen umgeben ist, wahrnehmen, verdeutlicht, dass räumliche Koordinatensysteme, je nach Betrachtung, auseinanderfallen.

Würde man bei einem Bild ausmessen, wo sich denn da der Hintergrund befindet, würden wir feststellen, dass er sich physikalisch genauso auf der Oberfläche eines auf Papier gezeichneten Bildes befindet, wie der Vordergrund. Der gestalterische Hintergrund ist das, was wir als weiter entfernt wahrnehmen, der Vordergrund ist uns nah.

Wird diese Überlegung auf das Hören übertragen, entsteht, vor allem bei mehrkanaligen Wiedergabeverfahren wie 5.1., die Frage, was für ein „Hinten“ nun gemeint ist: Das Hinten des Hörbildhintergrunds oder das Hinten innerhalb des Wiedergaberaumes. Beide Hinten fallen bei mehrkanaligen Wiedergabesystemen auseinander. Der Hörbildhintergrund wird in der Regel von den vorderen und hinteren Lautsprechern wiedergegeben. Bei 2.0-Aufnahmen ist der Tonbildhintergrund immer vorne, da die Lautsprecher ja alle vor uns stehen.

Der dargestellte Hintergrund einer 5.1.-Mischung wird sich in optimaler Abhörposition, also im Sweet-Spot kreisförmig um uns herum befinden, der Vordergrund des Hörbildes befindet sich vor uns.

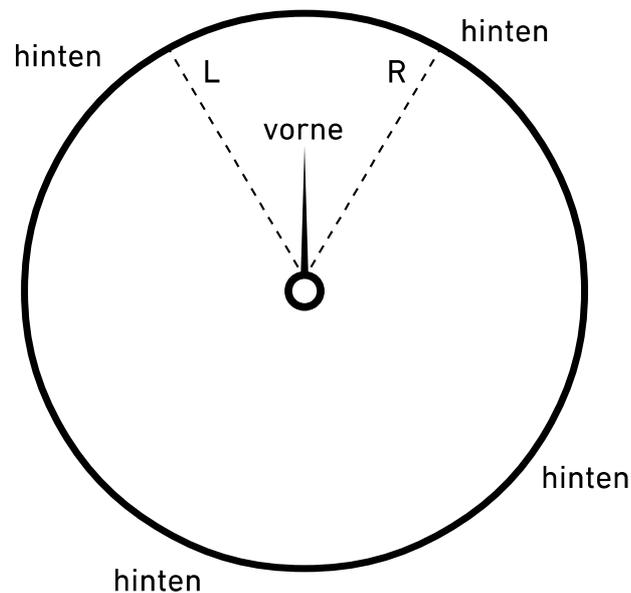


Abbildung 12: Wo ist Hinten? (Eigene Grafik)

Es ist kein Zufall, dass die rückwärtigen Lautsprecher vorwiegend zur Umhüllung verwendet werden, sei es als Verstärkung diffuser Hallanteile bei Klassikmischungen, oder für Sound-Atmosphären in Kinoproduktionen.

An dieser Stelle sei auf einen Gedanken von Lenelis Kruse verwiesen:

Das Wahrgenommene verweist auf das Wahrnehmbare, der aktuelle Wahrnehmungsraum auf den Potentiellen. (Kruse 117)

Raumwahrnehmung kann demnach auch als ein Von-Hier-Nach-Da verstanden werden. „Vorne“ wäre hier und „Hinten“ später. Ausgehend davon, dass dieses „Später“ vom jetzigen Standpunkt, egal in welche Richtung ich geistig blicke, seine grundsätzliche Eigenschaft nicht verändert, kann abgeleitet werden, dass jeder wahrgenommene Raum rund ist. Der „Hintergrund“ ist das Entfernte, das uns umgibt, und für uns dennoch nie greifbar ist.

Hall

Bei der Darstellung und Gestaltung von Räumlichkeit ist auch der auch Nachhall zu berücksichtigen. Man unterscheidet zwischen Direktschall, frühen Reflexionen und dem diffusen Nachhallfeld. Frühe Reflexionen sind die ersten an Wänden und Gegenständen im Raum zurückgeworfenen Schallinformationen und geben entscheidend über die wahrgenommene Raumgröße Auskunft. Ob Musikdarbietungen als angenehm empfunden

werden, hängt vor allem von den frühen seitlichen Reflexionen ab. Diese beschreiben die räumliche Ausdehnung der Hörereignisse (Bruhn und Michel 653). Die frühen Reflexionen treffen in etwa spätestens bis 100 ms nach dem Direktschall beim Hörer oder am Mikrofon ein (Görne 88).

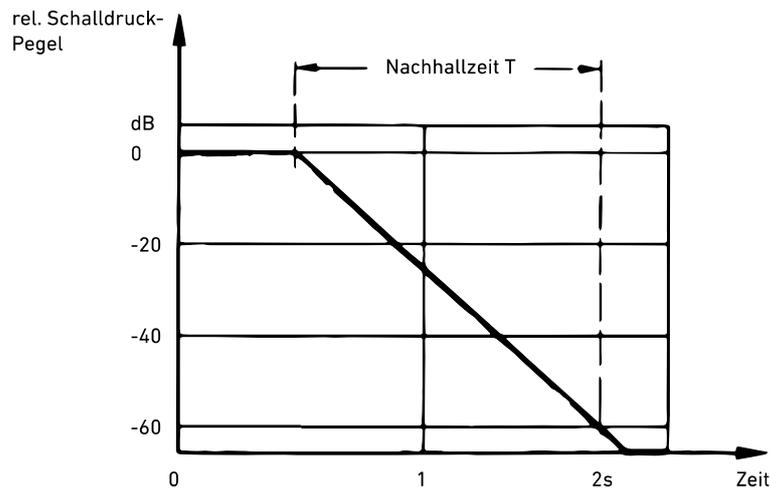


Abbildung 13: Definition der Nachhallzeit durch den Abfall des Schalldruckpegels nach Abschalten der Schallquelle (Dickreiter, Michael, Seite 31)

Auf die frühen Reflexionen folgt ein diffuses Nachhallfeld, das je nach Raumbeschaffenheit kürzer oder länger sein kann. Auch das Frequenzspektrum des Nachhalls gibt Informationen darüber, wie der dargestellte Raum beschaffen ist. Die Nachhallzeit wird häufig in T_{60} angegeben, der Zeit (T =time) in der die Schallenergie auf ein Millionstel, bzw. der Schallpegel um 60 dB - einem Tausendstel des ursprünglichen Wertes - gefallen ist (Görne 83).

Bei monauralen Schallaufnahmen ist das Verhältnis von direktem Schall zu diffusem Schall ein wichtiger Indikator für den Abstand zur Schallquelle. Mit zunehmender Entfernung nimmt der Anteil an diffusem Schall gegenüber dem Direktschall zu. Sind in einer Audiodarbietung zwei in dem Anteil der jeweiligen Diffusität unterscheidbare Signale zu hören, wird derjenige als entfernter eingeschätzt, der diffuser wahrgenommen wird (Aschoff 34). Dinge im Raum können als „entfernt“ oder „nah“ wahrgenommen werden. Voraussetzung ist, dass die Diffusität deutlich unterschiedlich ist (Jecklin 11). Selbst bei einer monophonen Audioproduktion können unterschiedliche Positionen im Raum dargestellt werden: Der Anteil an Diffusität dient als ein Merkmal der relativen Entfernung, die Klangfarbe des Nachhalls gibt Auskunft über die Raumbeschaffenheit und die frühen Reflexionen, wie bei mehrkanaligen Produktionen auch, dienen als Merkmal der Raumgröße.

Hall ist ein gestalterisches Mittel, um Vordergrund und Hintergrund voneinander abzugrenzen.

Laut Jecklin bestimmt Halligkeit den Raumeindruck und ermöglicht, in einen anderen Raum „hineinzuhören“ (Jecklin 12–13). Wenn Immersion bedeutet, bewusst von einem, dem architektonischen Raum, in einen anderen Raum - dem Hörbild - zu wechseln, ist Hall auch ein immersives Gestaltungsmittel.

Bei 5.1.-Wiedergabeverfahren werden über die Lautsprecher, die sich im Abhörraum entgegen Sichtrichtung hinter dem Zuhörenden befinden, insbesondere bei Klassikaufnahmen, Rauminformationen wiedergegeben. Bei Klassikaufnahmen sind die 5.1.-

Mikrofonanordnungen nicht mitten im Orchester angebracht, sondern hängen, ähnlich wie bei 2.0, immer noch räumlich davor. Selbst bei 5.1. Kinomischungen werden die Surround-Lautsprecher nicht für die Haupttoninformationen verwendet, sondern dienen dem Zweck der Verräumlichung des Höreindrucks.

Wahrnehmungshierarchie

Sowohl bei Surround-Kinomischungen als auch bei 5.1-Klassikaufnahmen findet die Wiedergabe in Sichtrichtung, entweder in Richtung der Leinwand, oder aber in Richtung der dargestellten Bühne, statt. Daraus folgernd besteht bei beiden Hauptanwendungen für Surround-Mischungen eine klare Wahrnehmungs-Hierarchie: Der Zuhörer hört „in Richtung“ einer Darstellung und befindet sich in der Regel nicht in einem Abhörraum, in dem er diese Richtung frei wählen kann.

Untersuchung

Um sowohl Unterschiede hinsichtlich der Immersion zwischen 2.0 und 5.1-Wiedergabesystemen zu untersuchen, als auch Hinweise darauf zu finden, inwiefern Gestaltwahrnehmungsunterschiede zwischen beiden Verfahren vorliegen, wurde eine Studie durchgeführt.

Konzept

Bei der Planung der Studie wurde berücksichtigt, dass physikalische Messungen allein nicht geeignet wären, um Effekte der Gestaltwahrnehmung zu erfassen – insbesondere dann nicht, wenn davon ausgegangen wird, dass zwischen physikalischen Gegebenheiten und der Wahrnehmung von Hördingen Unterschiede bestehen.

Weitere Vorüberlegungen und Annahmen waren:

- Wenn Immersion den Übertritt von einem Raum in den anderen bedeutet und Bewegung und Zeit Grundkomponenten des Raumes sind, sind Unterschiede der Immersion Resultat von Qualitätsunterschieden innerhalb des wahrgenommenen Hörbildes.
- Es wird vermutet, dass eine Änderung der Geometrie des dargestellten Raumes eine Änderung der Immersion bedeutet. Inkonsistenzen der Form, Paradoxien und Verzerrungen des wahrgenommenen Raumes erschweren die Immersion.
- Physikalische Unterschiede zwischen 2.0 und 5.1 sind bekannt. Dazu gehört u.a. die Verschiedenartigkeit von Phantomschallquellen und diskreten Schallquellen.
- Es ist bekannt, dass der Sweet-Spot bei 2.0 klein ist (ca. 20cm breit) aber es existieren kaum Untersuchungen darüber, was ausserhalb des Sweet-Spots mit dem Hörbild geschieht bzw. wie sich die Form des Raumes verändert, wenn Zuhörer sich ausserhalb des Sweet-Spots befinden.

Ausserdem wird angenommen, dass für die Entstehung einer Raumvorstellung nicht jede Information diskret vorliegen muss, da bei der Wahrnehmung - auf Grund der Wirkung der Gestaltgesetze - Lücken in der Darstellung sinnvoll ergänzt werden.

Da die Auswahl an Studien, die sich mit der Gestaltwahrnehmung bei 2.0 und 5.1-Wiedergabeverfahren in Abhängigkeit der Position zum Sweet-Spot auseinandersetzen, sehr gering ist, stellt die durchgeführte Untersuchung keine Weiterentwicklung etwa schon bekannter Forschungsvorhaben dar, sondern wurde ausgehend der Kombination von Kenntnissen der Gestalttheorie und vorhandenem Wissen im Bereich Tontechnik entwickelt.

Beschreibung der Untersuchung, Untersuchungsmethodik

In einem Vergleich zwischen der Wiedergabe 2.0 und 5.1 wurde untersucht, wie sich – in Abhängigkeit zum Wiedergabeverfahren – das Hörbild und die Geometrie einer wahrgenommenen Szene, sowie das Gefühl der Eingebundenheit des Zuhörers, verändert, wenn sich dieser nicht mehr im Sweet-Spot befindet, sondern auf einer Achse frontal vor dem rechten bzw. linken Lautsprecher.

Vorstudie

Vor der Untersuchung wurde ein Vorstudie in kleinem Rahmen durchgeführt. Die Vorstudie fand in einem privaten Tonstudio statt. Die Ergebnisse der Vorstudie flossen in das Studiendesign ein. Im Unterschied zur Vorstudie wurde die Gestaltung der Szene für die Hauptstudie optimiert.

Beschreibung

Für die Hauptstudie wurde ein Hörbild mit eindeutiger Trennung zwischen Figur und Grund verwendet: Die Aufnahme eines Sprechers, der von 1 bis 10 zählt und in gleichmässigem Tempo und gleicher Intensität von links nach rechts bzw. rechts nach links gepannt wird, während, mit vermindertem Pegel und 100 Prozent Hall-Anteil, auf den vorderen beiden Lautsprechern (2.0) bzw. den vorderen und hinteren Lautsprechern (5.1), Signale wiedergegeben wurden, welche eine sich im Raum befindende Menschengruppe darstellten.

Die Versuchsteilnehmer sollten ausgehend von der Sitzposition (links, mitte und rechts) und Wiedergabeverfahren (2.0 und 5.1) entscheiden, welchen Weg die gehörte Zahlenreihe nimmt, und wo die einzelnen Zahlen auf der Stereobasis zu hören sind. Zusätzlich sollten die Teilnehmer Angaben darüber machen, ob ihnen die Beantwortung leicht oder schwer gefallen ist, sie das Gefühl hatten, in die Szene eingebunden zu sein und entscheiden, ob Vordergrund und Hintergrund sich in einem gemeinsamen Hörbild befinden.

Versuchsordnung

Der Versuch wurde im Regieraum A im Tonstudio U54 der Hochschule der Medien in Stuttgart durchgeführt. Der Raum ist 16,34 m² gross. In der Regie befindet sich eine Studer Vista 7 Mixingkonsole, ein Geräterack, eine DAW mit den Systemen ProTools HD (Avid) und Sequoia (Magix). Im hinteren Bereich des Regieraumes ist ein Sofa aufgebaut.

Als Abhöre für die Untersuchung kamen 5 aktive ADAM S2.5A Nahfeldmonitore zum Einsatz, welche gemäss ITU-R BS.775 angeordnet sind. Da die Szene keine tieffrequenten Special-FX enthält, wurde auf den Einsatz eines LFE verzichtet. Der Abstand zwischen den beiden Lautsprechern L und R beträgt 2,10 m.

