TECHNISCHE UNIVERSITÄT BERLIN

FAKULTÄT I – GEISTES- UND BILDUNGSWISSENSCHAFTEN
INSTITUT FÜR SPRACHE UND KOMMUNIKATION

FACHGEBIET AUDIOKOMMUNIKATION

Zum Einfluss von Differenzen im Frequenzgang von Lautsprechern auf die stereofone Abbildung

Masterarbeit

vorgelegt von

Ajan-Alexander Hannemann

[redacted]

[redacted]	Prof. Dr. Stefan Weinzierl Technische Universität Berlin
Zweitgutachter:	Prof. Dr.-Ing. Anselm Goertz Technische Universität Berlin
Praktische Betreuung:	Martin Lorenz
Abgabe:	19. Dezember 2017

Eidesstattliche Erklärung

Hiermit erkläre ich an Eides statt gegenüber der Fakultät I der Technischen Universität Berlin, dass die vorliegende, dieser Erklärung angefügte Arbeit selbstständig und nur unter Zuhilfenahme der im Literaturverzeichnis genannten Quellen und Hilfsmittel angefertigt wurde. Alle Stellen der Arbeit, die anderen Werken dem Wortlaut oder dem Sinn nach entnommen wurden, sind kenntlich gemacht. Ich reiche die Arbeit erstmals als Prüfungsleistung ein. Ich versichere, dass diese Arbeit oder wesentliche Teile dieser Arbeit nicht bereits dem Leistungserwerb in einer anderen Lehrveranstaltung zugrunde lagen.

Mit meiner Unterschrift bestätige ich, dass ich über fachübliche Zitierregeln unterrichtet worden bin und verstanden habe. Die im betroffenen Fachgebiet üblichen Zitiervorschriften sind eingehalten worden. Eine Überprüfung der Arbeit auf Plagiate mithilfe elektronischer Hilfsmittel darf vorgenommen werden.

Berlin, den 19. Dezember 2017 _____

Ajan-Alexander Hannemann

Inhaltsverzeichnis

1 Einleitung ... 1

2 Forschungsstand und Fragestellungen ... 3

3 Theoretische Grundlagen ... 6

3.1 Psychoakustik ... 6

3.1.1 Lokalisation natürlicher Schallquellen ... 6

3.1.2 Lokalisation im überlagerten Schallfeld ... 7

3.1.3 Klangfarbenwahrnehmung ... 8

3.1.4 Verdeckung und Lautheit ... 9

3.2 Raumakustik ... 10

3.3 Akustische Messtechnik ... 12

3.4 Lautsprecher ... 13

4 Methodik ... 16

4.1 Auswahl und Generierung der Stimuli ... 16

4.1.1 Auswahl der Frequenzpunkte der Differenzen im Frequenzgang ... 16

4.1.2 Bandbreite der Differenz im Frequenzgang ... 17

4.1.3 Verwendete Signalarten ... 18

4.1.4 Erzeugung der Differenzen im Frequenzgang ... 19

4.2 Auswahl der Versuchslautsprecher ... 22

4.3 Beschreibung der Versuchsräume ... 25

4.3.1 Reflexionsarmer Raum ... 26

4.3.2 Hörstudio ... 26

4.3.3 Messtechnischer Vergleich der Versuchsräume ... 27

4.4 Stichprobenumfang ... 29

4.5 Beschreibung der Versuchsverfahren ... 29

4.5.1 Verfahren für die Schätzung der Wahrnehmungsschwellen ... 29

4.5.2 Verfahren für das Messen des Einflusses auf Wahrnehmungsqualitäten ... 30

4.6 Versuchsaufbau ... 32

4.7 Versuchsdurchführung ... 36

4.7.1 Vorversuch ... 36

4.7.2 Hauptversuch ... 40

5 Ergebnisse ... 42

5.1 Wahrnehmungsschwellen ... 42

5.1.1 Haupteffekte ... 43

5.1.2 Interaktionseffekte ... 44

5.1.3 Abhängigkeit der Wahrnehmungsschwellen vom Alter oder der Hörerfahrung ... 47

5.2 Einfluss auf Wahrnehmungsqualitäten ... 48
5.2.1 Aufbereitung der Daten ... 48
5.2.2 Korrelationen der Wahrnehmungsqualitäten ... 49
5.2.3 Haupt- und Interaktionseffekte ... 50

