TECHNISCHE UNIVERSITÄT BERLIN
Fakultät I - Geisteswissenschaften
Fachgebiet Audiokommunikation

Magisterarbeit

Audiovisuelle Raumwahrnehmung

vorgelegt von:

Student: Matthias Jentsch
Studiengang: Kommunikationswissenschaft
[redacted]

Berlin, den 30. September 2012

Betreuer:

Dr. Hans-Joachim Maempel

Danksagung

Ich möchte mich für die Unterstützung von Dr. Hans-Joachim Maempel bedanken, der mir bei allen inhaltlichen Fragen zur Seite stand.
Auch danke ich Alexander Lindau, Frank Schulz, Fabian Brinkman, Michael Horn und Dr. Steffen Lepa für ihre bereitwillige Unterstützung bei der Akquise der Daten und der technischen Umsetzung des Hörversuchs.
Ebenfalls danke ich Tobias Marx, der trotz engem Zeitmanagement immer ein offenes Ohr für mich hatte.
Mein aufrichtiger Dank gilt auch meinen Eltern, die mich bei allem was ich tue unterstützen.
Schließlich danke ich meiner Freundin Christine Worth für die andauernde Motivierung und nicht zuletzt für ihre Geduld und ihr Verständnis.

Inhalt

1 EINLEITUNG 1

2 STAND DER FORSCHUNG 2

2.1 MULTIMODALE WAHRNEHMUNG 2

2.2 AUDIOVISUELLE INTERAKTION 3

2.3 AUDITIVE WAHRNEHMUNG VON RÄUMEN 6

2.4 AUDIOVISUELLE WAHRNEHMUNG VON RÄUMEN 9

3 ZIELSETZUNG 13

4 MATERIAL UND METHODE 15

4.1 ÜBERSICHT 15

4.2 VERSUCHSDESIGN 15

4.2.1 Variation der optischen und akustischen Raumpräsenz 15

4.2.2 Konfliktreize - Variation der optischen und akustischen Raumeigenschaften 16

4.3 STIMULI 17

4.3.1 Auswahlkriterien und Beschreibung der simulierten Räume 17

4.3.2 Darbietungsinhalte der Stimuli 19

4.4 ERHEBUNGSINSTRUMENT 20

4.5 DIE STICHPROBE 21

4.6 STATISTISCHE ANALYSEVERFAHREN 21

4.7 AKUSTISCHE SIMULATION 22

4.7.1 Relative Lautheit 23

4.8 OPTISCHE SIMULATION 24

4.9 TECHNISCHER VERSUCHSAUFBAU 27

4.9.1 Datenerhebung - Limesurvey 27

4.9.2 Stimulus-Wiedergabe 28

4.9.3 Technische Realisation 28

4.9.4 Versuchsraum des Fachgebiets Audiokommunikation 30

5 AUSWERTUNG 31

5.1 VARIATION DER OPTISCHEN UND AKUSTISCHEN RAUMPRÄSENZ 31

5.1.1 Variation des Faktors Raum 32

5.1.2 Variation des Faktors Präsentationsmodus 34

5.1.3 Interaktionseffekte zwischen den Faktoren Präsentationsmodus und Raum 35

5.2 KONFLIKTREIZE - VARIATION DER OPTISCHEN UND AKUSTISCHEN RAUM 38

5.2.1 Variation des Faktors Darbietungsinhalt 38

5.2.2 Variation des Faktors optischer Raum 40

5.2.3 Variation des Faktors akustischer Raum 43

5.2.4 Interaktionseffekte 48

5.3 METHODISCHE SONDERFRAGE: ZUSAMMENHANG DER PERZEPTIVEN VOLUMENSKALEN 50

5.3.1 Perzeptive Raumgröße 50

5.3.2 Zusammenhang der perzeptiven und realen Volumina 53

5.4 ZUSAMMENHANG PHYSIKALISCHER UND PERZEPTIVER MERKMALE 56

6 DISKUSSION 66

6.1 UNIMODALE VISUELLE MERKMALE 66

6.2 UNIMODALE AUDITIVE MERKMALE....68
6.3 MULTIMODALE MERKMALE....73

7 SCHLUSSFOLGERUNG....79

8 ANHANG....82
8.1 FRAGEBOGEN....82
8.2 OPERATIONALISIERUNG UND KODIERUNG DER MERKMALE....84
8.3 PHYSIKALISCHE KENNWERTE DER RÄUME....85
8.4 STATISTISCHE KENNZAHLEN....86
8.4.1 Variation der optischen und akustischen Raumpräsenz....86
8.4.2 Konfliktreize – Variation der optischen und akustischen Raumeigenschaften....87
8.5 ABBILDUNGEN - ZUSAMMENHANG DER PERZEPTIVEN UND REALEN VOLUMENSKALEN 90

9 LITERATURVERZEICHNIS....91

Verzeichnis der Abbildungen

Abbildung 4.1: Grundrisse der Räume.....19
Abbildung 4.2: Messroboter FABIAN.....22
Abbildung 4.3: Sichtwinkel: a.) Sichtwinkel eines Objektivs (Brennweite 27 mm) b.) Sichtwinkel beim Betrachten des Bildschirms c.) perspektivische Verzerrung des Versuchsaufbaus.....25
Abbildung 4.4: jeweils linkes Bild der im Versuch verwendeten stereoskopischen Fotos. Oben links: Raum Kt Oben rechts: Raum Kh. Unten links: Raum Gt, Unten rechts: Raum Gh. 26
Abbildung 4.5: FWONDER Instanz.....28
Abbildung 4.6: Schematischer Versuchsaufbau.....30
Abbildung 5.46: Regressionsmodell der Merkmale Volumen (Skala) und Volumen (berechnet).....51
Abbildung 5.47: Vergleich der Mittelwerte: Volumen (Skala), Volumen (berechnet)..52
Abbildung 5.48: Physikalisches und perzeptives Volumen für die optoakustischen Stimuli mit sprachlichem Inhalt.....53

Verzeichnis der Tabellen

Tabelle 1:Variation der optischen und akustischen Raumpräsenz: optisch (o), optoakustisch (oa), optisch (o).....15
Tabelle 2: Variation der optischen und akustischen Raumeigenschaften (k -Konfliktreiz, p – passender Reiz).....16
Tabelle 3: Variation: Volumen, Absorptions- koeffizienten.....18
Tabelle 4: Bezeichnung der Räume.....18
Tabelle 5: Perzeptive Maße.....20
Tabelle 6: Auswertung perzeptiver Merkmale für die Variation der optischen und akustischen Raumpräsenz.....31
Tabelle 7: Physikalische und perzeptive Volumina.....53

Inhalt der Daten-CD

- Pdf-Version der vorliegenden Arbeit
- Für die Simulation verwendete akustische und optischen Raumdaten
- Pure Data-Skripte
- Empirische Daten
- Statistische Auswertungen (SPSS-Format)
- Verwendete Literatur

1 Einleitung

Diese Arbeit entstand im größeren Zusammenhang der SEACEN- Forschungsgruppe (*Simulation and Evaluation of Acoustical Environments*), die sich mit der Erforschung und Weiterentwicklung von Auralisation befasst. Sie dient als Vorstudie für das Forschungsprojekt „Audio-visual perception of acoustical environments“, bei dem mittels stereoskopischer Projektionen von Konzertsälen und simultaner Binauralsynthese derselben die audiovisuelle Raumwahrnehmung erforscht werden soll.

Die audiovisuelle Wahrnehmung von Räumen ist ein wichtiger Aspekt sowohl simulierter, als auch realer akustischer Umgebungen. Bisherige Studien sind geprägt durch einen hochgradig technischen Ansatz, wie die Verwendung modellbasierter Raumsimulationen, verschiedenste Auralisationsverfahren und hochgradig spezifischer unabhängiger und abhängiger Variablen. Zudem werden eine Vielzahl verschiedener empirischer Paradigmen und Methoden verwendet, wodurch die bisherigen Ergebnisse Konsistenz und Adaptierbarkeit vermissen lassen. Eine einheitliche Forschungsstrategie lässt sich zu diesem Zeitpunkt noch nicht erkennen. Aus diesem Grund wurde für die vorliegende Studie ein explorativer Forschungsansatz gewählt.

Ein wichtiger Aspekt dieser Studie ist die Gewährleistung einer größtmöglichen externen und ökologischen Validität, welche durch eine datenbasierte Simulation der Räume sichergestellt werden soll. Zudem kann eine allgemeine Quantifizierung des Beitrags optischer und akustischer Stimuluskomponenten auf die Beurteilung perzeptiver Raummerkmale nur dann sinnvoll erfolgen, wenn eine maximale Entsprechung von optischen und akustischen Stimuluskomponenten gewährleistet ist.

Eine Neuerung dieser Studie besteht in der Anwendung eines Konflikreizparadigmas, in dem optische und akustische Stimuluskomponenten kategorial variiert werden. Mittels anschließender Varianzanalyse kann somit der Einfluss akustischer und optischer Rauminformationen auf die Beurteilung verschiedener wahrnehmungsbezogener Merkmale quantifiziert werden. Außerdem wurde die optische und akustische Raumpräsenz variiert, wobei den Probanden akustische, optische und optoakustische Stimuli präsentiert wurden.

Die gewonnenen Erkenntnisse sollen zu einem tieferen Verständnis multimodaler Prozesse der Musik- und Sprachwahrnehmung in Räumen beitragen.

2 Stand der Forschung

2.1 Multimodale Wahrnehmung

Im Alltag werden Informationen aus der Umwelt durch die verschiedenen Sinne wie Hören, Sehen, Geruchssinn oder Tastsinn aufgenommen. Dabei können Informationen aus verschiedenen Sinneskanälen demselben Ereignis zugeordnet werden. Eine bekannte Person können wir beispielsweise anhand des Aussehens, am Klang der Stimme oder sogar am Geruch erkennen. Die Informationen aus verschiedenen Sinneskanälen werden zu einem einzigen Ereignis kombiniert, welches auch Perzept genannt wird. Die Sinneskanäle werden auch als Modalitäten bezeichnet. Für den gesamten Prozess des Zusammenführens verschiedener Eindrücke hat sich der Begriff multimodale (auch crossmodale) Integration etabliert. Dieser multimodale Wahrnehmungsprozess stellt eine komplexe Anforderung an das menschliche Wahrnehmungssystem dar. Dazu gehört ebenso die Entscheidung, ob Eindrücke verschiedener Modalitäten einem einheitlichen Ursprung zugeordnet werden können, oder getrennt wahrgenommen werden. Diese Entscheidung fällt vorbewusst anhand verschiedener Kriterien. Im Bereich der audiovisuellen Wahrnehmung hat sich unter anderem die zeitliche und räumliche Nähe als ein wichtiges Kriterium herausgestellt [Slutsky2001]. Ein eindrückliches Beispiel audiovisueller zeitlicher Diskrepanz ist die Wahrnehmung von Blitz und Donner. Aufgrund der physikalischen Tatsache, dass Licht sich mit einer höheren Geschwindigkeit (ca. 300000 km/s) ausbreitet als Schall (ca. 343 m/s) treffen optische und akustische Wahrnehmungsinhalte oft mit einer zeitlichen Differenz von mehreren Sekunden beim Beobachter ein. Auch wenn dem Betrachter bekannt ist, dass Blitz und Donner im selben Augenblick entstehen, werden die auditiven und visuellen Wahrnehmungsinhalte dieses Ereignisses als getrennt erlebt. Die Kombination multimodaler Ereignisse zu einem einzigen Perzept ist aber nicht nur durch zeitliche Merkmale, sondern auch durch räumliche [Vroomen2004], semantische [Calvert2004] und stimulusspezifische Kriterien [Stein1996] bedingt.

Multimodale Integration versetzt den Menschen in die Lage Informationen verschiedener Sinne zu vereinen und fehlende Informationen gegebenenfalls zu ergänzen. Kann man die Lippenbewegungen eines Sprechers beim Zuhören mitverfolgen,

verbessert sich die Sprachverständlichkeit [Rossing2007]. Stein und Meredith [1993] konnten nachweisen, dass auditive und visuelle Reize geringer Intensität eher entdeckt werden, wenn sie gemeinsam präsentiert werden.

Wie aber interagieren die verschiedenen Sinnesmodalitäten miteinander? Im gegebenen Kontext ist die Interaktion von Hören und sehen von Bedeutung und soll im folgenden Kapitel näher beleuchtet werden.

2.2 Audiovisuelle Interaktion

Audiovisuelle Interaktionseffekte sind schon seit einigen Jahren Gegenstand der Forschung. Meist werden basale Stimuli verwendet um einfache Zusammenhänge zwischen den Modalitäten aufzudecken.
Audiovisuelle Wahrnehmung kann definiert werden als gleichzeitiges Sehen und Hören, wobei sich auditiver und visueller Modus gegenseitig beeinflussen. Schon in den 1960er Jahren wurde belegt, dass Menschen die Reizintensität in einer Modalität in Beziehung zur Reizintensität einer anderen Modalität setzen können. In einer Studie von Steven et al. [1965] gelang es den Testpersonen der Studie die Helligkeit eines Lichtes der Lautheit eines Geräusches zuzuordnen und anders herum [Steven1965]. Das spricht dafür, dass es einen übergeordneten Prozess gibt, der es ermöglicht Stimulusintensitäten unabhängig von der Modalität zu bewerten. Eine intermodale Repräsentation des Ereignisses macht es einerseits möglich visuelle und auditive Perzepte zu vergleichen und führt anderseits zu Interaktionen der Modalitäten. Eine Interaktion liegt vor, wenn entweder das Perzept einer Modalität das Perzept einer anderen Modalität beeinflusst. Eine Interaktion liegt ebenfalls vor, wenn das gemeinsame multimodale (z. b. audiovisuelle) Perzept durch eine der beiden Modalitäten beeinflusst wird. Fastl konnte in einer Untersuchung nachweisen, dass ein roter Zug 15 % lauter wahrgenommen wird als ein grüner [Fastl2004]. Ähnliche Experimente zum Einfluss der visuellen Modalität auf die Intensitätswahrnehmung der auditiven Modalität zeigten, dass weißes Rauschen positiver bewertet wird, wenn es vom Bild eines Wasserfalles begleitet wird [Viollon1998] oder, dass Motorengeräusche zusammen mit einem Bild präsentiert 10 dB lauter sein können um den gleichen Belästigungsgrad zu erzielen [Hashimoto2001]. Diese Studien belegen einen audiovisuellen Inter-

aktionseffekt, bei dem die visuelle Modalität die wahrgenommene Intensität des auditiven Perzepts beeinflusst.

Einen eindrücklichen Beweis für audiovisuelle Interaktion bei der phonetischen Lauterkennung, lieferte McGurk [McGurk1976]. Er präsentierte den Testpersonen eine optoakustische Aufzeichnung eines Sprechers. Der optische Bestandteil des Stimulus zeigte den Kopf des Sprechers, wie er die Silbe „Ga“ formulierte. Der akustische Stimulusbestandteil bestand aus der gesprochene Silbe „Ba“. Das im Nachhinein als McGurk-Effekt bekannt gewordene Phänomen bestand darin, dass die Testpersonen die Silbe „Da“ wahrnahmen. Auf bemerkenswerte Weise verschmelzen die divergierenden Stimuluskomponenten zu einem multimodalen Perzept, dass in diesem Fall eine Mischung der auditiven und visuellen Information darstellt. Grund dafür ist, dass der Artikulationsort der Silbe „Da“ (alveolar) zwischen dem der Silbe „Ga“ (velar) und „Ba“ (labial) liegt.

McGurks Studie ist insofern exemplarisch für andere Studien der audiovisuellen Interaktionsforschung, als dass sie das Konfliktreizparadigma anwendet. Dabei werden Stimuli verwendet optoakustisch divergente Stimuli. Diese zeichnen sich dadurch aus, dass sich die Stimulusinhalte der verschiedenen Modalitäten widersprechen. Im Falle des McGurk-Effekts führt der Konfliktreiz zu einem fusionierten Perzept. Je nach Grad der Divergenz kann ein Konfliktreiz jedoch auch zu mehrdeutigen oder getrennten (segregierten) Perzepten führen.

Prägnante Beispiele für letzteres finden sich in Studien zur zeitlichen Relation zwischen optischen und akustischen Stimuluskomponenten. Eine zeitliche Diskrepanz (Asynchronität) besteht, wenn der Ton eines Ereignisses dem Bild desselben voraus oder hinterher läuft. Verschiedene Studien konnten belegen, dass Asynchronität nur erkannt wird, wenn ein gewisser Grad an Bild-Ton-Versatz überschritten wird [Dixon1980][Rudloff1997]. Die Ergebnisse von Dixon und Rudloff zeigen die allgemeine Tendenz, dass ein vorauslaufender Ton früher als asynchron erkannt wird, als ein nachlaufender Ton. Ein Unterschied zeigt sich jedoch bei der Art des Inhaltes. Asynchronität wird bei singulären Ereignissen, etwa einem Schlag mit einem Hammer, bei vorauslaufendem Ton bei 60-70 ms erkannt, bei nachlaufendem Ton bei 80-180 ms. Werden sprachliche Stimuli verwendet ist die Toleranz gegenüber dem Tonversatz größer. Hier wird Asynchronität bei vorauslaufendem Ton bei ca. 80-130 ms, bei nachlaufendem Ton bei 140-250 ms erkannt. Allgemein kann festgestellt werden, dass auditive und visuelle Wahrnehmungsinhalte bei Überschreitung eines gewissen

Divergenzgrades nicht mehr zu einem einheitlichen Perzept integriert werden. Es entsteht ein mehrdeutiges oder getrenntes Perzept.

Ein weiterer Aspekt audiovisueller Interaktionseffekte ist die Integration räumlich divergenter Reize. Im Alltag begegnen sie uns beispielsweise beim Fernsehen oder im Kino. Die Lautsprecher sind meistens seitlich angebracht, also oft räumlich divergent zu den Bildern bzw. Sprechern. Trotzdem scheint gesprochene Sprache aus dem Mund eines Sprechers zu kommen. Die Zuordnung eines Reizes zu einer bestimmten Position wird auch Lokalisation genannt. Wird eine gewisse räumliche Divergenz zwischen optischen und akustischen Schlüsselmerkmalen nicht überschritten, verschmelzen diese bei der audiovisuellen Lokalisation des Ereignisses zu einer Kompromissposition, die sich näher oder sogar vollständig am Ort des optischen Reizes befindet. Auf Grund der Dominanz der visuellen Modalität bei der audiovisuellen Lokalisation, wird dieses Phänomen auch als Visual-Capture-Effekt bezeichnet [Mershon1980]. Ein Spezialfall ist der sogenannte Bauchredner- oder Ventriloquismus-Effekt [Welch1980, Jack1973]. Er bezeichnet die Verschiebung der Lokalisation der Stimme eines Bauchredners hin zur sichtbaren, seitlich versetzen Puppe. Gardner [Gardner1968] konnte einen ähnlichen Effekt für eine Divergenz in der Entfernung der optischen und akustischen Quelle feststellen. Er kam zu dem Ergebnis, dass Testpersonen in einem reflektionsarmen Raum dazu tendieren den Klang an einem nahe gelegenen Lautsprecher zu lokalisieren, auch wenn die eigentliche Klangquelle ein weiter entfernter, unsichtbarer Lautsprecher ist. Dieses Phänomen bezeichnet er als Proximity-Effekt. Mershon et al. [1980] bestätigen den Effekt auch für normale Räume, in denen der Testperson durch die Raumakustik mehr auditive Entfernungsmerkmale, beispielsweise das Verhältnis von Direkt- zu Diffusschall, zur Verfügung stehen. Zudem stellt er fest, dass der Effekt bei einem zu großen Entfernungsunterschied nicht auftritt. Da der beobachtete Effekt sich nicht auffallend von anderen Interaktionseffekten der audiovisuellen Lokalisation unterscheidet, schlägt er vor, keinen eigenen Begriff (Proximity-Effekt) zu verwenden, sondern ihn als weiteres Beispiel für den allgemeineren Visual-Capture-Effekt zu behandeln. Zahorik bildete den Versuch in einem normalen Raum, mit einer Anordnung fünf sichtbarer, in verschiedenen Entfernungen positionierter Lautsprecher nach [Zahorik2001]. Mit diesem Versuchsaufbau konnte kein Visual-Capture-Effekt nachgewiesen werden. Zahorik führt dies auf die Anwesenheit auditiver Schlüsselmerkmale durch den Raum und die Sichtbarkeit der fünf potentiellen Klangquellen zurück.

Der Frage, inwiefern verschiedene Merkmale von Räumen die auditive Modalität beeinflussen, wird im nächsten Kapitel nachgegangen.

2.3 Auditive Wahrnehmung von Räumen

Viele Studien beschäftigen sich mit der auditiven Wahrnehmung von Räumen. Dabei liegt der Fokus im Allgemeinen darauf, Zusammenhänge zwischen technisch messbaren (objektiven) Merkmalen und wahrnehmungsbezogenen (subjektiven) Merkmalen von Räumen zu identifizieren. Wichtige perzeptive Merkmale sind unter anderem: Raumeindruck (*spatial impression*), perzeptive Raumgröße (*perceived room size*), Präsenz (*presence*), Wärme (*warmth*) oder Lautheit (*loudness*).

Der Begriff Raumeindruck wurde in den 1970 Jahren geprägt. Verschiedene Studien bestätigen einen linearen Zusammenhang des frühen Seitenschallgrades zum Raumeindruck [Marshall1967, Barron1971, Barron1981]. Blauert et al. [1986] zeigen, dass der Raumeindruck einen starken Zusammenhang mit der Präferenz von Konzertsälen aufweist und damit auch als Gütekriterium für die Raumakustik fungieren kann. Die Ergebnisse deuten darauf hin, dass die spektralen Komponenten der frühen seitlichen Reflektionen unter 3 kHz zur Wahrnehmung der Bildtiefe beitragen, während die spektralen Komponenten über 3 kHz zu einer Verbreiterung des Klangbildes führen. Dieses Ergebnisse gaben erste Anhaltspunkte dafür, dass eine Unterteilung des Raumeindrucks in verschiedene untergeordnete Aspekte notwendig ist. Folgerichtig wurde das Konzept des perzeptiven Raumeindrucks durch Bradley und Soulodre erweitert, indem sie zwischen scheinbarer Quellbreite und Umhüllung differenzierten [Bradley1995a, Bradley1995b].

Die scheinbare Quellbreite (ASW – *apparent source width*) wird auch als Räumlichkeit (*spaciousness*) bezeichnet. Nach Beranek spricht man von Räumlichkeit, wenn: „die aufgeführte Musik für den Zuhörer von einer Quelle zu stammen scheint, die breiter ist, als die optische Breite der eigentlichen Quelle“ [Beranek1996]. Long [2006] definiert sie ähnlich als: „Wahrgenommene Verbreiterung der Quelle, über die sichtbaren Grenzen hinaus “.

Ein robuster Prädiktor für die scheinbare Quellbreite ist der frühe interaurale Korellationskoeffizient ($IACC_E$), welcher die Ähnlichkeit der Schallsignale an beiden Ohren innerhalb der ersten 80 ms beschreibt [Beranek1996]. Eine geringe Korrelation

zwischen den Signalen, hervorgerufen durch viele starke seitliche Schallreflektionen, erweckt beim Hörer den Eindruck einer breiteren Schallquelle. Einen weiterer Prädiktor ist die beim Zuhörer eintreffende Schallenergie des Diffusschalls innerhalb der ersten 80 ms nach Eintreffen des Direktschalls dar [Soulodre2003].

Die Umhüllung (LEV – *listener envelopment*) kann definiert werden als: „der Eindruck sich innerhalb eines diffusen Schallfeldes zu befinden, bzw. davon umgeben zu sein“ [Rossing2007]. Nach Bradley und Soulodre bezieht sich die Umhüllung auf das Gefühl des Zuhörers, von Schall umgeben und eingehüllt zu sein [Soulodre1995]. Durch künstlich erzeugte Schallfelder mit drei oder fünf Lautsprechern in einem reflektionsarmen Raum, identifizierten sie als Einflussgröße auf die subjektive Wahrnehmung der Umhüllung den Grad seitlich eintreffender Schallenergie, die den Zuhörer nach den ersten 80 ms nach Eintreffen des Direktschalls erreicht. Als weitere Einflussgrößen wurden die Nachhallzeit und das Klarheitsmaß (C_{80}) identifiziert.

Ein weiterer Aspekt der Wahrnehmung von Räumen ist die perzeptive Raumgröße. McGrath et al. [1996] fanden heraus, dass blinde Menschen und Menschen mit verbundenen Augen durch das Erzeugen von Sprachlauten in der Lage sind, auditiv zwischen verschieden großen Räumen zu unterscheiden, sowie die eigene Position im Raum zu beurteilen. Um relevante Schlüsselmerkmale für die Wahrnehmung der akustischen Raumgröße zu identifizieren, führte Hameed et al. [2004] ein Experiment mit modellbasiert simulierten, akustischen Räumen durch. Das Schallfeld wurde in einem reflektionsarmen Raum mit Hilfe eines 16-kanaligen Lautsprecheraufbaus realisiert. In einem Hörtest wurden die Parameter Nachhallzeit und das Verhältnis zwischen Direkt- und Diffusschall variiert. Die Nachhallzeit stellte sich als hoch signifkanter Prädiktor für die perzeptive Raumgröße heraus. Die Variation des Direkt- / Diffusschallverhältnis hingegen führt zu keinem signifikanten Effekt, was darauf hindeutet, dass es nicht als Schlüsselmerkmal für die perzeptive Bestimmung der Raumgröße verwendet wird. Sandvad hingegen stellte in einer Untersuchung fest, dass einige Versuchspersonen das Direkt-/Diffusschallverhältnis als Schlüsselmerkmal zur Bestimmung der Raumgröße verwenden, andere die Nachhallzeit [Sandvad 1999].

Cabrera et al. [2005] verwendeten eine Kombination aus modellbasierten und datenbasierten Raumimpulsantworten, wiedergegeben durch statische Binauralsynthese, um Einflussfaktoren auf die auditiv wahrgenommene Raumgröße zu bestimmen. Sie variierten das Volumen, die Entfernung zwischen Klangquelle und Empfänger und die Nachhallzeit. Die Ergebnisse deuten auf eine Beziehung zwischen perzeptiver und

physikalischer Raumgröße hin. Als stärkster Einflussfaktor auf die wahrgenommene Raumgröße stellt sich die Nachhallzeit heraus. Wurde das Volumen des Raumes konstant gehalten und nur die Entfernung zwischen Klangquelle und Empfänger variiert, zeigte das Verhältnis zwischen früh und spät einfallender Schallenergie einen engen Zusammenhang mit der perzeptiven Raumgröße. In einem Folgeversuch wurden auditive Raumgrößenschätzungen in realen Räumen untersucht, wobei Sprachstimuli über einen Lautsprecher in drei verschiedenen Räumen wiedergegeben wurden [Pop2005]. Es zeigte sich auch dort ein signifikanter Zusammenhang zwischen perzeptiver und physikalischer Raumgröße, sowie ein positiver Zusammenhang zwischen Entfernung und perzeptiver Raumgröße.

Ein wichtiges Schlüsselmerkmal für die auditive Schätzung von Entfernungen ist der Schalldruckpegel [Ashmead1990]. Bronkhorst et al. [1999] konnten nachweisen, dass auch das Verhältnis von Direkt- zu Diffusschall als Schlüsselmerkmal dient. Zahorik. konnte die Ergebnisse bestätigen und stellte zudem fest, dass bei Sprachstimuli der Schallpegel als Schlüsselmerkmal dominiert, bei komplexeren Stimuli hingegen das Verhältnis zwischen Direkt- und Diffusschall [Zahorik2002]. Zahorik vermutet eine Gewichtung der Schlüsselmerkmale in Abhängigkeit von der Vertrautheit der Klangquelle. Da Sprache in der Regel ein vertrautes Signal ist, bei dem übliche Schallpegel durch Erfahrungen bekannt sind, ist der Schallpegel in diesem Fall das robustere Schlüsselmerkmal.

Zusammenfassend kann man sagen, dass die Erforschung der auditiven Wahrnehmung von Räumen schon einen umfangreichen Kenntnisstand erreicht hat. Wie aber steht es um die audiovisuelle Wahrnehmung von Räumen?

2.4 Audiovisuelle Wahrnehmung von Räumen

Der Überblick in den Kapiteln 2.1und 2.2 zeigt, dass audiovisuelle Interaktionseffekte alle Stadien des Wahrnehmungsprozesses betreffen. Der Aspekt der audiovisuellen Lokalisation ist am besten erforscht. Kapitel 2.3 gibt einen Überblick über Studien, die sich zum Ziel gesetzt haben, Korrelationen zwischen objektiven raumakustischen Merkmalen und der subjektiven Wahrnehmung aufzudecken, meist mit dem Anspruch die akustische Güte eines Konzertsaales vorherzusagen und objektiv messen zu können. Die Zusammenhänge erweisen sich in den meisten Fällen als komplex, da oft mehrere Faktoren die subjektive Bewertung beeinflussen. Es liegt nahe, dass audiovisuelle Interaktionseffekte auch bei der Wahrnehmung von Räumen zu erwarten sind. Hinweise auf solche Effekte konnten schon durch einige Studien geliefert werden.

Larsson et al. [2001] untersuchen den Einfluss visueller und auditiver Information in einer virtuellen Umgebung. In einem ersten Experiment wurde die Leistung der Versuchspersonen bei der Durchführung einer Aufgabe (Auffinden von vier Bällen und merken der darin enthaltenen Phrasen) unter einer unimodalen (optischen) und einer bimodalen (optoakustischen) Bedingung getestet. Unter der bimodalen Bedingung war die Erinnerungsleistung der Probanden höher, als unter der unimodalen Bedingung. Zudem wurden die Präsenz, das Gefallen und der Grad der Konzentration abgefragt. Alle drei Merkmale zeigen unter der bimodalen Bedingung eine signifikant höhere Ausprägung. In einem zweiten Experiment wurde der Einfluss visueller und auditiver Information auf die Bewertung raumakustischer Merkmale untersucht. Verglichen wurden eine unimodale (akustische) Bedingung und drei optoakustische Bedingungen. Der akustische Stimulusanteil war in allen Stimuli derselbe. In der ersten optoakustischen Bedingung wurde der Raum durch eine Foto repräsentiert, in der zweiten durch ein computergeneriertes virtuelles Modell. In der dritten befanden sich die Versuchspersonen in den realen Räumen, wobei auch dort der Klang über Kopfhörer wiedergegeben wurde. Es zeigte sich, dass die Entfernung zur Quelle, die perzeptive Raumgröße und die scheinbare Quellbreite unter der Bedingung des realen Raumes signifikant von den anderen Bedingungen abwichen. Larsson et al. schließen daraus zum Einen, dass die optische Information die Wahrnehmung auditiver Raummerkmale beeinflusst. Zum Anderen beeinflusst auch die Güte der Simulation bzw. der Stimuli die Wahrnehmung. Obwohl Auralisations- und Visualisierungstechniken

forschungsökonomisch große Vorteile gegenüber Versuchen in realen Situationen haben, ist nicht eindeutig geklärt, inwieweit virtuelle Reproduktionen realer Räume die gleichen Ergebnisse erzielen, wie Versuche in realen Situationen [Larrson2002]. Larrson et al. schlagen daher vor, Maße für die Güte von Simulationen zu entwickeln.

Sandvad [1999] fertigte binaurale Aufnahmen verschiedener Räume an und bat Versuchspersonen diese akustischen Repräsentationen zu Fotos aus denselben Räumen zuzuordnen. Die meisten Personen waren in der Lage die Aufgabe zu lösen, der Prozentsatz korrekter Zuordnungen lag bei 70 %. Sanvad schloss daraus, dass der Nachhall Raummerkmale enthalten muss, die vom Hörer extrahiert werden können. Zudem waren die Personen in der Lage diese auditiven Informationen zu den visuellen Informationen in Beziehung zu setzen.