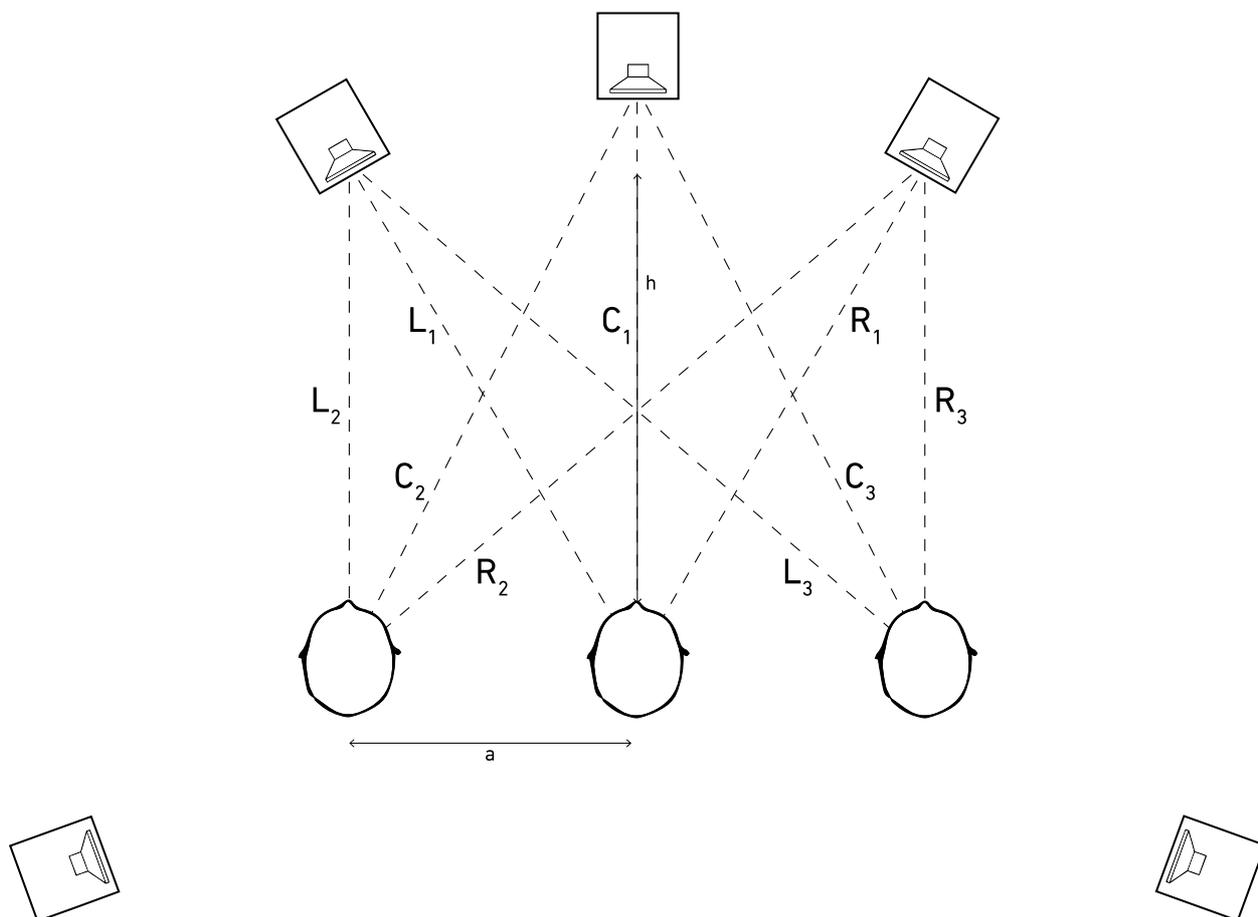


Abbildung 14: Versuchsaufbau, Draufsicht (eigene Grafik)

Es wurden drei unterschiedliche Sitzpositionen festgelegt:

- Mitte: Im Sweet-Spot
- Links: Auf einer Achse frontal vor dem linken Lautsprecher
- Rechts: Auf einer Achse frontal vor dem rechten Lautsprecher

Die Szenen wurden mit dem Produktionsprogramm Sequoia wiedergegeben und im Studer Vista-Pult so geroutet, dass jeweils 2 Lautsprecher (Stereo), oder 5 Lautsprecher (5.1) angesprochen wurden.

Die Oberfläche des Studer-Mischpults war während der Untersuchung abgeschaltet, um keine visuellen Hinweise darauf zu geben, ob eine Stereo- oder Surroundmischung zu hören ist. Gleichzeitig wurde damit eine unbeabsichtigte Veränderung des Mischungsverhältnisses, z.B. auf Grund des Verstellens eines Faders durch die Probanden, verhindert.

Das Szenendesign

Beim Design der Szene wurden die Gestaltgesetze berücksichtigt, insbesondere die Aufteilung in Figur und Grund, sowie das Gesetz der Nähe, Ähnlichkeit und des gemeinsamen Schicksals.

Die Szene bestand aus folgenden Elementen:

1. Figur

Eine Stimme war zu hören, die folgendes sagt:

„1-da-da-2-da-da-3-da-da-4-da-da-5-da-da-6-da-da-7-da-da-8-da-da-9-da-da-10“

Der Abstand zwischen den Zahlen 1-10 und dem „da“ wurde gleich gross gewählt. Der sich ergebende Rhythmus entspricht einem $\frac{3}{4}$ -Takt. Um keine semiotische Fehlverknüpfung zu erzeugen, wurde die Silbe „da“ als Füllwort ausgewählt, welche im Zweifelsfall entweder als „dort“ decodiert worden wäre, was der Aufgabenstellung - der Lokalisierung - entsprochen hätte, oder als „dada“, einem Hinweis auf die Kunstrichtung des frühen zwanzigsten Jahrhunderts, welche den Anspruch hatte, Nicht-Sinnhaftes hervorzubringen. Das Füllwort „da“ eignete sich für die Verwendung als Füllwort auf Grund der relativ sinnfreien oder stimmigen semiotischen Deutungsmöglichkeit mehr, als beispielsweise „pi“ oder „pa“.

Die Füll-Laute zwischen den Zahlen ergaben ein zusammenhängendes Hörding ausreichender Grösse, was die Bewegung von einer Seite bis zur anderen Seite des Hörbildes in ausreichend langsamer Geschwindigkeit – ohne allzu grosse Lücken – ermöglichte. Gleichzeitig wurde dadurch sichergestellt, dass die Probanden sich nur auf die Position von 10 Zahlen innerhalb des Bildes konzentrieren mussten und nicht, wären keine Fülllaute verwendet worden, auf 28.



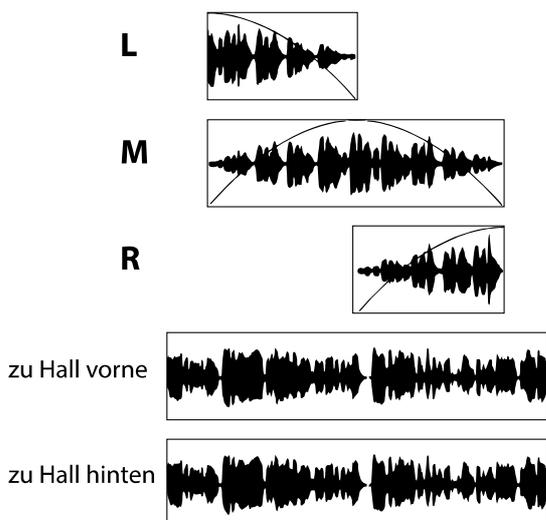
Abbildung 15: Abfolge der Zahlen und der "Da Da" (eigene Grafik)

- Das Gesetz der Nähe: Abstand zwischen Zahlen und „da“ gering
- Gesetz der Ähnlichkeit: Eine Silbe, Zahlen und „da“ gleich gesprochen
- Gesetz des gemeinsamen Schicksals: Bewegung von einer Seite zur anderen

Die Stimme wurde durch Dazumischen eines zum (Hinter-)Grund passenden Mono-Halls (90% dry / 10% wet) etwas mehr in das Bild integriert.

Die Figur kam in der Untersuchung in der Reihenfolge links → rechts und gespiegelt von rechts → links vor. Als Panninglaw wurde „konstante Leistung“ gewählt, welches bei 2.0 innerhalb des Sweet-Spots eine gleichförmige Bewegung von einer Seite zur anderen Seite ergeben sollte. Die Mischung bei 5.1. orientierte sich an derselben Idee: Hier wurde das Signal von Links zur Mitte und von der Mitte nach Rechts gleichförmig gepannt.

5.1



2.0

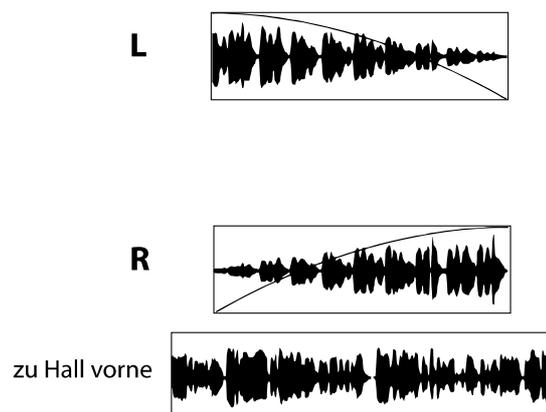


Abbildung 16: Spuren der 2.0 und 5.1-Produktion (eigene Grafik)

Bei einer gleichförmige Bewegung von einer zur anderen Seite im Stereo-Panorama ist eine Mittendämpfung von 3 dB erforderlich.

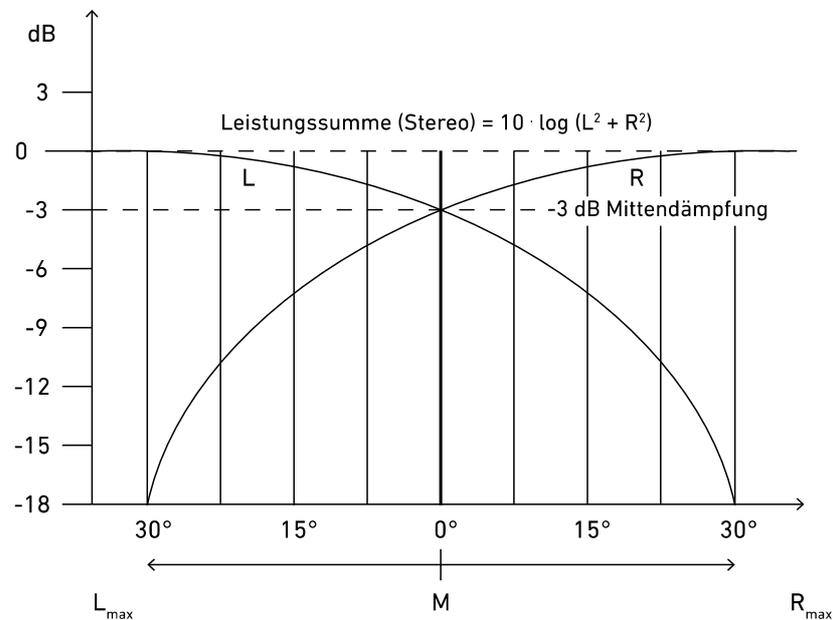


Abbildung 17: Dämpfungsverlauf, konstante Leistung im Stereo-Panorama
(eigene Grafik, nach Sengpiel "Das theoretische Panpot")

Die Stimme wurde mit einem de-esser bearbeitet um störende und hervorspringende Zischlaute zu eliminieren.

2. Grund

Für den Hintergrund wurde, wie für den Vordergrund, eine Sprachaufnahme gewählt. Es handelte sich um ein Gedicht. Das Signal dieser Aufnahme wurde durch ein Hallgerät (plugin) geschickt und der wet-Anteil auf 100% eingestellt, so dass die Verständlichkeit vollständig verloren ging. Um eine ausreichende Dekorrelation der wiedergegebenen Signale von vorderen und hinteren Lautsprechern bei der 5.1.-Wiedergabe zu gewährleisten, wurden für vorne und hinten sowohl zwei gesonderte gleiche Hallplugins verwendet, als auch die Einstellungen unterschiedlich gewählt.

- Stereo-Hall vorne:
pre-delay 32 ms, 1st/2nd early reflections 10/11 ms, reverb 2 sec
- Stereo-Hall hinten:
pre-delay 42 ms, reverb 2 sec, modulation 50%

Hall vorne



Hall hinten



Erste frühe Reflexionen = 0dB

Modulation der Hallfahne

Abbildung 18: Einstellungen des Halls vorne und hinten (Screenshot)

Der Pegel des verhallten Signals, das über die hinteren Lautsprecher wiedergegeben wurde, wurde gegenüber dem Pegel des Halls, der auf die vorderen Lautsprecher geschickt wurde, bei der Wiedergabe um 6dB abgesenkt, um zu verhindern, dass die Signale der hinteren Lautsprecher den Probanden eine deutliche Auskunft darüber geben, welches

Wiedergabeverfahren sie gerade hören. Die Stärke der Umhüllung von Hinten sollte sich gerade an der Grenze der Wahrnehmbarkeit befinden.

Variationen der Szenen

Von der Szene existierten 4 verschiedene Variationen: Stereo Panning von links nach rechts (2LR), 5.1Panning von links nach rechts (5LR), Stereo Panning von rechts nach links (2RL), 5.1Panning von rechts nach links (5RL).

Fragebogen

Zu jeder gehörten Szene existierte genau ein Versuchsbogen, der in mehrere Abschnitte aufgeteilt war. Der erste Abschnitt behandelte die Form der wahrgenommenen Wegstrecke, welche die Zahlenreihe vom rechten Rand des Bildes bis zum linken Rand nimmt. Um eindeutiger Ergebnisse zu erhalten, konnten die Probanden zwischen 7 vorgegebenen Streckenverläufen wählen. Die Form der vorgegebenen Kurven gehen auf Ergebnisse des Vorversuchs zurück.

Zu klären war vorab, ob die Übersetzung der Wahrnehmung von Hördingen ins Bildliche grundsätzlich erfolgen kann. Hilfreich waren Überlegungen aus der Synästhesie-Forschung.

Ein Begriff aus der Synästhesieforschung lautet „Notations-Synästhesie“. Darunter wird verstanden, dass ein Synästhetiker musikalische Parameter unmittelbar in visuelle Parameter umsetzt (Wohler 165). Laut Arnold Wohler gehen auf diesen Zusammenhang die Ursprünge der Notensetzung zurück.

Die Struktur der Notationssynästhesie ist ein Hinweis darauf, dass das historisch entstandene Notationssystem der Musik weniger ein Erfundenes, als vielmehr ein Entdecktes ist. Sie ist das Explizitwerden eines Zuordnungsprinzips, das im kognitiven System angelegt ist und wirksam ist, wenn bewusstseinspezifische Repräsentationen zueinander ins Verhältnis treten. (Wohler 165)

Tiefe Töne werden bei der Notations-Synästhesie beispielsweise als weiter unten liegend, hohe Töne als weiter oben empfunden, leise Töne als klein, laute als gross.

Untersuchungen über die bildliche Darstellung von Rhythmen bei Kindern zeigen, dass bei einem Rhythmus der aus Viertelnoten und Achtelnoten besteht, diese in einen sinnvollen Zusammenhang gesetzt werden und dementsprechend notiert werden.