6 Diskussion ... 57
6.1 Wahrnehmungsschwellen von Differenzen im Frequenzgang ... 57
6.2 Einfluss von Differenzen im Frequenzgang auf die stereofone Abbildung ... 58
6.3 Limitationen und Fehlerbetrachtung ... 59
6.4 Implikationen und Ausblick ... 60
6.5 Einfluss der Erkenntnisse in die Qualitätssicherung der Lautsprecher-Fertigung ... 61

Zusammenfassung ... I

Literaturverzeichnis ... II

Tabellenverzeichnis ... VII

Abbildungsverzeichnis ... VIII

Anhang ... X

1 EINLEITUNG

Hersteller in der High-End-Audiobranche verfügen durch die Höhe der erzielbaren Verkaufspreisepreise über entsprechende Ressourcen, den klangtechnisch anspruchsvollen Kunden optimierte Audiotechnik bereitzustellen. Neben hochwertigen Konstruktionsmaterialien, wird auch Wert auf die Auswahl hochwertiger Bauteile gelegt. In der Lautsprecherentwicklung umfasst letzteres die Frequenzweichenschaltung und die Chassis[1] selbst. Unvermeidbare Exemplarstreuungen der Chassis resultieren allerdings in messtechnisch erfassbaren Abweichungen innerhalb der gefertigten Serien. Bei der Verwendung zweier ungleicher Chassis in der Stereofonie entstehen dadurch, unabhängig von räumlichen Einflüssen, spektrale Differenzen zwischen den Schallquellen. Bei inkohärenten Signalen in der Stereofonie, wird das Hörereignis diffuser lokalisiert und verliert somit an räumlicher Schärfe, wodurch die Umhüllung des Hörers zunimmt (Blauert und Braasch, 2008, S.106). In der Stereofonie bevorzugen die Zuhörer bei Sprache eine lokalisierungsscharfe und bei Musik eine räumlichere Darbietung (Klippel, 1988). Die Auswirkungen einer geringeren Abbildungsschärfe sind demzufolge nicht per se als schlecht einzustufen (Wittek, 2000, S.126).

Genormte Anforderungen zur akustischen Gleichheit der Lautsprecher in der Stereofonie gibt es keine. Der 1962 veröffentlichte Qualitätsstandard für Audio-Wiedergabegeräte, die DIN 45500, zeigte erstmals Grenzen für die Übertragungsfunktion von Lautsprechern auf, die das Prädikat HiFi tragen durften. Aufgrund der heutigen technischen Möglichkeiten, ist das Einhalten der damals festgelegten Grenzen für anspruchsvolle Hersteller unproblematisch, weshalb die Anforderungen nicht mehr Teil der ersetzenden DIN EN 61305 sind. Es wurden aber auch keine neuen, strengeren Anforderungen definiert. Aus eigenem Anspruch selektieren Hersteller von HiFi- und Studio-Lautsprechern Chassis oder ganze Lautsprecher bereits. Doch wo liegen hier sinnvolle Toleranzgrenzen? Unnötig enge Toleranzen erzeugen einen höheren Ausschuss, was wiederum die Kosten der Herstellung erhöht.

Ziel dieser Arbeit ist es festzustellen, in welcher Höhe die Wahrnehmungsschwellen spektraler Differenzen zwischen den Lautsprechern in Abhängigkeit von Frequenz, Signalart und akustischer Umgebung liegen und zu untersuchen, in wie weit sie auf klangliche und insbesondere geometrische Wahrnehmungsqualitäten Einfluss nehmen. Dazu wurde ein zweigeteilter

[1] Als (Lautsprecher-)Chassis gilt im folgendem die Komponente einer Lautsprecherbox, welche für die Schallwandlung zuständig ist.

Hörversuch durchgeführt. Die daraus gewonnenen Erkenntnisse sind vor allem für Lautsprecher in High-End- und Studioumgebungen relevant, da hier auf eine neutrale, der aufgenommenen Darbietung möglichst ähnlichen Wiedergabe, Wert gelegt wird.