Valente und Braasch [2008] untersuchen Einflüsse des optischen Raumes auf die erwarteten akustischen Raumeigenschaften. Den Versuchspersonen wurden Bilder von fünf realen Räumen mit Hilfe einer zweidimensionalen Videoprojektion präsentiert. Als Stimulusinhalte wurden ein Harfen- und ein Trommel-Sample verwendet. Die Musiker wurden aufgezeichnet und mit Hilfe des Blue-Box-Verfahrens in die optischen Räume eingefügt. Die akustische Simulation des Raumes basierte nicht auf realen Daten, sondern wurde modellbasiert mit Hilfe eines ViMiC-Systems (Virtual Microphone Contol) realisiert. Die Wiedergabe erfolgte über ein konzentrisch angeordnetes Lautsprecher-Array. Außerdem wurde die Helligkeit der optischen Räume variiert. Während der Präsentation der optoakustischen Stimuli hatten die Probanden die Aufgabe die Nachhallzeit und das Verhältnis zwischen frühem und spätem Diffusschallpegel so einzustellen, dass ein mit der optischen Simulation korrespondierender akustischer Raumeindruck entstand. Es zeigte sich ein signifikanter Effekt der Raumgröße auf das erwartete Verhältnis zwischen frühem und spätem Diffusschallpegel. Je größer das Raumvolumen, desto geringer war das Verhältnis. Die Variation der Helligkeit und des musikalischen Inhaltes zeigten keinen signifikanten Effekt. Die erwartete Nachhallzeit wurde signifikant sowohl von der Raumgröße als auch vom musikalischen Inhalt beeinflusst. Je größer das Volumen des Raumes war, desto größer war die erwartete Nachhallzeit. Der Trommelstimulus ergab geringere erwartete Nachhallzeiten als der Harfenstimulus. Valente et al. führen dies zurück, auf Unterschiede in der Impulshaftigkeit der musikalischen Stimulusinhalte. Die kurzen Attack-Zeiten des Trommelstimulus führten zu einem geringeren Maskierungseffekt, als der legato gespielte Harfenstimulus, wodurch der Nachhall beim Trommelstimulus deutlicher

wahrzunehmen war und die eingestellten Nachhallzeiten geringer ausfielen. Zudem trat ein in Bezug auf die Nachhallzeit auch ein signifikanter Interaktionseffekt zwischen der Raumgröße und dem musikalischen Stimulusinhalten auf. Je größer das Raumvolumen war, desto größer war der Unterschied zwischen den erwarteten Nachhallzeiten der beiden musikalischen Inhalte. Die Variation der Helligkeit zeigte auch für die Nachhallzeit keinen signifikanten Effekt. Trotz der anfänglichen audiovisuellen Konfliktsituation, wurde die Anpassung der akustischen Parameter von den Versuchspersonen in signifikant ähnlicher Weise vorgenommen. Der Versuch lässt somit darauf schließen, dass Personen in der Lage sind aus visuellen Repräsentationen von Räumen und musikalischen Aufführungen Annahmen über den Klang der akustischen Umgebung zu treffen. In einem weiteren Versuch [Valente2009] wurden optoakustische Modelle von fünf verschieden großen Räumen erstellt. Ein Gesangsensemble wurde in drei verschieden breiten Aufstellungen gefilmt und per Blue-Box-Verfahren in die Räume projiziert. Die drei unabhängigen Variablen waren die Größe des Raumes, die Entfernung zum Ensemble und die Breite der Aufstellung des Ensembles. Die Versuchspersonen wurden gebeten die auditive Quellbreite jeweils so zu justieren, dass sie dem optischen Stimulus entspricht. Die Auswertung ergibt, dass die Fähigkeit die auditive Quellbreite korrekt zu justieren in einem umgekehrten Verhältnis zur Entfernung zur Quelle und zur Nachhallzeit steht.

Vorländer et al. [2003]sammelten Beurteilungen der akustischen Qualität eines computergenerierten, optoakustisch simulierten Theatersaales. Die Decke des Theatersaales war beweglich und konnte geschlossen, halboffen oder offen sein. Entsprechende akustische Impulsantworten wurden mit Hilfe der Software CESAR generiert. In einem Konfliktreizparadigma wurden den Versuchspersonen alle optoakustischen Kombinationen der drei optischen und der drei akustischen Reize präsentiert. Der akustische offene Raume bekam in Kombination mit dem optischen geschlossenen Raum schlechte Bewertungen. Zusammen mit dem optischen offenen Raum präsentiert, wurde der akustische Stimulus besser bewertet. Die beste Bewertung bekam die Kombination geschlossener optischer und geschlossener akustischer Raum. Auch unter der optischen, offenen Bedingung wurde der akustische Stimulus des geschlossenen Raumes präferiert. Vorländer et al. bestätigen durch ihre Studie die Existenz audiovisueller Interaktionseffekte bei der Wahrnehmung von Raumakustik.

Larsson et al. [2007]. untersuchen die audiovisuelle Wahrnehmung der Raumgröße in virtuellen Umgebungen. Sie überprüften die These, dass die tatsächliche Raumgrö-

ße bei optoakustischer Repräsentation besser geschätzt wird, als bei ausschließlich optischer oder ausschließlich akustischer Repräsentation. Als Indikator für die wahrgenommene Raumgröße wurden die Versuchspersonen gebeten, die Breite des Raumes zu schätzen. Die optischen und akustischen Stimuluskomponenten für drei Räume verschiedener Größe wurden modellbasiert generiert und mit Hilfe zweidimensionaler Videoprojektion und eines Ambisonic-Lautsprecherarrays wiedergegeben. Die Versuchspersonen waren in der Lage zwischen den verschiedenen Räumen zu unterscheiden. Zudem führten die verschiedenen Modalitäten zu unterschiedlichen Schätzungen der Breite. Unter der akustischen Bedingung wurde die Breite der Räume überschätzt, unter der optischen unterschätzt. Die Schätzungen unter der optoakustischen Bedingung lagen jeweils zwischen den beiden unimodalen Bedingungen. Für den kleinen und den mittleren Raum war die Schätzung unter der optoakustischen Bedingung am genauesten. Larrson et al. kommen zu dem Schluss, dass es unter der optoakustischen Bedingung zu einer Integration der visuellen und auditiven Perzepte kommt.

Côté et al. [2011] untersuchten den Einfluss audiovisueller Schlüsselmerkmale auf die Entfernungsschätzung in einer virtuellen Umgebung. Dabei variierten sie die Entfernung zur Klangquelle (virtueller Lautsprecher) und die Anwesenheit akustischer und optischer Stimulusanteile. Es zeigte sich, dass die Entfernung unter allen Bedingungen unterschätzt wurde. Dabei besteht unter der optischen Bedingung ein linearer Zusammenhang zwischen tatsächlicher und wahrgenommener Entfernung. Unter der akustischen Bedingung kann der Zusammenhang durch eine Potenzfunktion beschrieben werden. Die Schätzungen unter der optoakustischen Bedingung zeigten in dieser Studie keinen signifikanten Unterschied zur optischen Bedingung.

3 Zielsetzung

Der in den vorangegangenen Kapiteln erarbeitete Überblick macht deutlich, dass die Erforschung von Aspekten der audiovisuellen Raumwahrnehmung noch ein junges Forschungsgebiet ist. Die bisherige Forschung ist geprägt, durch einen hochgradig technischen Ansatz wie die Verwendung von Raummodellen, durch verschiedenste technische Auralisationsverfahren erzeugte Klangfelder und die Berücksichtigung weniger, hochgradig spezifischer unabhängiger und abhängiger Variablen sowie durch die Anwendung verschiedener empirischer Paradigmen und Methoden. In Bezug auf die Frage nach audiovisuellen Interaktionseffekten haben sich noch keine einheitlichen Forschungsstrategien entwickelt und die Ergebnisse lassen Konsistenz und Adaptierbarkeit vermissen. Aus diesen Gründen erscheint für die vorliegende Studie ein explorativer Forschungsansatz angebracht.

Viele Studien ignorieren die Tatsache, dass der Beitrag optischer und akustischer Stimuluskomponenten auf die Ausprägung perzeptiver Raummerkmale nur dann allgemein quantifiziert werden kann, wenn eine vollkommene Entsprechung der Stimuluskomponenten gewährleistet ist. Für modellbasiert simulierte, virtuelle Umgebungen kann der Grad der Entsprechung nicht beziffert werden. Hingegen ist er bei datenbasierten Raumsimulationen maximal. Aus diesem Grund wurde die Simulation der Räume datenbasiert vorgenommen, wobei optische und akustische Stimuluskomponenten jeweils aus realen Räumen stammen.

Die Neuerung der vorliegenden Studie besteht in der Anwendung eines Konfliktreizparadigmas, indem optische und akustische Stimuluskomponenten kategorial variiert werden. Zudem wird die optische und akustische Raumpräsenz variiert, wobei die Stimuli optisch, akustisch und optoakustisch präsentiert werden. Die Simulation der Räume wurde datenbasiert, mittels stereoskopischer Abbildungen und dynamischer Binauralsynthese realisiert, um eine größtmögliche externe und ökologische Validität zu gewährleisten..

Der empirische Versuch wird innerhalb einer Testsession als Messwiederholung (Repeated-Measures-Design) realisiert. Für ausgewählte uni- und multimodale Merkmale wurde der Einfluss akustischer und optischer Stimuluskomponenten durch eine anschließende Varianzanalyse quantifiziert.

Die gewonnenen Erkenntnisse dieser Studie tragen bei zum Verständnis multimodaler Prozesse der Musik- und Sprachwahrnehmung in Räumen und sind von Interesse sowohl für die Kognitionspsychologie als auch für die systematische Musikwissenschaft. Weiterhin sind die Ergebnisse relevant für die Verbesserung von Simulationstechniken wie Virtual-Reality-Systeme. Außerdem können bereits bestehende Erkenntnisse perzeptiver Raumakustik anhand der neuen Daten gegebenenfalls reinterpretiert und bewertet werden.

4 Material und Methode

4.1 Übersicht

In diesem Kapitel wird die Konzeption, Erstellung und Durchführung der empirischen Untersuchung erläutert. Es werden die Kriterien dargelegt, nach denen die zu simulierenden Räume ausgesucht wurden. Außerdem wird beschrieben, auf welche Weise optische, akustische und strukturelle Daten in diesen Räumen erfasst wurden und wie die für den Versuch verwendeten Stimuli generiert wurden. Ferner wird der verwendete Fragebogen und der Ablauf des Experiments sowie die zur Realisierung gebrauchte Hard- und Software dokumentiert.

4.2 Versuchsdesign

4.2.1 Variation der optischen und akustischen Raumpräsenz

Modalität		akustischer Raum				
		kein	Kt	Kh	Gt	Gh
optischer Raum	kein	-	a	a	a	a
	Kt	o	oa			
	Kh	o		oa		
	Gt	o			oa	
	Gh	o				oa

Tabelle 1:Variation der optischen und akustischen Raumpräsenz: optisch (o), optoakustisch (oa), optisch (o).

In diesem Experiment wird die optische und die akustische Raumpräsenz variiert. Dabei wird zwischen drei, sich bezüglich der angesprochenen Modalitätskanäle unterscheidenden Stimulusgruppen differenziert: optische, optoakustische und akustische Stimuli. Eine schematische Darstellung der sich bei vier Räumen ergebenden Stimuli ist Tabelle 1 zu entnehmen. Diagonal ergeben sich auf diese Weise vier optoakustische (multimodale) Stimuli, horizontal vier akustische und vertikal vier optische (jeweils unimodal). Diese Matrix gilt jeweils für die Darbietungsinhalte Sprache und Musik, woraus sich eine Gesamtanzahl von 24 Stimuli ergibt. Da als optische Klangquelle ein Lautsprecher anstelle der eigentlichen Klangquelle (Sprecher, Cellospieler) zum Einsatz kam, unterscheiden sich die optischen Stimuli nicht bezüglich ihres Darbietungsinhalts. Die Gesamtzahl der für die

Variation der optischen und akustischen Raumpräsenz benötigten Stimuli sinkt damit auf 20.

4.2.2 Konfliktreize - Variation der optischen und akustischen Raumeigenschaften

		akustischer Raum			
		Kt	Kh	Gt	Gh
optischer Raum	Kt	p	k	k	k
	Kh	k	p	k	k
	Gt	k	k	p	k
	Gh	k	k	k	p

Tabelle 2: Variation der optischen und akustischen Raumeigenschaften (k -Konfliktreiz, p – passender Reiz).

In diesem Experiment werden unabhängig voneinander die akustischen und optischen Raumeigenschaften variiert. Alle Stimuli sind in diesem Fall immer optoakustisch, jedoch entstehen sowohl Konfliktreize, als auch Stimuli die eine optoakustische Passung aufweisen. Die Matrix in Tabelle 2 zeigt alle durch Kombination der akustischen und der optischen Räume entstehenden Stimuli, wobei je ein Kästchen einen Stimulus repräsentiert. Bei den auf der Diagonalen liegenden Stimuli stammen akustischer und der optischer Stimulusanteil von demselben Raum. Diese werden als Stimuli mit optoakustischer Passung bezeichnet. Alle anderen Stimuli in diesem Experiment sind Kombinationen der akustischen und optischen Stimulusanteile ohne Passung und werden als Konfliktreize bezeichnet. Der Vorteil dieser Vorgehensweise ist, dass die Variation der optischen und akustischen Eigenschaften in absolut gleichem Maße stattfindet, da optischer und akustischer Raum anhand vor Ort akquirierter Daten simuliert sind und sich bei jeweils einem Stimulus vollkommen entsprechen. Wären akustischer und optischer Raum modellbasiert simuliert, könnte nie von einer faktischen, allenfalls von einer theoretischen Passung ausgegangen werden.

Aus der Matrix in Tabelle 2 ergeben sich 2 x 16 Stimuli für die Darbietungsinhalte Musik und Sprache. Die Stimuli mit optoakustischer Passung sind schon in A enthalten müssen demnach nicht ein weiteres Mal abgefragt werden. Für beide Versuche ergibt sich eine Gesamtanzahl von 44 Stimuli: 2 x 12 Konfliktreize für die Variation der optischen und akustischen Raumeigenschaften, 12 unimodale für die Variation der optischen und akustischen Raumpräsenz und 2 x 4 mit optoakustischer Passung, die für beide Auswertungen verwendet werden. Durch das Fehlen der Inhaltsvariation in der optischen Raumrepräsentation entstehen nur 12, anstatt 16 unimodale Stimuli.

Sowohl die Variation der optischen und akustischen Raumpräsenz, als auch die Variation der optischen und akustischen Raumeigenschaften wurden in einem Versuch vereint, indem alle 44 Stimuli nacheinander in einem Messwiederholungsdesign abgefragt wurden. Aus einer durchschnittlichen Bearbeitungszeit von 2:15 Minuten, ergab sich, inklusive zweier Testduchläufe, eine durchschnittliche Versuchsdauer von 105 Minuten. Die Probanden konnten eine Pause (5 – 10 Minuten) zu einem selbstgewählten Zeitpunkt einlegen, was alle Probanden in Anspruch nahmen.

4.3 Stimuli

4.3.1 Auswahlkriterien und Beschreibung der simulierten Räume

Um den Umfang der empirischen Untersuchung in einem durchführbaren Rahmen zu halten, wurde für den Hörversuch vier verschiedene Räume simuliert. Die Kriterien für die Auswahl der Räume waren ihre Primär- und Sekundärstruktur. Die Primärstruktur umfasst die geometrischen Parameter: Größe und Form. Die Sekundärstruktur umfasst die Parameter: Oberflächenstruktur, -material und Bepolsterung. Diese Parameter eignen sich besonders für die Selektion, da sie maßgeblich die Nachhallzeit, einen der wichtigsten physikalischen Prädiktoren für die Schätzung der Raumgröße [Cabrera 2005], beeinflussen.

Die Größe eines Raumes lässt sich numerisch auf einer eindimensionalen metrischen Skala durch die Variable Volumen beschreiben. Je größer das Volumen, desto länger die Nachhallzeit. Die Sekundärstruktur lässt sich indirekt durch ihren akustischen Effekt, die Schallabsorption beschreiben. Das dazu genutzte Maß ist der mittleren Absorptionskoeffizienten ($\alpha_{\varnothing}$). Dieser gibt an, wie stark die Oberflächen eines Raumes den Schall absorbieren bzw. reflektieren. Je kleiner der Absorptionskoeffizient, desto länger die Nachhallzeit. Für beide Variablen wurden jeweils zwei Ausprägungen gebildet; großes Volumen vs. kleines Volumen und niedriger mittlerer Absorptionskoeffizient vs. hoher mittlerer Absorptionskoeffizient. Die Kombination dieser Ausprägungen ergibt vier Räume, mit vier voneinander verschiedenen Nachhallzeiten. (Tabelle 3). Gesucht wurden demnach zwei kleine Räume ähnlicher Größe mit jeweils kurzem und längerem Hall und zwei große Räume mit den gleichen Eigenschaften.

Bei den kleinen Räumen fiel die Wahl zum Einen auf einen Seminarraum der TU Berlin (EN 190), welcher eine verhältnismäßig lange Nachhallzeit aufweist. Zum An-

deren auf das Elektronische Studio der TU Berlin (EN 324), das eine verhältnismäßig geringere Nachhallzeit aufweist, da dort durch raumakustische Maßnahmen die Nachhallzeit reduziert wurde. Die Räume mit großem Volumen waren der Wellenfeldsynthesesaal H 104 der TU Berlin, welcher für seine Größe eine geringe Nachhallzeit aufweist und der UdK Konzertsaal Bundesallee mit einer langen Nachhallzeit.

Name Volumen $\alpha_{ø}$ Nachhallzeit		Volumen [m³]	
		klein	groß
Absorpionskoeffizient $\alpha_{ø}$	hoch	EN 324 230 m³ 0,36 (0,36 s)	H 104 3300 m³ 0,28 (1,07s)
	gering	EN 190 190 m³ 0,17 (0.71 s)	Konzertsaal (UdK) 4200 m³ 0,17 (1,79 s)

Tabelle 3: Variation: Volumen, Absorptionskoeffizienten.

Damit die Räume unmissverständlich und knapp benannt werden können, werden für den weiteren Verlauf der Arbeit die in Tabelle 4 dargestellten Bezeichnungen eingeführt. Die Buchstaben „K“ und „G“ codieren die Größe des Volumens (klein – groß), die Buchstaben „t“ und „h“ codieren die Halligkeit (trocken – hallig).

Raum Kt	EN 324	elektronisches Studio	klein	trocken
Raum Kh	EN 190	Seminarraum	klein	hallig
Raum Gt	H 104	Hörsaal / Wellenfeldsynthese	groß	trocken
Raum Gh	Konzertsaal der UdK	musikalische Aufführungen	groß	hallig

Tabelle 4: Bezeichnung der Räume.

Da auch die Form eines Raumes eine Variable in Bezug auf die Raumwahrnehmung darstellt, sollte diese ebenfalls über alle Räume systematisch variiert werden oder aber unverändert bleiben. In diesem Versuch wird die Form nicht systematisch variiert. Raum Ht weicht bezüglich der Form von den anderen Räumen. Während Raum Kt, Kh und Gh einen rechteckigen Grundriss mit ebenem Boden aufweisen, weist Raum Gt einen diamantförmigen Grundriss und eine zum hinteren Ende ansteigende Bestuhlung auf. Die Größenverhältnisse und Formen der Räume können Abbildung 4.1 entnommen werden.

Die strukturellen Abmessungen wurden mit einem Laserdistometer erfasst. Der mittlere Absorptionskoeffizient konnte anhand dieser Daten und der jeweiligen Nachhallzeit durch Umstellung der Sabinschen Nachhallgleichung nach $\alpha_{ø}$ approximiert werden:

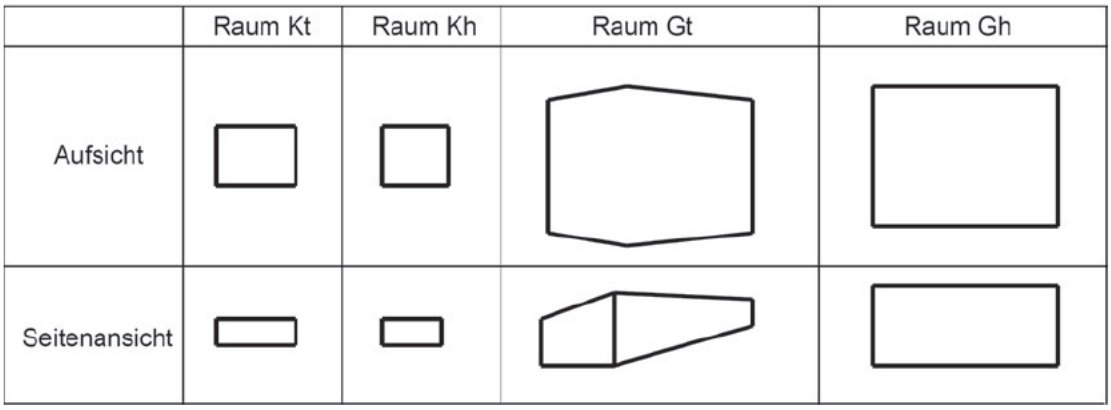


Abbildung 4.1: Grundrisse der Räume.

T: Nachhallzeit
V: Volumen
A: Oberfläche
$\alpha_{\varnothing}$: durchschnittlicher Absorptionskoeffizient

Eine Übersicht der optischen, akustischen und strukturellen Kennwerte der einzelnen Räume findet sich im Anhang 8.3 (S. 85).

4.3.2 Darbietungsinhalte der Stimuli

Die Darbietungsinhalte der Stimuli waren eine nachhallfreie Aufnahme einer englischen Sprecherin und eine ebenfalls nachhallfreie Aufnahme einer Solocellopassage. Der gesprochene Text lautete:

„In language infinitely many words can be written with a small set of letters. In arithmatic infinitely many numbers can be composed from just a few digits, with the help of the symbol zero, the principle of position and the concept of base.“

Der Musikausschnitt entstammte der „Variation und Thema Nr.2“ von Carl Maria von Weber. Sowohl der Musik-, als auch der Sprachausschnitt hatten eine Länge von 16 Sekunden. Die Aufnahmen wurden der CD „Bang & Olufsen - Music for Archimedes,“ von 1989 entnommen. Aufgenommen wurden sie im refektionsarmen Raum der Technischen Universität Dänemark mit einem omnidirektionalen Mikrofon (B&K 4003), das aus einer Entfernung von 1 m auf die Quelle gerichtet war. Die Aufnahmen lagen im WAV-Format (16 Bit ; 44,1 kHz ; monophon) vor und wurden in Echtzeit mit den jeweiligen Impulsantworten gefaltet.

Es ist zu beachten, dass die Hörbeispiele eine unterschiedliche Lautstärke aufweisen. Die Sprecherin ist bei einer Entfernung von einem Meter leiser als das Cello. Da

dieser Lautstärkeunterschied jedoch natürlich ist, wurde aus Gründen der ökologischen Validität von einer Manipulation abgesehen.

4.4 Erhebungsinstrument

In Anlehnung an [Rumsey 2002], [Berg 2006], [Berg 2009] und [Schubert 2001] wurden 20 perzeptive Maße abgefragt. Alle Maße sind eindimensional und beziehen sich auf verschiedene Modalitäten. In Tabelle 5 sind alle abgefragten perzeptiven Maße dargestellt, geordnet nach Modalität und Referenzobjekt. Der Begriff Referenzobjekt bezieht sich auf die Klangquelle, den Raum und den Gesamteindruck [Berg2001]. Die Frage nach der Lautheit beispielsweise wurde für alle drei Referenzobjekte gesondert gestellt.

		Modalität		
		visuell	auditiv	audio-visuell
Referenzobjekt	Klangquelle	Helligkeit	Lautheit Wärme Brillianz Zeittransparenz	Entfernung
	Raum	Helligkeit Farbintensität Kontrast Mittlerer Farbton	Lautheit Wärme Brillianz Halligkeit Umhüllung	Breite Höhe Länge Volumen
	Gesamteindruck		Lautheit	audio-visuelle Passung Anwesenheit Aktivierung Potenz Gefallen

Tabelle 5: Perzeptive Maße.

Operationalisiert wurden die Maße durch jeweils nur eine Indikatorvariable, um den Fragebogen in einem vertretbaren Umfang zu halten. Die Entfernungs- und Ausdehnungsangaben wurden in Metern und mit Hilfe eines Schiebereglers erfasst. Der Proband hatte jeweils die Wahl, eine Strecke zwischen einem und 30 Metern in Einmeter-Schritten anzugeben. Die maximale Distanz von 30 Metern für die Schieberegler wurde festgelegt, da die größte gemessene Strecke 26,4 Meter (Raum GT) betrug und die Beschriftung der Schieberegler einheitlich gestaltet werden sollte. Die Erfassung des mittleren Farbtons wurde kategorial mit Hilfe einer im Fragebogen integrierten Farbskala realisiert, die alle Primär-, Sekundär- und Tertiärfarben umfasste. Alle anderen Daten wurden über eine siebenstufige Rating-Skala, so dass dich ein spezifisches semantisches Differential ergibt. Im Anhang 8.2 (S. 84) gibt eine Übersichttabelle Auskunft über die Operationalisierung und Kodierung der einzelnen Merkmale.

Eine Besonderheit ergab sich bei der Operationalisierung des Merkmals Zeittransparenz. Die Operationalisierung musste mit unterschiedlicher Formulierung für die verschiedenen Darbietungsinhalte erfolgen. Bei Erklingen des Sprachsamples wurde das

Merkmal durch das begriffliche Gegensatzpaar *Sprache gut verständlich – Sprache nicht gut verständlich* abgefragt, bei Erklingen des Musiksamples hingegen wurde es durch *Musik gut durchhörbar – nicht gut durchhörbar* operationalisiert. Der Begriff *durchhörbar* wurde in der Einführung zum Versuch gesondert erläutert, als: „Wie deutlich aufeinander folgende Töne voneinander getrennt wahrgenommen werden können.“ Die Gestaltung des Fragebogens findet sich im Anhang 8.1 (S.82).

4.5 Die Stichprobe

Die Stichprobe umfasste 35 Versuchspersonen und bestand zum überwiegenden Teil aus Studenten des Studiengangs Audiokommunikation der TU Berlin. Bei den meisten Teilnehmern kann daher ein akustisches Grundwissen und ein Verständnis der im Fragebogen verwendeten Begrifflichkeiten vorausgesetzt werden. Das Alter der Versuchspersonen lag zwischen 23 und 55 Jahren, es betrug durchschnittlich 28 Jahre. Von den 35 Versuchspersonen waren 8 weiblich und 27 männlich. Die Frage, ob eine Einschränkung der Hör- oder Sehfähigkeit vorliege, wurde von allen Teilnehmern verneint. Acht Probanden waren keine deutschen Muttersprachler. Je nach Sprachkenntnis war eine genauere Einweisung in die im Fragebogen verwendeten Begriffe notwendig. Die Erhebung erstreckte sich über einen Zeitraum von 6 Wochen.

4.6 Statistische Analyseverfahren

Unter Annahme von Intervallskalierung sind alle erhobenen perzeptiven Maße metrisch. Dies erlaubt die Berechnen arithmetischer Mittelwerte und die Anwendung parametrischer statistischer Verfahren zur Auswertung.

Durch eine dreifaktorielle Varianzanalyse kann für die Variation der optischen und akustischen Raumpräsenz bestimmt werden, ob diese einen signifikanten Einfluss auf die jeweiligen perzeptiven Merkmale ausübt und welcher Art dieser Einfluss ist.
Durch eine dreifaktorielle Varianzanalyse mit den Faktoren *akustischer Raum*, *optischer Raum* und *Darbietungsinhalt* (Cello/Sprecher) lässt sich bei der Variation der optischen und akustischen Raumeigenschaften zum Einen bestimmen, ob die Variation der Eigenschaften überhaupt einen Effekt auf die durch den Fragebogen gemessenen perzeptiven Merkmale ausübt. Zum Anderen lässt sich bestimmen, ob ein perzeptives Merkmal stärker von der akustischen oder von der optischen Modalität beeinflusst wird.

4.7 Akustische Simulation

Der akustischen Räume wurden durch dynamische Binauralsynthese simuliert. Bei der Binauralsynthese wird die mit Hilfe eine Kunstkopfs binaural aufgenommene Impulsantwort eines Raumes, auch BRIR (binaural room impulse response) genannt, mit einem beliebigen Quellsignal gefaltet. Das auf diese Weise synthetisierte binaurale Signal wird über Kopfhörer wiedergegeben. Dieses Verfahren ermöglicht eine sehr plausible Simulation der akustischen Umgebung. Die dynamische Binauralsynthese erreicht einen noch höheren Grad an Realismus, in dem die Blickrichtung des Kopfes mit Hilfe eines Trackers in Echtzeit berücksichtigt wird. Das räumliche Klangbild passt sich der jeweiligen Blickrichtung dynamisch an. Hierfür ist eine einzelne BRIR nicht mehr ausreichend, sondern es wird für jede Kopfposition eine eigene BRIR benötigt.

Im Versuch wurde eine dynamischen Binauralsynthese mit einem Freiheitsgrad von 160° in der horizontalen Ebene und mit einer Winkelauflösung von 1° verwendet. Das bedeutet, die Versuchsperson konnte ihren Kopf 80° zur linken und ebenfalls 80° zur rechten Seite wenden. Da im Hörversuch nur eine Senderposition simuliert wird, wurde aus jedem der vier Räume genau ein binauraler Datensatz für die entsprechende Sender-Empfängerstrecke benötigt. Ein Datensatz bestand aus 161 Impulsantworten (160° + Mittelposition).

Für die Räume Kt, Kh und Gt, konnten die Binauraldatensätze einer bestehenden Datenbank entnommen werden, die im Rahmen der Studie [Kosanke2010] gesammelt wurden. Für Raum Gh existierte kein Datensatz, weshalb dieser erst akquiriert werden musste. Dies geschah mit dem Messroboter FABIAN [Lindau 2006], mit welchem auch die bereits bestehenden Datensätze erzeugt wurden (Abbildung 4.2). Die Ohren des Roboters befanden sich auf einer Höhe von 1,2 m. Als Klangquelle diente ein frequenzgangentzerrter 3-kanaliger Dodekaederlautsprecher mit kugelförmigem Abstrahlverhalten. Dieser wurde in der Mitte der Bühne platziert. Dies gilt ebenso für die bestehenden Datensätze. In den Räumen ohne Bühne würde der Lautsprecher vorn mittig platziert.

Abbildung 4.2: Messroboter FABIAN.

Ebenfalls musste zur Anpassung der Lautstärken der simulierten Räume untereinander das Stärkemaß (G) (siehe Kapitel 4.7.1) eines jeden Raumes ermittelt werden. Alle weiteren raumakustischen Kennwerte konnten aus diesen Messsignalen mit Hilfe der Software *Monkey Forrest* berechnet werden. Die Entfernung des Messroboters zur Klangquelle betrug in jedem Raum in etwa den doppelten Hallradius. Die strukturellen und raumakustischen und optischen Kennwerte sowie die Sender-Empfänger-Distanz der jeweiligen Räume sind im Anhang 8.3 (S. 85) zusammengefasst.

Der beim Hörversuch zur Wiedergabe verwendete Kopfhörer war ein frequenzgangentzerrter Staxx SR 202 mit angebrachtem Tracker zur Bestimmung der Kopfausrichtung.

4.7.1 Relative Lautheit

Die Lautheit ist ein wichtiges Schlüsselmerkmal bei der subjektiven Beurteilung von Raummerkmalen, beispielsweise für die Schätzung von Entfernungen [Zahorik 2002]. Aus diesem Grund wurde für die korrekte Reproduktion der akustischen Umgebung die relative Lautheit der vier Räume zueinander wiederhergestellt.

Durch Schallreflektionen an den Oberflächen eines Raumes mischt sich zum Direktschall einer Klangquelle auch ein Anteil Diffusschall. In Abhängigkeit von Größe und Absorptionseigenschaften des Raumes trägt der reflektierte Schall somit zu einer Anhebung des Gesamtpegels bei. Ein Maß für den energetischen Anteil des Raumes ist das Stärkemaß (G). Im übertragenen Sinne kann man sich einen Raum als ein den Direktschall verstärkendes System begreifen, wobei das Stärkemaß als Wirkungsgrad des Raumes betrachtet werden kann, da es zugeführte und abgegebene Leistung des Systems in ein Verhältnis setzt. Im Gegensatz zum Wirkungsgrad jedoch, welcher in der Regel eine negative Energiebilanz aufweist, zeigt das Stärkemaß eine Verstärkung des Ursprungssignals an. Das Stärkemaß berechnet sich nach der Formel:

$$G = 10\,log\,\frac{\int_0^{\infty} p^2(x,t)dt}{\int_0^{7ms} \gamma \cdot p^2(s,t)dt} - 10\,log\left(4\pi\frac{s^2}{m^2}\right)dB$$

s: Bezugsentfernung (10 m)

x: Entfernung des Messplatzes von der Schallquelle in m

γ: Bündelungsgrad der Quelle

Der am Hörerplatz gemessene Schallpegel eines Kugelstrahlers wird in Bezug gesetzt zum Schallpegel desselben Strahlers im Freifeld in einer Entfernung von 10 Metern. Die daraus resultierende Angabe in Dezibel (dB) beschreibt die Verstärkung, die das Signal durch den Raum erfährt. Für die korrekte Abbildung der Lautstärkerelationen zwischen den Räumen wurde für jedem Raum das Stärkemaß ermittelt.

Anschließend wurde der Pegel der binauralen Impulsantworten der Räume anhand der 0°-Position normalisiert und um den Wert des entsprechenden Stärkemaßes verstärkt. Somit ist eine den realen Bedingungen entsprechende relative Lautheit der Räume untereinander sichergestellt.

4.8 Optische Simulation

Im Gegensatz zur akustischen war die optische Simulation statisch. Sie bestand aus einem stereoskopischen Foto, welches den betreffenden Raum aus der Empfängerposition zeigte, die zuvor der Messroboter eingenommen hatte. Als optische Klangquelle diente der für die akustischen Messungen verwendete Dodekaeder-Lautsprecher. Wie zuvor der Messroboter, befand sich auch das Objektiv der Kamera in einer Höhe von 1,2 Metern. Die Bilder wurden aufgenommen mit einer Pentax K-x Spiegelreflexkamera. Um eine hohe Schärfentiefe zu erzielen wurde eine kleine Blende gewählt (f / 22). Die Kamera wurde auf eine Stereoschiene montiert, die es erlaubte, sukzessive das rechte und das linke Bild zu schießen. Der Abstand zwischen den beiden Objektivpositionen (Augenabstand) betrug 7 cm.

Um bei der Darstellung der Bilder in der Versuchsumgebung eine realistische Perspektive zu erzeugen müssen ein paar Überlegungen bezüglich der Brennweite ange-

stellt werden. Objektive mit Weitwinkeleffekt (Brennweiten < 35 mm) bspw. erreichen einen Sichtwinkel über 115°. Dies geschieht auf Kosten einer leichten perspektivischen Verzerrung, da das Sichtfeld vergrößert, aber auf gleicher Fläche abgebildet wird. Die Verzerrung bewirkt im Vergleich zur Realität ein größeres Sichtfeld. Teleobjektive dagegen (Brennweite > 85 mm) haben durch den geringen Sichtwinkel einen gegensätzlichen Effekt.