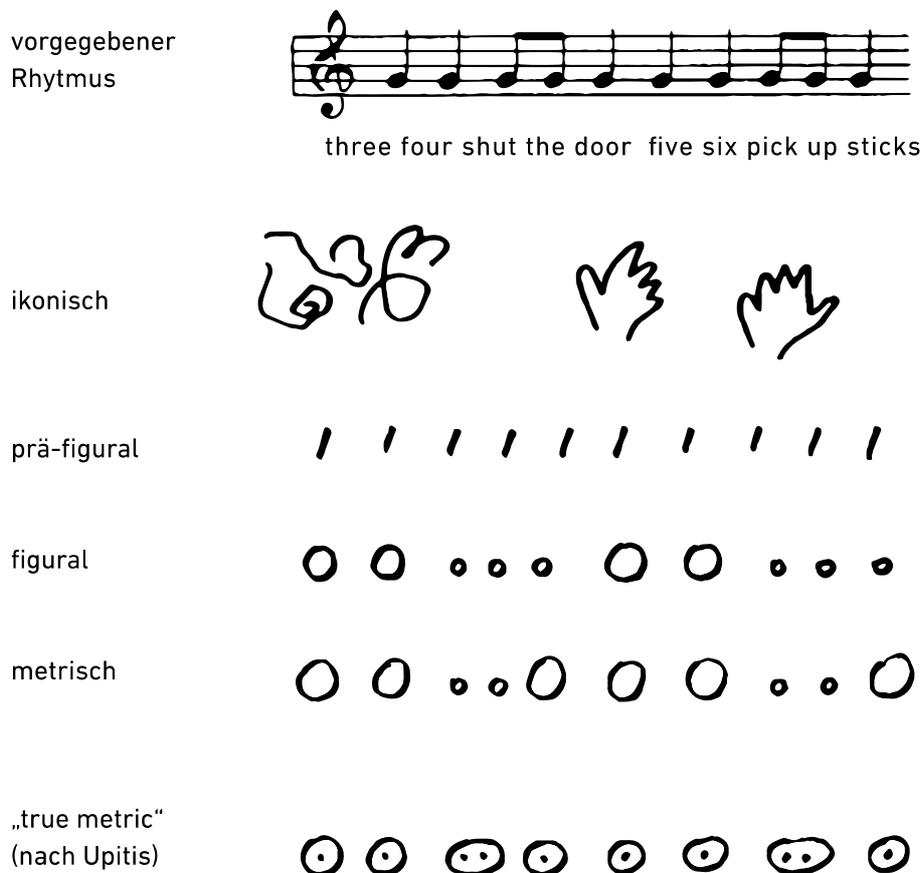
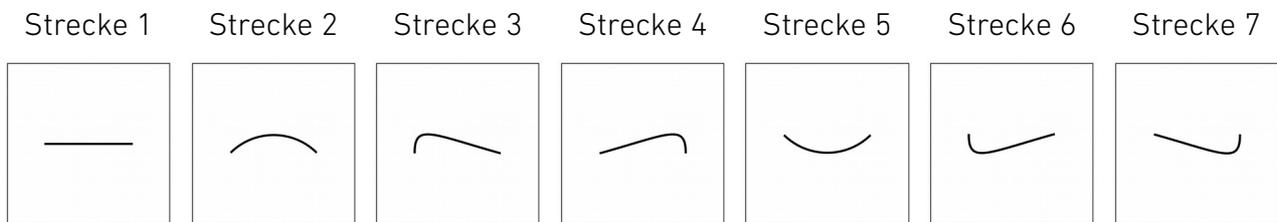


Abbildung 19: Entwicklungsstufen der bildlichen Darstellung von Rhythmen bei Kindern
(Rösing, Helmut und Bruhn, Herbert, Seite 22, nach Hildebrandt und Uptis)

Es kann davon ausgegangen werden, dass Visualisierungen von Gehörtem ähnlichen Regeln folgen, wie bereits die Wahrnehmung von Bild und Ton, in ihren spezifischen Wahrnehmungsräumen. Der wahrgenommene Raum von Hördingen wird von Menschen ähnlich in sinnvollen Zusammenfassungen strukturiert, wie für diese Zusammenfassungen sinnvolle, möglichst einfache Übersetzungen, existieren. Es wird vorausgesetzt, dass zusammenhängende Hör-Dinge nach der Übersetzung zu zusammenhängenden Seh-Dingen werden, welche die Hördinge adäquat repräsentieren.

Ein Streckenverlauf von links nach rechts in einem Hörraum wird beispielsweise nicht als Strecke von rechts nach links visualisiert und derart auf einem Blatt Papier notiert (was, auch wenn dies nicht der Fokus der hier in dieser Arbeit durchgeführten Studie gewesen ist, durch die Untersuchungsergebnisse bestätigt werden konnte).

Die 7 Streckenverläufe, die zur Auswahl standen, und von den Versuchsteilnehmern angekreuzt werden konnten:



Unterhalb der 7 Streckenverläufe befanden sich auf dem Versuchsbogen qualitative 3 Fragen zur Eingebundenheit des Zuhörers im dargestellten Raum, zur Schwierigkeit der Höraufgabe und zur Einschätzung über die gestalterische Verbindung zwischen Vorder- und Hintergrund.

Im zweiten Abschnitt des Versuchsbogens, auf dem unteren Teil, befand sich ein waagerechter Strich, bei dem der linke Rand, die Mitte und der rechte Rand hervorgehoben waren. Auf diesem Strich sollten die Versuchsteilnehmer die wahrgenommene Position der gehörten Zahlen auf der Stereo-Basis bzw. im Hörbild abtragen.



Ablauf der Studie

Jeweils drei Versuchspersonen nahmen gleichzeitig an der Untersuchung teil. Die Teilnehmer wurden gebeten, eine von drei möglichen Sitzpositionen einzunehmen: Mittig im Sweet-Spot, oder links bzw. rechts in frontaler Achse vor den Lautsprechern L oder R.

Jeder Versuchsdurchlauf dauerte 45 Minuten und enthielt 2 Pausen von 5 Minuten. In jeder dieser Pausen wurden die Probanden, welche entweder auf dem linken Platz oder in der Mitte saßen, dazu aufgefordert, den Sitzplatz rechts neben sich einzunehmen. Der Teilnehmer, der zuvor den Platz ganz rechts einnahm, wechselte ganz nach links.

An einem Tag fanden 6 Durchläufe statt, unterbrochen von einer Stunde Mittagspause. Der Untersuchungszeitraum umfasste drei Tage.

Die Probanden, die an der Untersuchung teilnahmen, wurden auf die Studie durch einen Aufruf in Facebook, in einer Rundmail über den HdM-Verteiler, durch persönliche Ansprache vor dem Hochschulgebäude und per Vorstellung in einer Vorlesung hingewiesen. Die möglichen Teilnehmer wurden nicht gezielt ausgewählt, da davon ausgegangen wurde, dass

Gestaltwahrnehmung keine tontechnischen Vorerfahrungen voraussetzt und eine Selektion der Probanden das Ergebnis nicht verfälschen sollte. Kein Teilnehmer war unter 20, der älteste Teilnehmer über 60.

Alter (in %):

<19	20-29	30-39	40-49	50-59	>60
0	60	20	8,57	5,71	5,71

Die Versuchsleitung begrüßte jeweils zur vollen Stunde eine 3er-Gruppe und erklärte dieser kurz den Ablauf der Untersuchung. Jeder Proband erhielt neben einem Bogen, auf dem die Untersuchung zusätzlich noch einmal erklärt wurde, einen zusammengehefteten Stapel mit 4 Untersuchungsbögen. Auf jedem einzelnen Bogen wurde die Teilnehmernummer, sowie der Buchstabe „A“ notiert, um den ersten von drei Teilen (A, B und C) des jeweiligen Durchlaufs auf dem Bogen festzuhalten um nachher eine Zuordnung zu ermöglichen. Die Probanden wurden darüber informiert, dass der Ablauf der 45 Untersuchungsminuten von der Versuchsleitung nicht unterbrochen werden kann, sondern vorproduziert ist.

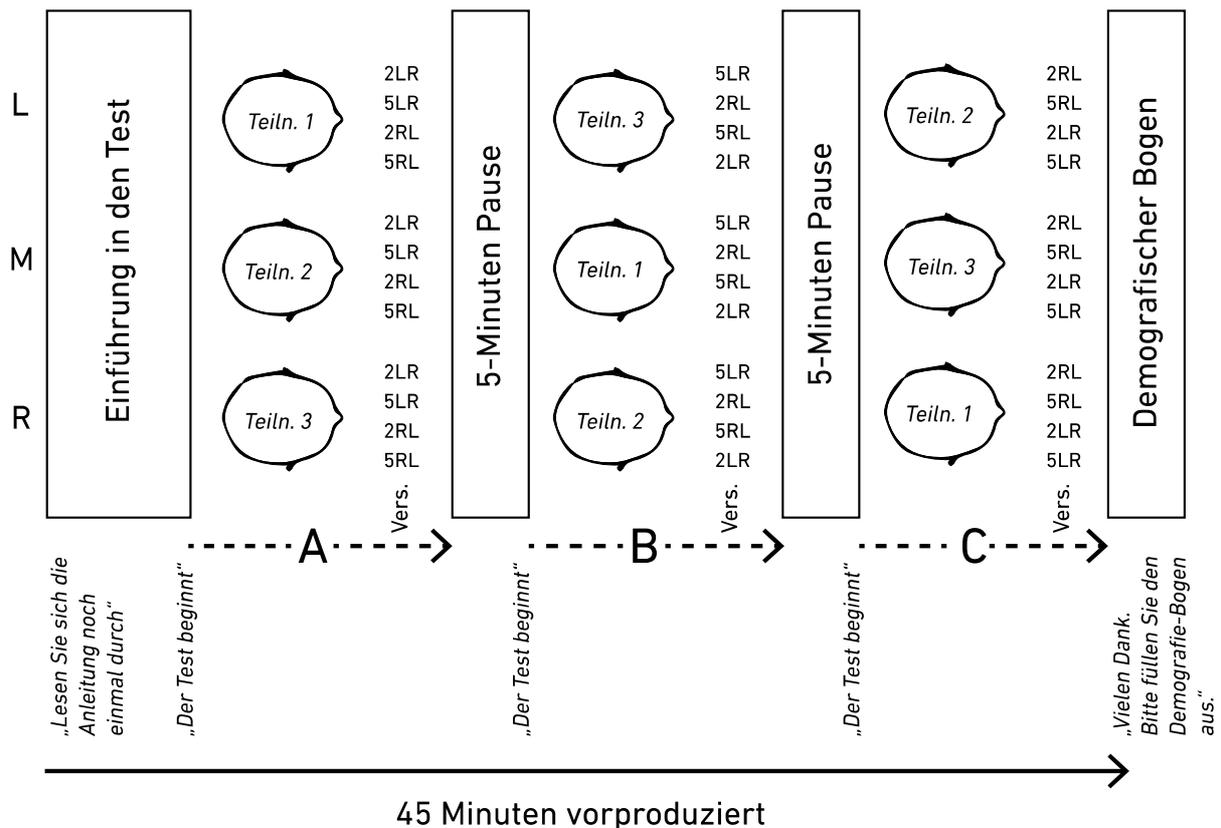


Abbildung 20: Ablauf der Studie (eigene Grafik)

In jedem Teil wurde den Probanden eine Abfolge der grundsätzlich gleichen Szenen vorgespielt, die unterschiedlich wiedergegeben wurden: 2LR, 5LR, 2RL, 5RL. In jeder Teilaufgabe wurde die Wiedergabe der Szene drei mal wiederholt, um den Teilnehmern zu ermöglichen, erst zuzuhören, dann anzukreuzen bzw. die Zahlen auf den Zahlenstrahl (Lokalisierungsposition) einzutragen und ihre Wahl dann noch einmal kontrollieren zu können.

Die Reihenfolge dieser 4 Szenen unterschied sich bei jeder 3er-Gruppe, sowie in jedem Teil eines Durchlaufs, um den Einfluss von Lern- und Ermüdungseffekten möglichst gleichmässig zu verteilen.

Das Verteilungsmuster:

	Teil A	Teil B	Teil C
1. 3er-Gruppe	2LR, 5LR, 2RL, 5RL	5LR, 2RL, 5RL, 2LR	2RL, 5RL, 2LR, 5LR
2. 3er-Gruppe	5LR, 2RL, 5RL, 2LR	2RL, 5RL, 2LR, 5LR	5RL, 2LR, 5LR, 2RL
3. 3er-Gruppe	2RL, 5RL, 2LR, 5LR	5RL, 2LR, 5LR, 2RL	2LR, 5LR, 2RL, 5RL
4. 3er-Gruppe	5RL, 2LR, 5LR, 2RL	2LR, 5LR, 2RL, 5RL	5LR, 2RL, 5RL, 2LR
n+4te 3er-Gruppe = nte 3er-Gruppe			

Durch die einzelnen Teile des Durchlaufs (A, B und C), die von 5 Minuten-Pausen unterbrochen wurden, um die Sitzposition wechseln zu können, führte eine vorab aufgenommene Ansage, die auf den Beginn der einzelnen Teile, sowie auf die unterschiedlichen Szenen hinwies. Vor Beginn der Teile B und C erhielten die Probanden jeweils erneut 4 Bögen, auf denen die Versuchsleitung die Nummer des Probanden, sowie den Buchstaben des Teils notierte.

Die Ansage „Der Test beginnt“ diente den Probanden zu Beginn des Tests, sowie nach den 5-Minuten-Pausen als Hinweis, um sich wieder auf die 4 Szenen des darauffolgenden Teils konzentrieren zu können. Die Versuchsleitung notierte nach dem Start der Vorführung in einem extra Bogen welche der 4 vorab produzierten Versionen welcher 3er-Gruppe vorgespielt wurde. Am Ende jedes Durchlaufs wurden die Probanden gebeten einen Bogen mit Angaben zur Demografie auszufüllen und wurden dann verabschiedet.

Verfahren der Datenauswertung

Nach den drei Tagen, an denen die Untersuchung stattfand, wurde – unter Zuhilfenahme des Versuchsleiterbogens - auf jedem einzelnen Bogen die jeweilige Version der Szene notiert und auf welcher Sitzposition diese Version abgehört wurde.

Die Beschriftungen:

L – 2LR	Sitzposition links, Stereo (2.0), Panning von links nach rechts
L – 5LR	Sitzposition links, Surround (5.1), Panning von links nach rechts
L – 2RL	Sitzposition links, Stereo (2.0), Panning von rechts nach links
L – 5RL	Sitzposition links, Surround (5.1), Panning von rechts nach links
M – 2LR	Sitzposition mitte, Stereo (2.0), Panning von links nach rechts
M – 5LR	Sitzposition mitte, Surround (5.1), Panning von links nach rechts
M – 2RL	Sitzposition mitte, Stereo (2.0), Panning von rechts nach links
M – 5RL	Sitzposition mitte, Surround (5.1), Panning von rechts nach links
R – 2LR	Sitzposition rechts, Stereo (2.0), Panning von links nach rechts
R – 5LR	Sitzposition rechts, Surround (5.1), Panning von links nach rechts
R – 2RL	Sitzposition rechts, Stereo (2.0), Panning von rechts nach links
R – 5RL	Sitzposition rechts, Surround (5.1), Panning von rechts nach links

Die Daten wurden in das Tabellenkalkulationsprogramm Calc übertragen und für jede der drei mal 4 Versionen ein gesondertes Tabellenblatt angelegt. Die ankreuzbaren Felder konnten auf Grund ihrer Eindeutigkeit direkt in die Auswertungstabelle übernommen werden. Die Angaben zu der Hörposition wurden mit einem Lineal ausgemessen und die jeweiligen Werte in das dazugehörige Tabellenblatt eingetragen. Die Hörpositionen wurden auf der rechten Seite des Tabellenblatts in prozentuale Werte umgerechnet.

Die Spalten der Tabellen der 12 Versionen:

	Anzahl der Tabellenspalten
Teilnehmernummer	1
Wegstrecke	7
Vordergrund innerhalb des Bildes?	2
Schwierigkeit leicht / schwer	2
Immersion, Eingebundenheit des Zuhörers	2
Positionen, absolut	10
Positionen, relativ	10
Tonerfahren?	1

Für die Auswertung der Daten wurden weitere Tabellenblätter angelegt:

- Durchschnittsposition/Lokalisation der Zahlen im Hörbild (Vergleich aller 12 Versionen)
- Gemittelte Position der Zahlen im Hörbild (Vergleich aller 12 Versionen)
- Standardabweichung der wahrgenommenen Positionen (Vergleich aller 12 Versionen)
- Abweichung der Position vom berechneten Ideal (Vergleich aller 12 Versionen)
- Vergleich \emptyset der Positionen bei 2.0 und 5.1. zum Ideal (Panningrichtung unberücksichtigt)
- Positionen der wahrgenommenen Zahlen auf den Aussensitzpositionen in Bezug auf die Panningrichtung (nähernd / entfernend)
- Standardabweichung der wahrgenommenen Zahlen auf den Aussensitzpositionen in Bezug auf die Panningrichtung (nähernd / entfernend)
- Angabe über Eingebundenheit des Vordergrunds in den Hintergrund des Hörbildes, abhängig von der Sitzposition (in Prozent)
- Angabe über den Leichtigkeitsgrad der Lokalisierung, abhängig von der Sitzposition (in Prozent)
- Angabe über die Eingebundenheit der Zuhörer in Abhängigkeit von der Sitzposition (in Prozent)

Die Daten wurden in RStudio importiert und grafisch aufbereitet. Weitere Grafiken wurden in einem Grafikprogramm erstellt.

Ergebnisse

Zusammenfassung der Ergebnisse

Unterschiede zwischen den Wiedergabeverfahren 2.0 und 5.1. zeigen sich, wie – ausgehend von den in der Literatur beschriebenen Effekten - zu erwarten war, hinsichtlich der Form des wahrgenommenen Hörbildes bei einer Hörposition ausserhalb der Mitte des Sweet-Spots. Die Probanden wählten bei 5.1. unabhängig der Sitzposition weitgehend symmetrische Wegstreckenverläufe, bei 2.0 nur auf der mittleren Sitzposition. Es zeigt sich auch, dass ausserhalb des Sweet-Spots die Schwierigkeit der Lokalisation bei beiden Wiedergabeverfahren unterschiedlich bewertet wurde. Bei Sitzpositionen ausserhalb des Sweet-Spots wurde die Lokalisierung bei 5.1 als leichter bewertet, als bei 2.0.

Unabhängig der Sitzposition haben sich mehr Probanden bei der 5.1-Wiedergabe als Teil der Szene begriffen.

Hinsichtlich der Eingebundenheit des Vordergrunds in die Szene, zeigt sich kein Unterschied zwischen 2.0 und 5.1. Bei beiden Wiedergabeverfahren haben die Versuchsteilnehmer auf den Sitzpositionen ausserhalb der Mitte den Vordergrund weniger als Teil der Szene (und mit weniger Bezug zum Hintergrund) wahrgenommen, als bei einer Hörposition im Zentrum des Sweet-Spots.

Auf den äusseren Sitzpositionen bei 2.0-Wiedergabe zeigen sich bei der Lokalisierung Unterschiede in der Standard-Abweichung, je nachdem, ob sich das Hörding auf die Probanden zubewegte oder wegbewegte. Bei einer Wegbewegung ist die Standardabweichung im Schnitt geringer als bei einer Hinbewegung.

Unabhängig des Wiedergabeverfahrens erreicht die Lokalisierung auf einer Sitzposition in der Mitte des Sweet-Spots eine Genauigkeit, die vom berechneten Ideal im Schnitt um weniger als 1% abweicht.

Insgesamt haben 35 Personen an der Untersuchung teilgenommen.

Die Ergebnisse im Detail

Kurvenverlauf/Wegstrecke des wahrgenommenen Hördings

Die meisten der 35 Probanden haben bei der Wiedergabe über 5.1. symmetrischere oder geradere Streckenverläufe angegeben, als bei 2.0. Am häufigsten wurde ein geradliniger, symmetrischer Verlauf bei 5.1 (Strecke 1) und einer mittleren Sitzposition wahrgenommen.

Unsymmetrische Kurvenverläufe, wie Strecke 3, Strecke 5, sowieso Strecke 6 und Strecke 7 (siehe Abbildung Streckenverläufe), wurden von den Studienteilnehmern vor allem auf den Aussensitzpositionen und einer Wiedergabe über 2.0 wahrgenommen.

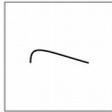
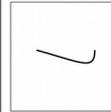
Auf der mittleren Sitzposition erreichte Strecke 5 sowohl bei 5.1, als auch bei 2.0 den höchsten Wert. Beim Panningverlauf rechts→links wurde Strecke 5 bei einer Sitzposition direkt im Sweet-Spot 14 mal (2.0) bzw. 17 mal gewählt (5.1).

Strecke 2 wurde bei keiner Sitzposition und keinem Wiederverfahren von mehr als 5 Personen gewählt.

Auf der rechten Sitzposition, 5.1 und einem Panningverlauf rechts→links wählten die Probanden Strecke 3 und Strecke 7 am häufigsten. In der spiegelbildlichen Situation, auf der linken Sitzposition und einem Panning links→rechts, wurde bei 5.1. die gerade Strecke (Strecke 1) am häufigsten gewählt. Vergleicht man die Form von Strecke 3 und Strecke 7, fällt auf, dass beide Kurven sich ähneln und sich nur an Ende und Anfang voneinander unterscheiden.

Die Streckenverläufe, die von den Teilnehmern angekreuzt wurden

in absoluten Zahlen, Anzahl der Teilnehmer

									
Sitzp.	Verf.	Pan.	Str.1	Str.2	Str.3	Str.4	Str.5	Str.6	Str.7
links	2.0	LR	7	1	1	4	4	16	2
		RL	8	1	13	2	1	9	0
	5.1	LR	12	4	2	3	5	7	2
		RL	8	1	5	3	7	7	4
mitte	2.0	LR	9	3	4	1	9	2	7
		RL	9	4	0	1	14	4	3
	5.1	LR	13	4	2	0	2	2	12
		RL	5	3	2	0	17	3	5
rechts	2.0	LR	6	0	3	11	2	0	12
		RL	2	0	4	8	3	0	18
	5.1	LR	9	5	3	5	2	2	9
		RL	4	5	11	0	2	2	11

Lokalisation der Zahlen auf der Stereobasis

Sowohl Lokalisationsgenauigkeit, als auch Sicherheit der Lokalisation sind in den Aussensitzpositionen (ausserhalb des Sweet-Spots) bei 5.1 grösser, als bei 2.0. Bei 2.0 unterscheidet sich die Lokalisation in den Aussensitzpositionen zudem je nach Panningrichtung (Hörding entfernt sich bzw. kommt näher).

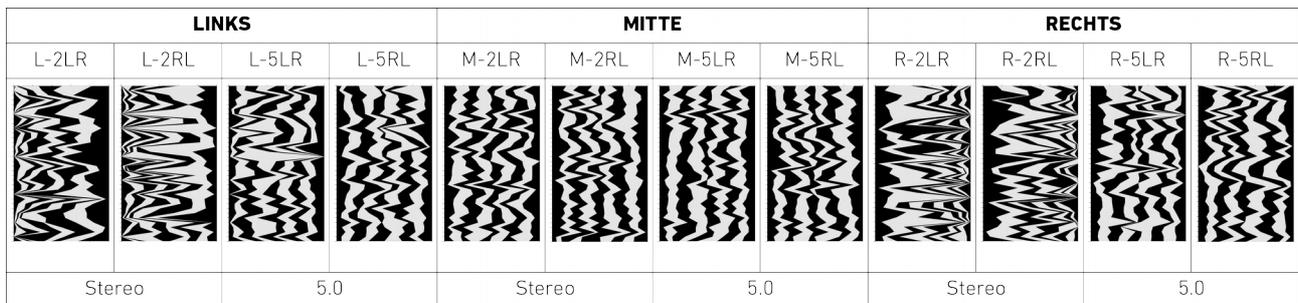


Abbildung 21: Übersicht über Lokalisierung, Hörposition L, M, R - horizontal: wahrgenommene Position, vertikal: einzelne Teilnehmer (eigene Grafik). Vertikale Achse: Die unterschiedlichen Teilnehmer.

Die Sicherheit der Lokalisation der wiedergegebenen Zahlenreihe ist in mittlerer Sitzposition, also inmitten des Sweet-Spots, bei 2.0 und 5.1. annähernd gleich.

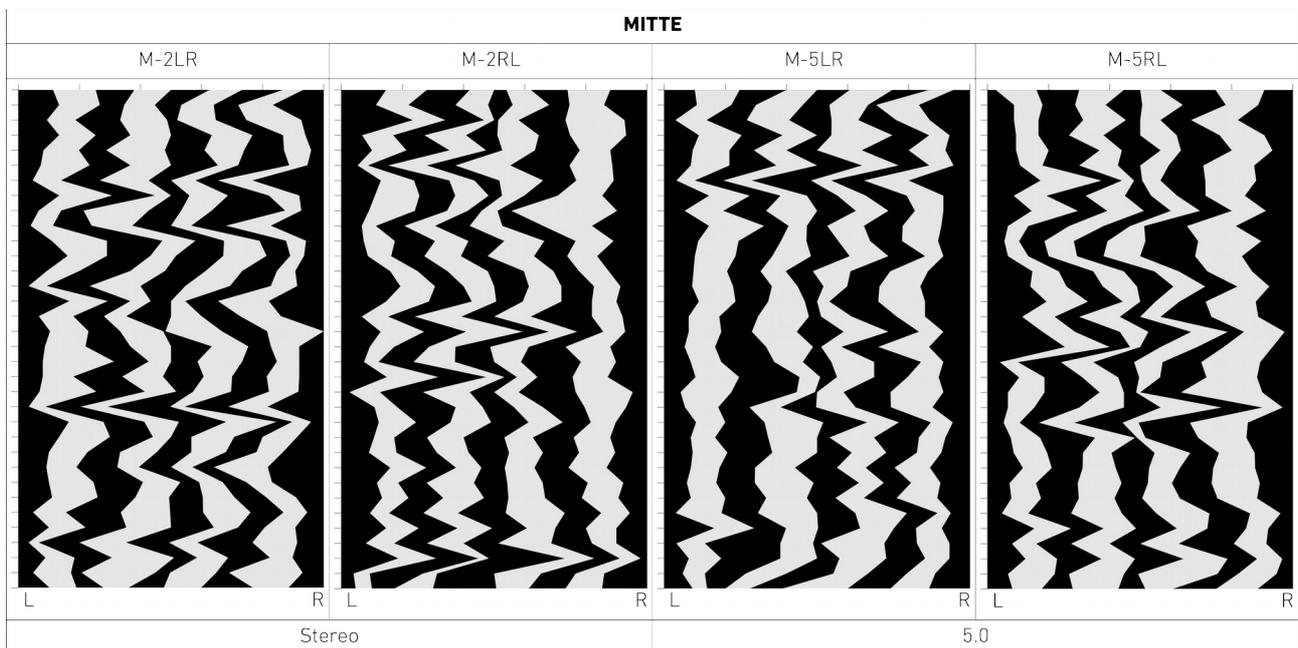


Abbildung 22: Lokalisierung bei Hörposition Mitte, „Sweet-Spot“ - horizontal: wahrgenommene Position, vertikal: einzelne Teilnehmer (eigene Grafik)

In den Aussensitzpositionen ist bei 2.0 erkennbar, dass das Hörding stärker am Lautsprecher festzukleben scheint, wenn es sich von der Sitzposition wegbewegt. Sowohl auf linker Sitzposition und Panning links→rechts (L-2LR), als auch auf der rechten Sitzposition und einem Panning rechts→ links (R-2RL) ist eine deutlich kompaktere Verteilung der unterschiedlichen Lokalisationswerte bei 2.0 zu sehen.