Die Brennweite am Objektiv betrug für alle Räume konstant 18 mm, um auch beim geringsten Abstand zur Klangquelle diese noch vollständig im Bildausschnitt zu erfassen. Da die Kamera einen im Verhältnis zum Kleinbildformat (35 mm) kleineren Bildsensor hat muss die Brennweite des Objektivs mit einem Faktor von 1,5 (Crop-Faktor) multipliziert werden um die Differenz zum Kleinbildformat auszugleichen. Je nach Größe des in der Kamera verbauten Bildsensors ergibt sich ein anderer Faktor. Die reale Brennweite der Kamera betrug demnach 27 mm. Der horizontale Öffnungs- bzw. Sichtwinkel bei dieser Brennweite beträgt 65° (Abbildung 4.3.a). Aber auch beim Betrachten der Fotos auf einem Bildschirm ergibt sich ein Sichtwinkel, der abhängig ist von der Breite des Bildes und der Entfernung des Betrachters vom Bild (Abbildung 4.3.b). Eine optimale, verzerrungsarme Darstellung kann erreicht werden, wenn der Sichtwinkel des Objektivs und der Sichtwinkel beim Betrachten identisch sind. Es entsteht der Eindruck, der Bildschirm sei ein Fenster im Raum, durch das ein dahinter liegender Raum betrachtet wird.

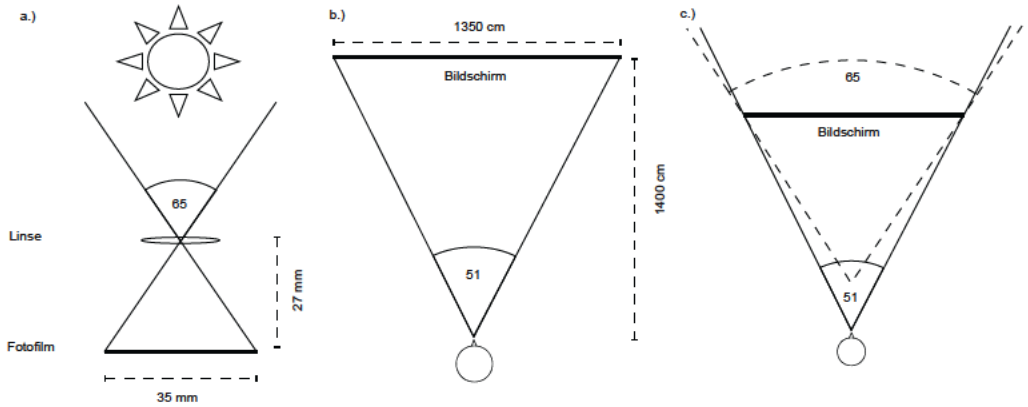


Abbildung 4.3: Sichtwinkel: a.) Sichtwinkel eines Objektivs (Brennweite 27 mm) b.) Sichtwinkel beim Betrachten des Bildschirms c.) perspektivische Verzerrung des Versuchsaufbaus.

Im vorliegenden Versuchsaufbau betrug die Bildbreite 1350 cm. Ein optimaler, den Fotos entsprechender Sichtwinkel von 65° wäre bei einem Betrachtungsabstand von 1050 cm entstanden. Jedoch war das Betrachten des Bildschirms aus dieser Nähe zu

anstrengend, weshalb sie auf 1400 cm erhöht werden musste. Die dadurch entstehende Verzerrung der Perspektive ist in Abbildung 4.3.c dargestellt.

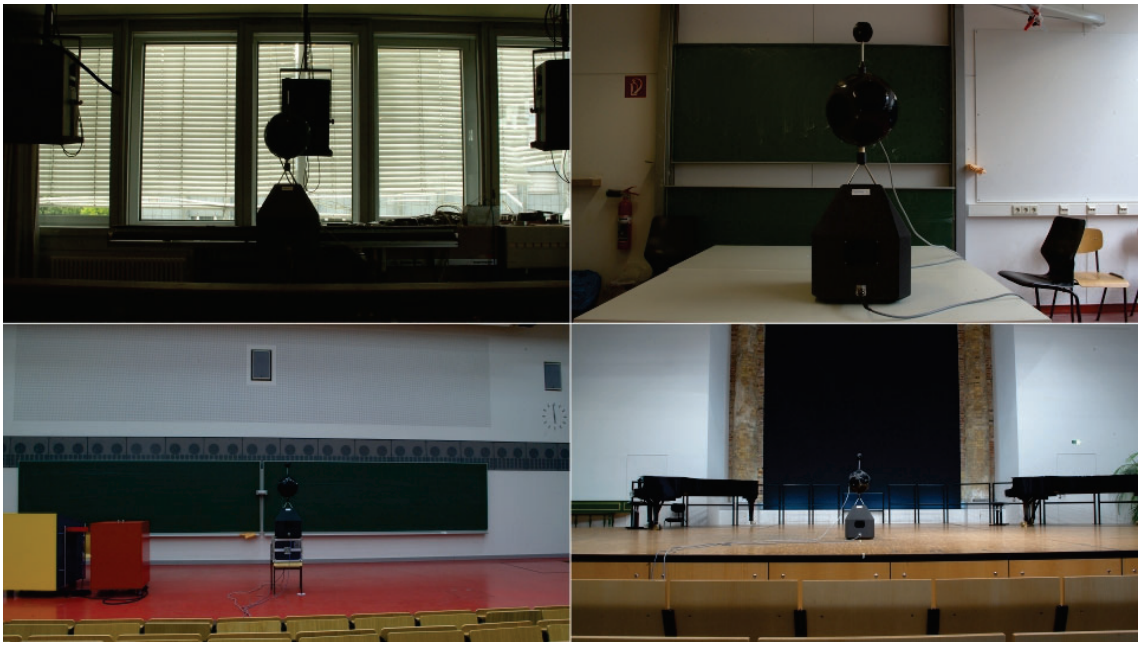


Abbildung 4.4: jeweils linkes Bild der im Versuch verwendeten stereoskopischen Fotos. Oben links: Raum Kt Oben rechts: Raum Kh. Unten links: Raum Gt, Unten rechts: Raum Gh.

Die Nachbearbeitung bestand im Zuschneiden und Konvertieren der Bilder in ein 16:9 Format, bei einer Auflösung von 1920x1080 Pixeln. Die Fotos wurden dem Probanden über einen 3D-fähigen DLP-Ferseher in FullHD-Auflösung (1920 x 1080 Pixel) präsentiert.

Um den 3D-Effekt nutzen zu können war das Tragen einer Shutterbrille notwendig. Es stellte sich heraus, dass der Fernseher einen relativ starken Ghosting-Effekt erzeugte. Ein Ghosting-Effekt tritt auf, wenn Teile des Bildes die nur für das linke Auge sichtbar sein sollten auch mit dem rechten zu sehen sind und umgekehrt. Da dieser Effekt am stärksten an dem abgebildeten Lautsprecher zu sehen war, der einen hohen Kontrast zu seinem Umfeld hatte, wurde die Parallaxe des Bildes soweit verschoben, bis der Lautsprecher im linken Bild sich mit dem im Rechten deckte. Der dreidimensionale Eindruck des Bildes wurde dadurch nicht nennenswert beeinträchtigt. Abbildung 4.4 zeigt die für die Stimuli verwendeten Fotografien aus der jeweiligen Empfängerposition.

4.9 Technischer Versuchsaufbau

Die Anforderungen an ein Programm zur Realisierung des Fragebogens umfassten die folgenden Punkte:

- eine grafische Eingabemaske zur Darstellung auf einem Bildschirm
- automatisches Abspeichern und Verwalten der eingegebenen Daten
- Exportmöglichkeit der Daten (EXCEL, SPSS)
- das Einbinden von Grafiken in den Fragebogen
- die Verwendung von Schiebereglern für die Meterangaben
- automatisches Randomisieren der Stimuli
- senden von OSC-Signalen beim Laden eines Fragebogens
- einfache Bedienung und Handhabung

4.9.1 Datenerhebung - Limesurvey

Die genannten Anforderungen wurden von der Open Source Software LimeSurvey (Version 1.87+ Build 8518) erfüllt. Die Software bietet die Möglichkeit, Fragebögen mit individuellem Design zu erstellen. Die Umfragen können anschließend mit Hilfe eines Internetbrowsers geöffnet und beantwortet werden. Die Antworten werden automatisch in einer Datenbank gespeichert und können nach Abschluss der Umfrage in verschiedenen Formaten (Excel, SPSS u.a.) exportiert werden. Limesurvey ermöglicht es, die Daten sicher auf einem Server abzulegen.

Das Senden von OSC-Signalen wird nicht nativ von Limesurvey unterstützt war aber von zentraler Bedeutung für die Steuerung des gesamten Versuchs. Durch das Einbinden eines von Dr. Steffen Lepa programmierten Java-Applets wurde dies jedoch ermöglicht. Somit wurde nach dem Laden eines Fragebogens ein stimulusspezifisches OSC-Signal gesendet, welcher die Wiedergabe des entsprechenden Stimulus auslöste.

Durch die flexiblen Konfigurationsmöglichkeiten des OSC-Protokolls, die Echtzeitfähigkeit und die hohe Kompatibilität zwischen verschiedenen Programmen und

Betriebssystemen war es für die Kommunikation zwischen der Erhebungssoftware, dem Audio- und dem Videoserver sehr gut geeignet.

4.9.2 Stimulus-Wiedergabe

Die für den Versuch verwendeten optoakustischen Stimuli setzen sich jeweils zusammen aus einer optischen und einer akustischen Simulation, sowie dem jeweiligen Darbietungsinhalt, Sprache oder Musik. Daher wurden die Stimuli nicht im vor hinein gerendert, sondern während des Versuchs in Echtzeit erzeugt.

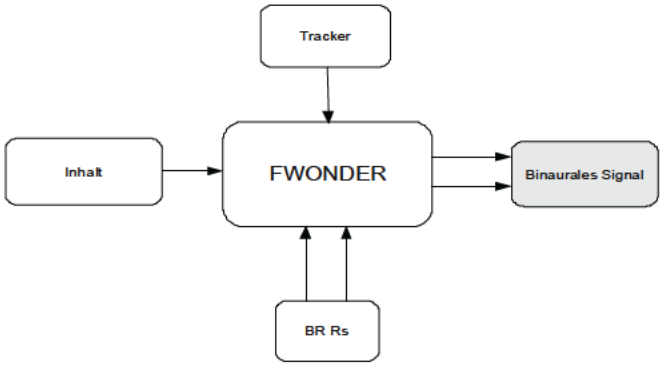


Abbildung 4.5: FWONDER Instanz

Die akustische Simulation wurde mit Hilfe eines linux-basierten Audioservers realisiert. Dieser war mit dem Head-Tracker verbunden, der die aktuelle Kopfausrichtung der Versuchsperson übermittelte. Für die Echtzeitfaltung des Darbietungsinhaltes mit den entsprechenden Impulsantworten wurde die an der TU Berlin entwickelte Software FWONDER verwendet. Diese erzeugt bei Eingabe eines einkanaligen Eingangssignals und unter Berücksichtigung der Tracker-Daten sowie des BRIR-Datensatzes ein binaurales Ausgabesignal (Abbildung 4.5). Für alle vier Räume liefen vier dieser FWONDER-Instanzen parallel um unerwünschte Nebengeräusche beim Ein- und Abschalten zu vermeiden.

4.9.3 Technische Realisation

Das Routing und die Kommunikation via lokalem Netzwerk wurde mit PureData realisiert. Der Audioserver empfing ein OSC-Signal mit der Information welche Binauralsynthese und welcher Darbietungsinhalt abgespielt werden sollten. PureData schickte das Audiosignal des entsprechenden Samples an alle FWONDER-Instanzen, hob aber nur die Lautstärke der gewünschten Instanz an. Auf diese Weise werden technischen Störgeräusche vermieden. Nach Abspielen des Samples wurde die Lautstärke wieder abgesenkt und ein OSC-Signal bezüglich der erfolgreichen Darbietung zurückgesendet.

Ein zweiter, windowsbasierter Videoserver, übernahm zugleich die Darbietung der Fragebögen mit Hilfe des Internetbrowsers *Firefox* der Firma Mozilla und die Darbietung des jeweiligen stereoskopischen Bildes mit dem *Stereoscopic Player* der Firma 3dtv.at. Die Ausgabe erfolgte über einen 61-Zoll großen, Full HD- und 3D-fähigen Fernseher mit Shutterbrillen-Technologie. Die optische Simulation und das Anzeigen und Ausfüllen des jeweiligen Fragebogens konnte so über denselben Bildschirm erfolgen.

Firefox und der Stereoscopic Player wurden ins Vollbild versetzt und mit Hilfe der Skriptsprache AutoIT konnte per OSC-Signal das jeweils benötigte Programmfenster in den Vordergrund geholt werden. Durch das Laden des jeweils folgenden Fragebogens im Browser wurde das dort implementierte OSC-Applet aktiviert, welches ein stimulusspezifisches OSC-Signal sendete. Dieses Signal enthielt die Information, aus welchem Bild, welcher Binauralsynthese und welchem Darbietungsinhalt der Stimulus bestand. Die Information zu Binauralsynthese und Darbietungsinhalt wurden von PureData empfangen und direkt an die PureData-Instanz des Audioservers weitergeleitet. In Bezug auf die Bildinformation löste PureData ein Skript aus, welches den Player das entsprechende Foto anzeigen ließ. Ein weiteres Skript holte den Player zeitgleich in den Vordergrund. Nach Abschluss des Wiedergabevorgangs sandte der Audioserver ein entsprechendes OSC-Signal, worauf wieder der Browser in den Vordergrund geholt wurde und der Proband mit dem Ausfüllen des Fragebogens beginnen konnte. Zum Ausfüllen des Fragebogens war nur die Bedienung einer Computermaus notwendig. Eine schematische Darstellung des Versuchsaufbaus zeigt Abbildung 4.6.

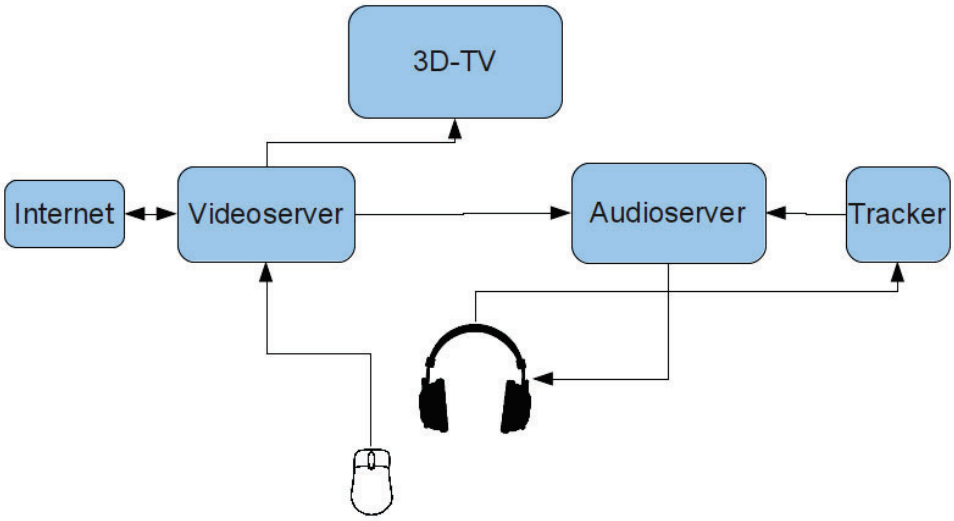


Abbildung 4.6: Schematischer Versuchsaufbau

4.9.4 Versuchsraum des Fachgebiets Audiokommunikation

Als Versuchsraum diente das MediaLab der TU Berlin. Die Wände des Raumes waren mit schweren schwarzen Vorhängen verhangen um den Raum abzudunkeln und die Halligkeit zu minimieren. Die Fenster waren mit schwarzer Folie verklebt, um Lichteinfall von außen zu vermeiden. Leider war der Raum nicht immer vollkommen still. Gelegentlich konnten leiser Straßenverkehr oder Geräusche aus dem Treppenhaus vernommen werden. Eine Auswirkung auf die Antworten der Probanden wird jedoch nicht angenommen, da durch das Abhören über Kopfhörer eventuelle Geräusche nur beim Ausfüllen des Fragebogens und nicht beim Erklingen der Stimuli hörbar waren. Zudem waren die Geräusche nur sporadisch und zufällig über alle Probanden verteilt.

Es war für die Probanden jederzeit möglich den Versuchsleiter bei Fragen durch das Drücken eines Knopfes zu rufen.

5 Auswertung

5.1 Variation der optischen und akustischen Raumpräsenz

Für die Variation der optischen und akustischen Raumpräsenz kommen für die Auswertung nur die multimodalen perzeptiven Merkmale in Frage, welche sich auf die optische und die akustische Modalität beziehen. Dazu gehören die in Tabelle 6 hervorgehobenen Merkmale. Um die Ergebnisse auszuwerten, wurde eine dreifaktorielle Varianzanalyse mit den Faktoren Raum (Faktorstufen: Kt, Kh, Gt, Gh), Darbietungsinhalt (Faktorstufen: Sprache, Musik) und Präsentationsmodus (Faktorstufen: optisch, akustisch, optoakustisch) durchgeführt. Als optoakustische Stimuli wurden jene mit optoakustischer Passung ausgewählt. Die optischen Stimuli bestanden aus dem stereoskopischen Bild des entsprechenden Raumes, die akustischen Stimuli aus der entsprechenden Binauralsynthese. Während des akustischen Präsentationsmodus blieb der Monitor grau und zeigte kein Foto.

		Modalität		
		visuell	auditiv	audio-visuell
Referenzobjekt	Klangquelle	Helligkeit	Lautheit Wärme Brillianz Zeittransparenz	Entfernung
	Raum	Helligkeit Farbintensität Kontrast Mittlerer Farbton	Lautheit Wärme Brillianz Halligkeit Umhüllung	Breite Höhe Länge Volumen
	Gesamteindruck	-	Lautheit	audio-visuelle Passung Anwesenheit Aktivierung Potenz Gefallen

Tabelle 6: Auswertung perzeptiver Merkmale für die Variation der optischen und akustischen Raumpräsenz

Da auf dem Foto anstelle eines Cellisten oder eines Sprechers ein Lautsprecher als Klangquelle abgebildet war, ergibt sich keine Variation des optischen Inhaltes. Der Faktor Inhalt kann daher nicht sinnvoll ausgewertet werden. Soweit nicht anders angegeben beträgt die Anzahl der Versuchspersonen 35.

Es wurden jene Effekte in die Auswertung aufgenommen, die ein Signifikanzniveau von 0,05 % erreichten und eine Testpower größer 0,8 aufwiesen. Eine tabellarische Übersicht der statistischen Kennzahlen, inklusive F-Statistik und Anzahl der Freiheitsgrade, findet sich im Anhang 8.4 (S.86). Im Text wir als Maß für die statistischen Signifikanz eines Effekts der p-Wert verwendet. Als Maß für die Effektstärke wird das Eta² verwendet. Es gibt den prozentualen Anteil der erklärten Varianz an. Erzielt ein Faktor beispielsweise ein Eta² von 0,25, dann werden 25 % der Gesamtvarianz

durch diesen Faktor erklärt werden. Daher wird die Effektstärke im folgenden Verlauf in Prozent angegeben.

5.1.1 Variation des Faktors Raum

Die Variation des Raumes führt zu hoch signifikanten Effekten auf die perzeptiven Merkmale: Entfernung zur Klangquelle, Volumen (Skala), Volumen (berechnet) und Gefallen (alle $p < 0{,}001$). Das Volumen (Skala) weist mit 69 % die größte Effektstärke auf. Die Effektstärke des berechneten Volumens ist mit 31 % weniger als halb so groß.

Abbildung 5.1 zeigt die Beurteilung der Volumina (Skala). Der Raum Kt wird am kleinsten geschätzt, der Raum Kh größer. Bedeutend größer werden die Räume Gh und Gt beurteilt. Der Raum Gh weist das größte perzeptive Volumen (Skala) auf.

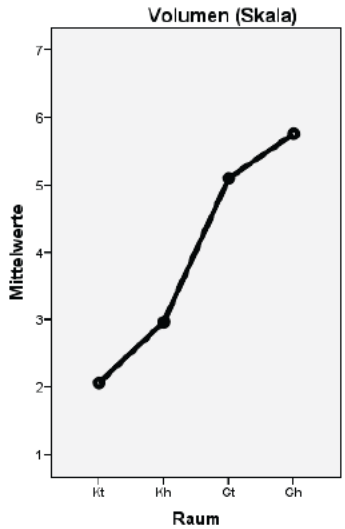


Abbildung 5.1

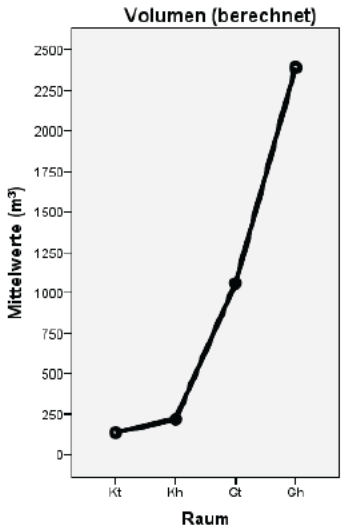


Abbildung 5.2

Die Rangfolge der Räume ist für das Volumen (berechnet) identisch (Abbildung 5.2). Gt und Gh erreichen auch hier deutlich größere perzeptive Volumina als Kt und Kh. Die Volumina der Räume Kt und Kh liegen näher beieinander. Ein Unterschied zum Volumen (Skala) besteht darin, dass der Raum Gh einen mehr als doppelt so hohen Wert erreicht als der Raum Gt. Der Unterschied zwischen den Räumen ist beim Volumen (Skala) weniger prägnant.

Für die Entfernung zur Klangquelle beträgt die Größe des Effektes 48 %. Die Entfernung zur Klangquelle wird für den Raum Kt auf 2,52 m, für Kh auf 2,82 m, für Gt auf 6,35 m und für Raum Gh auf 8,04 m geschätzt (Abbildung 5.3).

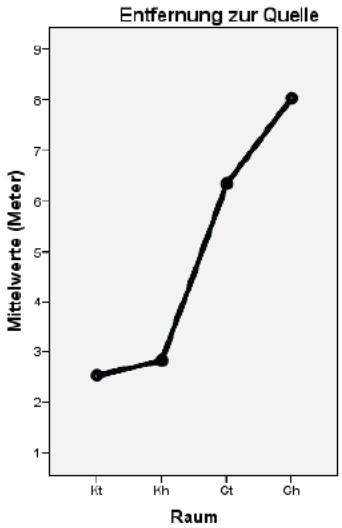


Abbildung 5.3

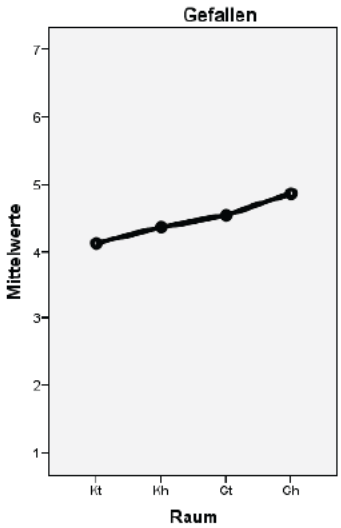


Abbildung 5.4

Das Merkmal Gefallen ist hoch signifikant durch die Variation des Raumes Beeinflusst ($p < 0{,}001$). Die Effektstärke beträgt 4 %. Abbildung 5.4 zeigt die Beurteilungen für das Gefallen der Räume in aufsteigender Reihenfolge: Kt, Kh, Gt, Gh.

Die Merkmale Aktivierung (1,7 %), Potenz (1,7 %) und Anwesenheit (1,4 %) erreichen nicht die nötige Testpower von 0,8.

5.1.2 Variation des Faktors Präsentationsmodus

Einen signifikanten Einfluss hat die Variation der optischen und akustischen Raumpräsenz auf die Merkmale Potenz, Aktivierung und Entfernung zur Klangquelle (alle $p < 0{,}001$).

Den stärksten Effekt übte die Variation der optischen und akustischen Raumpräsenz auf das Merkmal Potenz, mit 14 % erklärter Varianz aus. Unter der akustischen und die optoakustischen Bedingung ist die Ausprägung ähnlich (Abbildung 5.5). Auffällig ist, dass die optische Bedingung deutlich weniger potent wahrgenommen wird.

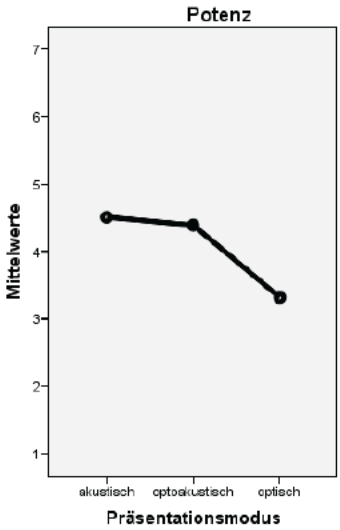


Abbildung 5.5

Die Aktivierung erreicht eine Effektstärke von 6 %. Die Art des Effektes ähnelt der Variable Potenz. Die optische Bedingung erreicht auch hier geringere Werte als die akustische und die optoakustische Bedingung (Abbildung 5.6). Bei den Variablen Potenz und Aktivierung ist die Tendenz zu erkennen, dass die optische Bedingung sowohl weniger potent, als auch weniger aktivierend wahrgenommen wird, als akustische und optoakustische Bedingung.

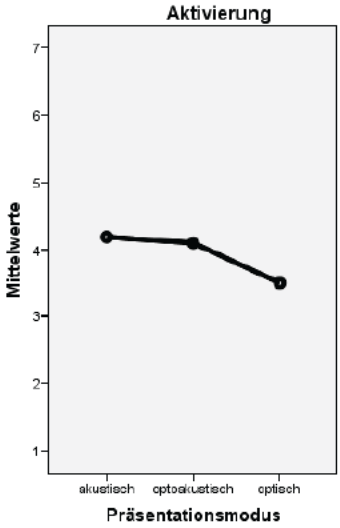


Abbildung 5.6

Die Variation der optischen und akustischen Raumpräsenz hat einen signifikanten Einfluss auf die geschätzte Entfernung zur Quelle ($p < 0{,}001$). Die Effektstärke beträgt 3,7 %. Unter der akustischen Bedingung wird die Entfernung bedeutend größer geschätzt als unter der optischen Bedingung (Abbildung 5.7). Der Wert der optoakustischen Bedingung liegt zwischen den beiden anderen.

Das Merkmal Gefallen (1,5 %) erreichte nicht die nötige Testpower von 0,8.

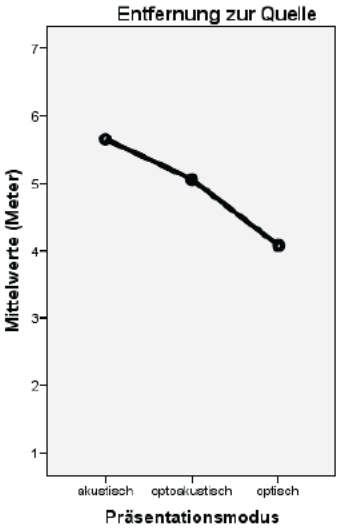


Abbildung 5.7

Keinen signifikanten Effekt erzielten das perzeptive Volumen (Skala / berechnet) und das Merkmal Anwesenheit.

5.1.3 Interaktionseffekte zwischen den Faktoren Präsentationsmodus und Raum

In Bezug auf das Merkmal Potenz ergibt sich ein hoch signifikanter Interaktionseffekt zwischen den Faktoren Präsentationsmodus und Raum ($p < 0{,}001$). Die Stärke des Effektes beträgt 5,5 %. In Abbildung 5.8 wird deutlich, dass die Variation der optischen und akustischen Raumpräsenz bei den unterschiedlichen Räumen einen unterschiedlichen Effekt hat. Zum Einen fällt auf, dass die Räume Kt und Kh unter der optischen Bedingung eine geringere Potenz aufweisen als die Räume Gt und Gh. Zum Anderen ist die Differenz zwischen akustischer und optischer Bedingung bei den Räumen Kt und Kh ausgeprägter. Die Räume Gt und Gh hingegen weisen unter der akustischen Bedingung die geringere Potenz auf. Zudem unterscheidet sich die optische Bedingung weniger stark vom der akustischen und der optoakustischen Bedingung.

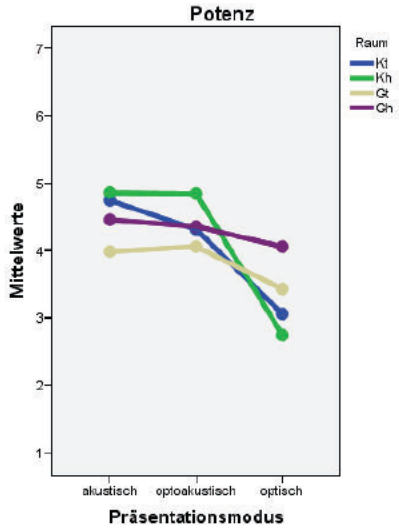


Abbildung 5.8

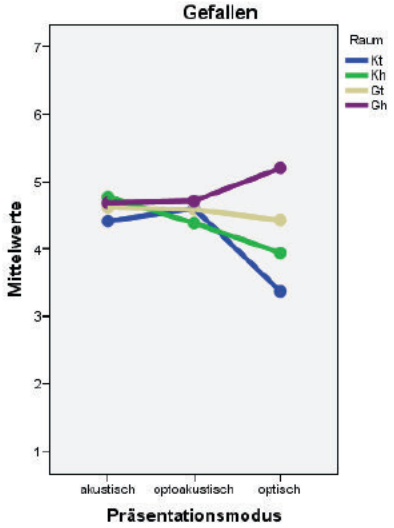


Abbildung 5.9

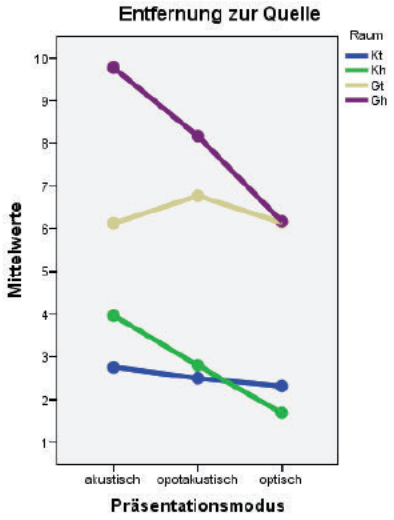


Abbildung 5.10

Bei dem Merkmal Gefallen tritt ein statistisch hoch signifikanter Effekt zwischen den Faktoren Raum und Präsentationsmodus auf ($p < 0{,}001$). Die Effektstärke beträgt 4,5 %. Die akustische und die optoakustische Bedingung üben kaum einen unterschiedlichen Effekt auf die verschiedenen Räume aus (Abbildung 5.9). Raum Kt gefällt unter der akustischen Bedingung am wenigsten, Raum Kh unter der optoakustischen. Der auffälligste Effekt ist die Differenz der Räume unter der optischen Bedingung. Raum Kt gefällt optisch am wenigsten, Raum Gh am besten. Von den dazwischen liegenden Räumen Gt und Kh, gefällt Raum Gt besser.

Die Entfernung zur Quelle wird hoch signifikant durch das Zusammenspiel der Faktoren Raum und Präsentationmodus beeinflusst ($p < 0{,}001$). Die Effektstärke beträgt 3,2 %. Abbildung 5.10 lässt erkennen, dass unter der optischen Bedingung die Entfernung zur Quelle für die Räume Gt und Gh auf je ca. 6,15 Meter, für Raum Kh auf 1,69 Meter und für Raum Kt auf 2,31 Meter geschätzt wird. Die akustische Bedingung zeigt den größten Effekt. Die geschätzten Entfernungen in den Räumen differieren stärker. Die Entfernung im Raum Gh wird mit 9,7 Metern mit Abstand am weitesten geschätzt, gefolgt von Gt mit 6,2 Metern. Unter der akustischen Bedingung wird die Entfernung in Raum Kh weiter geschätzt als in Raum Kt. Die optoakustischen Mittelwerte liegen jeweils zwischen den akustischen und den optischen. Mit Ausnahme von Raum Gt. Hier wird die Entfernung unter der optoakustischen Bedingung am weitesten geschätzt.

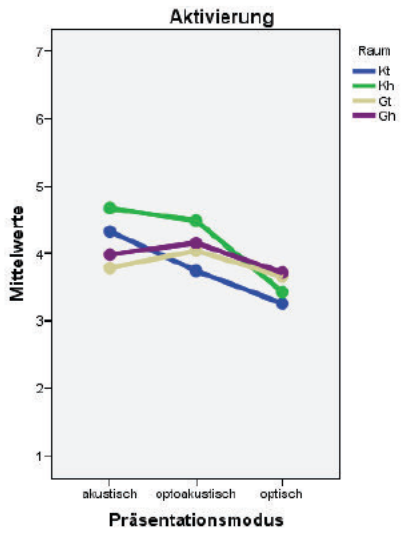


Abbildung 5.11

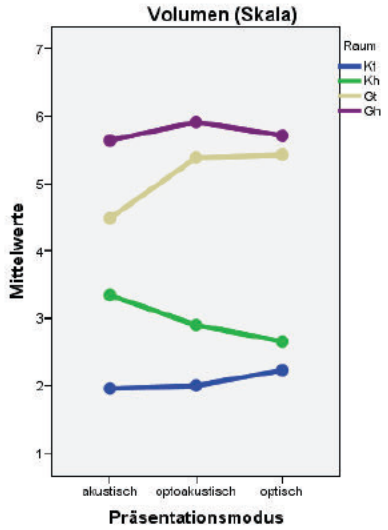


Abbildung 5.12

Ein hoch signifikanter Interaktionseffekt tritt bei dem Merkmal Aktivierung auf. Die Art des Effektes ist vergleichbar mit dem der Variable Potenz. Mit 3 % ist die Stärke des Effektes jedoch geringer. Unter der optischen Bedingung haben die Räume Gt und Gh eine aktivierende Wirkung, unter der akustischen Bedingung die Räume Kt und Kh (Abbildung 5.11). Bei den Räumen Gt und Gh ist die Tendenz zu erkennen, dass die optoakustische Bedingung am stärksten aktiviert. Bei den Räumen Kt und Kh hingegen liegt der optoakustische Mittelwert jeweils zwischen akustischem und optischem Mittelwert.

Das perzeptive Volumen (Skala) erzielt einen statistisch signifikanten Effekt ($p < 0{,}001$). Die Stärke des Effektes beträgt 2 %. Ähnlich wie bei dem Merkmal Entfernung zur Klangquelle sind die Differenzen unter der akustischen Bedingung zwischen den Räumen am größten (Abbildung 5.12). Unter der optischen Bedingung weisen jeweils die Räume Gt und Gh sowie Kt und Kh ähnliche Ausprägungen auf.

Die Merkmale Volumen (berechnet) und Anwesenheit zeigten keine signifikanten Interaktionseffekte.

5.2 Konfliktreize - Variation der optischen und akustischen Raum

5.2.1 Variation des Faktors Darbietungsinhalt

Die Variation des Inhaltes übt den stärksten Effekt auf die perzeptive Lautheit aus. Diese ist für alle drei Referenzobjekte hoch signifikant ($p < 0{,}001$). Ein Unterschied zwischen den drei Lautheitsbeurteilungen zeigt sich in der Effektstärke. Das Referenzobjekt Quelle zeigt eine Effektstärke von 25 %, gefolgt vom Gesamteindruck mit 17 %. Den mit 10 % geringsten aber noch immer deutlichen Effekt übten die verschiedenen Darbietungsinhalte auf die Lautheit des Referenzobjektes Raum aus. Abbildung 5.13 zeigt die Beurteilung der Lautheit in Bezug auf die verschiedenen Darbietungsinhalte für das Referenzobjekt Quelle. Der musikalische Stimulus wird deutlich lauter

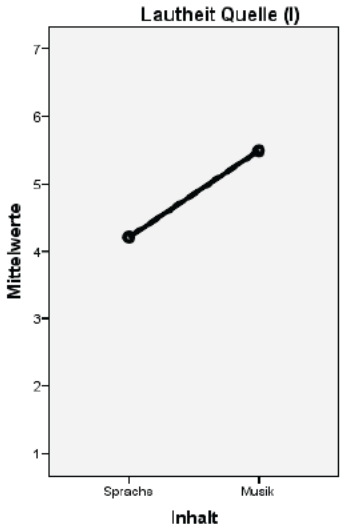


Abbildung 5.13

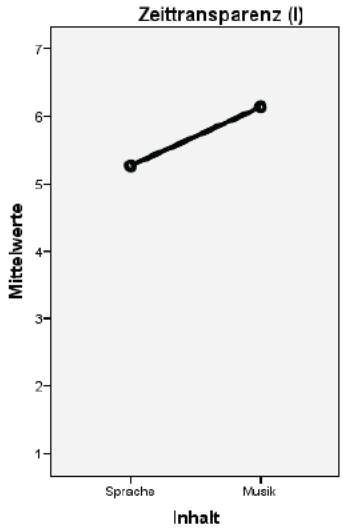


Abbildung 5.14

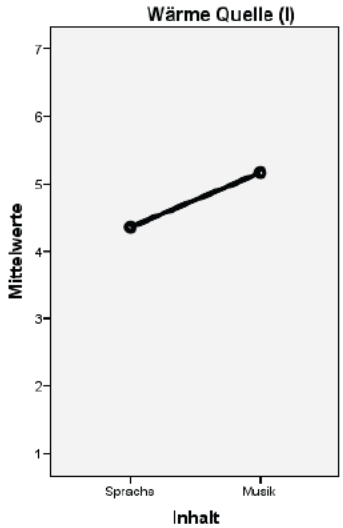


Abbildung 5.15

bewertet als der sprachliche Stimulus. Dieser Zusammenhang zeigt sich bei allen drei Referenzobjekten. Bei den Referenzobjekten Raum und Gesamteindruck ist der Mittelwert des musikalischen Stimulus mit 5,1 niedriger als beim Referenzobjekt Quelle mit 5,5.

Die Zeittransparenz zeigt auch einen hoch signifikanten Effekt ($p < 0{,}001$). Die Effektstärke für den Faktor Inhalt liegt bei 15 %. Die zeitliche Transparenz ist für den musikalischen Stimulus größer als für den sprachlichen Stimulus (Abbildung 5.14).

Die Wärme wurde für die Referenzobjekte Quelle und Raum abgefragt, beide weisen einen signifikanten Effekt auf (Quelle $p < 0{,}000$; Raum $p = 0{,}002$). Die Effekt-

stärke für die Wärme der Quelle beträgt 9 %. Für den Raum ist der Wert mit 4 % deutlich geringer. Die Art des Zusammenhangs ist bei beiden ähnlich. Der musikalische Stimulus erzeugt in beiden Fällen eine größere Wärme. Für die Quelle wird der musikalische Stimulus mit einem Mittelwert von 5,2 (Abbildung 5.15) wärmer bewertet als für den Raum mit 4,9.

Die Brillanz ist ebenfalls signifikant vom Darbietungsinhalt beeinflusst (Quelle $p < 0{,}000$; Raum $p = 0{,}001$). Die Effektstärke beträgt für die Quelle 7 % und für den Raum 4 %. Die Brillanz ist für den musikalischen Stimulus größer (Abbildung 5.16). Ein Vergleich der Referenzobjekte zeigt eine insgesamt geringfügig höhere Brillanz für die Quelle. Die Differenz der Mittelwerte beträgt 0,2 für den sprachlichen und 0,3 für den musikalischen Stimulus.

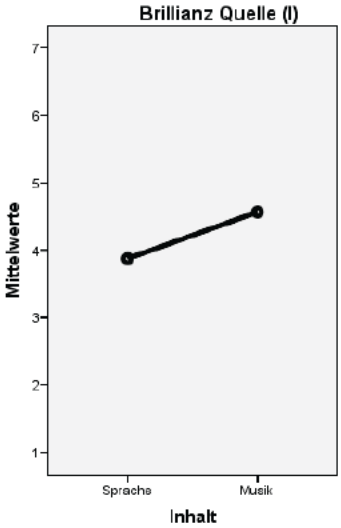


Abbildung 5.16

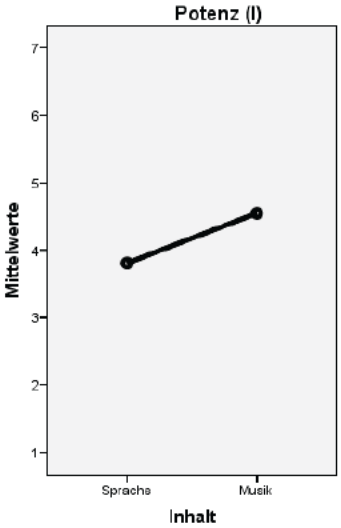


Abbildung 5.17

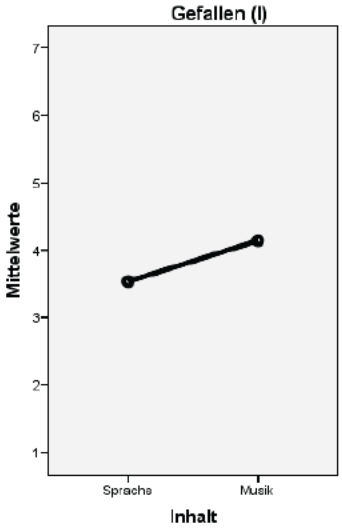


Abbildung 5.18

Auch auf die Beurteilung es Merkmals Potenz hat die Variation des Darbietungsinhaltes einen hoch signifikanten Effekt ($p < 0{,}001$). Die Effektstärke liegt bei 7 %. Der musikalische Stimulus wird mit einem Mittelwert von 4,5 als kraftvoller wahrgenommen als der sprachliche Stimulus, welcher einen Mittelwert von 3,8 erreicht (Abbildung 5.17).

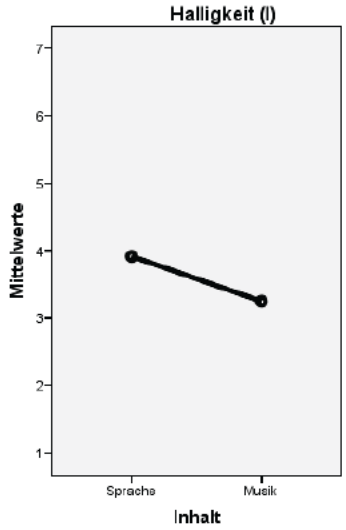


Abbildung 5.19

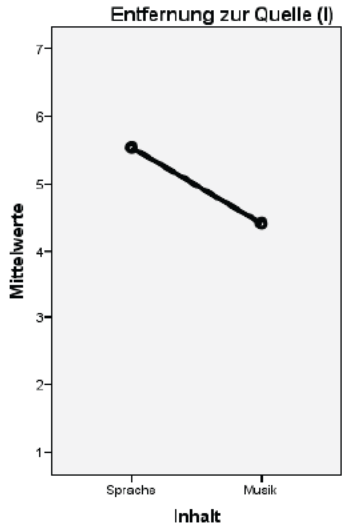


Abbildung 5.20

Das Merkmal Gefallen wird hoch signifikant durch die Variation des Darbietungsinhaltes beeinflusst ($p < 0{,}001$). Sie Stärke des Effektes liegt bei 3,8 %. Der musikalische Inhalt gefällt den Versuchspersonen besser, als der sprachliche (Abbildung 5.18).

Die perzeptive Halligkeit wird je nach Darbietungsinhalt signifikant unterschiedlich beurteilt ($p < 0{,}001$). Die Effektstärke liegt bei 3,4 %. Abbildung 5.19 zeigt, dass die der sprachliche Inhalt zu einer erhöhten perzeptiven Halligkeit führt.

Die geschätzte Entfernung zur Quelle wird signifikant durch die Darbietungsinhalte beeinflusst ($p < 0{,}001$). Die Effektstärke beträgt 3,2 %. Die Entfernung wird für den musikalischen Stimulus geringer geschätzt, als für den sprachlichen Stimulus (Abbildung 5.20).

Die Merkmale Aktivierung (1,8 %), Anwesenheit (1 %), Passung (0,9 %), Volumen (Skala) (0,7 %) und Volumen (berechnet) (0,6 %) werden ebenfalls signifikant, erreichen jedoch eine Testpower von weniger als 0,8. Auf die Merkmale Umhüllung, Helligkeit, Farbintensität und Kontrast hatte die Variation des Darbietungsinhaltes keinen signifikanten Effekt.

5.2.2 *Variation des Faktors optischer Raum*

Den stärksten Effekt erzielt die Variation des optischen Raumes auf die perzeptive Helligkeit ($p < 0{,}001$). Das gilt sowohl für das Referenzobjekt Quelle, als auch für den Raum. Die Effektstärke in Bezug auf die Quelle beträgt 57 %, in Bezug auf den Raum sogar 59 %. In der Art des Effektes ist zwischen den Referenzobjekten kein bemerkenswerter Unterschied zu entdecken. Abbildung 5.21 zeigt die Art des Effektes am Beispiel des Raumes. Der optische Raum Kt wird mit Abstand am dunkelsten

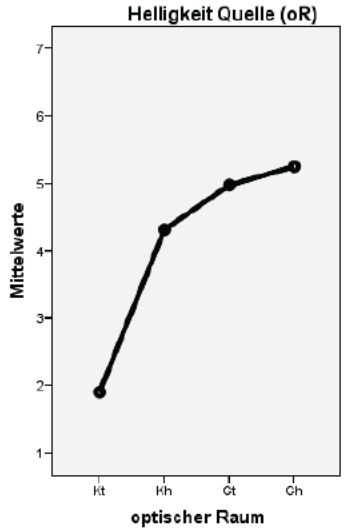


Abbildung 5.21

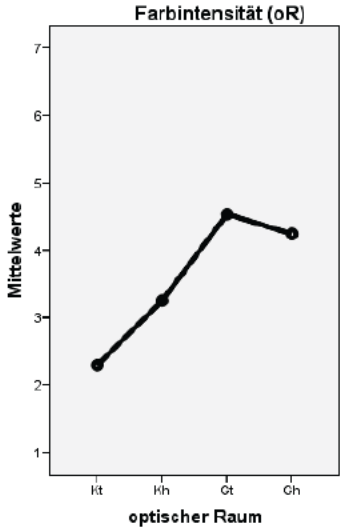


Abbildung 5.22

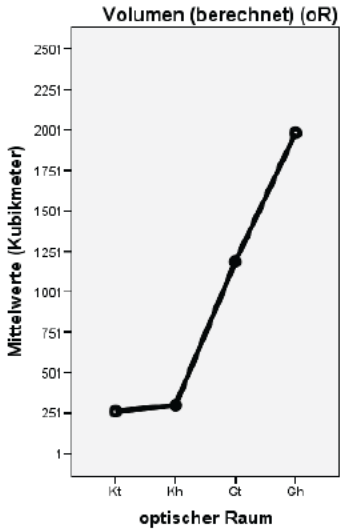


Abbildung 5.23

wahrgenommen. Raum Kh nimmt eine mittlere Position ein. Die optischen Räume Gt und Gh weisen die größte perzeptive Helligkeit auf, wobei Gh am hellsten wahrgenommen wird.

Einen ebenfalls hoch signifikanten Effekt erzielt der optische Raum auf die Farbintensität ($p < 0{,}001$). Die Stärke des Effektes beträgt 36 %. Die Ausprägung des Effektes in Bezug auf die einzelnen optischen Räume ist in Abbildung 5.22 dargestellt. Der optische Raum Gt zeigt die stärkste Ausprägung der perzeptiven Farbintensität, dicht gefolgt von Gh. Deutlich weniger farbintensiv wird Raum Kh wahrgenommen. Eine noch geringere Ausprägung weist Raum Kt auf.

Die Merkmale Volumen (Skala) und Volumen (berechnet) werden hoch signifikant durch den optischen Raum beeinflusst ($p < 0{,}001$). Die Effektstärke beträgt für Volumen (Skala) 31 % und für Volumen (berechnet) 20 %. Abbildung 5.24 zeigt die Art des Einflusses der optischen Räume auf das perzeptive Volumen (Skala). Die optischen Räume Gt und Gh werden als größer beurteilt, als die beiden anderen Räume. Zudem weisen sie keinen nennenswerten Unterschied untereinander auf. Das Volumen der optischen Räume Kt und Kh wird bedeutend kleiner geschätzt, dabei weist der optische Raum Kh ein etwas größeres perzeptives Volumen auf. Das Volumen (berechnet), welches sich aus der Multiplikation der Merkmale Länge, Breite und Höhe errechnet ist in Abbildung 5.23 abgebildet. Durch die Angabe der Raumdimensionen in Metern sind die berechneten Mittelwerte auf der Y-Achse als perzeptive Kubikmeter zu verstehen. Die Räume Kt und Kh ergeben sich auch beim berech-

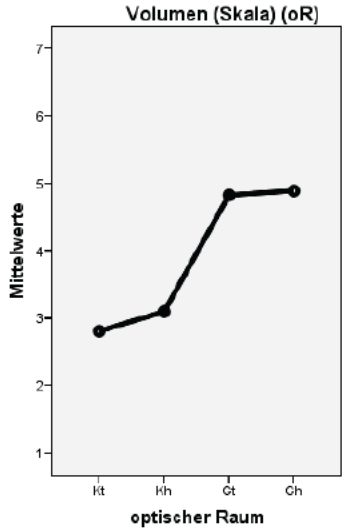


Abbildung 5.24

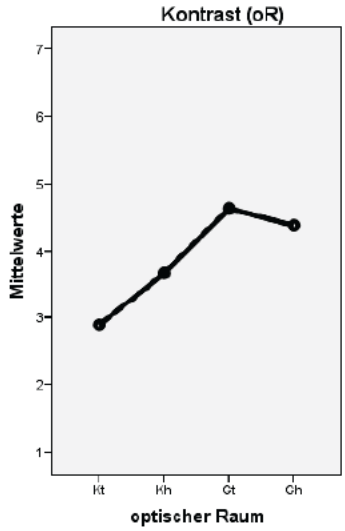


Abbildung 5.25

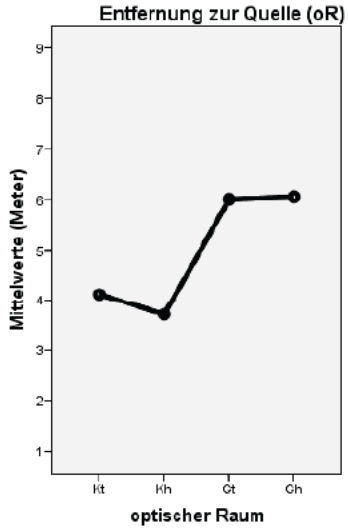


Abbildung 5.26

neten Volumen als die Kleineren. Der Raum Kh zeigt auch hier ein etwas größeres Volumen. Im Unterschied zum Merkmal Volumen (Skala) unterscheiden sich die optischen Räume Gt und Gh beträchtlich voneinander, wobei Gh das größte perzeptive Volumen (berechnet) aufweist. Der Zusammenhang der beiden perzeptiven Volumenskalen wird in Kapitel 5.3 (S. 50) näher untersucht.

Der perzeptive Kontrast erzielt ebenfalls einen hoch signifikanten Effekt ($p < 0{,}001$). Die Effektstärke beträgt 22 %. Das Verhältnis der Mittelwerte untereinander zeigt Abbildung 5.25. Der optische Raum Gt zeigt die stärkste Ausprägung, dicht gefolgt von Gh. Deutlich weniger farbintensiv wird Raum Kh wahrgenommen. Eine noch geringere Ausprägung weist Raum Kt auf. Dieser Zusammenhang ist nahezu identisch mit der Variable Farbintensität (Abbildung 5.22).

Die Entfernung zur Klangquelle wurde durch die Variation des optischen Raumes signifikant beeinflusst ($p < 0{,}001$). Der optische Raum trägt mit einer Effektstärke von 12 % zur Entfernungsschätzung bei. Die Entfernung zur Klangquelle ist für die optischen Räume Gt und Gh nahezu identisch (Abbildung 5.26). Für den optischen Raum Gt wird sie auf 6,01 m geschätzt, für Gh auf 6,06 m. Deutlich geringer ist die wahrgenommene Entfernung bei den Räumen Kt und Kh. Die geschätzte Entfernung für den optischen Raum Kt beträgt 4,13 m, für den Raum Kh 3,71 m.

Die Variation des optischen Raumes übt einen signifikanten Effekt auf das Merkmal Gefallen aus ($p = 0{,}001$). Die Effektstärke beträgt 1,3 %. Die optischen Räume Gt und Gh werden am höchsten beurteilt und

weisen eine ähnliche Ausprägung auf, während der optische Raum Kt die geringste und Kh eine etwas höhere Bewertung erhält.

Der Effekt trifft in gleicher Art und Weise, lediglich mit geringerer Stärke, bei den Merkmalen Anwesenheit (Abbildung 5.29) und Potenz (Abbildung 5.28) auf. Die Effektstärke beträgt für die Anwesenheit 1,1 % ($p < 0{,}001$), und für die Potenz 0,8 % ($p = 0{,}001$).

Auf folgende Merkmale übte die Variation des optischen Raumes keinen signifikanten Effekt aus: Lautheit, Wärme, Brillanz, Zeittransparenz, Halligkeit, Umhüllung, Aktivierung.

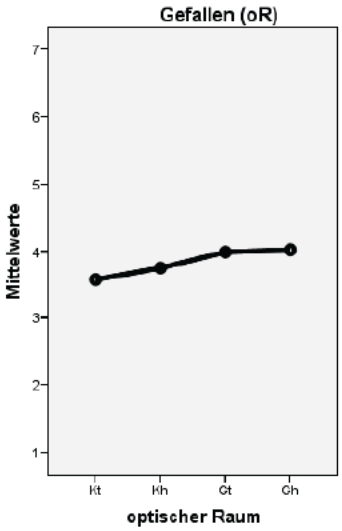


Abbildung 5.27

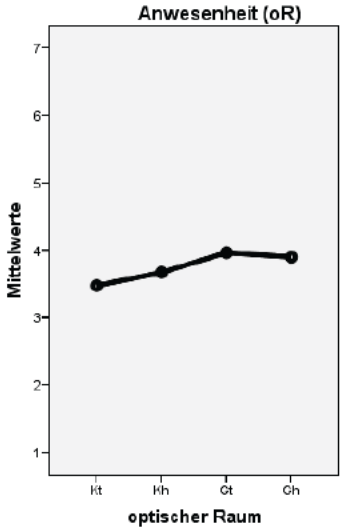


Abbildung 5.29

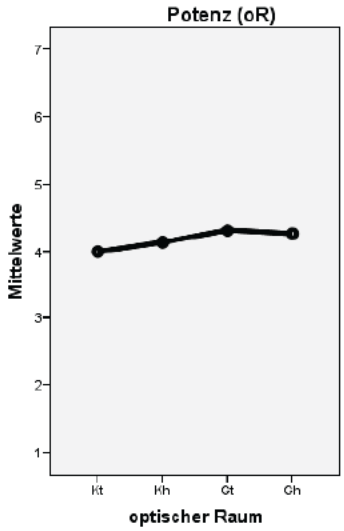


Abbildung 5.28

5.2.3 Variation des Faktors akustischer Raum

Die Variation des akustischen Raumes übt einen hoch signifikanten Effekt ($p < 0{,}001$) auf das Merkmal Halligkeit aus. Mit 53 % ist die Stärke des Effektes außerordentlich hoch. Dies zeigt sich auch in Abbildung 5.30. Es treten keine Gruppierungen der akustischen Räume auf, alle vier unterscheiden sich voneinander. Von den Räumen mit geringerem Volumen weist Raum Kt die geringste, Raum Kh eine größere perzeptive Halligkeit auf. Die Räume mit größerem Volumen werden als halliger bewertet, wobei Raum Gh als am halligsten bewertet wird.

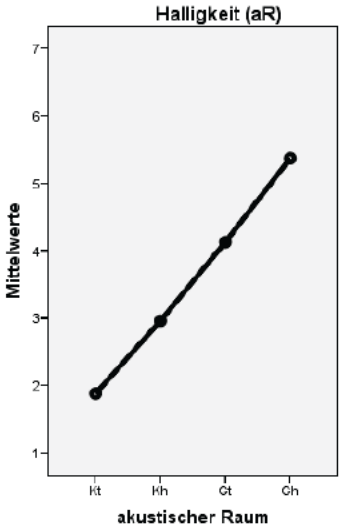


Abbildung 5.30

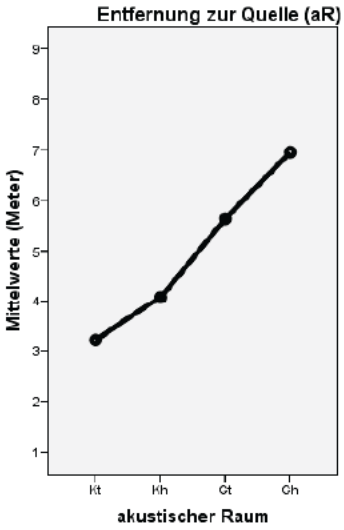


Abbildung 5.31

Die wahrgenommen Entfernung zur Quelle wird hoch signifikanten durch die Variation des akustischen Raumes beeinflusst ($p < 0{,}001$). Die Effektstärke liegt bei 21 % erklärter Varianz. Es zeigt sich keine Gruppierung von Räumen. Die geschätzte Entfernung variiert von Raum zu Raum (Abbildung 5.31). Sie beträgt für den akustischen Raum Kt 3,22 m, für Raum Kh 4,08 m, für Raum Gt 4,65 m und für Raum Gh 6,96 m.

Die perzeptive Lautheit erzielt in Bezug auf alle drei Referenzobjekte einen hoch signifikanten Effekt ($p < 0{,}001$). Ein Unterschied besteht in der Stärke des Effektes. Die größte Effektstärke ergibt sich mit 21 % für das Referenzobjekt Quelle, die zweit höchste mit 14 % für den Gesamteindruck. Den geringsten Effekt auf die perzeptive Lautheit hat die Variation des akustischen Raumes in Bezug auf das Referenzobjekt Raum. Hier ist die Effektstärke mit 10 % nur etwa halb so groß wie für das Referenzobjekt Quelle. Die Art des Effektes ist für die Quelle (Abbildung 5.32) und Gesamteindruck (Abbildung 5.33) ähnlich. Für das Referenzobjekt Quelle werden die Räume Kt und

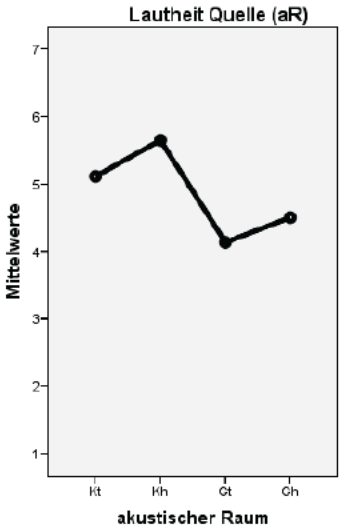


Abbildung 5.32

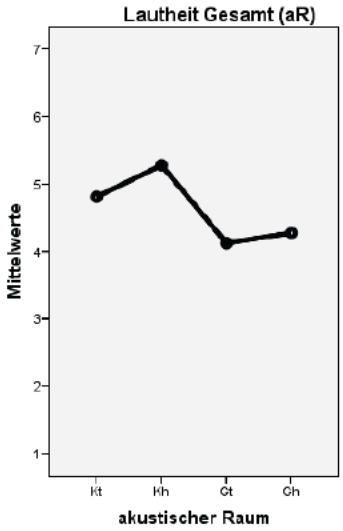


Abbildung 5.33

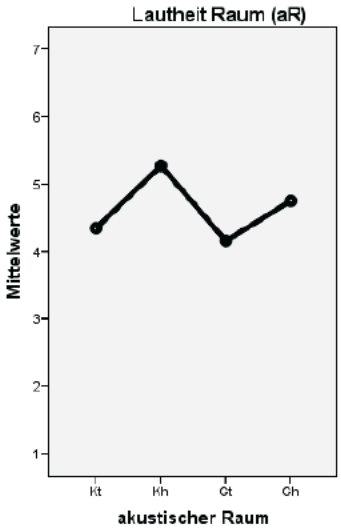


Abbildung 5.34

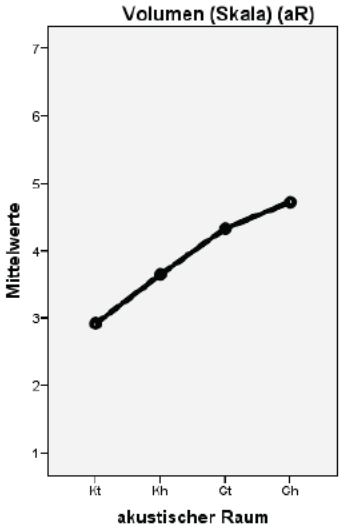


Abbildung 5.35

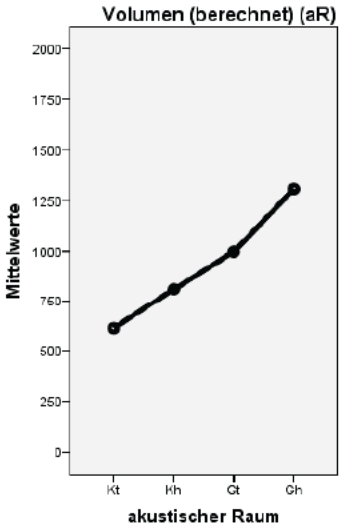


Abbildung 5.36

Kh lauter wahrgenommen, als die Räume Gh und Gt, wobei Raum Kh der lautere von beiden ist. Von den Räumen Gt und Gh ist Gh der lautere. Beim Gesamteindruck werden die Räume Kt, Kh und Gh im Verhältnis leiser bewertet, wodurch Raum Gt und Gh nur noch eine geringe Differenz aufweisen. Abgesehen davon bleibt das Lautheitsverhältnis zwischen den Räumen gleich. In dieser Hinsicht unterscheidet sich die Lautheit des Raumes von den beiden anderen Referenzobjekten (Abbildung 5.34). Hier weisen die Räume Kt und Gt die geringste Lautheit auf, die zudem noch von ähnlicher Ausprägung ist. Raum Gh wird hier deutlich lauter bewertet und bewegt sich in der Größenordnung von Raum Kh der auch in diesem Fall die größte perzeptive Lautheit aufweist.

Beide perzeptiven Volumenskalen erreichen eine hohe Signifikanz ($p < 0{,}001$). Ein bedeutender Unterschied zeigt sich in der Effektgröße. Während das Merkmal Volumen (Skala) eine Effektgröße von 16 % aufweist, beträgt diese für die aus den Angaben Höhe, Länge und Breite berechnete Volumenskala nur 2,6 %. Das perzeptive Volumen (Skala) des akustischen Raumes Kt ist am geringsten, gefolgt von Raum Kh (Abbildung 5.35). Noch größer wird Raum Gt geschätzt. Das größte perzeptive Volumen (Skala) erreichte in diesem Zusammenhang der akustische Raum Gh. Trotz der geringen Effektstärke des Merkmales Volumen (berechnet), sei die Art des Effektes hier der Vollständigkeit halber erwähnt. In Abbildung 5.38 ist zu sehen, dass die Mittelwerte der perzeptiven Volumina (berechnet) für alle vier Räume verschieden sind. Die Rangfolge der akustischen Räume entspricht der des Merkmales Volumen (Skala). Die Y-Achse kann als perzeptive Kubikmeterangabe verstanden werden.

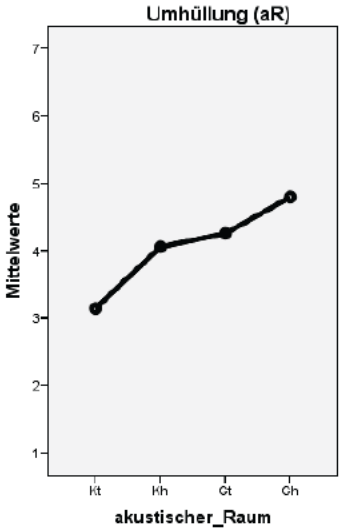


Abbildung5.37

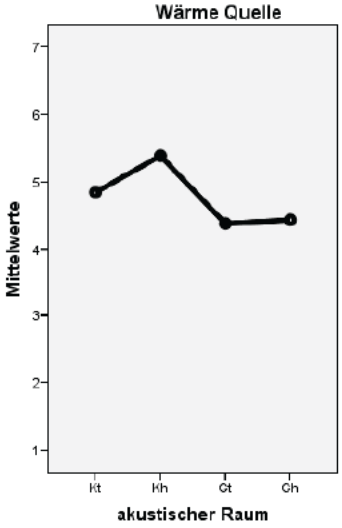


Abbildung 5.38

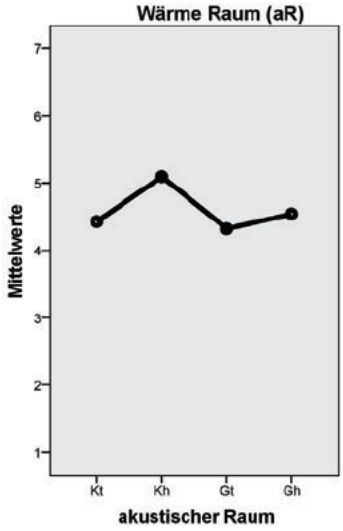


Abbildung 5.39

Das Merkmal Umhüllung wird hoch signifikant ($p < 0{,}001$) vom akustischen Raum beeinflusst. Die Stärke des Effektes beträgt 14 %. Am wenigsten umhüllend wird der akustische Raum Kt wahrgenommen, am umhüllendsten der akustische Raum Gh (Abbildung 5.38). Die Räume Kh und Gt befinden sich dazwischen und weisen eine ähnliche Ausprägung auf, wobei Gt als etwas umhüllender wahrgenommen wird.

Auch die Wärme erzielt einen signifikanten Effekt. Während die Wärme in Bezug auf des Referenzobjekt Quelle ein Signifikanzniveau von $p < 0{,}001$ erreichte, betrug dieses für das Referenzobjekt Raum $p = 0{,}002$. Zudem ist die Effektstärke in Bezug auf die Quelle mit 9 % größer als in Bezug auf den Raum (5 %). Am wärmsten wird der akustische Raum Kh empfunden, gefolgt von Kt. Die Räume Gt und Gh weisen eine geringere perzeptive Wärme auf und zeigen eine ähnliche Ausprägung (Abbildung 5.38). Im Verhältnis dazu senkt sich die Wärme beim Referenzobjekt Raum für die Räume Kt und Kh ab. Raum Gh hingegen wirkt wärmer (Abbildung 5.39). Raum Kt und Gh zeigen im Gegensatz zum Referenzobjekt Quelle eine ähnliche Ausprägung.