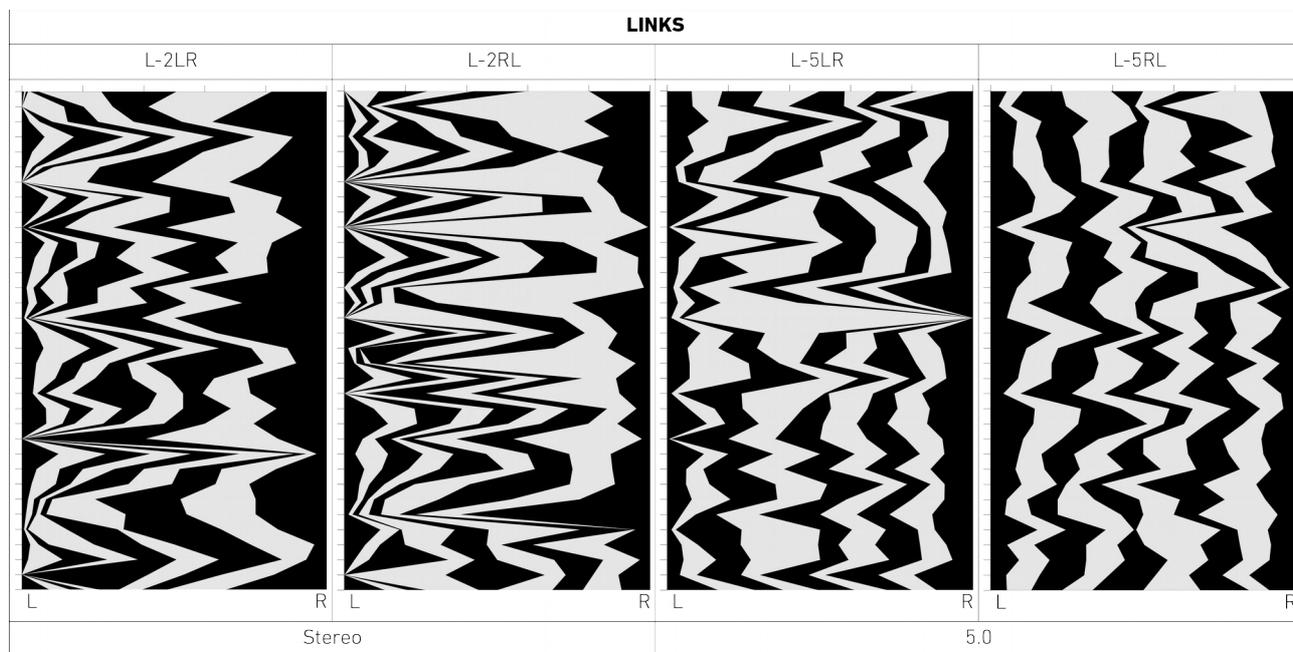


Abbildung 23: Lokalisierung bei Hörposition Links - horizontal: wahrgenommene Position, vertikal: einzelne Teilnehmer (eigene Grafik)

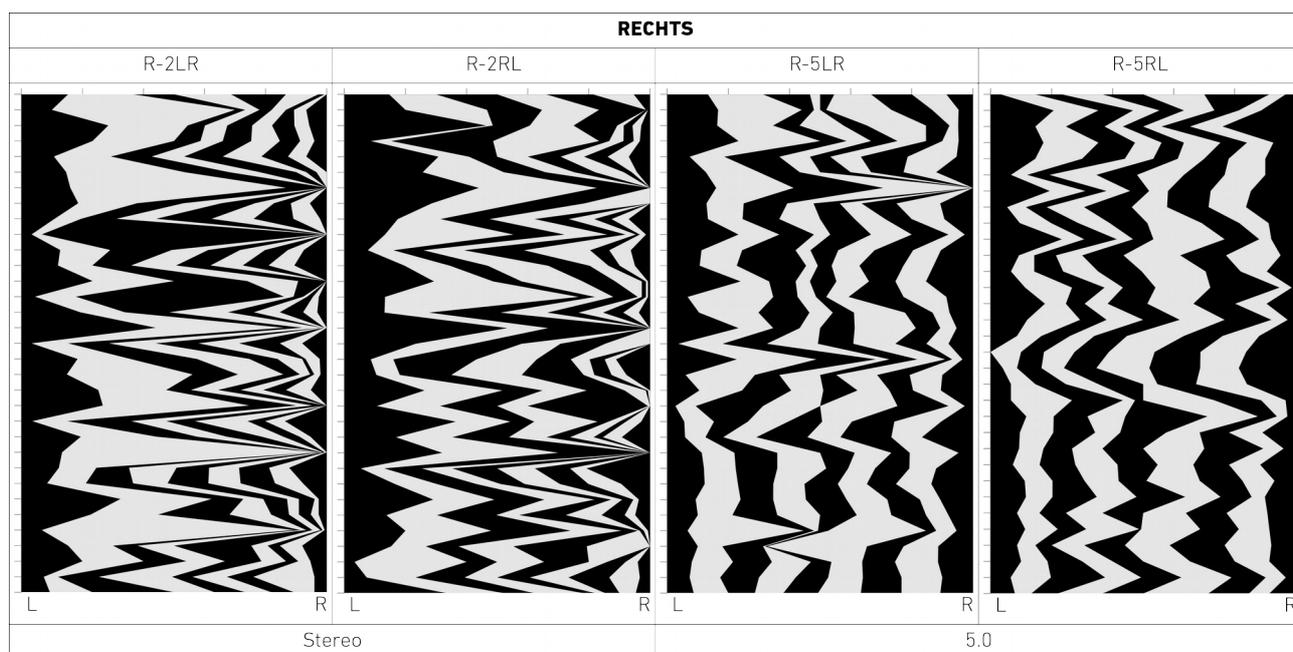


Abbildung 24: Lokalisierung bei Hörposition Rechts - horizontal: wahrgenommene Position, vertikal: einzelne Teilnehmer (eigene Grafik)

Durchschnittliche Lokalisation auf der Stereo-Basis (in Prozent rechts)

in Abhängigkeit von Sitzposition (L, M oder R), Wiedergabeverfahren (2.0 oder 5.1) und Richtung des Pannings (LR oder RL)

Sitzp.	Verf.	Pan.	L	L -1	L -2	L -3	L -4	R -4	R -3	R -2	R -1	R
links	2.0	LR	0	3,69	8,47	13,42	19,92	30,04	42,92	60,74	78,86	100
		RL	0	5,14	9,38	15,12	22,93	32,94	46,31	60,37	86,74	100
	5.1	LR	0	6,63	15,32	27,34	46,27	58,23	69,57	79,19	88,02	100
		RL	0	9,26	19,42	32,81	44,47	53,51	66,2	77,89	90,18	100
mitte	2.0	LR	0	9,17	21,22	33,72	46,25	56,93	68,13	79,45	90,07	100
		RL	0	11	20,75	32,67	43,4	53,76	64,98	77,07	89,53	100
	5.1	LR	0	8,8	20,26	33,51	45,66	55,5	67,47	79,02	89,66	100
		RL	0	10,06	20,14	31,78	43,02	52,32	62,38	75,18	89,49	100
rechts	2.0	LR	0	15,95	42,95	60,15	70,81	79,67	86,27	91,26	95,81	100
		RL	0	21,06	41,28	58	71,93	82,73	89,42	93,13	96,22	100
	5.1	LR	0	11,94	27,5	40,51	50,31	58,47	71,57	82,09	91,44	100
		RL	0	10,38	20,33	30,48	40,72	52,23	67,44	80,7	90,58	100

Die durchschnittliche Lokalisation der einzelnen Zahlen („1“, „2“, ... „10“) ist inmitten des Sweet-Spots bei 2.0 und 5.1. sehr ähnlich. In den Aussen-Sitzpositionen (links und rechts) sind bei 2.0 ungleichmässiger Lokalisierungs-Abstände zu den Nachbarpositionen erkennbar.

Auf der Sitzposition links liegen die Werte bei 2.0 im Bereich L (L, L-1, L-2,...) und auf der Sitzposition rechts im Bereich R (R, R-1, R-2,...) näher zusammen.

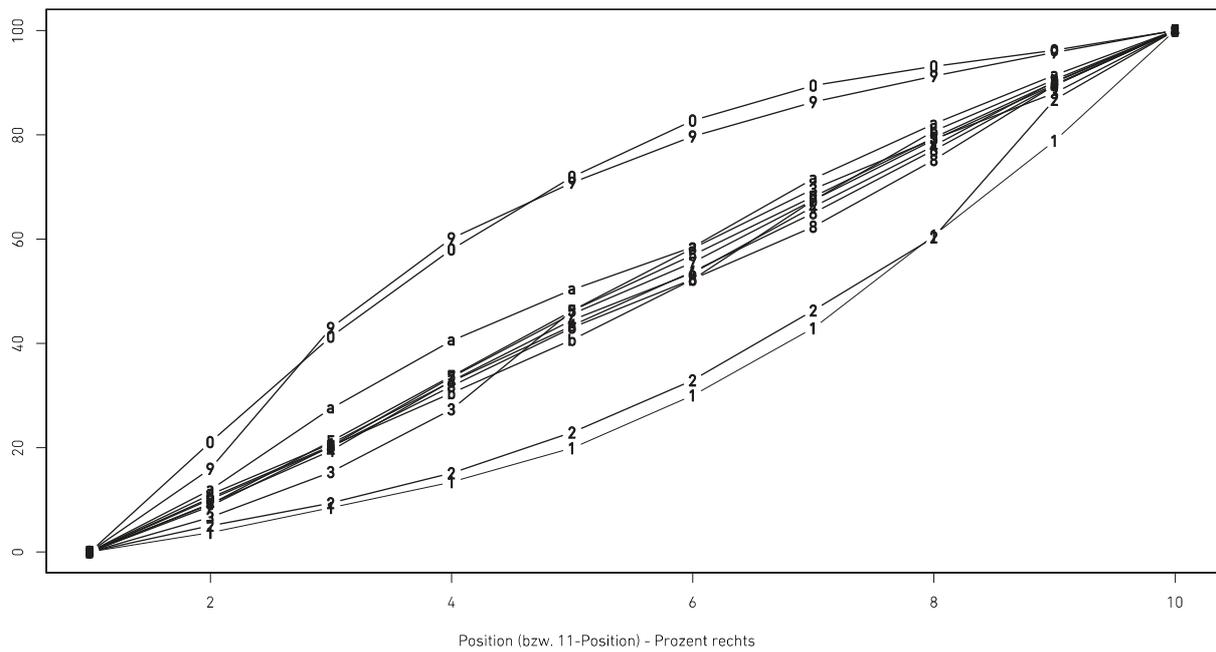


Abbildung 25: Lokalisation der Zahlen 1-10. Horizontal: Zahlen, Vertikal: % rechts im Stereopanorama, 1+2 Hörposition links 2.0, 9+0 Hörposition rechts (eigene Grafik)

Stereo



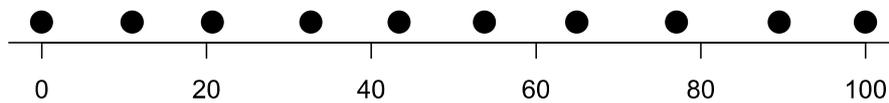
L-2LR (Ø Position)



L-2RL (Ø Position)



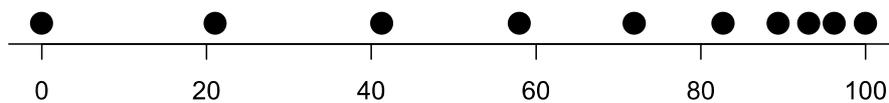
M-2LR (Ø Position)



M-2RL (Ø Position)



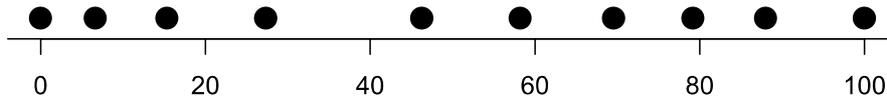
R-2LR (Ø Position)



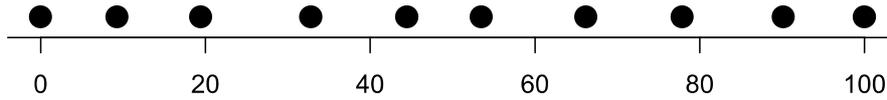
R-2RL (Ø Position)

Abbildung 26: Lokalisierung auf der Stereobasis, % rechts - 2.0 in Abhängigkeit von der Sitzposition (L, M, R)

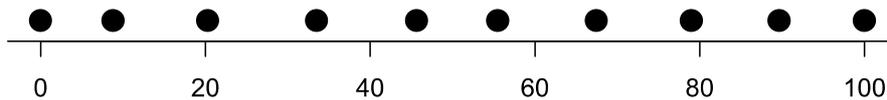
5.1



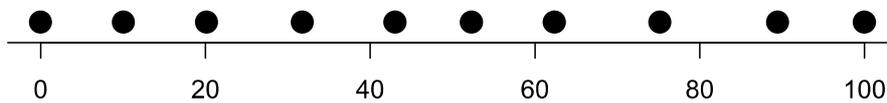
L-5LR (∅ Position)



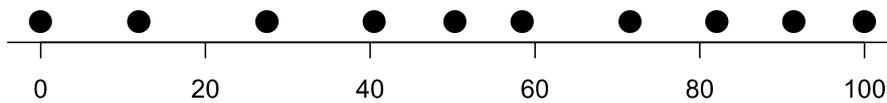
L-5RL (∅ Position)



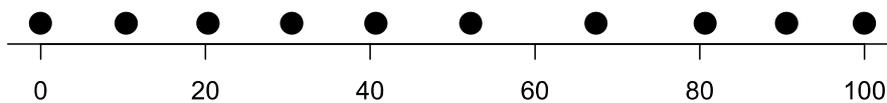
M-5LR (∅ Position)



M-5RL (∅ Position)



R-5LR (∅ Position)



R-5RL (∅ Position)

Abbildung 27: Lokalisierung auf der Stereobasis, % rechts - 5.1 in Abhängigkeit von der Sitzposition (L, M, R)

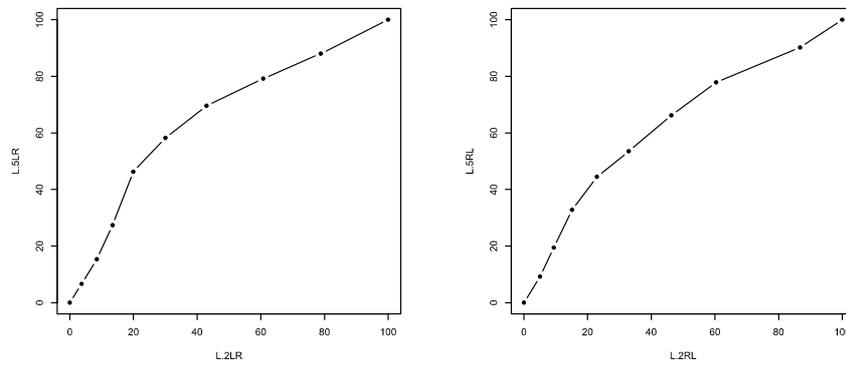


Abbildung 28: Sitzposition links, Vergleich 5.1 und 2.0, % rechts (eigene Grafik)

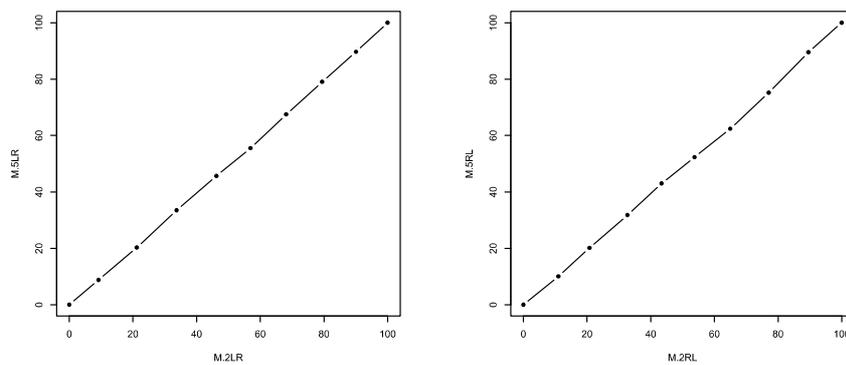


Abbildung 29: Sitzposition mitte, Vergleich 5.1 und 2.0, % rechts (eigene Grafik)

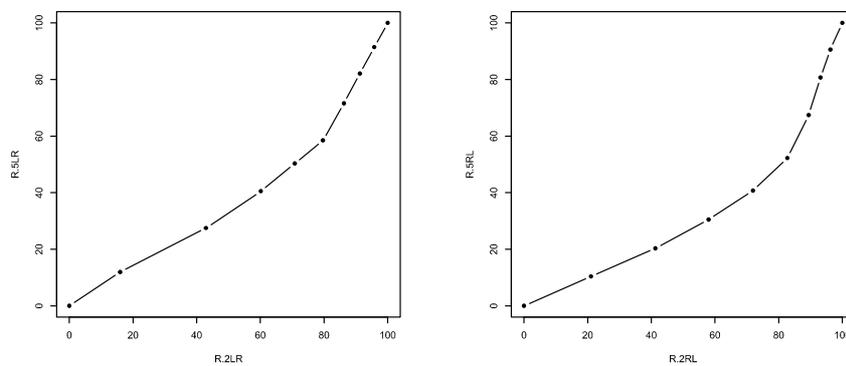
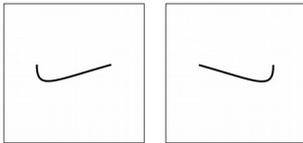


Abbildung 30: Sitzposition rechts, Vergleich 5.1 und 2.0, % rechts (eigene Grafik)

2.0 ausserhalb des Sweet-Spots, Vorschlag für eine Visualisierung

Die wahrgenommenen Hördinge hängen bei 2.0 stark von der Richtung des Pannings ab. Teilnehmer, die rechts oder links ausserhalb der Mitte des Sweet-Spots sassen, wählten bei einem sich von dieser Aussensitzposition wegbewegenden Hörding folgende Streckenverläufe:



L-2LR	R-2RL
45,7 %	51,4 %

Bei einem sich nähernden Hörding wählten die Probanden mehrheitlich folgende Wegstrecken:



L-2RL	R-2LR	R-2LR
37,1 %	31,4%	34,2 %

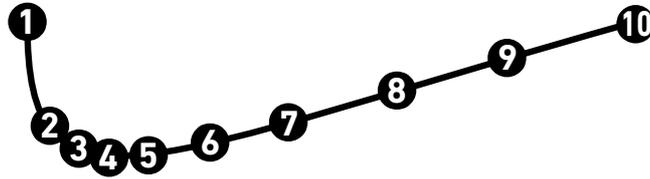
Auf der rechten Sitzposition wurden bei 2.0 und sich wegbewegendem Hörding zwei verschiedene Streckenverläufe etwa gleich häufig gewählt.