Der Effekt auf das Merkmal Zeittransparenz ist mit einem Signifikanzniveau von $p < 0{,}001$ hoch signifikant. Die Größe des Effektes beträgt 8 %. Wie in Abbildung 5.40 zu sehen, ist die Zeittransparenz generell hoch. Die akustischen Räume Kt und Kh weisen die höchste Zeittransparenz auf und befinden sich auf gleichem Niveau. Die Räume Gt und Gh fallen weniger zeittransparent aus, wobei Raum Gh die geringste Zeittransparenz aufweist.

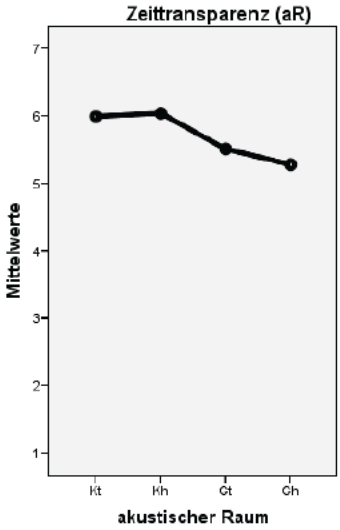


Abbildung 5.40

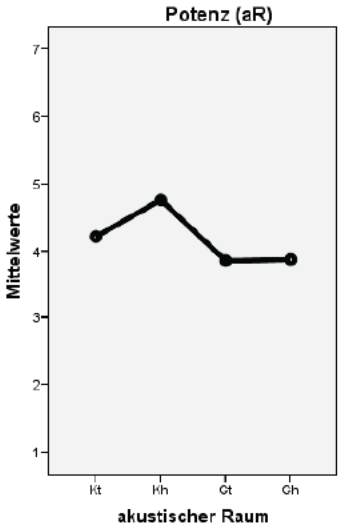


Abbildung 5.41

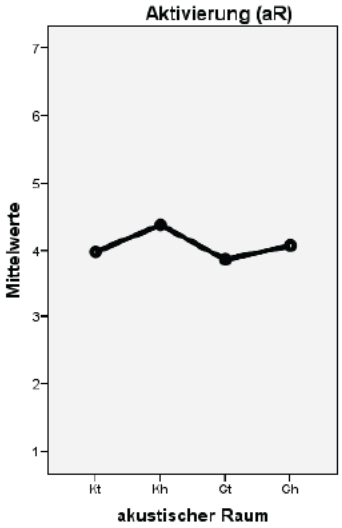


Abbildung 5.42

Die Potenz wird ebenfalls hoch signifikant beeinflusst ($p < 0{,}001$). Die Stärke des Effekts liegt bei 7 %. Die akustischen Räume Gt und Gh zeigen eine ähnliche Ausprägung und sind zugleich die am wenigsten potenten Räume (Abbildung 5.41). Die Räume Kt und Kh führen zu einer höheren Bewertung der Potenz. Der akustischen Raum Kh weist den höchsten Wert auf.

Das Merkmal Aktivierung wird signifikant durch den akustischen Raum beeinflusst ($p = 0{,}005$). Die Effektstärke beträgt 2,7 %. Abbildung 5.42 zeigt die Art des Effekts. Für den akustischen Raum Kh wird die Aktivierung am höchsten beurteilt. Die Räume Kt, Gt und Kt wirken weniger aktivierend, wobei Raum Gt die geringste Ausprägung aufweist.

Die Merkmale Anwesenheit (2 %) und Gefallen (1,3 %) weisen ebenfalls einen signifikanten Effekt auf, dieser wird jedoch durch einen starken Interaktionseffekt zwischen optischem und akustischem Raum erklärt, welcher im folgenden Abschnitt besprochen wird.

Das Merkmal Brillanz des Raumes erreichte nicht die notwendige Testpower von 0,8. Die Merkmale Brillanz der Quelle, Helligkeit, Farbintensität und Kontrast wurden nicht signifikant durch den akustischen Raum beeinflusst.

5.2.4 Interaktionseffekte

Es tritt ein prominenter Interaktionseffekt zwischen optischem und akustischem Raum auf, der sich insbesondere auf die drei Merkmale optoakustische Passung, Anwesenheit und Gefallen auswirkt. Der Interaktionseffekt erzielt in allen drei Fällen ein Signifikanzniveau von $p < 0{,}001$. Die Effektstärke ist für die Passung mit 33 % am größten, gefolgt von Anwesenheit (31 %) und Gefallen mit (16 %).

Abbildung 5.43 zeigt die Art der Interaktion. Auf der X-Achse sind die akustischen Räume Kt bis Gt abgetragen. Die Mittelwerte sind in dieser Abbildung nach den optischen Räumen aufgeschlüsselt, daher gibt es pro akustischem Raum nicht nur einen Mittelwert, sondern vier. Jeder optische Raum hat eine eigene Färbung, die in der Legende eingesehen werden kann. Bei Mittelwerten, die mit einem Quadrat gekennzeichnet sind, liegt eine physikalische optoakustische Passung des optischen und akustischen Raumes vor. Wie zu erkennen ist, weisen diese Stimuluskombinationen jeweils die größte perzeptive Passung auf. Je schlechter die reale Passung der Stimuluskombinationen, desto schlechter werden sie perzeptiv beurteilt. Eine besonders schlechte Passung erzielen die optischen Räume Gt und Gh mit dem akustischen Raum Kt und die optischen Räume Kt und Kh mir dem akustischen Raum Gh. Zudem fällt auf, das bei den optischen Räume Kh und Gt bei realer Passung zwar die perzeptive Passung am größten ist, jedoch ist der Unterschied zu den akustischen Räumen ähnlicher Größe jeweils geringer als bei den optischen Räumen Kt und Gh.

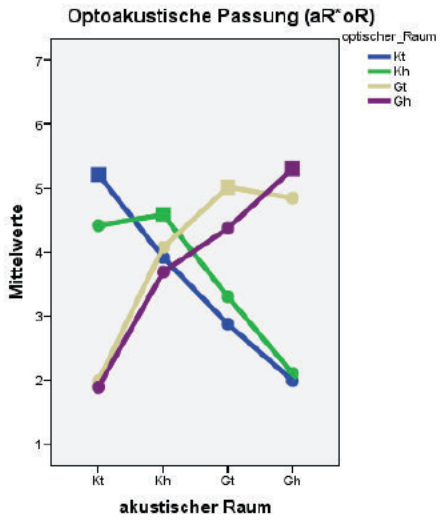


Abbildung 5.43

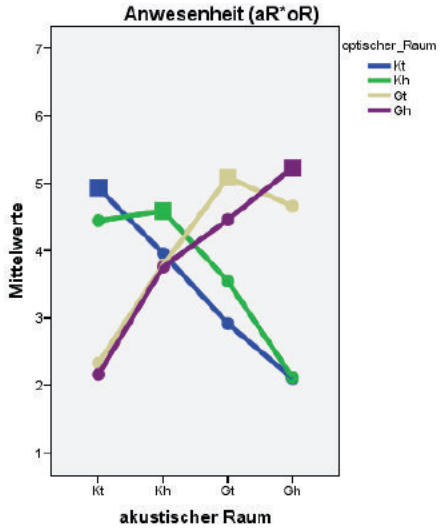


Abbildung 5.44

Dieser Interaktionseffekt zeigt sich mit ähnlich hoher Effektstärke bei dem Merkmal Anwesenheit (31 %). Die Anwesenheit ist bei optoakustischer Passung der Stimuli am größten (Abbildung 5.44).

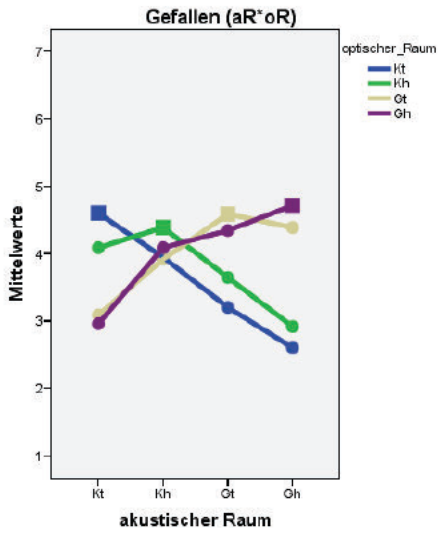


Abbildung 5.45

Wenn auch mit geringerer Effektstärke, so weist die optoakustische Passung dennoch einen maßgeblichen Effekt auf das Merkmal Gefallen (16 %) aus (Abbildung 5.45).

Ebenfalls signifikant aber mit einer Effektstärke kleiner 4 % kann dieser Interaktionseffekt zwischen optischem und akustischem Raum auch bei den Merkmalen Potenz (2,4 %) und Lautheit (Gesamteindruck) (1,8 %) beobachtet werden.

5.3 Methodische Sonderfrage: Zusammenhang der perzeptiven Volumenskalen

5.3.1 Perzeptive Raumgröße

Das physikalische Maß für die Raumgröße ist dessen räumlicher Inhalt, der als Volumen bezeichnet wird. Das Konzept des Volumens ist allgemein bekannt und der Begriff in der Alltagssprache gebräuchlich. Daher wurde er für die Erhebung der perzeptiven Raumgröße im Fragebogen genutzt.

Zwar gehört der Begriff zum allgemeinen Sprachgebrauch, das Schätzen von Volumina dagegen kommt im Alltag nur selten vor. Zieht man in Betracht, dass es vielen Menschen schon nicht leicht fällt, eindimensionale Größen wie Entfernungen valide zu schätzen, dann scheint es plausibel, dass die direkte Schätzung eines Raumvolumens, welches im Verhältnis zu seinen eindimensionalen Maßen Höhe, Länge und Breite mit der 3. Potenz wächst, ohne Training nahezu unmöglich ist.

Von der Angabe einer Kubikmeterzahl als Fragebogen-Item wurde aus diesen Gründen abgesehen. Um trotzdem ein perzeptives Maßes für die Raumgröße zu erhalten wurden zwei verschiedene Ansätze verfolgt. Zum einen wurde die Probanden gebeten das Volumen auf einer siebenstufigen, äquidistanten Skala mit den Endpunkten „klein“ und „groß“ zu schätzen. Das auf diese Weise erhobene perzeptive Volumenmaß wird im weiteren Verlauf der Arbeit als „Volumen (Skala)“ bezeichnet.

Um ein weiteres Maß zu erhalten wurden die Probanden gebeten Breite, Länge und Höhe jedes Raumstimulus in Metern zu schätzen. In Analogie zur Volumenberechnung in realen Räumen, wurde durch Multiplikation dieser drei Merkmale ein weiteres Maß für das perzeptive Raumvolumen berechnet, das im weiteren Verlauf als „Volumen (berechnet)“ bezeichnet wird.

Es soll zum Einen die Frage beantwortet werden, auf welche Weise die beiden perzeptiven Volumenskalen zusammenhängen und zum Anderen wie das Merkmal Volumen (berechnet) mit den tatsächlichen Volumina der Räume zusammenhängt. Als Resultat soll entschieden wie die perzeptive Raumgröße in zukünftigen Studien am sinnvollsten operationalisiert werden kann.

Der Zusammenhang der Merkmale Volumen (Skala) und Volumen (berechnet) wurde auf zwei Arten bestimmt. Da die Schätzung von Größe ist ein hoch kognitiver

Prozess. Die Vorstellung von der Größe eines Raumes wird durch verschiedene Modalitäten gespeist und anhand dieser Vorstellung wird eine Schätzung abgegeben. Aufgrund des Grades an Abstraktion dieses Prozesses geht der Autor davon aus, dass weder die Art der Präsentation der Stimuli (optisch, akustisch, optoakustisch), noch das Verwenden von Konfliktreizen oder verschiedenen akustischen Darbietungsinhalten einen Einfluss auf den Zusammenhang der beiden Merkmale hat. Aus diesem Grund sind die Messungen aller 44 Stimuli gleichermaßen geeignet den Zusammenhang aufzudecken.

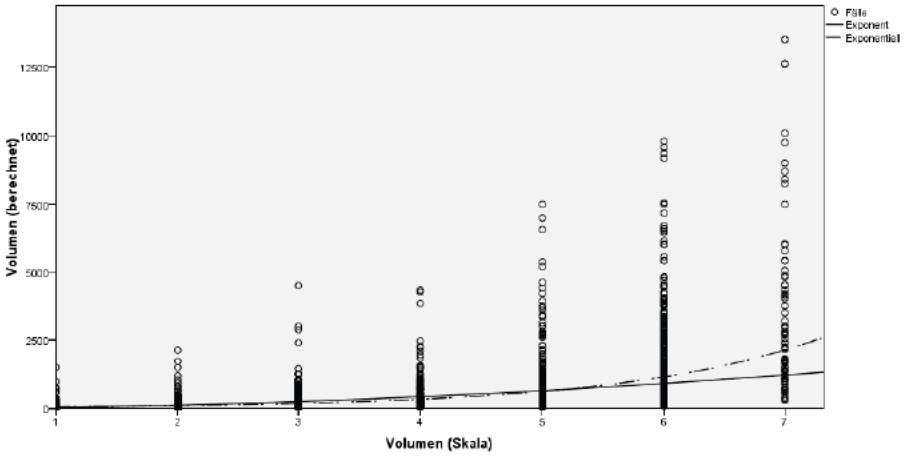


Abbildung 5.46: *Regressionsmodell der Merkmale Volumen (Skala) und Volumen (berechnet)*

Dazu wurde eine Kurvenanpassug zwischen den beiden Merkmalen durchgeführt. Als Fälle dienten alle Merkmals-Paare. 44 Stimuli mal 35 Versuchspersonen ergaben nach Abzug ungültiger Angaben und drei extremen Ausreißern 1509 Fälle. Grafik Abbildung 5.46 zeigt die Verteilung dieser Merkmalspaare. Je größer die Werte des Merkmals Volumen (Skala), desto mehr Fälle weisen auch ein größeres Volumen (berechnet) auf. Das Regressionsmodell ergab als passendste Funktion eine Exponentialfunktion mit den Koeffizienten:

$$29{,}428 \cdot e^{0{,}612\,x} \quad (R^2 = 0{,}551;\ F = 1474{,}265;\ p = {,}000).$$

Die zweitbeste Passung erzielte eine Potenzfunktion mit den Koeffizienten:

$$33{,}140 \cdot x^{1{,}855} \quad (R^2 = 0{,}495;\ F = 1851{,}352;\ p = {,}000)\ .$$

Aufgrund der hohen interpersonellen Varianz erreicht das Fitmaß R^2 in beiden Fällen eher geringe Werte, die einen Zusammenhang jedoch belegen.

Ein zweiter Ansatz, den Zusammenhang der Skalen darzustellen bestand darin die Interpersonenvarianz vor dem Vergleich herauszumitteln. Da extreme Ausreißer vor kamen, die den Mittelwert im Verhältnis zur Grundgesamtheit stark verzerren, wurden diese von der Auswertung ausgeschlossen. Um dies auf systematische Weise zu tun wurden jeweils jene Fälle ausgeschlossen, die sich außerhalb des 10ten und 90ten Perzentil befanden. Für die verbleibenden Fälle wurden die Werte der Merkmale für jeden Stimulus über die entsprechenden Versuchspersonen gemittelt. Das Ergebnis ist jeweils ein konkreter Wert für das Merkmal Volumen (Skala) und das Merkmal Volumen (berechnet) pro Stimulus. In Abbildung 5.47 sind die 44 Mittelwerte gegeneinander aufgetragen. Jeder Punkt repräsentiert einen Stimulus.

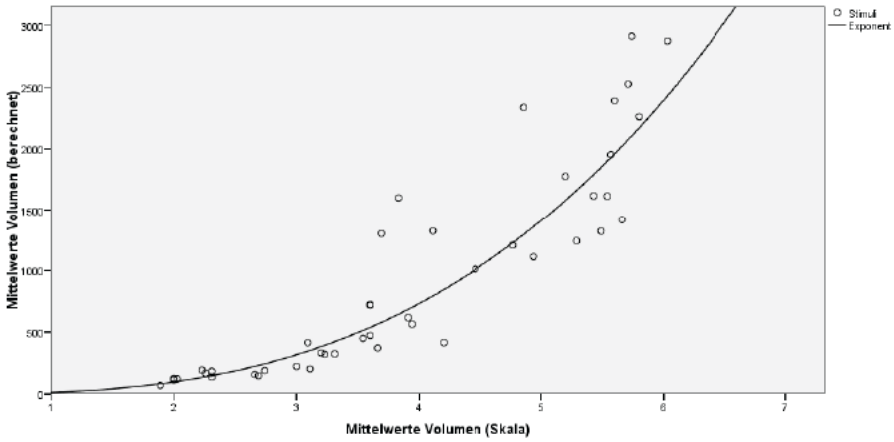


Abbildung 5.47: *Vergleich der Mittelwerte: Volumen (Skala), Volumen (berechnet)*

Ein Zusammenhang ist schon mit dem Auge deutlich zu erkennen. Je kleiner die Werte des Merkmals Volumen (Skala), desto kleiner auch die Werte des Merkmals Volumen (berechnet), wobei die Streuung bei größeren Volumina zunimmt. Wie erwartet stellt sich der Zusammenhang ohne die Interpersonenvarianz als viel ausgeprägter dar. Das Regressionsmodell ergab als beste Passung zur Verteilung eine Potenzfunktion mit der Formel:

$$12{,}911 \cdot x^{2{,}915} \quad (R^2 = 0{,}900;\ F = 378{,}831;\ p = {,}000)\ .$$

Diese ermöglicht es aus den Werten der siebenstufigen, äquidistanten Skalen den entsprechenden Wert d Volumen (berechnet) vorherzusagen. Die Annahme, dass beide Skalen dasselbe messen scheint auf Grund des klaren Zusammenhanges berechtigt.

5.3.2 Zusammenhang der perzeptiven und realen Volumina

Raum	Physikalisches Volumen (m³)	Perzeptives Volumen (berechnet)				
Präsentationsmodus Inhalt		optisch -	optoakustisch Sprache	optoakustisch Musik	akustisch Sprache	akustisch Musik
Kh	190	164	224	164	522	470
Kt	230	222	142	132	138	78
Gt	3300	1758	1445	1438	1427	468
Gh	4200	2851	3478	2633	3171	1841

Tabelle 7: Physikalische und perzeptive Volumina

Ein weiterer interessanter Gesichtspunkt ist, wie gut die perzeptiven Volumina mit der Realität übereinstimmen, also wie exakt die Versuchspersonen das physikalische Volumen bestimmen konnten. Jeder Raum hat jeweils ein physikalisches Volumen. Bedingt durch das Versuchsdesign ergeben sich jedoch fünf perzeptive Volumina, die sich aus dem Zusammenspiel von akustischem Darbietungsinhalt und Präsentationsmodus ergeben. Bei den optoakustischen Stimuli wurden nur jene mit optoakustischer Passung berücksichtigt. Bei der Berechnung der Mittelwerte wurden ebenfalls nur Fälle berücksichtigt, die innerhalb der zehnten und neunzigsten Perzentile lagen, um Verzerrungen durch Ausreißer systematisch zu vermeiden.

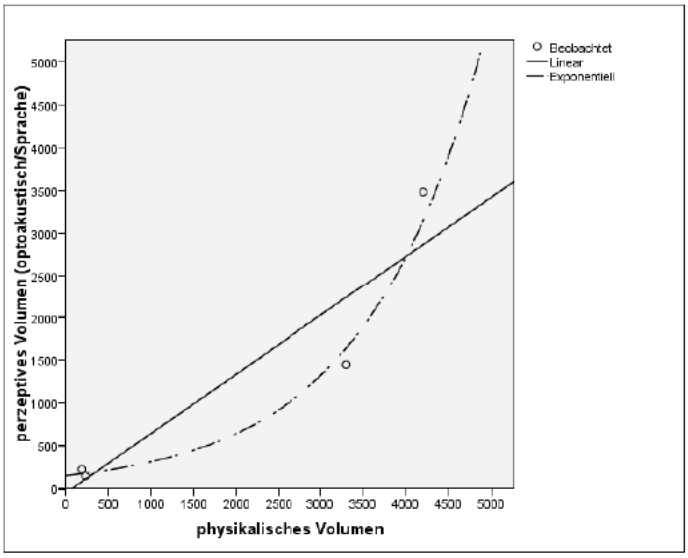


Abbildung 5.48: *Physikalisches und perzeptives Volumen für die optoakustischen Stimuli mit sprachlichem Inhalt*

Tabelle 7 gibt einen Überblick über die physikalischen und perzeptiven Volumina. Es lässt sich erkennen, dass bis auf Raum Kh die Volumina generell unterschätzt werden. Zudem führt der lauter e musikalische Stimulus generell zu geringeren Schätzungen. Da Werte nur für vier Räume vorliegen lässt sich hier lediglich ein Trendtest durchführen. Exemplarisch sind in Abbildung 5.48 physikalisches Volumen und per-

zeptives Volumen der Stimuluskombination optoakustisch/Sprache gegeneinander aufgetragen. Eine Kurvenanpassung ergab als passendste Funktion eine Exponentialfunktion mit den Koeffizienten:

$151{,}243 \cdot e^{(0{,}000722\ x)}$ (korrigiertes R^2 =,969; F =93,671 ; p =,011).

Da die Annahme naheliegend ist, dass der Zusammenhang linearer Natur ist, wurde auch eine lineare Funktion in die Analyse eingeschlossen. Besteht tatsächlich eine linearer Zusammenhang wird er durch eine lineare Funktion mit den folgenden Koeffizienten charakterisiert:

$0{,}694 \cdot x - 51{,}953$ (korrigiertes R^2 = 0,787; F = 12,105; p = ,074).

Das R^2 ist in diesem Fall verhältnismäßig gering, da bei Raum Kh das perzeptive Volumen größer ist als das Physikalische. Bei der optischen Stimuluskombination, erreicht die lineare Funktion ein bedeutend höheres R^2 (korrigiertes R^2 = 0,955; F = 65,261; p = ,015).

Die Steigung der linearen Funktion liegt bei alles Stimuluskombinationen zwischen 0,56 und 0,68. Einzig bei der Stimuluskombination akustisch/Musik weist die Steigung einen deutlich geringeren Wert von 0,288 auf (siehe Anhang 8.5, S. 90). In jedem Fall bestätigt die Tatsache, dass die Steigungen immer kleiner eins sind den anfänglichen Verdacht, das die physikalischen Volumina systematisch unterschätzt werden. Aufgrund der geringen Fallanzahl ist eine gesicherte Angabe darüber, ob ein lineares oder ein exponentioelles Modell die systematische Verschätzung der tatsächlichen Volumina besser beschreibt leider nicht möglich. Dafür müsste die Anzahl der untersuchten Räume deutlich höher sein, was forschungsökonomisch einen immensen Aufwand bedeuten würde.

In der Zusammenschau zeigt sich, dass sowohl das Merkmal Volumen (Skala) als auch das Merkmal Volumen (berechnet) stark mit dem physikalischen Volumen zusammenhängen. Dies gilt unabhängig von Präsentationsmodus und akustischem Darbietungsinhalt. Damit sind die Erhebungsarten gleichermaßen für die Messung des perzeptiven Volumens geeignet.

Da der Aufwand für das Merkmal Volumen (berechnet) durch die dreifache Dimensionsschätzung (Höhe, Breite, Länge) für die Versuchsperson sehr hoch ist, der Informationsgewinn hingegen relativ gering, sollte in zukünftigen Versuchen das Erheben des perzeptiven Volumens mit Hilfe einer einzigen äquidistanten Skala erfolgen.

5.4 Zusammenhang physikalischer und perzeptiver Merkmale

Dieses Kapitel befasst sich damit, Beziehungen zwischen den physikalischen Merkmalen der Räume und den perzeptiven Merkmalen aufzudecken. Als Daten für die perzeptiven Merkmale dienen die Mittelwerte der Stimuli mit optoakustischer Passung. Die Auswertung wurde für den musikalischen und den sprachlichen Darbietungsinhalt vorgenommen. Wenn keine grundlegenden Unterschiede zwischen den Inhalten festgestellt werden konnte, dann werden die Auswertungen des musikalischen Inhaltes präsentiert.

Da als Fälle nur vier Räume vorliegen kann keine inferenzstatistisch aussagekräftige Auswertung vorgenommen werden. Als Alternative werden Trendtests durchgeführt, um tendenzielle Zusammenhänge aufzudecken. Dafür wird ein physikalisches Merkmal eines Raumes einem perzeptiven zugeordnet und grafisch dargestellt. Auf der Y-Achse befindet sich jeweils das perzeptive Merkmal, auf der X-Achse das physikalische Merkmal. Die vier Punkte repräsentieren die Räume und sind entsprechend gekennzeichnet. Die Trendtests werden in erster Linie für Variablen durchgeführt, bei denen aus theoretischen Überlegungen heraus ein Zusammenhang zu erwarten ist. Dies dient zum Einen der Überprüfung, ob entsprechende physikalische Merkmale als Prädiktoren geeignet sind und zum Anderen, ob die jeweiligen Items die gewünschten perzeptiven Merkmale adäquat messen.

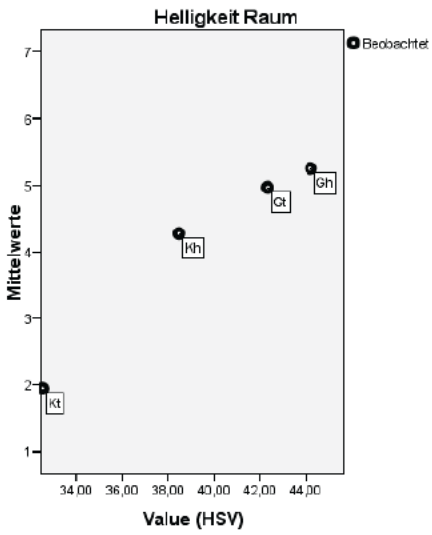


Abbildung 5.49

Helligkeit

Als physikalischer Prädiktor für die Helligkeit dient die Value-Komponente des HSV-Farbraums, die aus den Fotos der Räume extrahiert wurde. Abbildung 5.49 zeigt beispielhaft den Zusammenhang der Variable Helligkeit des Raumes mit dem Value-Wert. Die Helligkeit der Quelle zeigt denselben Zusammenhang. Ein Zusammenhang der perzeptiven Helligkeit und der physikalischen Prädiktorvariable kann für das Merkmal Helligkeit angenommen werden.

Farbintensität

Als physikalischer Prädiktor für die perzeptive Farbintensität wurde die Farbsättigungs-Komponente (*Saturation*) des HSV-Farbraums angenommen. Der Zusammenhang ist in Abbildung 5.50 grafisch dargestellt. Ein eindeutiger Zusammenhang ist nicht gegeben, da Raum Kt eine hohe Sättigung, jedoch eine nur geringe perzeptive Farbintensität aufweist.

Ein stärkere Zusammenhang scheint in Bezug auf die Streuung des Hue-Wertes zu bestehen (Abbildung 5.51), welches als optisches Maß für die Anzahl verschiedener Farbkomponenten eines Bildes dienen kann.

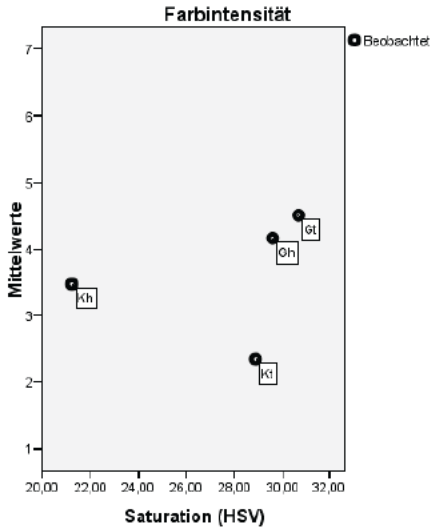


Abbildung 5.50

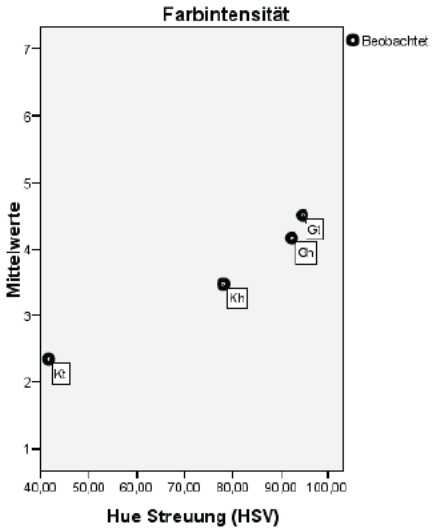


Abbildung 5.51

Kontrast

Als physikalisches Maß für den Kontrast wird das Maß der Streuung der Value-Komponente des HSV-Farbraums angenommen. Der Trendtest in Abbildung 5.52 lässt jedoch keinen eindeutigen Zusammenhang erkennen. Tendenziell scheint der Zusammenhang sogar negativ zu sein.

Ein besserer Zusammenhang zeigt sich, wie bei der Farbintensität, für das Maß der Streuung der Hue-Komponente.

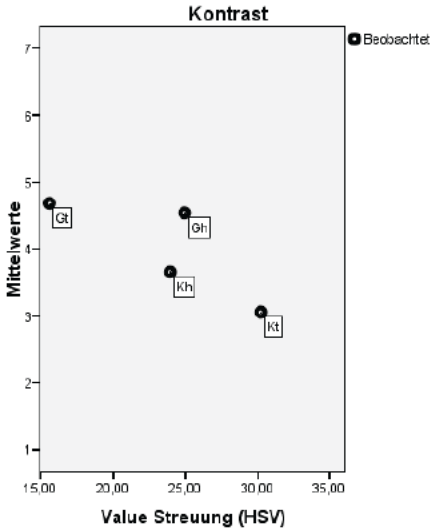


Abbildung 5.52

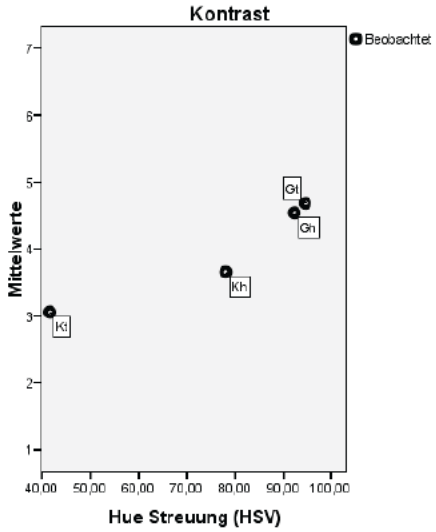


Abbildung 5.53

Lautheit

Der Zusammenhang zwischen den physikalischen Merkmalen der Räume und der perzeptiven Lautheit, unterscheidet sich für die verschiedenen Referenzobjekte. In Bezug auf den Gesamteindruck und die Klangquelle steht die perzeptive Lautheit in einem negativen Zusammenhang zur physikalischen Entfernung der Klangquelle (Abbildung 5.54). Ein positiver Zusammenhang, ist zum Stärkemaß (G) zu zu erkennen (Abbildung 5.55). Bei der Lautheit des Raumes, sind diese Zusammenhänge nicht zu finden.

Das Referenzobjekt Quelle zeigt einen weiteren Zusammenhang. In Abbildung 5.56 ist die Tendenz eines negativen Zusammenhanges zwischen der Lautheit der Quelle und der physikalischen Nachhallzeit zu erkennen. Diese Tendenz findet sich nicht bei den Referenzobjekten Raum und Gesamteindruck.

Der Zusammenhang zu Entfernung und Stärkemaß findet sich in schwachem Maße auch bei der Lautheit des Raumes.

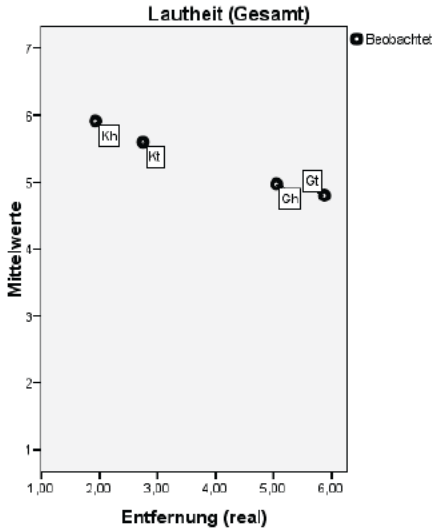


Abbildung 5.54

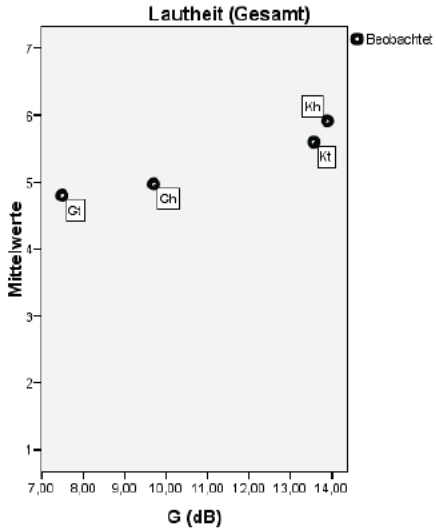


Abbildung 5.55

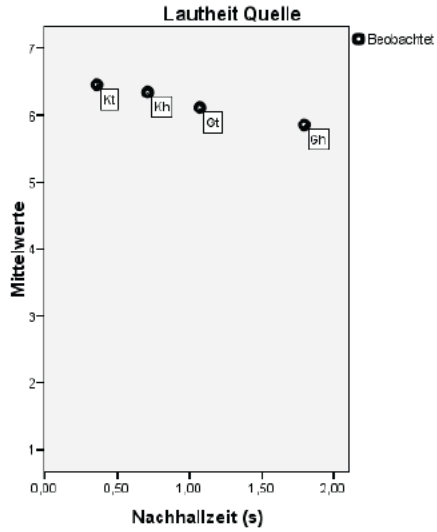


Abbildung 5.56

Wärme

Die Wärme eines Klangs ist bestimmt durch das Verhältnis zwischen hoch- und niederfrequenter Schallenergie. Je geringer der Anteil hoher Frequenzen im Verhältnis zu den tiefen Frequenzen, desto wärmer wird ein Klang empfunden. Als physikalischer Prädiktor wurde daher das Bassverhältnis angenommen. Der Trendtest zeigt hingegen keinen Zusammenhang der perzeptiven Wärme zum Bassverhältnis (Abbildung 5.57). Jedoch zeigen sich Zusammenhänge zum Stärkemaß (G) (Abbildung 5.58), und zur Entfernung der Quelle (Abbildung 5.59). Wie sie auch für die Lautheit festgestellt werden konnten. Dies gilt in gleichem Maße für die Wärme der Quelle und die Wärme des Raumes.

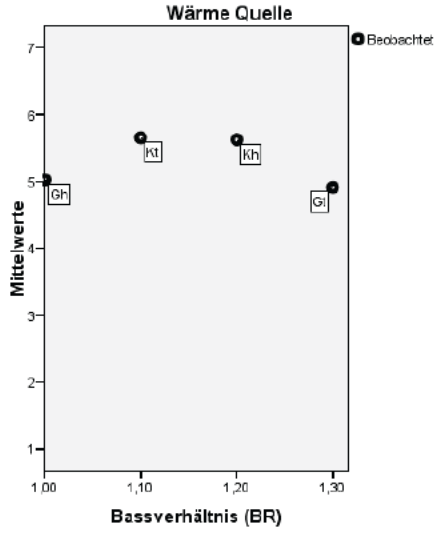


Abbildung 5.57

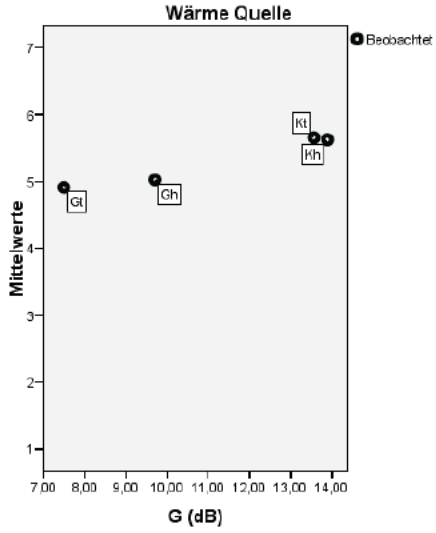


Abbildung 5.58

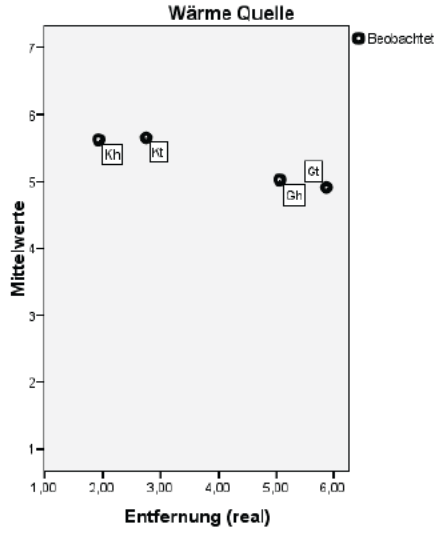


Abbildung 5.59

Brillanz

Für die Brillanz dient ebenfalls das Bassverhältnis als Prädiktor, jedoch wird ein negativer Zusammenhang angenommen. Der Trendtest in Abbildung 5.60 zeigt die Tendenz eines negativen Zusammenhangs, der jedoch nicht eindeutig ist. Die Brillanz des Raumes Kt wird geringer beurteilt, als das Bassverhältnis es vorhersagt. Für die Brillanz des Raumes ist der Zusammenhang zum Bassverhältnis nicht zu erkennen. Dafür zeigt sich ein positiver Zusammenhang zur Nachhallzeit (Abbildung 5.61), der sich bei der Brillanz der Quelle hingegen nicht findet.

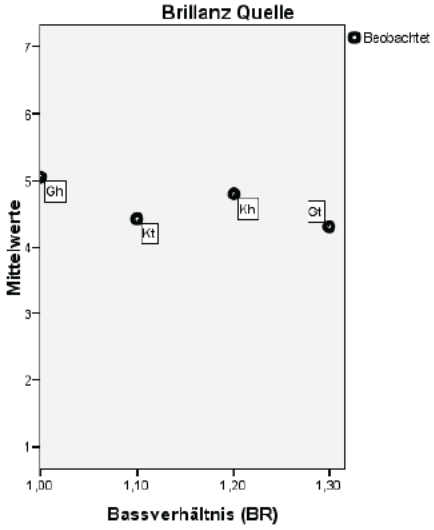


Abbildung 5.60

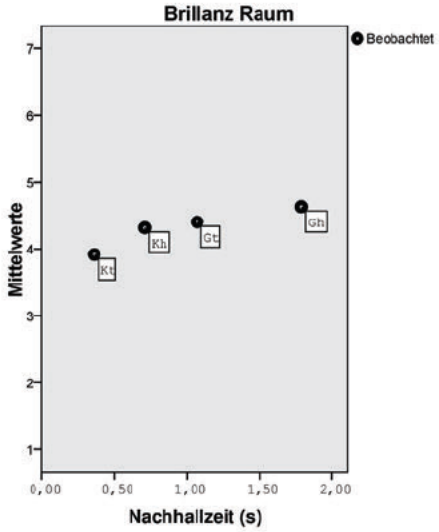


Abbildung 5.61

Zeittansparenz

Der Trendtest zeigt keine eindeutige Beziehung zum Klarheitsmaß C_{80}, welches als Prädiktor angenommen wurde. Abbildung 5.62 zeigt, dass die Räume Kh und Gt, die ein fast identisches Klarheitsmaß aufweisen, sehr unterschiedlich beurteilt werden. Statt dessen zeigt sich ein klarer negativer Zusammenhang zur Entfernung (Abbildung 5.63) und ein positiver zum Stärkemaß (G) (Abbildung 5.64).

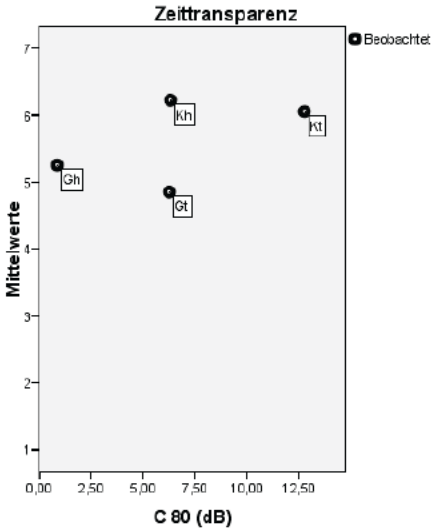


Abbildung 5.62

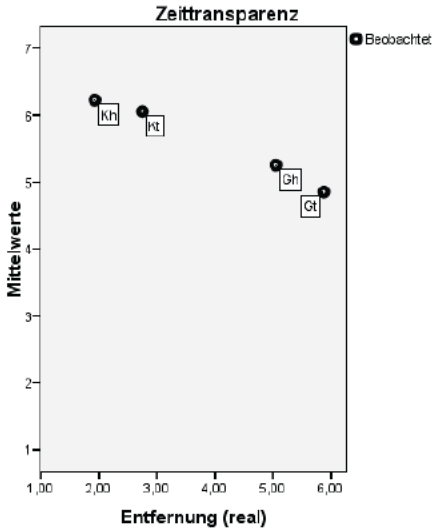


Abbildung 5.63

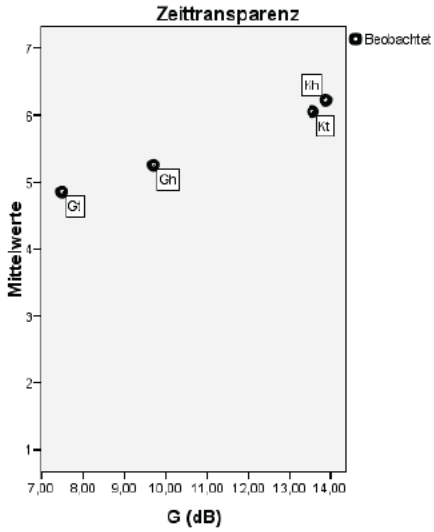


Abbildung 5.64

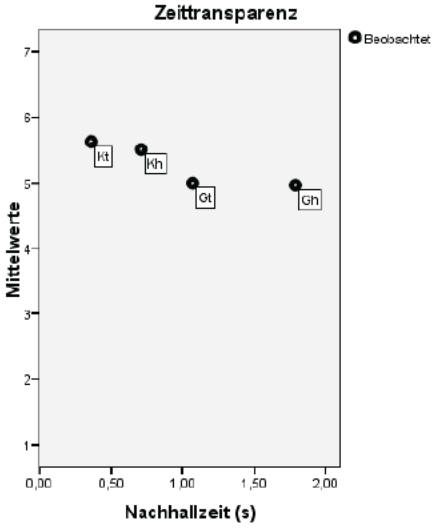


Abbildung 5.65

Da die Variation des akustischen Raumes in Kapitel 5.1.1 jedoch eine Verteilung der Mittelwerte aufweist, die sich von der Verteilung der Mittelwerte der perzeptiven Lautheit unterscheidet ist anzunehmen, dass der sprachliche Inhalt hier einen maßgeblichen Effekt ausübt. Aus diesem Grund wurden Trendtests auch für den sprachlichen Darbietungsinhalt durchgeführt.

Ein Zusammenhang zum Klarheitsmaß ist auch für den sprachlichen Darbietungsinhalt nicht zu erkennen. Der Zusammenhang zur Entfernung und zum Stärkemaß, der sich für den musikalischen Inhalt ergibt, zeigt sich für den sprachlichen Inhalt nicht. Dafür lässt Abbildung 5.65 einen negativen Zusammenhang zur Nachhallzeit erkennen.

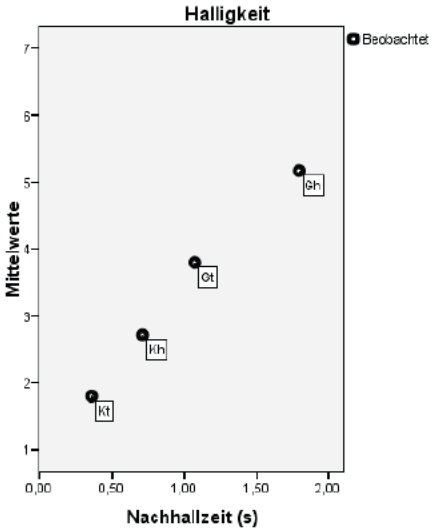


Abbildung 5.66

Halligkeit

Für das perzeptive Merkmal Halligkeit lässt sich erwartungsgemäß ein starker positiver Zusammenhang zur physikalischen Prädiktorvariable Nachhallzeit feststellen (Abbildung 5.66).

Umhüllung

Als Prädiktoren für die das Merkmal Umhüllung haben unter anderem Bradley und Soulodre [Soulodre1995] die Nachhallzeit und das Klarheitsmaß C_{80} identifiziert. Im Trendtest bestätigt sich diese Annahme. Der Grad der Umhüllung weist einen deutlichen positiven Zusammenhang zur Nachhallzeit (Abbildung 5.67) und einen negativen zum Klarheitsmaß auf.

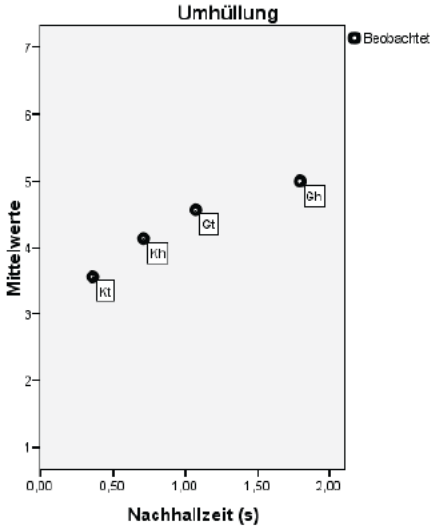


Abbildung 5.67

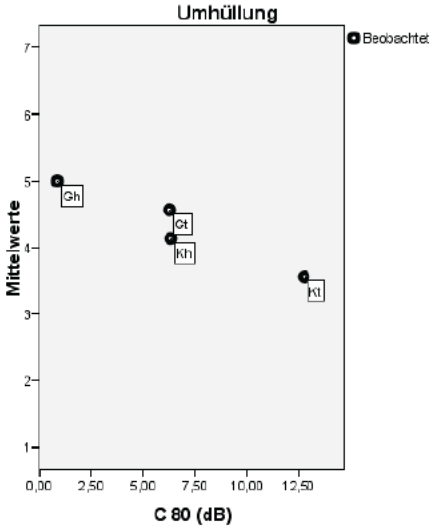


Abbildung 5.68

Entfernung zur Klangquelle

Abbildung 5.69 zeigt den Zusammenhang zwischen perzeptiver und physikalischer Entfernung zur Klangquelle. Ein eindeutiger Zusammenhang lässt sich nicht feststellen. Die Größenordnung der Schätzungen liegt jedoch jeweils nah bei den tatsächlichen Entfernungen. Besonders die Entfernungen in den Räumen Kh und Gt wurden sehr exakt geschätzt. Es fällt auf, dass die Entfernung in Raum Gh deutlich überschätzt wurde. Eine Erklärung hierfür findet sich in Abbildung 5.70, die den Zusammenhang der geschätzten Entfernungen zur Nachhallzeit zeigt. Die Abbildung legt die Vermutung nahe, dass bei größeren Entfernungen die Nachhallzeit an Einfluss

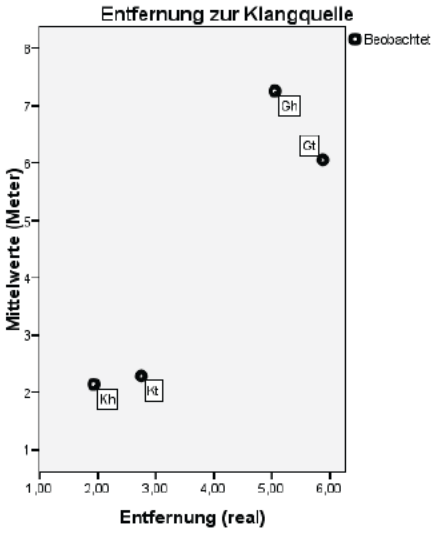


Abbildung 5.69

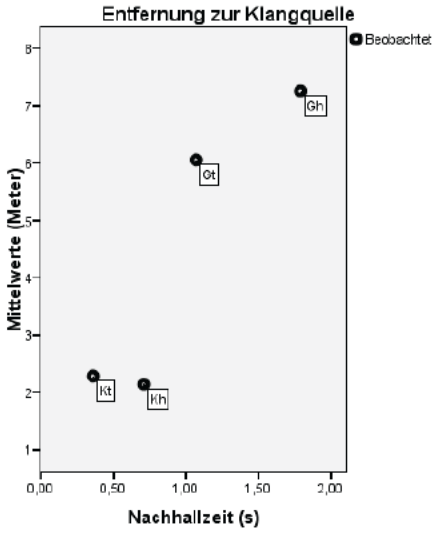


Abbildung 5.70

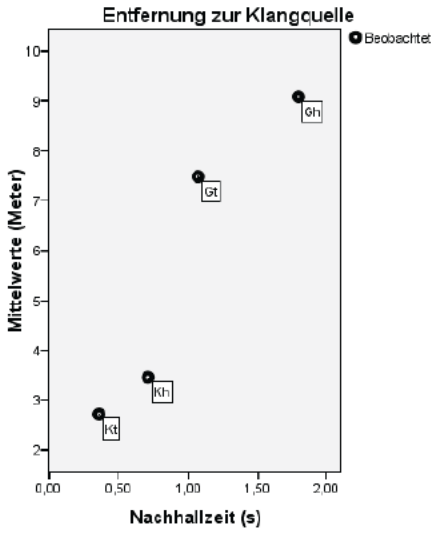


Abbildung 5.71

gewinnt. Um dieser Vermutung nachzugehen wurde ein Trendtest für die geschätzte Entfernung und die Nachhallzeit auch für den sprachlichen Stimulus durchgeführt. In Abbildung 5.71 zeigt sich, dass der Zusammenhang für den sprachlichen Stimulus noch deutlicher hervortritt, als für den musikalischen.

Volumen

In Kapitel 5.3 (S. 50)konnte nachgewiesen werden, dass das Merkmal Volumen (Skala) die Volumenschätzung adäquat beschreibt, weshalb sich die Auswertung auf diese Skala beschränkt. Der Trendtest in Abbildung 5.72 lässt einen tendenziellen Zusammenhang des perzeptiven Volumens zum realen Volumen erkennen, wobei das Volu-

men des Raumes Kh deutlich überschätzt wird. Wie bei dem Merkmal Entfernung zur Klangquelle deutet diese Tatsache auf einen Einfluss der Nachhallzeit hin. Diesen Zusammenhang bestätigt der entsprechende Trendtest in Abbildung 5.73.

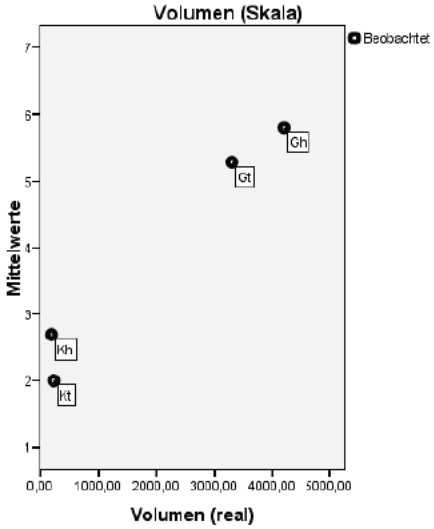


Abbildung 5.72

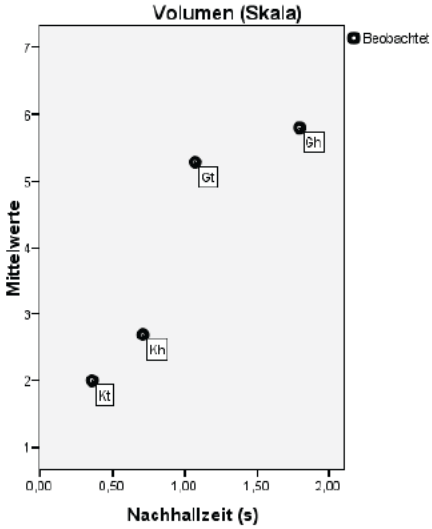


Abbildung 5.73

Potenz

Das Merkmal Potenz lässt sich die Tendenz eines negativen Zusammenhanges zur Entfernung feststellen (Abbildung 5.68). Die Tendenz eines positiven Zusammenhanges zum Stärkemaß G zeigt sich in Abbildung 5.53. diese Ergebnisse geben Grund zu der Annahme, dass die Lautheit einen maßgeblichen Einfluss auf die Beurteilung der Potenz ausübt.

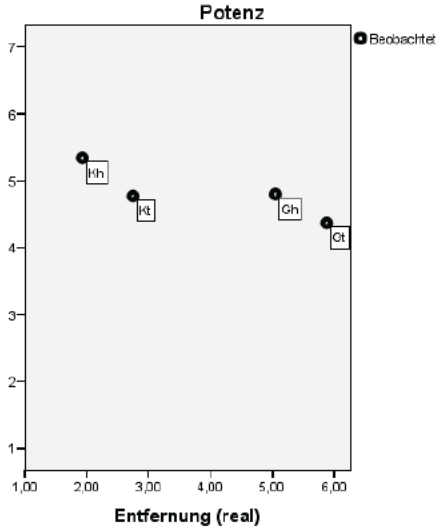


Abbildung 5.74

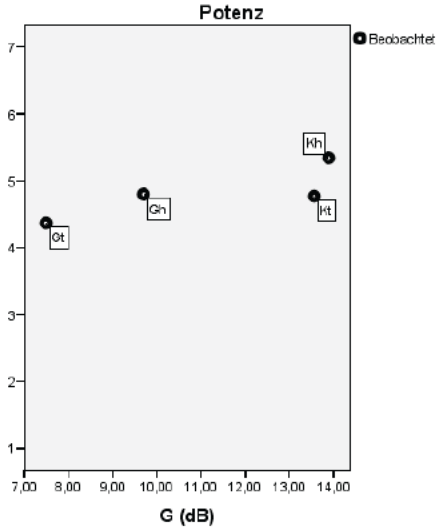


Abbildung 5.75

Die Trendtestes für die Merkmale Aktivierung, Anwesenheit, Gefallen und optoakustische Passung zeigten keine auffälligen Zusammenhänge zu den physikalischen Eigenschaften.

6 Diskussion

6.1 Unimodale visuelle Merkmale

Helligkeit

Der Trendtest für das perzeptive Merkmal Helligkeit lässt die Vermutung zu, dass die Versuchspersonen in der Lage sind, zwischen verschiedenen Graden der Helligkeit in komplexen Abbildungen, zu differenzieren. Zudem kann man darauf schließen, dass die Value-Komponente des HSV-Farbraums einer grafischen Abbildung als Prädiktor für die perzeptive Helligkeit geeignet ist.

Die Auswertung der Konfliktreize zeigte keinen Einfluss des akustischen Raumes auf die perzeptive Helligkeit. Ausschließlich die Variation des optischen Raumes zeigte einen Effekt. Dieser ist mit einer Effektstärke von 59 % (Raum) bzw. 57 % (Quelle) zudem sehr hoch. Dieses Ergebnis deutet darauf hin, dass die perzeptive Helligkeit ausschließlich durch die visuelle Modalität beeinflusst wird.

Die Referenzobjekte Raum und Quelle wiesen sowohl im Trendtest, als auch für die Konfliktreize keine bemerkenswerten Unterschiede in der Ausprägung der Mittelwerte auf. Für das vorliegende Versuchsdesign hätte auf diese Differenzierung verzichtet werden können. Allerdings kann nicht ausgeschlossen werden, dass bei Konzertsälen, die eine gesonderte Bühnenbeleuchtung aufweisen, nicht doch Unterschiede auftreten.

Farbintensität

Die Auswertung der Trendtests gibt einen Hinweis darauf, dass das Fragebogen-Item zur Farbintensität unter Umständen etwas anderes gemessen hat. Beim Vergleich der perzeptiven Farbintensität mit der Sättigung zeigt der Raum Kt eine starke Abweichung von den anderen Räumen. Dies kann damit zusammenhängen, dass dieser Raum im Vergleich zu den anderen eine besonders geringe physikalische Helligkeit aufweist. Dies führt zu der These, dass Personen dazu tendieren bei geringer physikalischer Helligkeit auch die Farbintensität gering zu beurteilen.

Dies erklärt jedoch nicht den deutlichen Zusammenhang der perzeptiven Farbintensität mit der Streuung der Hue-Komponente. Da dieser Wert als Maß für die Anzahl verschiedener Farbkomponenten eines Bildes gedeutet wird, ist davon auszugehen, dass die Versuchspersonen mit diesem Item eher die Farbigkeit, also das Vorhandensein verschiedener Farbwerte, als die Intensität der Farben beurteilt haben.

Die Auswertung der Konfliktreize ergab eine Effektstärke von 36 % für die Variation des optischen Raumes. Weder die Variation des akustischen Raumes, noch die Variation des Darbietungsinhaltes übt einen signifikanten Effekt auf die Beurteilung der Farbintensität aus. Diese Ergebnisse führen zu den Schluss, dass die Wahrnehmung der Farbintensität maßgeblich, im vorliegenden Versuchsdesign sogar ausschließlich, durch die visuelle Modalität beeinflusst wird.

Kontrast

Der perzeptive Kontrast zeigt im Trendtest einen tendenziell negativen Zusammenhang zum physikalischen Kontrast. Es kann demnach nicht ausgeschlossen werden, dass dieses Item entweder von den Versuchspersonen nicht korrekt verstanden wurde, oder es den Versuchspersonen nicht möglich war, den Kontrast der Abbildungen einzuschätzen.

Dieser Verdacht wird durch die Tatsache gestützt, dass die Art des Effektes auf das perzeptive Merkmal Kontrast fast identisch ist mit dem der Farbintensität. Dies gilt sowohl für den Trendtest in Bezug auf die Streuung der Hue-Werte, die mit dem Kontrast in keinem Bezug stehen sollten, als auch für die Variation des optischen Raumes. Da im Fragebogen diese beiden Items direkt aufeinander folgen, scheint die Vermutung berechtigt, dass sich die Versuchspersonen bei der Beurteilung des Kontrast an der vorangegangenen Variable orientiert haben. In diesem Zusammenhang sollte das Item für zukünftige Studien überdacht werden. Entweder muss eine alternative Operationalisierung des Merkmals im Fragebogen erfolgen, oder den Versuchspersonen muss eingehender erklärt werden, eventuell anhand von Beispielfotos, wie dieses Merkmal zu beurteilen ist.

Ein Einfluss der akustischen Modalität auf die Beurteilung des Kontrastes konnte nicht nachgewiesen werden.

Zusammenfassend kann festgehalten werden, dass bei keinem der unimodalen, visuellen Merkmale ein signifikanter Einfluss des akustischen Raumes oder des Darbietungsinhaltes nachgewiesen werden konnte. Die Merkmale scheinen im vorliegenden Versuchsdesign ausschließlich durch die visuelle Modalität beeinflusst zu sein.

6.2 Unimodale auditive Merkmale

Lautheit

Die perzeptive Lautheit in Bezug auf den Gesamteindruck und die Klangquelle, zeigen im Trendtest einen negativen Zusammenhang zur Entfernung und einen positiven Zusammenhang zum Stärkemaß. Bei zunehmender Entfernung zur Quelle nimmt der Schalldruckpegel des Direktschalls ab, weshalb eine abnehmende perzeptive Lautheit plausibel erscheint. Auch der positive Zusammenhang zum Stärkemaß, als Maß für die Lautstärke eines Raumes ist nachvollziehbar. Dies weist darauf hin, dass die Versuchspersonen mit dem Fragebogen-Item Lautstärke des Gesamteindrucks und Lautstärke der Quelle in der Lage sind, die an der Empfängerposition eintreffende Schallenergie zu beurteilen. Zudem kann die Funktion des Stärkemaßes als Prädiktor für die perzeptive Lautheit bestätigt werden.

Bei dem Merkmal Lautheit der Quelle, zeigte der Trendtest einen negativen Zusammenhang zur Nachhallzeit. Dieser lässt sich nicht durch den Schalldruck erklären, da die Länge des Nachhalls nicht in einem direkten Zusammenhang zum Schalldruck steht. Eine mögliche Erklärung ist, dass eine kognitive Komponente ins Spiel kommt, indem die Versuchspersonen bei zunehmender Nachhallzeit einen geringeren Schalldruck der Quelle annehmen. Trifft dies zu, wäre beim Trendtest zum Merkmal Lautheit des Raumes ein gegensätzlicher Trend für das physikalische Merkmal Nachhallzeit zu erwarten. Der entsprechende Trendtest konnte diesen Zusammenhang jedoch nicht bestätigen und die Annahme somit nicht stützen.

Die Auswertung der Konfliktreize zeigte einen starken Effekt durch die Variation des Darbietungsinhaltes. Der lautere, musikalische Stimulus wurde von den Probanden erwartungsgemäß als lauter beurteilt. Die größte Effektstärke zeigte sich beim Referenzobjekt Quelle mit 25 %, die geringste beim Referenzobjekt Raum mit 10 %. Diese Verteilung ist plausibel, da der Inhalt eher mit der Klangquelle, als mit dem

Raum verbunden ist. Dass das Referenzobjekt Raum in Bezug auf die Variation des Darbietungsinhaltes trotzdem noch eine Effektstärke von 10 % erreicht deutet darauf hin, dass die Probanden nicht in der Lage sind die Lautheit des Inhaltes gänzlich von der Lautheit des Raumes trennen zu können.

Die Erwartung, dass die Lautheit des Raumes stärker durch die Variation der akustischen Räume beeinflusst wird, kann hingegen nicht bestätigt werden. Die Effektstärke für das Referenzobjekt Quelle ist mit 21 % zwar geringer als bei der Variation des Darbietungsinhaltes aber noch immer mehr als doppelt so stark als für das Referenzobjekt Raum (10 %). Dies ist ein weiterer Hinweis darauf, dass die Probanden bei der Beurteilung der Lautheit der Quelle stark durch die Lautheit des Raumes beeinflusst werden, bzw. diese nicht getrennt voneinander beurteilen können.

Die Verteilung der Mittelwerte der Lautheit des Raumes in (Abbildung (5.34 (Lautheit Raum/akustischer Raum)) weicht von der anderen Referenzobjekte ab. Im Besonderen wird die Lautheit des Raumes Kt deutlich geringer und die Lautheit des Raumes Gh deutlich höher beurteilt. Es entsteht der Eindruck, dass der Absorptionsgrad der Räume in das Urteil einfließt. Dies führt zu der These, dass eine längere Nachhallzeit zu einer höheren Beurteilung der Lautheit eines Raumes beiträgt.

Der Gesamteindruck der Lautheit liegt in Bezug auf die Effektstärke und die Mittelwerte zwischen den Referenzobjekten Raum und Quelle. Die Vermutung liegt daher nahe, dass es sich bei diesem Merkmal um eine Kombination der beiden anderen Lautheitsmerkmale handelt. In zukünftigen Versuchen kann aus diesem Grund eventuell auf das Referenzobjekt Gesamteindruck verzichtet werden. Die Lautheit des Raumes und der Klangquelle hingegen zeigen genügend Unterschiede um die Verwendung beider Merkmale zu rechtfertigen.

Ein Einfluss des optischen Raumes auf die perzeptive Lautheit konnte für keines der Referenzobjekte nachgewiesen werden.

Wärme

Der Trendtest zeigt keinen Zusammenhang der perzeptiven Wärme zum physikalischen Prädiktor (Bass Ratio). Statt dessen konnte, in Analogie zur perzeptiven Lautheit, ein Zusammenhang zum Stärkemaß und zur Entfernung der Klangquelle festgestellt werden. Zwar kann nicht ausgeschlossen werden, dass die Lautheit einen Einfluss auf die wahrgenommene Klangfarbe ausübt und ebenfalls als Prädiktor für die

perzeptive Wärme angenommen werden kann, wahrscheinlicher ist jedoch, dass die Probanden nicht die Klangfarbe, sondern die Lautheit bewertet haben.

Dieser Verdacht erhärtet sich durch die Auswertung der Konfliktreize. Die Verteilung der Mittelwerte gleicht sehr stark den Mittelwerten der Lautheit. Es findet sich sogar dieselbe Verschiebung der Mittelwerte der Räume Kt und Gh zwischen den Referenzobjekten Raum und Quelle (Abbildung 5.38 und Abbildung 5.39). Auch die Effektstärkendifferenz zwischen den Referenzobjekten findet sich wieder.

Da die Probanden grundsätzlich in der Lage sein sollten verschiedene Klangfarben zu unterscheiden und die Operationalisierung der Wärme (voll – dünn) plausibel erscheint, muss diesem Phänomen eine andere Ursache zugrunde liegen. Eine mögliche Erklärung ist vermutlich im Versuchsdesign zu finden. Da der Fragebogen erst nach Ablauf der 16-sekündigen Stimuli ausgefüllt wurde, spielt auch der Faktor Gedächtnis eine Rolle. Eventuell werden komplexere Merkmale wie die Klangfarbe weniger gut von den Probanden memoriert als grundlegendere Merkmale, wie die Lautheit. Nach Ablauf einer gewissen Zeit dominiert dann ein Merkmal die Erinnerung. Um solche Effekte zu vermeiden wäre ein Versuchsdesign mit persistenten Stimuli erforderlich, bei dem das Ausfüllen des Fragebogens zeitgleich mit dem Abspielen der Stimuli erfolgen kann.

Ein Einfluss des optischen Raumes auf das Merkmal Wärme konnte nicht nachgewiesen werden.

Brillanz

Die Trendtests zur Brillanz der Quelle zeigen für die Referenzobjekte Quelle und Raum unterschiedliche Zusammenhänge. Eine Beziehung der Brillanz zum Bassverhältnis als physikalischem Prädiktor lässt sich nur für das Referenzobjekt Quelle annehmen. Der negative Zusammenhang deutet darauf hin, dass ein geringerer energetischer Anteil hochfrequenter Schallenergie geringere Brillanzbeurteilungen hervorruft. Die Versuchspersonen scheinen demnach in der Lage zu sein, mit Hilfe dieses Items die Brillanz zu beurteilen.

Für die Brillanz des Raumes kann davon nicht ausgegangen werden. Im Trendtest lässt sich kein Zusammenhang zum Bassverhältnis feststellen. Stattdessen zeigt sich ein positiver Zusammenhang zur Nachhallzeit. Die Beurteilung der Brillanz für das

Referenzobjekt Raum scheint den Versuchspersonen nicht zu gelingen und wird durch das Merkmal Nachhallzeit dominiert.

Die Auswertung der Konfliktreize ergab eine sehr geringe Testpower für die Variation des optischen und akustischen Raumes, weshalb eine Interpretation der Ergebnisse nicht möglich ist. Einzig die Variation des Darbietungsinhaltes erzielt einen signifikanten Effekt mit hoher Testpower, lässt jedoch keine weiteren Schlussfolgerungen zu.