Ausgehend davon, dass sich die Formen der wahrgenommenen Streckenverläufe des Pannings bei 2.0 auf der linken Sitzposition zu der rechten Sitzposition symmetrisch verhalten müssten, hiesse das, dass bei einer Wegbewegung das Hörding erst am Ort des Lautsprechers wahrgenommen wurde, dann nach „vorne“ (in Richtung des Zuhörers) vorgewölbt, um sich danach nach rechts oder links zu entfernen.

Auf einer Hinbewegung (bei 2.0 und Aussensitzposition) wurde das Hörding als sich von links oder rechts nähernd wahrgenommen, um sich danach nach vorne zu wölben.

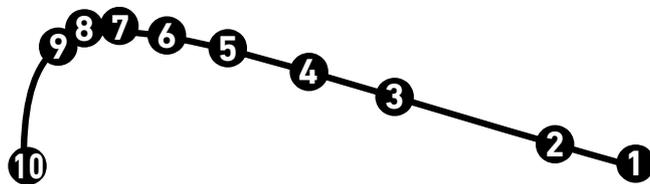
Es bietet sich folgende Visualisierung an:

Sitzposition links, Panning Links → Rechts



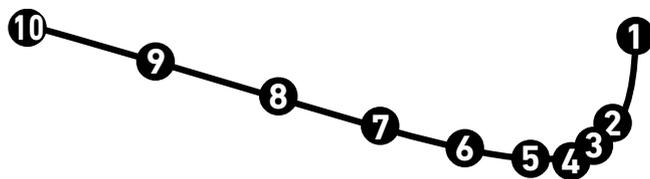
L-2LR

Sitzposition links, Panning Rechts → Links



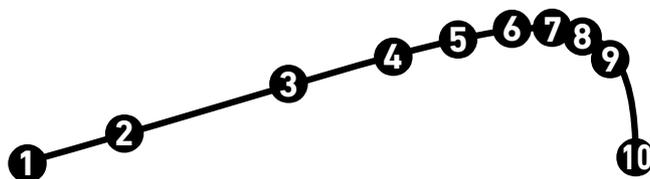
L-2RL

Sitzposition rechts, Panning Rechts → Links



R-2RL

Sitzposition rechts, Panning Links → Rechts



R-2LR

Genauigkeit der Lokalisation im Vergleich zum Ideal

Idealerweise würde sich bei einer Lokalisation der Zahlenreihe 1-10 auf der Stereobasis, die dem Ideal entspricht, eine Näherung zu folgenden Prozentwerten ergeben (0% entspricht ganz links, 100% ganz rechts):

0%, 11,11%, 22,22%, 33,33%, 44,44%, 55,56%, 66,67%, 77,78%, 88,89%, 100%
(rechts auf der Stereobasis)

In der mittleren Sitzposition (Abhöre innerhalb der Mitte des Sweet-Spots) erreicht die Lokalisation sowohl bei 2.0, als auch bei 5.1 gute Werte. Die minimale Abweichung der durchschnittlichen Lokalisation auf der Stereobasis vom Ideal beträgt 0,06 $\Delta\%$ bei 5.1 und 0,11 $\Delta\%$ bei 2.0. Die maximale Abweichung beträgt in der mittleren Sitzposition 4,29 $\Delta\%$ bei 5.1 und 1,94 $\Delta\%$ bei 2.0.

In den Aussensitzpositionen ist die durchschnittliche Abweichung der Lokalisation zum Ideal bei 2.0 deutlich grösser als bei 5.1. Der Maximalwert der durchschnittlichen Abweichung bei linker Sitzposition beträgt bei 2.0 25,52 $\Delta\%$ und bei rechter Sitzposition 27,49 $\Delta\%$, während bei 5.1 nur Werte von 5,99 $\Delta\%$ (linke Abhörposition), sowie 7,18 $\Delta\%$ (rechte Sitzposition) erreicht werden. Die maximale „Verzerrung“ (die Lokalisation, die am wenigsten dem Ideal entspricht) des Hörbildes bei 2.0 in den Aussensitzpositionen (direkt vor dem rechten oder linken vorderen Lautsprecher) beträgt bei gleichem Abstand zur Lautsprecherebene, wie bei einer Sitzposition inmitten des Sweet-Spots, etwas über 25 Prozent in Bezug zur Stereobasis.

Bei 2.0 hat bei einer Aussensitzposition auch die Richtung des Pannings einen Einfluss auf die Abweichung der Lokalisierung zum Ideal. Bei linker Sitzposition befindet sich die durchschnittliche minimale Abweichung zum Ideal bei Panning links \rightarrow rechts auf der linken Seite (2. Position von links) und bei Panning rechts \rightarrow links auf der rechten Seite (2. Position von rechts). Bei rechter Sitzposition verhält es sich genau umgekehrt (spiegelverkehrt).

Die erste Position ausserhalb der Extremposition (eins neben ganz links bzw. ganz rechts, also jeweils die 2. Zahl) wurde bei 2.0 und bei sich nähernden Hördingen im Verhältnis zur Sitzposition eher dem Ideal entsprechend lokalisiert (2,15 $\Delta\%$ bei L2RL und 4,84 $\Delta\%$ bei R2LR), als im Falle der Wegbewegung des Pannings (7,42 $\Delta\%$ bei L2LR und 7,33 $\Delta\%$ bei R2RL).

Abweichung der durchschnittlichen Lokalisation (Stichprobe) auf der Stereo-Basis vom Ideal (Differenz der Prozent r auf der Stereobasis, $\Delta\%$). Ideal = Berechnete Werte, gleichmässig verteilt

in Abhängigkeit von Sitzposition (L, M oder R), Wiedergabeverfahren (2.0 oder 5.1) und Richtung des Pannings (LR oder RL), Minima >0 hervorgehoben

Sitzp.	Verf.	Pan.	L	L -1	L -2	L -3	L -4	R -4	R -3	R -2	R -1	R
links	2.0	LR	0	<u>7,42</u>	13,75	19,91	24,52	25,52	23,75	17,04	10,03	0
		RL	0	6,07	12,84	18,21	21,51	22,62	20,36	17,41	<u>2,15</u>	0
	5.1	LR	0	4,48	6,9	5,99	1,83	2,67	2,9	1,41	<u>0,87</u>	0
		RL	0	1,85	2,8	0,52	<u>0,03</u>	2,05	0,47	0,11	1,29	0
mitte	2.0	LR	0	1,94	1	<u>0,39</u>	1,81	1,37	1,46	1,67	1,18	0
		RL	0	<u>0,11</u>	1,47	0,66	1,04	1,8	1,69	0,71	0,64	0
	5.1	LR	0	2,31	1,96	0,18	1,22	<u>0,06</u>	0,8	1,24	0,77	0
		RL	0	1,05	2,08	1,55	1,42	3,24	4,29	2,6	<u>0,6</u>	0
rechts	2.0	LR	0	<u>4,84</u>	20,73	26,82	26,37	24,11	19,6	13,48	6,92	0
		RL	0	9,95	19,06	24,67	27,49	27,17	22,75	15,35	<u>7,33</u>	0
	5.1	LR	0	<u>0,83</u>	5,28	7,18	5,87	2,91	4,9	4,31	2,55	0
		RL	0	<u>0,73</u>	1,89	2,85	3,72	3,33	0,77	2,92	1,69	0

Die Probanden, die sich jeweils im Sweet-Spot (in der mittleren Sitzposition) befunden haben, konnten die Zahlenreihe relativ genau lokalisieren, Verrechnet man die Panning-Bewegungen links→rechts und rechts→links symmetrisch, entstehen sowohl bei 2.0, als auch bei 5.1 durchschnittliche Lokalisierungs-Werte, die nahezu dem berechneten Ideal entsprechen.

Durchschnittlichen Lokalisation auf der Stereo-Basis, symmetrisch verrechnet, im Vergleich zum Ideal (berechnet), in Prozent

Sitzposition: Mitte, unabhängig der Richtung des Pannings

Sitzp.	Verf.	L	L -1	L -2	L -3	L -4	R -4	R -3	R -2	R -1	R
mitte	2.0	0	10,09	20,99	33,2	44,83	55,35	66,56	78,26	89,8	100
	5.1	0	9,43	20,2	32,65	44,34	53,91	64,93	77,1	89,58	100
	ideal	0	11,11	22,22	33,33	44,44	55,56	66,67	77,78	88,89	100

Ausgehend von diesen symmetrisch verrechneten (von der Bewegungsrichtung des Pannings bereinigten) durchschnittlichen Werten der Lokalisierung, erreicht 2.0 in mittlerer Sitzposition einen besseren Wert als 5.1.

Wert der durchschnittlichen Abweichung (\emptyset der Abweichung der durchschnittlich lokalisierten Position vom Ideal):

2.0: 0,448 Prozent

5.1: 0,924 Prozent

Sicherheit und Unsicherheit in Abhängig von der Panningrichtung

Die Standardabweichung ist das Mass der Unsicherheit. Ein hoher Wert lässt darauf schliessen, dass sich die Probanden weniger sicher waren, wo sie die jeweilige Zahl („1“, „2“, ...“10“) auf der Stereobasis lokalisierten.

Auf der mittleren Hörposition erreicht die Standardabweichung (Stichprobe) bei 2.0 einen Maximalwert von 7,98 und bei 5.1. einen Maximalwert von 8,39.

Auf den Aussenpositionen existieren sowohl bei 2.0, als auch bei 5.1. höhere Standardabweichungen im Vergleich zur Sitzposition innerhalb der Mitte des Sweet-Spots. Der Maximalwert bei 2.0 beträgt 25,16 (L-2RL), bei 5.1. beträgt er 13,86 (L-5LR). Am unsichersten waren die Probanden, die ausserhalb des Sweet-Spots sassen, beim Wiedergabeverfahren 2.0 und den Szenen, in denen sich das Hörding genähert hat: L2RL und R2LR.

Der Peak der grössten Unsicherheit unterscheidet sich in den Aussensitzpositionen bei 2.0 je nach Panningrichtung. Er befindet sich bei einer Wegbewegung (von der Sitzposition weg) bei der jeweils 7. genannten Zahl (entspricht 2/3 der Zahlenreihe), bei einer Hinbewegung (auf die Abhörposition zu) bei der 3. Zahl (entspricht etwas weniger als 1/4 der Zahlenreihe).

Auch auf der mittleren Sitzposition zeigen sich Unterschiede in Bezug auf die Richtung der Bewegung des Hördings. Sowohl bei 2.0, als auch bei 5.1. beträgt die Standardabweichung bei der zweiten gehörten Zahl („2“) den geringsten Wert der jeweiligen Reihe. Bei 2.0 sind dies die Werte 3,71 (2. Zahl von links bei links→rechts) und 3,36 (2. Zahl von rechts bei rechts→links). Bei 5.1. betragen diese Werte 2,97 (2. Zahl von links bei links→rechts) und 4,06 (2. Zahl von rechts bei rechts→links).

Auf der mittleren Sitzposition liegen die Werte der Standardabweichung der jeweiligen links→rechts-Panningbewegungen und rechts→links-Panningbewegungen bei beiden Wiedergabeverfahren (2.0 und 5.1.) näher zusammen, als dies in den Aussensitzpositionen der Fall ist. Die Bewegungsrichtung des Hördings hat somit in den Aussensitzpositionen einen grösseren Einfluss auf die Sicherheit der Lokalisierung, als dies in der Mittensitzposition der Fall ist.

Da sich je nach Panningrichtung Unterschiede der Standardabweichung zeigen, deutet dies auf einen grossen Einfluss der Gestaltwahrnehmung hin. Als Gestaltgesetz wird dies als „Gesetz des gemeinsamen Schicksals“ bezeichnet.

Standardabweichung der Lokalisation (Stichprobe) auf der Stereo-Basis (in Prozent r)

in Abhängigkeit von Sitzposition (L, M oder R), Wiedergabeverfahren (2.0 oder 5.1) und Richtung des Pannings (LR oder RL), Maxima hervorgehoben

Sitzp.	Verf.	Pan.	L	L -1	L -2	L -3	L -4	R -4	R -3	R -2	R -1	R
links	2.0	LR	0	3,14	6,84	10,72	13,73	16,54	<u>16,68</u>	13,28	10,77	0
		RL	0	4,65	8,01	12,46	15,9	19,27	22,01	<u>25,16</u>	11,03	0
	5.1	LR	0	4,29	9,84	11,81	<u>13,86</u>	11,4	9,61	7,46	5,88	0
		RL	0	4,17	6,7	7,38	7,26	7,23	8,24	<u>8,87</u>	4,32	0
mitte	2.0	LR	0	3,71	6,58	7,53	6,25	6,23	<u>7,65</u>	7,57	5,15	0
		RL	0	4,6	6,92	<u>7,98</u>	7,04	6,63	6,56	6,03	3,36	0
	5.1	LR	0	2,97	4,78	5,74	5,76	5,87	<u>6,62</u>	6,04	3,55	0
		RL	0	4,88	7,16	7,07	6,04	6,1	<u>8,39</u>	6,73	4,06	0
rechts	2.0	LR	0	9,06	<u>24,23</u>	21,96	17,92	14,17	10,12	6,77	3,72	0
		RL	0	12,34	17,92	<u>18,9</u>	16,94	13,29	9,86	6,77	4,42	0
	5.1	LR	0	5,94	10,24	10,74	<u>11,93</u>	11,25	9,71	6,86	3,9	0
		RL	0	5,15	7,26	<u>8,81</u>	8,04	7,74	8,09	7,84	5,13	0

- 1 L.2LR
- 2 L.2RL
- 3 R.2LR
- 4 R.2RL

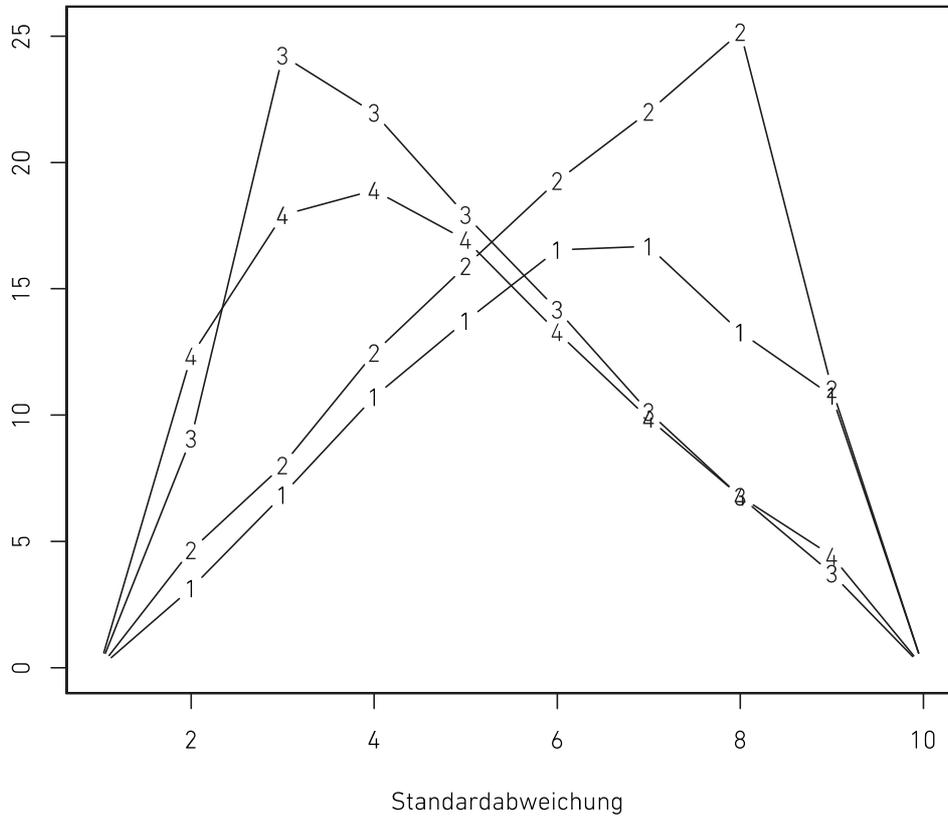


Abbildung 31: Standardabweichung (Mass der Unsicherheit) bei 2.0 auf den Aussensitzpositionen (links bzw. rechts des Sweet-Spots). Links nach rechts (LR), sowie rechts nach links (RL).

Qualitative Beurteilung

Die Teilnehmer der Untersuchung wurden befragt, ob sie die Zahlen (den Vordergrund der Szene) eingebettet in den Hintergrund wahrgenommen haben, ob sie es als „leicht“ empfanden, Angaben über den Bewegungsverlauf des Hördings zu machen und ob sie sich selbst als Teil der Szene empfunden haben.

Auf der mittleren Sitzposition waren unabhängig des Wiedergabeverfahrens (2.0 oder 5.1) mehr Probanden der Ansicht, dass der Vordergrund sich „in der Szene“ befindet, als in den äusseren Sitzpositionen. Ungefähr 1/3 der Teilnehmer meinten auf der Abhörposition inmitten des Sweet-Spots, dass der Vordergrund zum Hintergrund gehöre, während auf den äusseren Sitzpositionen (links und rechts) nur etwa 1/4 der Teilnehmer dieser Ansicht waren.

Bei 2.0 werden in den Aussenhörpositionen Vorder- und Hintergrund von durchschnittlich 23,53 Prozent der Teilnehmer als zusammengehörig erachtet, bei 5.1 sind dies 27,21 Prozent.

Unabhängig des Wiedergabeverfahrens zeigt sich, dass auf den Aussensitzpositionen weniger Teilnehmer eine Verbindung zwischen Vorder- und Hintergrund herstellen konnten, als auf der mittleren Abhörposition.

Auf der rechten Sitzposition zeigt sich bei 2.0 ein Unterschied von 5,88 Prozent zur linken Sitzposition (26,47 Prozent links, 20,59 Prozent rechts).

Teilnehmer, welche die Zahlen eingebettet in den Hintergrund wahrgenommen haben

in absoluten Zahlen und Prozent, unabhängig der Richtung des Pannings

Verfahren	Sitzposition	absolut	Prozent
Stereo	links	9	26,47
	mitte	11	<u>32,35</u>
	rechts	7	20,59
		\emptyset l+r	<u>23,53</u>
5.1	links	9	26,47
	mitte	11,5	<u>33,82</u>
	rechts	9,5	27,94
		\emptyset l+r	<u>27,21</u>

Bei der Wahl der Wegstrecken zeigt sich eine hauptsächliche Abhängigkeit vom Wiedergabeverfahren auf den Aussensitzpositionen. Während bei 5.1 sowohl inmitten des Sweet-Spots, als auch in den äusseren Abhörpositionen ähnlich viele Teilnehmer die Auswahl aus 7 vorgegebenen Wegstrecken als „leicht“ empfanden (jeweils um 63 Prozent), sind beim 2.0-Wiedergabeverfahren inmitten des Sweet-Spots 69 Prozent, aber nur 55 Prozent der Probanden auf den äusseren Sitzpositionen dieser Ansicht.

Beide äussere Sitzpositionen getrennt betrachtet, existiert bei 5.1 zwischen rechter und linker Abhörposition ein Unterschied von 7,36 Prozent. Auf rechter Sitzposition haben weniger Probanden die Aufgabe als leicht empfunden. Auch bei 2.0 haben weniger Untersuchungsteilnehmer, die rechts vom Sweet-Spot sassen, „leicht“ angekreuzt, als links. Der Unterschied beträgt hier 4,41 Prozent.

Teilnehmer, welche die Auswahl einer aus 7 Beispielstrecken als „leicht“ empfanden

in absoluten Zahlen und Prozent, unabhängig der Richtung des Pannings

Verfahren	Sitzposition	absolut	Prozent
Stereo	links	19,5	57,35
	mitte	23,5	<u>69,12</u>
	rechts	18	52,94
		\emptyset l+r	<u>55,15</u>
5.1	links	22,5	66,18
	mitte	21,5	<u>63,24</u>
	rechts	20	58,82
		\emptyset l+r	<u>62,5</u>

Ungefähr 5 Prozent mehr Teilnehmer der Untersuchung haben sich beim Wiedergabeverfahren 5.1. im Vergleich zu 2.0 mehr in die Szene eingebunden gefühlt, egal ob sie sich auf der Aussensitzposition befanden, oder in der Mitte des Sweet-Spots. Die Werte sind bei 5.1. ungefähr 6-7 Prozent höher.

Auf der rechten Sitzposition ist sowohl bei 2.0, als auch bei 5.1. der Prozentwert niedriger (bei Stereo 4,41 Prozent, bei 5.1 um 11,76 Prozent geringer als auf der linken Abhörposition).

Teilnehmer, die sich selbst als Teil der Szene empfunden haben

in absoluten Zahlen und Prozent, unabhängig der Richtung des Pannings (Hervorhebung des ausreissenden Wertes)

Verfahren	Sitzposition	absolut	Prozent
Stereo	links	9	26,47
	mitte	9	26,47
	rechts	7,5	22,06
		∅ l+r	24,26
5.1	links	11,5	33,82
	mitte	11	32,35
	rechts	7	<u>20,59</u>
		∅ l+r	27,205

Einfluss der Position der Untersuchungsleitung im Raum

Bei allen drei qualitativen Fragestellungen (Leichtigkeitgrad der Lokalisation, Verbindung Vorder- zu Hintergrund und Eingebundenheit des Zuhörers in die Szene) ist auffällig, dass viele Werte auf der rechten Sitzposition geringer ausfallen, als im Vergleich zu der linken Sitzposition.

Die „private Zone“ bzw. „persönliche Zone“ wird in der Psychologie als der Bereich zwischen einem Abstand von 50 cm und 1,5 Metern beschrieben (Gerber 262). Die Probanden auf der rechten Hörposition befanden sich ca. 1,5 Meter von der Versuchsleitung entfernt.

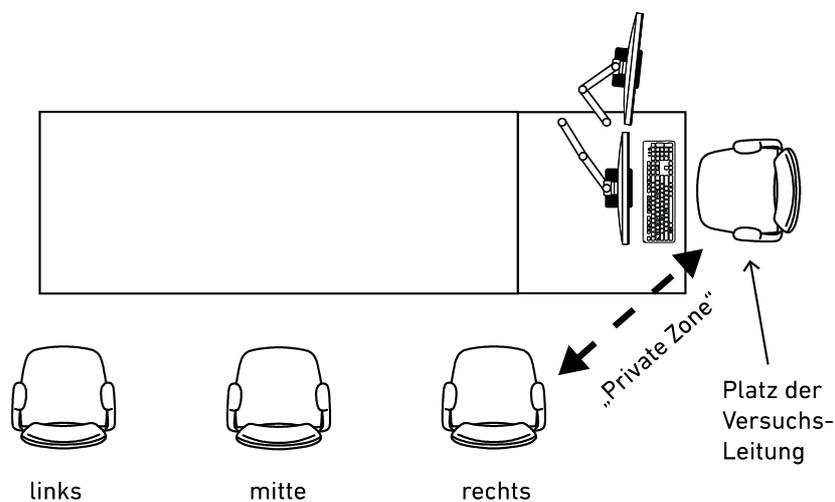


Abbildung 32: Abstand der Versuchsleitung zu den Probanden (eigene Grafik)

Als eine Erklärung für die abweichenden Werte auf der rechten Seite kann die geringere Distanz der Probanden zur Versuchsleitung herangezogen werden.

Demografie

An der Untersuchung waren die Altersgruppen ab 20-29 und älter vertreten. Die Gruppe der 20-29-Jährigen macht mit 60 Prozent den Hauptteil der Probanden aus. 20 Prozent gehörten der Altersgruppe 30-39 an, 8,6 Prozent der Gruppe der 40-49-Jährigen und jeweils ,7 Prozent der Gruppe 50-59 und 60+.

65,7 Prozent der Teilnehmer gaben „männlich“, 31,4 Prozent „weiblich“ und eine Person „anderes“ an.

Demografische Daten

Geschlecht		Tonerfahren		Medienpräferenz			Alter				
w	m	ja	nein	Bild	Ton	beides	20-29	30-39	40-49	50-59	60+
31,4	65,7	60	40	34,3	37,1	28,6	60	20	8,6	5,7	5,7

Als „tonerfahren“ schätzten sich 60 Prozent der Teilnehmer ein. Diejenigen mit Tonerfahrungen gaben u.a. folgendes an: Theorie aus Ton-Vorlesungen, Praxis bei Studioproduktion Ton (HdM), Ausbildung Mediengestalter Bild+Ton, Ausbildung Physikalisch-Technischer Assistent, Erfahrungen in Sounddesign für Film, Live-Mixer, Homerecording, Praktikum im Radio, Musiker.

Zur Medienpräferenz befragt, gaben jeweils etwas mehr als ein Drittel der Probanden eine Präferenz ausschliesslich für Bildmedien oder ausschliesslich für Tonmedien an. 28,6 Prozent der Teilnehmer sind Bild- und Tonmedien zugeneigt.

Zusammenfassung

Untersucht wurde, ob die Gestaltwahrnehmung einen Einfluss auf die Immersion hat und wie sich die Wiedergabeverfahren 2.0 und 5.1. qualitativ unterscheiden.

Befindet sich ein Zuhörer nicht inmitten des Sweet-Spots, zeigen sich Unterschiede zwischen beiden Wiedergabeverfahren hinsichtlich wahrgenommener Form eines Hörbildes und der Sicherheit der Lokalisierung. Der „Präzedenzeffekt“ ausserhalb des Sweet-Spots, bei dem sich die kürzere Laufzeit des Signals des näheren Lautsprechers bei Summenlokalisierung bemerkbar macht, ist bei 2.0 dafür verantwortlich, dass ein technisch gleichförmiges Panning eines Signals von einer Seite zur anderen, je nach Panningrichtung, als nach vorne oder hinten gewölbtes und nach links bzw. rechts verzogenes Hörding wahrgenommen wird. Die Lokalisationsgenauigkeit ist ausserhalb des Sweet-Spots abhängig von der Bewegungsrichtung des Pannings. Es zeigt sich, dass die Gestaltwahrnehmung ausserhalb des Sweet-Spots bei 2.0 eine grosse Auswirkung auf die Lokalisationsgenauigkeit hat. Der Einfluss des Gestaltgesetzes des „gemeinsamen Schicksals“ zeigt sich deutlich. Bei 2.0 ist die Lokalisationsgenauigkeit bei einer nicht-mittigen Abhöre grundsätzlich geringer als bei 5.1.

Unabhängig des Wiedergabeverfahrens (2.0 oder 5.1) zeigt sich, dass die Konsistenz des Hörbildes bei einer nicht-mittigen Sitzposition als geringer eingeschätzt wurde als direkt im Sweet-Spot. Die Eingebundenheit des Vordergrunds in das Hörbild ist vor allem abhängig von der Sitzposition, nicht vom Wiedergabeverfahren.

Inmitten des Sweet-Spots reicht die durchschnittliche Lokalisierungsgenauigkeit sowohl bei 2.0, als auch bei 5.1 sehr gut an das berechnete Ideal heran.

Unabhängig der Sitzposition – ob direkt in der Mitte oder ausserhalb des Sweet-Spots - ist die Eingebundenheit in die Szene bei 5.1. grösser als bei 2.0.

Zusammenfassend kann vermutet werden, dass die Lokalisationsgenauigkeit einen geringeren Einfluss auf die Immersion einer Audiowiedergabe hat, als die Umhüllung mit „Klang von Hinten“. Die immersive Grenze zwischen architektonischem und Wahrnehmungsraum hängt weniger von der Konsistenz des wahrgenommenen Bildes und seiner in ihm enthaltenen Hördingen ab, sondern vielmehr von der „Umrahmung des Bildes“ bzw. dem Anteil an „Hintergrund“, der uns umgibt.

Dass die Lokalisation von Hördingen auf Grund der Gestaltwahrnehmung Abweichungen zu theoretischen, an physikalischen Gesetzmässigkeiten orientierten Überlegungen aufzeigt, wurde deutlich: Die Untersuchung hat gezeigt, dass Gestaltgesetze einen Einfluss auf die Lokalisationsgenauigkeit haben. Eine direkte Wirkung auf die Immersion konnte aber nicht beobachtet werden.