Ein signifikanter Einfluss durch die Variation des optischen Raumes konnte nicht nachgewiesen werden.

Zeittransparenz

Die Auswertung der Trendtests für den musikalischen Inhalt führt zu der Annahme, dass die Versuchspersonen sich bei der Beurteilung der Zeittransparenz an der Lautheit orientiert haben. Dafür spricht die negative Beziehung zur Entfernung und die positive Beziehung zum Stärkemaß. Dieselben Zusammenhänge konnten auch für die perzeptive Lautheit festgestellt werden. Natürlich sind die Entfernung zu Klangquelle und die Lautheit ebenfalls Einflussfaktoren in Bezug auf die Zeittransparenz. Der fehlende Zusammenhang zum Klarheitsmaß C_{80}, welches als Prädiktor angenommen wurde, erhärtet jedoch den Verdacht, dass die Versuchspersonen sich bei der Beurteilung maßgeblich an der Lautheit der Stimuli orientiert haben.

Ein anderes Bild ergibt sich für den sprachlichen Darbietungsinhalt. Der einzig erkennbare Zusammenhang ergab sich für die Nachhallzeit. Eine längere Nachhallzeit führt bei sprachlichen Stimuli zu einer geringeren Beurteilung der Zeittransparenz. Einen Erklärungsansatz für diesen Effekt ist die Impulshaftigkeit der Stimuli. Valente et al. (2008) stellten eine unterschiedliche Beurteilung der Nachhallzeit für einen Harfen- und einen Trommelstimulus fest. Für einen legato gespielten Harfenstimulus erwarteten die Versuchspersonen eine längere Nachhallzeit, als für einen Trommelstimulus bei gleich bleibender Umgebung. Die Impulshaftigkeit des Trommelstimulus führte zu einem geringeren Maskierungseffekt, wodurch der Nachhall deutlicher wahrzunehmen ist. Ein ähnlicher Effekt kann auch für die verwendeten Musik- und Sprachstimuli in dieser Studie angenommen werden. Für den legato gespielten musikalischen Stimulus kann ein stärkerer Maskierungseffekt angenommen werden, als

für den Sprachstimulus. Dies führt dazu, dass die Nachhallzeit auf den Sprachstimulus einen Effekt ausübt, während der Musikstimulus davon unbeeinflusst bleibt.

Die Auswertung der Konfliktreize zeigte einen Effekt durch die Variation des Darbietungsinhaltes mit einer Effektgröße von 15 %. Dieser starke Effekt ist nicht nur auf die Reize selbst zurückzuführen, sondern auch auf die Fragebogen-Items. Diese mussten für den sprachlichen und den musikalischen Darbietungsinhalt verschieden formuliert werden. Die Zeittransparenz wurde generell sehr hoch beurteilt, wobei sie für den musikalischen Inhalt am höchsten ausfiel.

Bei der Variation des optischen und akustischen Raumes konnte ein signifikanter Effekt für den akustischen Raum mit einer Effektgröße von 8 % festgestellt werden. Es zeigte sich eine größere Zeittransparenz für die Räume mit kleinem Volumen.

Ein Einfluss der optischen Modalität auf die Zeittransparenz konnte nicht nachgewiesen werden.

Halligkeit

Durch die Auswertung der Trendtests konnte ein eindeutiger positiver Zusammenhang zwischen dem Merkmal Halligkeit und der Prädiktorvariable Nachhallzeit nachgewiesen werden. Die Versuchspersonen sind demnach in der Lage die verschiedenen Nachhallzeiten mit Hilfe des Fragebogen-Items perzeptiv zu beurteilen.

Die Variation des Darbietungsinhaltes zeigte, dass der sprachliche Inhalt im Vergleich zum musikalischen Inhalt zu einer erhöhten perzeptiven Halligkeit führt. Analog zur Zeittransparenz, lässt sich die These aufstellen, dass dieses Ergebnis mit einem geringeren Maskierungseffekt des sprachlichen Inhaltes zu begründen ist.

Die Variation des akustischen Raumes beeinflusste die perzeptive Halligkeit mit einer Effektstärke von 53 %. Damit ist die Halligkeit dasjenige Merkmal, dessen Beurteilung mit Abstand am stärksten durch den akustischen Raum beeinflusst wurde. Den nächst kleineren Effekt auf ein unimodales auditives Merkmal übte die Variation des akustischen Raumes auf das Merkmal Lautheit der Quelle mit 21 % aus. Die Halligkeit scheint das dominante Merkmal des akustischen Raumes zu sein.

Ein signifikanter Einfluss des optischen Raumes auf die Halligkeit konnte nicht nachgewiesen werden.

Umhüllung

Die Trendtests ergaben für die Umhüllung einen positiven Zusammenhang zur physikalischen Nachhallzeit und einen negativen Zusammenhang zum Klarheitsmaß C_{80}. Dies bestätigt die Ergebnisse von Bradley und Soulodre [Soulodre1995], welche beide Maße als Prädiktoren für die Umhüllung identifizierten.

Die Auswertung der Konfliktreize zeigte einen Effekt durch die Variation des akustischen Raumes mit einer Effektstärke von 14 %. Weder die Variation des Darbietungsinhaltes noch die des optischen Raumes übte einen signifikanten Effekt auf die Umhüllung aus.

Zusammenfassend kann festgehalten werden, dass bei keinem der unimodalen, auditiven Merkmale ein signifkanter Einfluss des optischen Raumes nachgewiesen werden konnte. Die Merkmale scheinen im vorliegenden Versuchsdesign ausschließlich durch die auditive Modalität beeinflusst zu sein.

6.3 Multimodale Merkmale

Entfernung zur Quelle

Der Trendtest zeigte keinen eindeutigen Zusammenhang zwischen geschätzter und physikalischer Entfernung. In den größeren Räumen Gt und Gh scheint die Nachhallzeit einen größeren Einfluss zu haben. Diese Tendenz ist für den sprachlichen Stimulus noch deutlicher zu erkennen als für den musikalischen. Dies führt zu der Annahme, dass die auditive Modalität einen starken Einfluss auf die Entfernungsschätzung ausübt.

Die Auswertung der Konfliktreize bestätigt diese Annahme. Die Variation des Darbietungsinhaltes führte dazu, dass die Klangquelle beim sprachlichen Stimulus deutlich entfernter wahrgenommen wird. Eine Erklärung für diesen Effekt ist, dass der sprachliche Stimulus einen geringeren Schallpegel aufweist und dieser einen wichtiges Schlüsselmerkmal für die Schätzung von Entfernungen ist [Ashmead1990, Bronkhorst1999, Zahorik2002].

Die Variation des optischen Raumes beeinflusste die Entfernungsschätzung mit einer Effektstärke von 12 %, während die Variation des akustischen Raumes eine Effektstärke von 21 % aufweist. Der Einfluss des akustischen Raumes ist damit fast doppelt so stark wie der des optischen Raumes. Die These, dass bei der audiovisuellen Lokalisation die visuelle Modalität die auditive Modalität dominiert [Mershon1980, Welch1980, Jack1973] kann für die Schätzung von Entfernungen nicht bestätigt werden.

Auch wenn die Relationen der geschätzten Entfernungen bei den optischen Räumen eher den physikalischen Gegebenheiten entsprechen, so kann trotzdem eine Dominanz der auditiven Modalität bei der Schätzung der Entfernung festgestellt werden.

Die Variation der optischen und akustischen Raumpräsenz erzielte eine Effektstärke von 3,7 %. Unter der akustischen Bedingung werden Entfernungen weiter geschätzt als unter der optischen Bedingung. Die geschätzte Entfernung liegt bei optoakustischer Bedingung zwischen den beiden unimodalen Bedingungen, was auf eine multimodale Integration der audiovisuellen Informationen schließen lässt. Ein Interaktionseffekt zwischen den Faktoren Präsentationsmodus und Raum (Abbildung 5.10) weist zudem darauf hin, dass der Einfluss des Präsentationsmodus sich für die vier Räume unterscheidet. Die Differenz der Entfernungsschätzungen zwischen der optischen und der akustischen Bedingung bei den Räumen mit hohem Absorptionskoeffizienten (Kt und Gt) unterscheiden sich nur geringfügig voneinander. Bei den Räumen mit geringem Absorptionskoeffizienten (Kh und Gh) hingegen ist die Differenz sehr ausgeprägt. Dies ist ein weiterer Hinweis für den starken Einfluss der auditiven Modalität.

Volumen

In Kapitel 5.3 konnte dargelegt werden, auf welche Weise die perzeptiven Merkmale Volumen (Skala) und Volumen (berechnet) miteinander in Beziehung stehen. Beide Skalen sind zur Erfassung des perzeptiven Volumens gleichermaßen geeignet und messen dasselbe Phänomen. Die Diskussion beschränkt sich daher auf das Merkmal Volumen (Skala).

Im Trendtest lässt sich ein tendenzieller Zusammenhang zum physikalischen Volumen erkennen. Dass das Volumen des Raumes Kt überschätzt wird, deutet jedoch auf einen weiteren Einflussfaktor bei der Beurteilung hin. Der Trendtest in Bezug auf die

Nachhallzeit legt die Vermutung nahe, dass auch sie einen maßgeblichen Einfluss ausübt.

Die Variation der optischen und akustischen Raumpräsenz zeigte keinen signifikanten Einfluss durch den Faktor Präsentationsmodus. Dieses Ergebnis erklärt sich durch den signifikanten Interaktionseffekt der Faktoren Präsentationsmodus und Raum Abbildung 4.4. Unter der optischen Bedingung gruppieren sich die Räume nach ihrem physikalischem Volumen. Die Räume mit großem Volumen Gt und Gh liegen nah bei einander, ebenso die Räume mit kleinem Volumen Kt und Kh. Unter der akustischen Bedingung hingegen ist keine Gruppierung zu erkennen, die Beurteilungen verteilen sich kontinuierlich. Für die Räume Gh und Kt ist kein signifikanter Effekt durch die Art der Präsentation zu erkennen. Bei ihnen führt der akustische und der optische Stimulusanteil zu den selben Beurteilungen. Die Beurteilungen Räume Gt und Kh hingegen unterscheiden sich stark zwischen der optischen und akustischen Bedingung. Dies ist vermutlich auf den relativ hohen Absorptionskoeffizienten des Raumes Gt zurückzuführen, welcher für seine Größe eine verhältnismäßig geringe Nachhallzeit aufweist. Bei Raum Kh verhält es sich entgegengesetzt. Die Verhältnismäßig lange Nachhallzeit führt zu einer deutlich höheren Beurteilung des Volumens unter der akustischen Bedingung. Diese Zusammenhänge lassen den Schluss zu, dass für eine entsprechende Raumgröße eine Erwartung der Versuchspersonen bezüglich der Nachhallzeit besteht. Zudem kann die Nachhallzeit als Prädiktor für die perzeptive Raumgröße bestätigt werden [Hameed2004].

Die Auswertung der Konfliktreize ergab eine Dominanz der optischen Modalität. Die Variation des optischen Raumes erklärt 31 % der Varianz, während die Variation des akustischen Raumes 16 % erklärt.

Audiovisuelle Passung

Die Auswertung der Konfliktreize für die optoakustische Passung ergab einen Interaktionseffekt zwischen optischem und akustischem Raum mit einer Effektstärke von 33 %. Dies war zu erwarten, da die Frage nur durch ein Zusammenspiel beider Faktoren beantwortet werden konnte. In der Auswertung ist zu erkennen (Abbildung 5.43), dass die optoakustisch passenden Stimuli jeweils am höchsten bewertet wurden. Damit kann die Beobachtung Sandvads [1999] bestätigt werden, dass Personen in der

Lage sind einen korrekten Zusammenhang zwischen optischem und akustischem Eindruck eines Raumes herzustellen.

Anwesenheit

Den größten Einfluss auf die Anwesenheit übte derselbe Interaktionseffekt aus, der für die optoakustische Passung beobachtet wurde. Auch die Effektstärke ist mit 31 % fast genauso groß. In diesem Zusammenhang kann das Merkmal Anwesenheit als Maß für die Güte der Simulation angenommen werden. Je größer die Diskrepanz zwischen visuellen und auditiven Informationen ist, desto weniger anwesend fühlen sich die Versuchspersonen.

Im Gegensatz dazu konnte für die Variation der optischen und akustischen Raumpräsenz kein Effekt festgestellt werden. Die Beurteilung der Anwesenheit wies keinen signifikanten Unterschied zwischen akustischer, optischer und optoakustischer Bedingung auf. Die Ergebnisse von Larrson et al. [2001], welche signifikant höhere Anwesenheitsbeurteilung unter der optoakustische Bedingung feststellten, kann im vorliegenden Versuchsdesign nicht bestätigt werden.

Gefallen

Die optoakustische Passung ist auch für das Merkmal Gefallen der stärkste Einflussfaktor. Die Auswertung der Konfliktreize ergab eine Effektstärke von 16 % für den Interaktionseffekt zwischen optischem und akustischem Raum. Mit einer Effektstärke von 3,8 % hat auch die Variation des Darbietungsinhaltes einen nicht unwesentlichen Effekt. Das Merkmal Gefallen scheint im vorliegenden Versuchsdesign im wesentlichen durch die Übereinstimmung der optischen und akustischen Stimuluskomponenten und die Art des Darbietungsinhaltes bestimmt zu sein.

Die Variation der optischen und akustischen Raumpräsenz lieferten durch eine zu geringe Testpower keine auswertbaren Daten. Die Beobachtung von Larsson et al. [2001], dass das Merkmal Gefallen unter der optoakustischen Bedingung höher gewertet wird kann daher nicht bestätigt werden. Ein Interaktionseffekt zwischen den Faktoren Präsentationsmodus und Raum mit einer Effektstärke von 4,5 % zeigt, dass die Beurteilungen sich für die verschiedenen Räume unter der akustischen und op-

toakustischen Bedingung kaum unterscheiden (Abbildung 5.9). Unter der optischen Bedingung hingegen gibt es klare Präferenzen unter den Räumen.

Potenz

Die Potenz wurde maßgeblich durch die Variation der optischen und akustischen Raumpräsenz beeinflusst. Die Effektstärke liegt bei 14 %. Die akustische und die optoakustische Bedingung weisen eine ähnliche Ausprägung auf. Unter der optischen Bedingung ist die Potenz hingegen viel geringer, was einen ersten Hinwies darauf gibt, dass die Potenz durch die auditive Modalität bestimmt wird.

Die Auswertung der Konfliktreize bestätigt diese Vermutung. Die Variation des akustischen Raumes erzielte eine Effektstärke von 7 %, die Variation des optischen Raumes hingegen nur 0,8 %. Die Potenz wird demnach maßgeblich durch die auditive Modalität geprägt.

Auch die optoakustische Passung übt einen Effekt auf die Potenz aus. Die Stärke des Interaktionseffektes beträgt 2,4 %. Der Einfluss ist zwar bedeutend geringer als bei den Merkmalen Anwesenheit und Gefallen, jedoch ist sein Auftreten beim Merkmal Potenz nicht plausibel zu erklären. Ein Erklärungsansatz für den Einfluss der optoakustischen Passung auf die Potenz liegt möglicherweise im Versuchsdesign selbst, genauer im Design des Fragebogens. Es besteht die Möglichkeit, dass die Beurteilung der optoakustischen Passung, allein durch die unmittelbare Nähe der Items im Fragebogen, die Beurteilung nachfolgender Items beeinflusst. Dieser Effekt ist vermutlich besonders groß, da die betreffenden Items sich am Ende des Fragebogens befinden. Zieht man in Betracht, dass die Beantwortung des Fragebogens im Durchschnitt 2 Minuten und 15 Sekunden dauerte, kann man davon ausgehen, dass der unmittelbare Eindruck der Stimuli bereits verblasste und die Versuchspersonen die Items aus dem Gedächtnis beantworteten. Ein weiteres Indiz für diesen Erklärungsansatz bietet die Tatsache, dass der Interaktionseffekt auch bei dem Merkmal Lautheit (Gesamteindruck) signifikant und mit einer Effektstärke von 1,8 % festgestellt werden kann. Das entsprechende Item befand sich ebenfalls im letzten Item-Block und der Effekt entzieht sich einer Erklärung.

Aktivierung

Im Trendtest lässt sich kein Bezug des Merkmals Aktivierung zu physikalischen Merkmalen erkennen. Den stärksten Effekt erzeugte die Variation der optischen und akustischen Raumpräsenz hatte zeigte mit einer Effektstärke von 6 % den stärksten Effekt auf das Merkmal Aktivierung. Wobei die optische Bedingung zu einer deutlich geringeren Beurteilung der Aktivierung führte als die akustische und optoakustische Bedingung.

7 Schlussfolgerung

Die vorliegende Studie verwendete eine audiovisuelle Umgebung in welcher die Versuchspersonen unimodale und multimodale Raummerkmale beurteilten. Ein besonders Anliegen bestand darin, eine möglichst reale, ökologisch valide, Simulation zu gewährleisten. Die Simulation erfolgte daher datenbasiert, durch Akquise akustischer und optischer Daten vor Ort.

Eine Vielzahl von Studien die sich mit multimodalen Interaktionen bei der Wahrnehmung von Räumen beschäftigen verwenden beschränkte Simulationstechniken. Simulierte Umgebungen werden dabei zum Teil stark vereinfacht reproduziert, etwa durch die Verwendung monoskopischer, nicht maßstabsgetreuer Schwarzweißaufnahmen bei der optischen Simulation, oder durch statische Binauralsynthese bei der akustischen Simulation. Zudem kann bei modellbasierten Raumsimulation die vollkommene Entsprechung akustischer und optischer Stimuluskomponenten nicht gewährleistet werden, was gesicherte Aussagen über den Einfluss der auditiven und visuellen Modalitäten auf perzeptive Raummerkmale nur eingeschränkt möglich macht. Im Gegensatz dazu gewährleistet das Versuchsdesign der vorliegenden Studie durch die datenbasierte Simulation eine größtmögliche Symmetrie von akustischen und optischen Raumeigenschaften.

Da die Erforschung audiovisueller Interaktionseffekte in der Raumwahrnehmung noch ein junges Forschungsgebiet ist, wurde ein explorativer Ansatz für die Studie gewählt.

Unter Anwendung eines Konfliktreizparadigmas, in welchem optische und akustische Stimulusanteile kategorial variiert wurden, konnte der Einfluss der visuellen und auditiven Modalitäten auf eine Vielzahl unimodaler und multimodaler Raummerkmale bestimmt werden. Die Kombination des Konfliktreizparadigmas mit einer ökologisch validen, datenbasierten Simulation wurde in dieser Form bisher nicht realisiert.

Für die unimodalen Merkmale konnten keine Interaktionseffekte nachgewiesen werden. Weder zeigten sich Einflüsse akustischer Stimuluskomponenten auf die visuelle Modalität, noch Einflüsse optischer Stimulusanteile auf die auditive Modalität. Da bei unimodalen Merkmalen wie der Lautheit oder der Helligkeit von der Dominanz einer Modalität ausgegangen werden kann, waren allenfalls subtile Effekte zu erwar-

ten. Es besteht Grund zu der Annahme, dass mit dem vorliegenden Versuchsdesign diese Effekte nicht nachzuweisen sind. Ein limitierender Faktor des Versuchsdesigns, welcher dafür verantwortlich sein könnte, besteht darin, dass die Stimuli für die Beantwortung des Fragebogens memoriert werden mussten. Die Reize hatten eine Dauer von jeweils 16 Sekunden. Die Beantwortung des Fragebogens erfolgte im Anschluss und dauerte durchschnittlich 2 Minuten und 15 Sekunden. Dies hat zur Folge, dass der unmittelbare Eindruck bei der Beantwortung der Fragebogen-Items nicht mehr vorhanden ist, was zur Minderung subtiler Effekte beitragen könnte. Um diese Effekte aufdecken zu können muss der Einfluss der Erinnerungsleistung minimiert oder sogar eliminiert werden. Ein Versuchsdesign mit persistenten Stimuli, bei dem die Präsentation der Stimuli und die Beantwortung des Fragebogens zeitgleich erfolgt, würde den Einfluss des Erinnerungsvermögens eliminieren und wäre möglicherweise in der Lage aussagekräftigere Ergebnisse zu liefern. Außerdem führt ein kürzerer Fragebogen, der sich in der Auswahl der Items auf wenige Merkmale beschränkt zu einer besseren Fokussierung der Versuchspersonen auf eine Aufgabenstellung und zur Minimierung des Einflusses des Erinnerungsvermögens.

Für die multimodalen Merkmale konnten Interaktionseffekte der auditiven und visuellen Modalität nachgewiesen werden. Interessanterweise zeigt sich ein starker Einfluss der auditiven Modalität bei denjenigen Merkmalen, die eine räumliche Dimensionsschätzung erforderten. Das Merkmal Entfernung zur Klangquelle wurde sogar dominant durch die auditive Modalität beeinflusst. Dieses Ergebnis widerspricht der These einer visuellen Dominanz bei der audiovisuellen Lokalisation und sollte in zukünftigen Studien genauer untersucht werden. Ursachen für die auditive Dominanz könnten im Versuchsdesign zu finden sein. Zum Einen ist die Qualität der optischen Simulation der Qualität der akustischen Simulation unterlegen. Die dynamische Binauralsynthese ermöglicht eine nahezu naturgetreue, 360° umfassende Simulation unter Berücksichtigung der Blickrichtung. Die optische Simulation hingegen beschränkt sich nur auf einen Teil des Sichtfeldes. Ein Einfluss dieser Differenz in der Güte der optischen und akustischen Simulation auf die Beurteilung ist nicht auszuschließen.

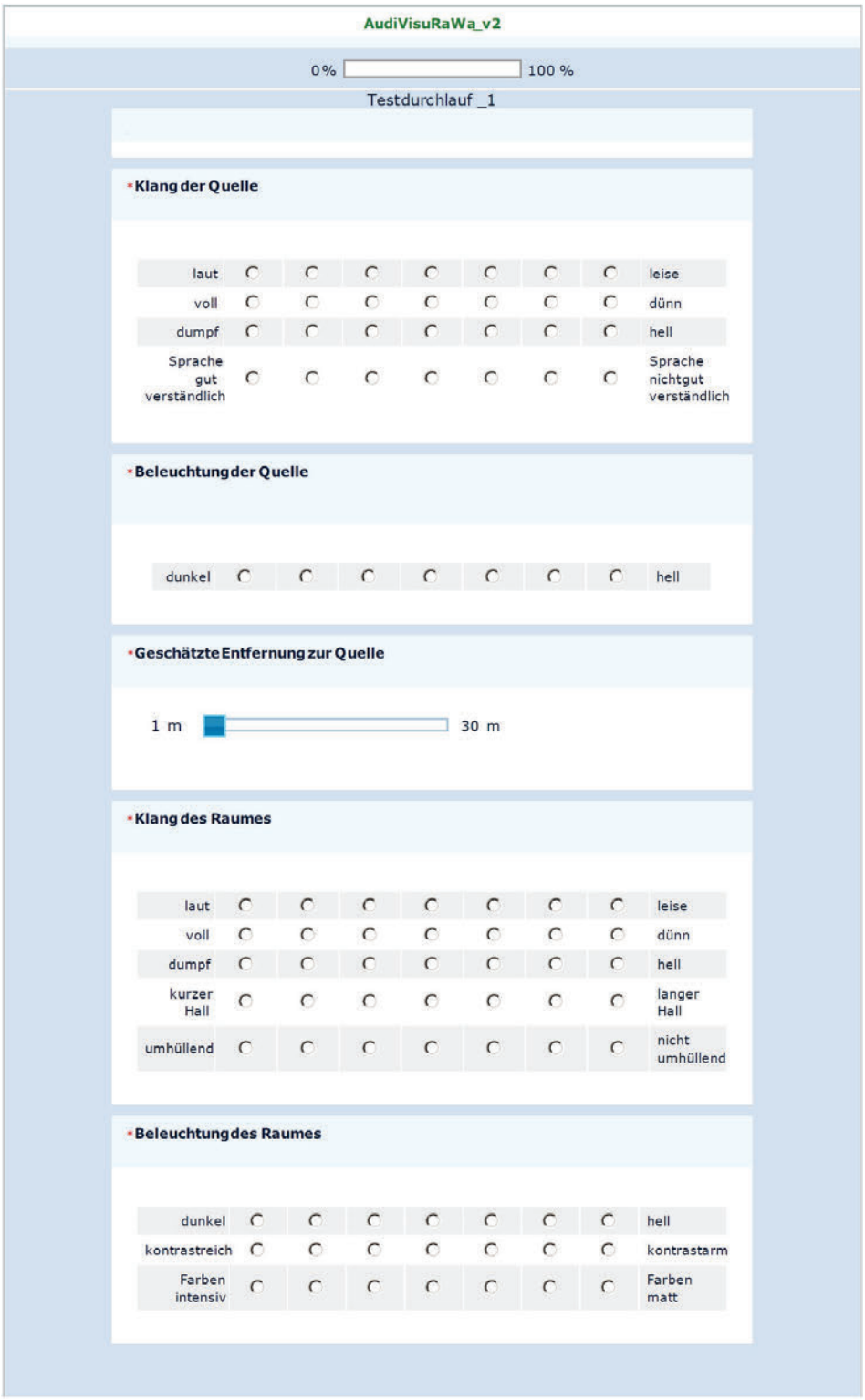
Zum Anderen war den Versuchspersonen durch die Frage nach der optoakustischen Passung bewusst, dass bei vielen Stimuli eine Differenz zwischen optischem und akustisch Eindruck bestand. Bei einigen Versuchspersonen kam die Frage auf, ob sie die Entfernung nach dem was sie sahen oder nach dem was sie hörten beurteilen

sollten. Dies gibt Anlass zu der Vermutung, dass die auditiven und visuellen Informationen bei zu großer Divergenz nicht als fusioniertes Perzept wahrgenommen wurden und die Versuchspersonen sich bei der Beurteilung bewusst auf eine Modalität beschränkten. Um diesen Effekt in zukünftigen Studien zu minimieren, sollte auf die Frage nach der optoakustischen Passung gegebenenfalls verzichtet werden. Zudem konnte in der vorliegenden Arbeit bestätigt werden, dass die Versuchspersonen in der Lage waren die optoakustische Passung der Stimuluskomponenten korrekt zu beurteilen.

Die Studie konnte zeigen, dass auditive und visuelle Informationen die Wahrnehmung von Räumen beeinflussen. Ein wichtiger Aspekt bei der Planung zukünftiger Studien ist die Güte der Simulation. Um eine Aussagekraft der Ergebnisse in Bezug auf die Wahrnehmung realer Umgebungen zu gewährleisten, müssen optische und akustische Simulationen eine größtmögliche Naturtreue aufweisen. Im akustischen Bereich bedeutet das beispielsweise die Anwendung individualisierter, dynamischer Binauralsynthese. Für optische Simulationen, sollte eine stereoskopische Präsentation zum Einsatz kommen, die das Gesichtsfeld idealerweise vollständig ausfüllt.

8 Anhang

8.1 Fragebogen



AudiVisuRaWa_v2

0 % 100 %

Testdurchlauf _1

*Klang der Quelle

laut	○	○	○	○	○	○	○	leise
voll	○	○	○	○	○	○	○	dünn
dumpf	○	○	○	○	○	○	○	hell
Sprache gut verständlich	○	○	○	○	○	○	○	Sprache nichtgut verständlich

*Beleuchtung der Quelle

dunkel	○	○	○	○	○	○	○	hell

*Geschätzte Entfernung zur Quelle

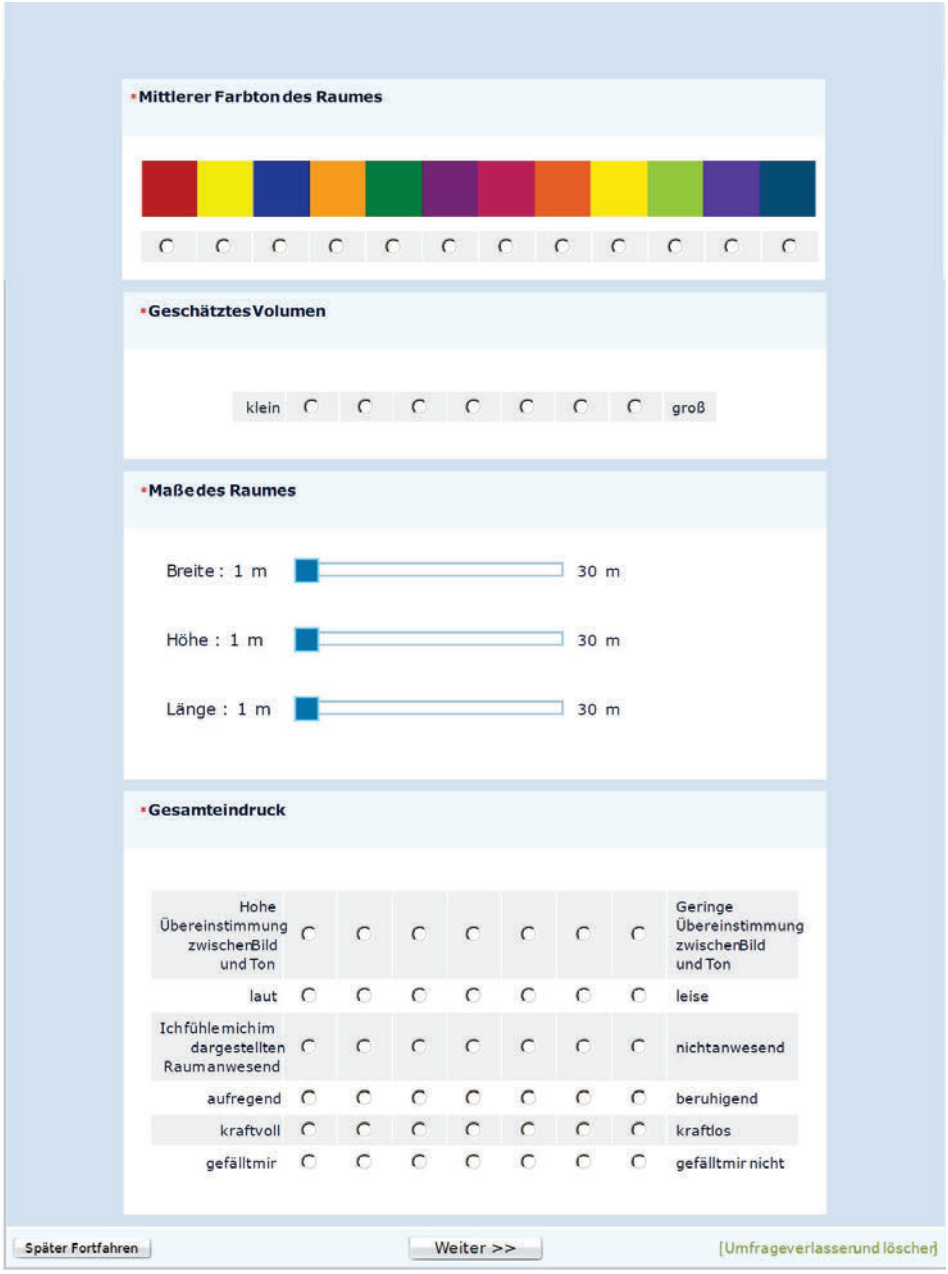
1 m 30 m

*Klang des Raumes

laut	○	○	○	○	○	○	○	leise
voll	○	○	○	○	○	○	○	dünn
dumpf	○	○	○	○	○	○	○	hell
kurzer Hall	○	○	○	○	○	○	○	langer Hall
umhüllend	○	○	○	○	○	○	○	nicht umhüllend

*Beleuchtung des Raumes

dunkel	○	○	○	○	○	○	○	hell
kontrastreich	○	○	○	○	○	○	○	kontrastarm
Farben intensiv	○	○	○	○	○	○	○	Farben matt



*Mittlerer Farbton des Raumes

*Geschätztes Volumen

klein | groß

*Maße des Raumes

Breite : 1 m — 30 m

Höhe : 1 m — 30 m

Länge : 1 m — 30 m

*Gesamteindruck

Hohe Übereinstimmung zwischen Bild und Ton								Geringe Übereinstimmung zwischen Bild und Ton
laut								leise
Ich fühle mich im dargestellten Raum anwesend								nicht anwesend
aufregend								beruhigend
kraftvoll								kraftlos
gefällt mir								gefällt mir nicht

Später Fortfahren | Weiter >> | [Umfrage verlassen und löschen]

8.2 Operationalisierung und Kodierung der Merkmale

Merkmal	Item		
Codierung	**1**	**-**	**7**
	unimodal optisch		
Farbintensität	Farben matt	-	Farben intensiv
Helligkeit Quelle	dunkel	-	hell
Helligkeit Raum	dunkel	-	hell
Kontrast	kontrastarm	-	kontrastreich
	unimodal akustisch		
Brillianz Quelle	dumpf	-	hell
Brillianz Raum	dumpf	-	hell
Halligkeit	kurzer Hall	-	langer Hall
Lautheit Gesamt	leise	-	laut
Lautheit Quelle	leise	-	laut
Lautheit Raum	leise	-	laut
Umhüllung	nicht umhüllend	-	umhüllend
Wärme Quelle	dünn	-	voll
Wärme Raum	dünn	-	voll
Zeittransparenz	Sprache nicht gut verständlich / Musik nicht gut durchhörbar	-	Sprache gut verständlich / Mus k gut durchhörbar
	multimodal		
Aktivierung	beruhigend	-	aufregend
Anwesenheit	Ich fühle mich im dargestellten Raum nicht anwesend	-	Ich fühle mich im dargestellten Raum anwesend
Gefallen	gefällt mir nicht	-	gefällt mir
Passung	Geringe Übereinstimmung zwischen Bild und Ton	-	Hohe Übereinstimmung zwischen Bild und Ton
Potenz	kraftlos	-	kraftvoll
Volumen (Skala)	klein	-	groß
Volumen (berechnet)	Breite x Höhe x Länge		
Breite	1 m	-	30 m
Höhe	1 m	-	30 m
Länge	1 m	-	30 m
Entfernung zur Quelle	1 m	-	30 m

8.3 Physikalische Kennwerte der Räume

Kennwerte	Raum Kh	Raum Kt	Raum Gt	Raum Gh
	Struktur			
Volumen ca. [m³]	190	230	3300	4200
Form	Rechteck	Rechteck	Diamant	Rechteck
max. Länge [m]	7	9,5	26,4	25
max. Breite [m]	8,5	7,7	20,8	18,1
max. Höhe [m]	3,2	3,1	8,8	9,6
Entfernung zur Klangquelle [m]	1,93	2,75	5,88	5,05
	Akustik			
Nachhallzeit T_{30} [s] (1000 Hz)	0.71	0,36	1,07	1,79
Mittl. Absorpionskoeffizient $\alpha_{ø}$	0,17	0,36	0,28	0,17
Early Decay Time EDT [s](1000 Hz)	0,68	0,44	0,95	1,73
Bass Ratio BR (1000 Hz)	1,2	1,1	1,3	1
Schwerpunktzeit t_s [ms] (1000 Hz)	46,01	25,27	47,74	103,57
Klarheitsmaß C_{80} [dB] (1000 Hz)	6,31	12,78	6,25	0,84
H [dB] (Liveness) (1000 Hz)	-2,9	-7,03	-4,58	-0,5
Stärkemaß G [dB] (1000Hz)	13,89	13,56	7,49	9,7
	Optik			
Hue Mittelwert [360°] (Zirkulär)	108	124	37	173
Hue Streuung [360°]	78	42	95	92
Saturation Mittelwert [0-100]	21	29	31	30
Saturation Streuung [0-100]	15	16	23	17
Value Mittelwert [0-100]	38	33	42	44
Value Streuung [0-100]	24	30	16	25
R Mittelwert [0-255]	95	81	97	103
R Streuung [0-255]	60	75	42	61
G Mittelwert [0-255]	93	81	91	99
G Streuung [0-255]	60	78	38	62
B Mittelwert [0-255]	85	68	87	95
B Streuung [0-255]	64	70	46	66

8.4 Statistische Kennzahlen

8.4.1 Variation der optischen und akustischen Raumpräsenz

Entfernung zur Quelle

Faktor	df korrigiert (Greenhouse-Geisser)	Quadratsumme vom Typ III	df	Mittel der Quadrate	F	Sig.	part. Eta²	Eta²	Test-power[a]
Präsenta ionsmodus	Nein	355,8	2	177,904	23,122	,000	,405	,037	1,000
Raum	Ja	4602,3	2	3004,491	133,887	,000	,797	,476	1,000
Raum * Präsentationsmodus	Ja	308,3	4	78,938	10,139	,000	,230	,032	1,000

Volumen (Skala)

Faktor	df korrigiert (Greenhouse-Geisser)	Quadratsumme vom Typ III	df	Mittel der Quadrate	F	Sig.	part. Eta²	Eta²	Test-power[a]
Raum	Ja	1914,8	2	808,627	243,537	,000	,877	,687	1,000
Raum * Präsentationsmodus	Ja	56,6	4	12,921	8,663	,000	,203	,020	,999

Volumen (berechnet)

Faktor	df korrigiert (Greenhouse-Geisser)	Quadratsumme vom Typ III	df	Mittel der Quadrate	F	Sig.	part. Eta²	Eta²	Test-power[a]
Raum	Ja	513311257,0	1	457246072,068	27,896	,000	,527	,311	1,000

Anwesenheit

Faktor	df korrigiert (Greenhouse-Geisser)	Quadratsumme vom Typ III	df	Mittel der Quadrate	F	Sig.	part. Eta²	Eta²	Test-power[a]
Raum	Nein	24,6	3	8,194	3,621	,016	,096	,014	,781

Aktivierung

Faktor	df korrigiert (Greenhouse-Geisser)	Quadratsumme vom Typ III	df	Mittel der Quadrate	F	Sig.	part. Eta²	Eta²	Test-power[a]
Präsenta ionsmodus	Ja	76,5	1	53,625	11,888	,000	,259	,058	,971
Raum	Ja	22,0	2	9,488	3,700	,024	,098	,017	,708
Raum * Präsentationsmodus	Nein	39,1	6	6,522	4,269	,000	,112	,030	,979

Potenz

Faktor	df korrigiert (Greenhouse-Geisser)	Quadratsumme vom Typ III	df	Mittel der Quadrate	F	Sig.	part. Eta²	Eta²	Test-power[a]
Präsenta ionsmodus	Ja	240,5	2	152,795	33,869	,000	,499	,136	1,000
Raum	Ja	24,4	2	10,645	3,273	,037	,088	,014	,647
Raum * Präsentationsmodus	Ja	96,9	4	23,951	8,761	,000	,205	,055	,999

Gefallen

Faktor	df korrigiert (Greenhouse-Geisser)	Quadratsumme vom Typ III	df	Mittel der Quadrate	F	Sig.	part. Eta²	Eta²	Test-power[a]
Präsenta ionsmodus	Nein	24,6	2	12,319	3,866	,026	,102	,015	,681
Raum	Nein	60,8	3	20,283	7,889	,000	,188	,037	,988
Raum * Präsentationsmodus	Ja	73,9	4	18,952	5,397	,001	,137	,045	,968

a. Unter Verwendung von Alpha = ,05 berechnet

8.4.2 Konfliktreize – Variation der optischen und akustischen Raumeigenschaften

Entfernung zur Quelle

Faktor	df korrigiert (Greenhouse-Geisser)	Quadratsumme vom Typ III	df	Mittel der Quadrate	F	Sig.	part. Eta²	Eta²	Test-power[a]
optischer Raum	Ja	1275,575	2,175	586,361	37,570	,000	,525	,115	1,000
akustischer Raum	Ja	2318,011	1,426	1626,016	44,935	,000	,569	,209	1,000
Inhalt	Nein	354,375	1,000	354,375	21,103	,000	,383	,032	,994
akustischer Raum * Inhalt	Ja	94,611	2,138	44,250	6,534	,002	,161	,009	,911

Lautheit Quelle

Faktor	df korrigiert (Greenhouse-Geisser)	Quadratsumme vom Typ III	df	Mittel der Quadrate	F	Sig.	part. Eta²	Eta²	Test-power[a]
akustischer Raum	Ja	375,217	2,216	169,312	69,122	,000	,670	,206	1,000
Inhalt	Nein	461,572	1,000	461,572	154,773	,000	,820	,253	1,000
akustischer Raum * Inhalt	Nein	17,017	3,000	5,672	5,503	,002	,139	,009	,932

Lautheit Raum

Faktor	df korrigiert (Greenhouse-Geisser)	Quadratsumme vom Typ III	df	Mittel der Quadrate	F	Sig.	part. Eta²	Eta²	Test-power[a]
akustischer Raum	Ja	204,353	1,600	127,750	13,136	,000	,279	,099	,989
Inhalt	Nein	208,294	1,000	208,294	33,938	,000	,500	,101	1,000
optischer Raum * akustischer Raum	Nein	17,429	9,000	1,937	2,082	,031	,058	,008	,868
akustischer Raum * Inhalt	Nein	19,810	3,000	6,603	5,217	,002	,133	,010	,918

Lautheit Gesamt

Faktor	df korrigiert (Greenhouse-Geisser)	Quadratsumme vom Typ III	df	Mittel der Quadrate	F	Sig.	part. Eta²	Eta²	Test-power[a]
akustischer Raum	Ja	233,553	2,211	105,612	30,376	,000	,472	,142	1,000
Inhalt	Nein	279,001	1,000	279,001	88,837	,000	,723	,170	1,000
optischer Raum * akustischer Raum	Nein	28,894	9,000	3,210	3,925	,000	,104	,018	,995
akustischer Raum * Inhalt	Nein	8,581	3,000	2,860	4,649	,004	,120	,005	,882

Gefallen

Faktor	df korrigiert (Greenhouse-Geisser)	Quadratsumme vom Typ III	df	Mittel der Quadrate	F	Sig.	part. Eta²	Eta²	Test-power[a]
optischer Raum	Nein	36,153	3,000	12,051	6,491	,000	,160	,013	,966
akustischer Raum	Ja	36,088	2,326	15,514	4,605	,009	,119	,013	,808
Inhalt	Nein	102,608	1,000	102,608	18,758	,000	,356	,038	,988
optischer Raum * akustischer Raum	Ja	422,994	5,546	76,274	21,398	,000	,386	,158	1,000

Anwesenheit

Faktor	df korrigiert (Greenhouse-Geisser)	Quadratsumme vom Typ III	df	Mittel der Quadrate	F	Sig.	part. Eta²	Eta²	Test-power[a]
optischer Raum	Nein	42,053	3,000	14,018	6,713	,000	,165	,011	,971
akustischer Raum	Ja	75,153	2,362	31,814	7,769	,000	,186	,020	,966
Inhalt	Nein	37,522	1,000	37,522	9,454	,004	,218	,010	,848
optischer Raum * akustischer Raum	Ja	1184,929	4,720	251,056	50,970	,000	,600	,314	1,000
akustischer Raum * Inhalt	Nein	19,724	3,000	6,575	3,280	,024	,088	,005	,736

Halligkeit

Faktor	df korrigiert (Greenhouse-Geisser)	Quadratsumme vom Typ III	df	Mittel der Quadrate	F	Sig.	part. Eta²	Eta²	Test-power[a]
akustischer Raum	Ja	1900,760	1,759	1080,577	171,696	,000	,835	,526	1,000
Inhalt	Nein	124,222	1,000	124,222	51,985	,000	,605	,034	1,000
akustischer Raum * Inhalt	Nein	28,138	3,000	9,379	5,129	,002	,131	,008	,913
optischer Raum * akustischer Raum * Inhalt	Nein	16,422	9,000	1,825	1,961	,044	,055	,005	,843

Optoakustische Passung

Faktor	df korrigiert (Greenhouse-Geisser)	Quadratsumme vom Typ III	df	Mittel der Quadrate	F	Sig.	part. Eta²	Eta²	Test-power[a]
optischer Raum	Nein	39,050	3,000	13,017	5,348	,002	,136	,009	,925
akustischer Raum	Nein	81,964	3,000	27,321	7,659	,000	,184	,019	,985
Inhalt	Nein	37,157	1,000	37,157	12,277	,001	,265	,009	,926
optischer Raum * akustischer Raum	Ja	1445,900	4,824	299,750	57,765	,000	,629	,334	1,000
akustischer Raum * Inhalt	Nein	33,736	3,000	11,245	6,269	,001	,156	,008	,960
optischer Raum * akustischer Raum * Inhalt	Nein	64,171	9,000	7,130	3,503	,000	,093	,015	,988

Potenz

Faktor	df korrigiert (Greenhouse-Geisser)	Quadratsumme vom Typ III	df	Mittel der Quadrate	F	Sig.	part. Eta²	Eta²	Test-power[a]
optischer Raum	Nein	16,574	3,000	5,525	4,805	,004	,124	,008	,893
akustischer Raum	Ja	148,810	2,087	71,310	15,483	,000	,313	,068	,999
Inhalt	Nein	152,294	1,000	152,294	19,984	,000	,370	,070	,991
optischer Raum * akustischer Raum	Nein	52,044	9,000	5,783	4,049	,000	,106	,024	,996

Umhüllung

Faktor	df korrigiert (Greenhouse-Geisser)	Quadratsumme vom Typ III	df	Mittel der Quadrate	F	Sig.	part. Eta²	Eta²	Test-power[a]
akustischer Raum	Ja	396,407	1,793	221,118	19,854	,000	,369	,140	1,000

Volumen Skala

Faktor	df korrigiert (Greenhouse-Geisser)	Quadratsumme vom Typ III	df	Mittel der Quadrate	F	Sig.	part. Eta²	Eta²	Test-power[a]
optischer Raum	Ja	1028,803	1,622	634,092	66,989	,000	,663	,307	1,000
akustischer Raum	Ja	529,288	1,462	362,014	48,778	,000	,589	,158	1,000
Inhalt	Nein	24,308	1,000	24,308	14,044	,001	,292	,007	,954
optischer Raum * akustischer Raum	Nein	22,301	9,000	2,478	2,399	,012	,066	,007	,919

Volumen berechnet

Faktor	df korrigiert (Greenhouse-Geisser)	Quadratsumme vom Typ III	df	Mittel der Quadrate	F	Sig.	part. Eta²	Eta²	Test-power[a]
optischer Raum	Ja	565608437,031	1,142	495121702,294	16,736	,000	,330	,200	,987
akustischer Raum	Ja	73355213,103	1,951	37602614,280	13,022	,000	,277	,026	,996
Inhalt	Ja	16069253,594	1,000	16069253,594	14,711	,001	,302	,006	,961
optischer Raum * Inhalt	Ja	15195972,417	1,789	8492648,071	5,508	,008	,139	,005	,804

Aktivierung

Faktor	df korrigiert (Greenhouse-Geisser)	Quadratsumme vom Typ III	df	Mittel der Quadrate	F	Sig.	part. Eta²	Eta²	Test-power[a]
akustischer Raum	Ja	40,839	2,212	18,464	5,330	,005	,136	,027	,852
Inhalt	Nein	26,414	1,000	26,414	6,630	,015	,163	,018	,706

Kontrast

Faktor	df korrigiert (Greenhouse-Geisser)	Quadratsumme vom Typ III	df	Mittel der Quadrate	F	Sig.	part. Eta²	Eta²	Test-power[a]
optischer Raum	Ja	518,346	2,096	247,327	21,286	,000	,385	,047	1,000

Farbintensität

Faktor	df korrigiert (Greenhouse-Geisser)	Quadratsumme vom Typ III	df	Mittel der Quadrate	F	Sig.	part. Eta²	Eta²	Test-power[a]
optischer Raum	Ja	872,496	2,364	369,093	48,793	,000	,589	,359	1,000

Zeittransparenz

Faktor	df korrigiert (Greenhouse-Geisser)	Quadratsumme vom Typ III	df	Mittel der Quadrate	F	Sig.	part. Eta²	Eta²	Test-power[a]
akustischer Raum	Ja	115,975	1,657	70,009	23,679	,000	,411	,010	1,000
Inhalt	Nein	210,889	1,000	210,889	49,074	,000	,591	,019	1,000
akustischer Raum * Inhalt	Nein	12,832	3,000	4,277	5,728	,001	,144	,001	,942

Wärme Quelle

Faktor	df korrigiert (Greenhouse-Geisser)	Quadratsumme vom Typ III	df	Mittel der Quadrate	F	Sig.	part. Eta²	Eta²	Test-power[a]
akustischer Raum	Ja	181,132	2,192	82,626	23,670	,000	,410	,016	1,000
Inhalt	Nein	182,414	1,000	182,414	19,949	,000	,370	,016	,991
akustischer Raum * Inhalt	Nein	21,221	3,000	7,074	3,944	,010	,104	,002	,819

Wärme Raum

Faktor	df korrigiert (Greenhouse-Geisser)	Quadratsumme vom Typ III	df	Mittel der Quadrate	F	Sig.	part. Eta²	Eta²	Test-power[a]
akustischer Raum	Ja	99,867	1,692	59,007	7,603	,002	,183	,047	,904
Inhalt	Nein	93,151	1,000	93,151	11,064	,002	,246	,044	,898
optischer Raum * akustischer Raum	Ja	26,308	6,038	4,357	2,683	,016	,073	,012	,861
akustischer Raum * Inhalt	Nein	35,503	3,000	11,834	7,487	,000	,180	,017	,983

Brillanz Quelle

Faktor	df korrigiert (Greenhouse-Geisser)	Quadratsumme vom Typ III	df	Mittel der Quadrate	F	Sig.	part. Eta²	Eta²	Test-power[a]
Inhalt	Nein	133,722	1,000	133,722	20,684	,000	,378	,073	,993
optischer Raum * Inhalt	Nein	11,296	3,000	3,765	2,925	,037	,079	,006	,680
akustischer Raum * Inhalt	Nein	23,474	3,000	7,825	6,379	,001	,158	,013	,963

Brillanz Raum

Faktor	df korrigiert (Greenhouse-Geisser)	Quadratsumme vom Typ III	df	Mittel der Quadrate	F	Sig.	part. Eta²	Eta²	Test-power[a]
akustischer Raum	Ja	41,732	1,931	21,614	3,541	,036	,094	,021	,629
Inhalt	Nein	86,914	1,000	86,914	13,147	,001	,279	,044	,941

Helligkeit Quelle

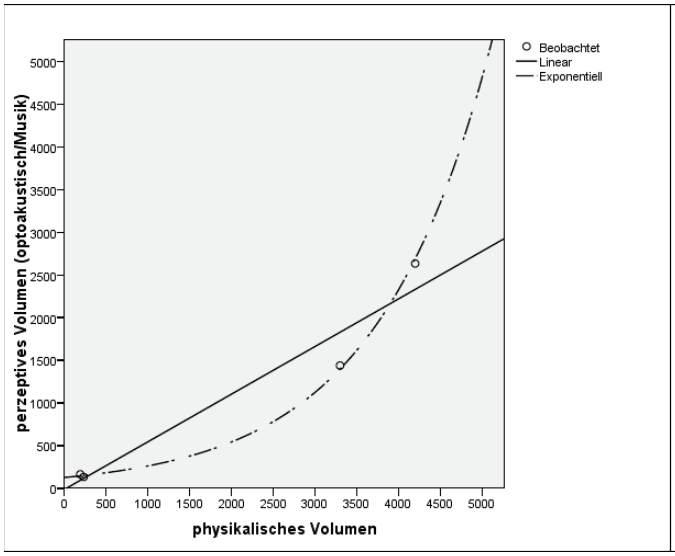
Faktor	df korrigiert (Greenhouse-Geisser)	Quadratsumme vom Typ III	df	Mittel der Quadrate	F	Sig.	part. Eta²	Eta²	Test-power[a]
optischer Raum	Ja	1853,396	1,673	1107,967	129,111	,000	,792	,571	1,000
akustischer Raum	Nein	9,517	3,000	3,172	3,174	,027	,085	,003	,720
optischer Raum * akustischer Raum	Ja	18,465	5,579	3,310	2,261	,044	,062	,006	,760

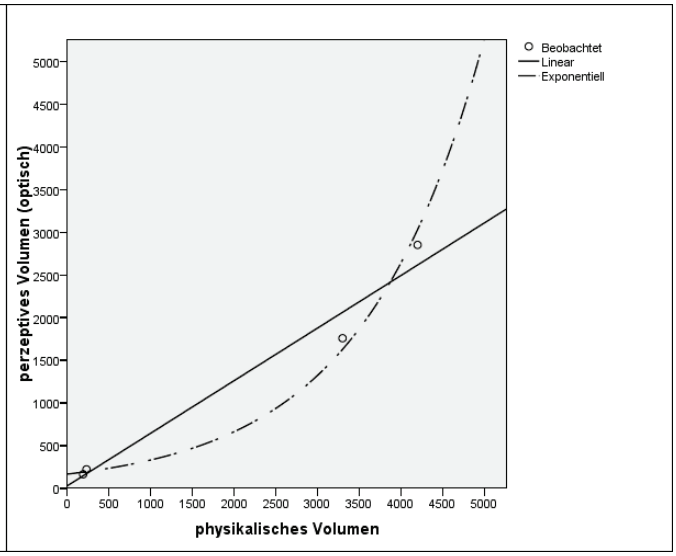
Helligkeit Raum

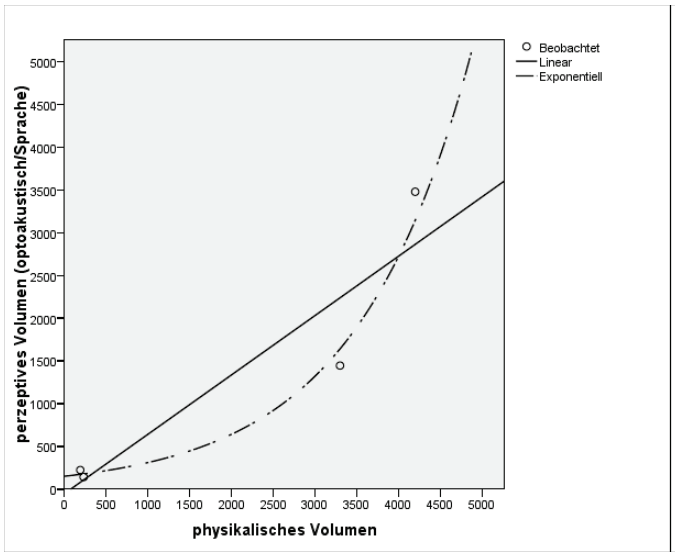
Faktor	df korrigiert (Greenhouse-Geisser)	Quadratsumme vom Typ III	df	Mittel der Quadrate	F	Sig.	part. Eta²	Eta²	Test-power[a]
optischer Raum	Ja	1749,264	2,048	854,203	125,927	,000	,787	,586	1,000

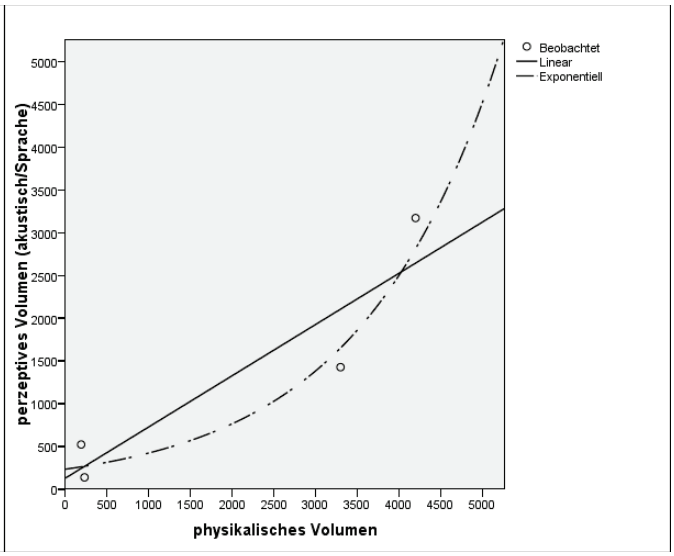
a. Unter Verwendung von Alpha = ,05 berechnet

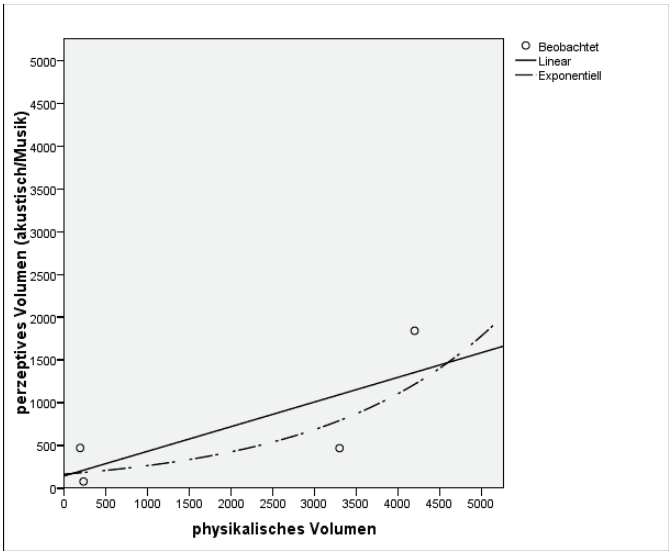
8.5 Abbildungen - Zusammenhang der perzeptiven und realen Volumenskalen











9 Literaturverzeichnis

[Ashmead1990] Ashmead DH, Le Roy D, Odom RD. Perception of relative distances of nearby sound sources. *Percept Psychophys*. 1990;47:326-331.

[Barron1971] Barron M. The subjective effects of first reflections in concert halls – the need for lateral reflections. *J. Sound Vib.* 1971;15:475-494.

[Barron1981] Barron M, Marshall AH. Spatial impression due to early lateral reflections in concert halls: the derivation of a physical measure. *J. Sound Vib.* 1981;77:211-232.

[Beranek1996] Beranek LL Concert and Opera Halls: How They Sound. Woodbury, NY, published for Acoustical Society of America through the American Institute of Physics. 1996.

[Berg2001] Berg J, Rumsey F. Verification and correlation of attributes used for describing the spatial quality of reproduced sound. *AES 19th International Conference*. 2001.

[Berg2006] Berg J, Rumsey F. Identification of Quality Attributes of Spatial Audio by Repertory Grid Technique. *J. Audio Eng. Soc.* 2006;54:365-379.

[Berg2009] Berg J. The contrasting and conflicting definitions of envelopment. *AES 126th Convention*, 2009, Munich, Germany. Convention Paper 7808.

[Blauert1986] Blauert J, Lindemann W. Auditory spaciousness: Some further psychoacoustic analyses. *J. Acoust. Soc. Am*. 1986;80:533-542.

[Bradley1995a] Bradley JS, Soulodre GA. The influence of late arriving energy on spatial impression. *J. Acoust. Soc. Am.* 1995;97:2263-2271.

[Bradley1995b] Bradley JS, Soulodre GA. Objective measures of listener envelopment. *J. Acoust. Soc. Am.* 1995;84:2590-2597.

[Bronkhorst1999] Bronkhorst AW, Houtgast T. Auditory distance perception in rooms. *Nature.* 1999;397:517–20.

[Cabrera2005] Cabrera D, Jeong D, Kwak HJ, Kim JY. Auditory room size perception for modelled and measured rooms. *Internoise, Congress and Exposition on Noise Control Engineering*. Brazil. 2005.

[Calvert2004] Calvert GA, Lewis JW. Hemodynamic Studies of Audiovisual Integration. *G. A. Calvert, C. Spence & B. E. Stein (Eds.), The Handbook of Mulisensory Processes*. Cambridge, Massachusetts. London: The MIT Press. 2004;483-502.

[Côté2011] Côté, N., Koehl, V., Paquier, M., & Devillers, F. Interaction between auditory and visual distance cues in virtual reality applications. *In Proceedings of the Forum Acusticum*. Denmark. 2011.

[Dixon1980] Dixon NF, Spitz S. The detection of auditory visual desynchrony. *Perception.* 1980;9:719-721.

[Fastl2004] Fastl H. Audio-visual interactions in loudness evaluation. *International Conference on Acoustics.* Kyoto. 2004. II-1161-II-1166.

[Gardner1968] Gardner MB. Proximity Image effect in sound localization. *J. Acoust. Soc. Am.*1968;43:163.

[Hashimoto2001] Hashimoto T, Hatano S. Effects of factors other than sound to the perception of sound quality. *17th ICA*. Rome. 2001.

[Hameed2004] Hameed S, Pakarinen J, Valde K, Pulkki V. Psychoacoustic Cues in Room Size Perception. *116th AES Convention.* Germany. 2004. Convention Paper 6084.

[Jack1973] Jack CE, Thurlow WR. Effects of degree of visual association and angle of displacement on the "ventriloquism" effect. *Percept. Mot. Skills* 1973;37:967-979.

[Kosanke2010] Lindau A, Kosanke L, Weinzierl S, Perceptual evaluation of physical predictors of the mixing time in binaural room impulse responses. *Audio Eng. Soc. Convention (128)*, London, United Kingdom. 2010;8089.

[Larsson2001] Larsson P. Vastfjall D, Kleiner M. Ecological Acoustics and the Multimodal Perception of Rooms: Real and Unreal Experiences of Auditory-Visual Virtual Environments. *Proc. International Conference on Auditory Display.* Espoo, Finland. 2001.

[Larsson2002] Larsson P, Västfjäll D, Kleiner M. Auditory-Visual interaction in real and virtual rooms. *Proc. Forum Acusticum.* Seville. 2002. Paper PSY-05-004-IP.

[Lindau2006] Lindau A,Weinzierl S. FABIAN - An instrument for the software-based measurement of binaural room impulse responses in multiple degrees of freedom, *Tonmeistertagungsbaende 24*. 2006. 1-5.

[Long2006] Long M. Architectural Acoustics. Elsevier Academic Press: London. 2006.

[Marshall1967] Marshall AH. A note on the importance of room cross-section in concert halls. *J. Sound Vib.* 1967;5:100-112.

[McGrath1996] McGrath R, Waldmann T, Fernström M. Listening to rooms and objects. *AES 16th International Conference*, Rovaniemi, Finland. 1996.

[McGurk1976] McGurk H, MacDonald J. Hearing lips and seeing voices. *Nature.* 1976;264:746-748.

[Mershon1980] Mershon DH, Desaulniers DH, Amerson TLJ, Kiefer SA. Visual capture in audition distance perception: proximity image effect reconsidered. *J. Aud Res.* 1980;20:129-36.

[Pop2005] Pop C, Cabrera D. Auditory Room Size Perception for real Rooms. *Proceedings of Acoustics.* 2005;115-121.

[Ross2007] Ross, LA, Saint-Amour D, Leavitt V M, Javitt DC, Foxe J J. Do you see what I am saying? Exploring visual enhancement of speech comprehension in noisy environments. *Cereb Cortex* 2007;17:1147–1153.

[Rossing2007] Rossing TD. Springer Handbook of Acoustics. Springer ScienceBusiness Media, LLC, New York. 2007.

[Rudloff1997] Rudloff I. Untersuchungen zur wahrgenommenen Synchronität von Bild und Ton bei Film und Fernsehen. Diplomarbeit. Psychologischen Instituts der Ruhr-Universität Bochum.1997.

[Rumsey2002] Rumsey F. Spatial quality evaluation for reproduced sound: terminology, meaning, and a scene-based paradigm. *J. Audio Eng. Soc.* 2002;50/9:651-666

[Sandvad1999] Sandvad J. Auditory perception of reverberant surroundings. *J. Acoust. Soc. Am.* 1999;105:2:11-93.

[Schubert2001] Schubert T, Friedmann F, Regenbrecht H. The experience of presence: Factor analytic insights. *Presence: Teleoperators and Virtual Environments.* 2001;10:266-281.

[Slutsky2001] Slutsky D, Recanzone G. Temporal and spatial dependency of the ventriloquism effect. *NeuroReport* 2001;12:7-10.

[Soulodre1995] Soulodre GA, Bradley JS. Subjective evaluation of new room acoustic measures. *J. Acoust. Soc. Am.* 1995;98:294-301.

[Soulodre2003] Soulodre GA, Lavoie MC, Norcross SG. Objective measurements of listener envelopement in mulitchannel surround systems. *J. Audio Eng. Soc*. 2003;51(9):826–840.

[Stein1993] Stein BE, Meredith MA. The Merging of the Senses. *MIT Cognitive neuroscience*. Cambridge, MA., US: The Mit Press.1993.

[Stein1996] Stein BE, London N, Wilkinson LK, Price DD. Enhancement of perceived visual intensity by auditory stimuli: a psychophysical analysis. *J Cogn Neurosci*. 1996;897-506.

[Stevens1965] Stevens JC, Marks LE. Cross-modality matching of brightness and loudness. *Proc. Natl. Acad. Sci U S A.* 1965;54:407-411.

[Valente2008] Valente DL, Braasch J. Subjective expectations adjustments of early-to-late reverberant energy ratio and reverberation time to match enviornmnetal cues of a musical performance. *J. Acoustica Special Issue: Virtual Acoustics.* 2008;94:840-855.

[Valente2009] Valente DL, Myrbeck SA, Braasch J. Matching perceived auditory width to the visual image of a performing ensemble in constrasting multi-modal environments. *127th AES Convention,* New York, USA. 2009;Convention Paper 7939

[Viollon1998] Viollon S, Lavandier C, Duval R. Development of an experimental procedure suitable to test audio-visual interactions in the complex urban sound environments. *Euro-noise*. 1998;2:1095-1100.

[Vorländer2003] Vorländer M, Abbou-Elleal E. Simultaneous visual and acoustical perception in room acoustics. *Techni Acustica*, Bilbao. 2003.

[Vroomen2004] Vroomen J, De Gelder B. Perceptual Effects of Cross-Modal Stimulation: Ventriloquism and the Freezing Phenomenon. *Calvert GA, Spence C, Stein BE (Eds.). The Handbook of Multisensory Processes*. Cambridge, Massachusetts, London. The MIT Press. 2004.

[Welch1980] Welch RB, Warren DH. Immediate perceptual response to intersensory discrepancy. *Psychol Bull.* 1980;88:638-667.

[Zahorik2001] Zahorik P. Estimating Sound Source Distance with and without Vision. *OptomVis Sci.* 2001;78:270-275.

[Zahorik2002] Zahorik P. Assessing auditory distance perception using virtual acoustics. *J Acoust Soc Am*. 2002;111:1832–1846

Eidesstattliche Erklärung

Hiermit erkläre ich, dass ich die vorliegende Arbeit selbstständig und ohne Benutzung anderer als der angegebenen Hilfsmittel angefertigt habe. Alle Stellen, die wörtlich oder sinngemäß aus veröffentlichten und nicht veröffentlichten Schriften entnommen wurden, sind als solche kenntlich gemacht. Die Arbeit ist in gleicher oder ähnlicher Form oder auszugsweise im Rahmen einer anderen Prüfung noch nicht vorgelegt worden.

Berlin, den 31.09.2012

Matthias Jentsch