Um einen möglichen Zusammenhang zwischen Gestaltwahrnehmung und Immersion besser beurteilen zu können, wären umfangreichere Untersuchungen mit weiteren Hörbildern (wie mehreren gleichzeitigen Pannings, Kreisbewegungen, Bewegungen in Richtung des Hintergrunds, usw.) nötig.

Literaturverzeichnis

- Arnheim, Rudolf, und Helmut H. Diederichs. *Rundfunk als Hörkunst: Und weitere Aufsätze zum Hörfunk*. 1. Aufl. Suhrkamp Verlag, 2001. Print.
- Arnheim, Rudolf, und Karl Prümm. *Film als Kunst*. 3. Aufl. Suhrkamp Verlag, 2002. Print.
- Aschoff, Volker. *Über das räumliche Hören / Biologische Periodik als selbsterregte Schwingung*. Springer-Verlag, 2013. Print.
- Balázs, Béla u. a. *Der Geist des Films*. 3. Aufl. Suhrkamp Verlag, 2001. Print.
- Bauer, Mathias. „Jahrbuch immersiver Medien 2011. Immersion: Abgrenzung, Annäherung, Erkundung., Im Auftrag des Fachbereichs Medien der Fachhochschule Kiel. by Institut für immersive Medien (Hg):: Marburg: Schüren, 9783894727451 17x24cm, kart. - Schroeders Sammler Bücher“. N.p., o. J. Web. 29 Sep. 2015.
- Bieger, Laura. *Ästhetik der Immersion: Raum-Erleben zwischen Welt und Bild. Las Vegas, Washington und die White City*. 1., Aufl. transcript, 2007. Print.
- „Bildgestaltung“. *Wikipedia* 1 Feb. 2014. *Wikipedia*. Web. 17 Nov. 2015.
- Blauert, Jens. „Untersuchungen zum Richtungshören in der Medianebene bei fixiertem Kopf -“. 1969. *publications.rwth-aachen.de*. Web. 14 Jan. 2016.
- Blauert, Jens, und Jonas Braasch. „Räumliches Hören“. *Handbuch der Audiotechnik*. Hg. von Stefan Weinzierl. Springer Berlin Heidelberg, 2008. 87–121. *link.springer.com*. Web. 29 Sep. 2015. VDI-Buch.
- Blauert, Jens, und Rolf-Dieter Dominicus. „Die Dinge, Gefühle und Gedanken der auralen Welt“. *Faszinosum „Klang“: Anthropologie - Medialität - kulturelle Praxis*. Hg. von Wolf Gerhard Schmidt. Walter de Gruyter GmbH & Co KG, 2015. Print.
- Bruhn, Herbert. „Tonpsychologie – Gehörpsychologie – Musikpsychologie“. *Musikpsychologie: Ein Handbuch*. Hg. von Rolf Oerter, Helmut Rösing, und Herbert Bruhn. 4. Aufl. Reinbek bei Hamburg: Rowohlt Taschenbuch, 2002. Print.
- Bruhn, Herbert, und Dieter Michel. „Hören im Raum“. *Musikpsychologie: Ein Handbuch*. Hg. von Herbert Bruhn, Rolf Oerter, und Helmut Rösing. 4. Aufl. Reinbek bei Hamburg: Rowohlt Taschenbuch, 2002. Print.
- Crow, David. *Zeichen: eine Einführung in die Semiotik für Grafikdesigner*. Stiebner Verlag GmbH, 2012. Print.
- Dickreiter, Michael. *Handbuch der Tonstudientechnik Bd. I. 6. verb. A. Saur, K G*, 1997. Print.
- Ebeling, Martin. „Die Ordnungsstrukturen der Töne“. *Faszinosum „Klang“: Anthropologie - Medialität - kulturelle Praxis*. Hg. von Wolf Gerhard Schmidt. Walter de Gruyter GmbH & Co KG, 2015. Print.
- Foucault, Michel. „Andere Räume“. *Aisthesis - Wahrnehmung heute oder Perspektiven einer anderen Ästhetik. Essais*. Hg. von Karlheinz Barck u. a. 1. Aufl. Leipzig: Reclam, 1990.

Print.

- Friedrich, Thomas, und Jörg H. Gleiter. *Einführung und phänomenologische Reduktion: Grundlagentexte zu Architektur, Design und Kunst*. LIT Verlag Münster, 2007. Print.
- Friesecke, Andreas. *Die Audio-Enzyklopädie: Ein Nachschlagewerk für Tontechniker*. Walter de Gruyter GmbH & Co KG, 2014. Print.
- Gerber, Markus. *Pädagogische Psychologie im Sportunterricht: Ein Lehrbuch in 14 Lektionen*. Meyer & Meyer Verlag, 2015. Print.
- Gibson, James J., Ivo Kohler Angelika Kohler (Übers.) u a PSYCHOLOGIE-. *Die Sinne und der Prozeß der Wahrnehmung*. Bern u.a., Hans Huber, 1973. Print.
- Görne, Thomas. *Tontechnik: Hören, Schallwandler, Impulsantwort und Faltung, digitale Signale, Mehrkanaltechnik, tontechnische Praxis*. 4., aktualisierte Auflage. Carl Hanser Verlag GmbH & Co. KG, 2014. Print.
- Günzel, Stephan. „Teil 2 - Phänomenologie der Räumlichkeit“. *Raumtheorie: Grundlagentexte aus Philosophie und Kulturwissenschaften*. Originalausgabe. Suhrkamp Verlag, 2006. Print.
- Haan, Gerhard de, und Tobias Rülcker. *Der Konstruktivismus als Grundlage für die Pädagogik*. Peter Lang, 2009. Print.
- Heer, Daniel. *Projektstudios für Filmsound: Audiopostproduktion und Studioteknik*. Diplomica Verlag, 2011. Print.
- Henle, Hubert. *Das Tonstudio Handbuch: Praktische Einführung in die professionelle Aufnahmetechnik. Grundlagen der Akustik. Analoge und digitale Audiotechnik. Auswahlkriterien ... Lautsprecher und Regieraum-Design*. 5. Auflage. GC Carstensen Verlag, 2001. Print.
- Herder, Johann Gottfried u. a. *Johann Gottfried von Herder's Sämmtliche Werke ...* In der C. Haas'schen Buchhandlung, 1819. Print.
- Herkner, Werner. *Psychologie*. Springer-Verlag, 2013. Print.
- Höfliger, Jean-Claude. *Jacques Derridas Husserl-Lektüren*. Königshausen & Neumann, 1995. Print.
- Husserl, Edmund. „Allgemeine Einführung in die reine Phänomenologie“. *Ideen zu einer reinen Phänomenologie und phänomenologischen Philosophie*. 1. Aufl. Hamburg: Meiner, F, 2009. Print.
- Husserl, Edmund, und Sebastian Luft. *Zur Phänomenologischen Reduktion: Texte aus dem Nachlass (1926–1935)*. Springer-Verlag, 2013. Print.
- Jecklin, Jürg. *Theorie der Tontechnik, 2. Gehör*. 2003.
- Kerins, Mark. *Beyond Dolby (Stereo): Cinema in the Digital Sound Age*. Indiana University Press, 2010. Print.
- Krämer, Oliver. *Strukturbilder, Sinnbilder, Weltbilder: Visualisierung als Hilfe beim Erleben und Verstehen von Musik*. 1. Aufl. Wißner-Verlag, 2011. Print.
- Kruse, Lenelis. *Räumliche Umwelt: Die Phänomenologie des räumlichen Verhaltens als*

- Beitrag zu einer psychologischen Umwelttheorie*. Walter de Gruyter, 1974. Print.
- Lang, Patrick u. a. *Über musikalische Phänomenologie: Ein Vortrag und weitere Materialien*. 2., überarbeitete und um Notenbeispiele erweiterte Auflage der Erstausgabe von 2001. Wißner-Verlag, 2008. Print.
- Peres, Constanze. „Raumzeitliche Strukturgemeinsamkeiten bildnerischer und musikalischer Werke“. *Zeit und Raum in Musik und Bildender Kunst*. Hg. von Tatjana Böhme-Mehner und Klaus Mehner. Köln: Böhlau Köln, 2000. Print.
- Schmicking, Daniel. *Hören und Klang: Empirisch phänomenologische Untersuchungen*. 1. Aufl. Würzburg: Königshausen u. Neumann, 2003. Print.
- Schmied, Gerhard. *Soziale Zeit*. Duncker & Humblot. Print.
- Sengpiel, Eberhard. „Praktische Daten zur Lokalisation von Phantomschallquellen bei Intensitäts- und Laufzeit-Stereophonie - PraktischeDatenZurStereo-Lokalisation.pdf“. N.p., 2009. Web. 18 Nov. 2015.
- Stadler, Michael, und Wolfgang Wildgen. „Semiotik und Gestalttheorie“. *Handbücher zur Sprach- und Kommunikationswissenschaft: Handbooks of linguistics and communication science. Semiotik : ein Handbuch zu den zeichentheoretischen Grundlagen von Natur und Kultur. ...* Hg. von Roland Posner und Herbert Ernst Wiegand. Walter de Gruyter, 2003. Print.
- Staubmann, Helmut. „Handlungstheoretische Systemtheorie: Talcott Parsons“. *Soziologische Theorie: Abriß der Ansätze ihrer Hauptvertreter*. Hg. von Julius Morel. Oldenbourg Verlag, 2001. Print.
- „Tongestaltung“. *Wikipedia* 3 Nov. 2015. *Wikipedia*. Web. 17 Nov. 2015.
- von Ehrenfels, Christian. „über gestaltqualitäten - Google-Suche“. *Über Gestaltqualitäten*. N.p., 1890. Web. 17 Nov. 2015.
- von Foerster, Heinz. „Wahrnehmen“. *Aisthesis - Wahrnehmung heute oder Perspektiven einer anderen Ästhetik. Essais*. Hg. von Karlheinz Barck u. a. 1. Aufl. Leipzig: Reclam, 1990. Print.
- Wohler, Arnold. *Synästhesie als ein strukturbildendes Moment in der Kunst des 20. Jahrhunderts: unter besonderer Berücksichtigung von Malerei und Musik*. Waxmann Verlag, 2010. Print.

Abbildungsverzeichnis

Abbildung 1: Simultanität der Räume (eigene Grafik).....	9
Abbildung 2: Sinnvolle Zusammenfassung von Parallelogramm und Quadrat zu einem Würfel (eigene Grafik).....	20
Abbildung 3: Ebenen räumlichen Hörens (Blauert, Jens, und Jonas Braasch, Seite 88).....	26
Abbildung 4: Schematic illustration of the localization of narrow-band sounds in the median plane, irrespective of the position of the sound source (Zwicker, Eberhard und Fastl, Hugo, Seite 319).....	27
Abbildung 5: Unterschiede der Schallwellen, die die beiden Ohren erreichen (Gibson, James J., Seite 112).....	28
Abbildung 6: Einohriges Richtungshören (Jecklin, Jürg, Seite 12).....	29
Abbildung 7: Physikalischer und wahrgenommener Raum sind zweierlei (eigene Grafik).....	31
Abbildung 8: Luftabsorption bei 50% Luftfeuchtigkeit und 20° Umgebungstemperatur (Andreas Friesecke, Seite 144).....	32
Abbildung 9: Summenlokalisationskurven für Pegel- bzw. Laufzeitdifferenzen der beiden Lautsprecher signale. Duchgezogene Linie: Zuhörer kopf fixiert; gestrichelte Linie: Kopf frei beweglich (Blauert, Jens, und Jonas Braasch, Seite 103, nach Wendt 1961).....	33
Abbildung 10: Wirkung der Zeitverzögerung zwischen identischen Signalen bei Stereowiedergabe (Dickreiter, Michael, Seite 129).....	34
Abbildung 11: Lautsprecheranordnung nach ITU-R BS.775-1 (Henle, Hubert, Seite 17).....	35
Abbildung 12: Wo ist Hinten? (Eigene Grafik).....	37
Abbildung 13: Definition der Nachhallzeit durch den Abfall des Schalldruckpegels nach Abschalten der Schallquelle (Dickreiter, Michael, Seite 31).....	38
Abbildung 14: Versuchsaufbau, Draufsicht (eigene Grafik).....	42
Abbildung 15: Abfolge der Zahlen und der "Da Da" (eigene Grafik).....	44
Abbildung 16: Spuren der 2.0 und 5.1-Produktion (eigene Grafik).....	44
Abbildung 17: Dämpfungsverlauf, konstante Leistung im Stereo-Panorama (eigene Grafik, nach Sengpiel "Das theoretische Panpot").....	45
Abbildung 18: Einstellungen des Halls vorne und hinten (Screenshot).....	46
Abbildung 19: Entwicklungsstufen der bildlichen Darstellung von Rhythmen bei Kindern (Rösing, Helmut und Bruhn, Herbert, Seite 22, nach Hildebrandt und Uptis).....	48

Abbildung 20: Ablauf der Studie (eigene Grafik).....	51
Abbildung 21: Übersicht über Lokalisierung, Hörposition L, M, R - horizontal: wahrgenommene Position, vertikal: einzelne Teilnehmer (eigene Grafik). Vertikale Achse: Die unterschiedlichen Teilnehmer.....	59
Abbildung 22: Lokalisierung bei Hörposition Mitte, „Sweet-Spot“ - horizontal: wahrgenommene Position, vertikal: einzelne Teilnehmer (eigene Grafik).....	59
Abbildung 23: Lokalisierung bei Hörposition Links - horizontal: wahrgenommene Position, vertikal: einzelne Teilnehmer (eigene Grafik).....	60
Abbildung 24: Abbildung 20: Lokalisierung bei Hörposition Rechts - horizontal: wahrgenommene Position, vertikal: einzelne Teilnehmer (eigene Grafik).....	60
Abbildung 25: Lokalisation der Zahlen 1-10. Horizontal: Zahlen, Vertikal: % rechts im Stereopanorama, 1+2 Hörposition links 2.0, 9+0 Hörposition rechts (eigene Grafik).....	62
Abbildung 26: Lokalisierung auf der Stereobasis, % rechts - 2.0 in Abhängigkeit von der Sitzposition (L, M, R).....	63
Abbildung 27: Lokalisierung auf der Stereobasis, % rechts - 5.1 in Abhängigkeit von der Sitzposition (L, M, R).....	64
Abbildung 28: Sitzposition links, Vergleich 5.1 und 2.0, % rechts (eigene Grafik).....	65
Abbildung 29: Sitzposition mitte, Vergleich 5.1 und 2.0, % rechts (eigene Grafik).....	65
Abbildung 30: Sitzposition rechts, Vergleich 5.1 und 2.0, % rechts (eigene Grafik).....	65
Abbildung 31: Standardabweichung (Mass der Unsicherheit) bei 2.0 auf den Aussensitzpositionen (links bzw. rechts des Sweet-Spots). Links nach rechts (LR), sowie rechts nach links (RL).....	73
Abbildung 32: Abstand der Versuchsleitung zu den Probanden (eigene Grafik).....	77

Abbildung Einband: Eigene Grafik.