2 FORSCHUNGSSTAND UND FRAGESTELLUNGEN

Die Wahrnehmungsschwellen von Anhebungen oder Absenkungen im Frequenzgang wurden hinreichend und unter verschiedenen Aspekten untersucht (Bücklein, 1962; Karjalainen et al., 1999; Klippel, 1988; Spiess 1990; Toole und Olive, 1987; Uprichard und Choisel, 2008). Neben der Frage ob das Gehör empfindlicher für An- oder Absenkungen im Frequenzgang ist, wurde auch die Beeinflussung der Schwellwerte durch die Frequenz, Bandbreite, der Signalart oder akustischen Umgebung unter verschiedensten Methoden ermittelt. Angaben zu Schwellwerten von spektralen Differenzen in der Stereofonie finden sich jedoch nicht. Mögliche Auswirkungen von der Verwendung zweier, in ihrem Übertragungsverhalten ungleicher, Lautsprecher führt Spiess (1990, S. 59) auf: „Exemplarstreuungen der Lautsprecher […] können dafür verantwortlich sein, wenn Stimmen und Instrumente scheinbar frequenzabhängig ihre Position ändern.“ Einen Beleg führt Spiess nicht auf, dies war allerdings auch nicht Gegenstand seiner Untersuchung.

In der ITU Empfehlung BS.644-1 (1990)[2], die für den professionellen Audiobereich ausgelegt ist, finden sich Grenzwerte zu frequenzabhängigen Pegeldifferenzen zwischen zwei Lautsprechern. Die Methodik und genauen Umstände unter denen die aufgeführten Grenzwerte ermittelt wurden, sind nicht aufgeführt oder referenziert. Aufgrund des Anwendungszwecks der Empfehlung sind diese vermutlich im Nahfeld in einer nachhallarmen Studioumgebung ermittelt worden. Auch nach Rücksprache mit Sean Olive (HARMAN), Günther Theile (IRT) und Jens Blauert (IEEE, AES) welche zu der Zeit als die Empfehlung veröffentlicht wurde zahlreiche perzeptive Versuche im Bereich der Elektroakustik vornahmen, bleiben die Umstände unter denen die aufgeführten Werte ermittelt worden sind unklar. Empfehlungen zu Toleranzgrenzen basierend auf Wahrnehmungsschwellen welche nicht im Nahbereich eines Lautsprecherpaares ermittelt wurden, konnten im Rahmen der Literaturrecherche nicht gefunden werden.

Der Einfluss verminderter Kohärenz zweier Schallsignale in der Stereofonie wurde, mit der zunehmenden Popularität der Stereo-Wiedergabe, bereits in den sechziger Jahren untersucht. Wie bei Jefress, Blodgett und Deatherage (1962) sowie Chernyak und Dubrovsky (1968) ersichtlich, resultiert eine sinkende interaurale Korrelation der Schallsignale in einer räumlicheren Ausdehnung des Hörereignisses. Der Grad der Korrelation der Schallsignale wurde bei den Arbeiten über die Verzögerung eines Kanals beeinflusst. Eine einfache Lokalisierbarkeit von Phantomschallquellen kann nur durch eine hohe Kohärenz der beiden Lautsprechersignale erreicht werden (Blauert und Braasch, 2008, S. 104). Zotter und Frank (2013) sowie Gribben und

[2] Audio quality parameters for the performance of a high-quality sound-programme transmission chain.

Lee (2016) zeigen in ihren Untersuchungen zur wahrgenommene Phantomschallquellenbreite, der Apparent Source Width (ASW), dass sich diese mit steigenden Phasen- oder Amplitudendifferenzen frequenzabhängig vergrößert. Zu hohen Frequenzen ist die ASW bei gleichem Korrelationsgrad geringer als zu tiefen. Bei geringer oder gar keiner Korrelation der Lautsprechersignale teilt sich die Phantomschallquelle räumlich auf, sie zerfällt („image splits"). Dies ist Abhängig von der Frequenz und der Lautstärke der Wiedergabe. Zu tiefen Frequenzen und höheren Lautstärken tritt das Zerfallen häufiger auf als zu hohen Frequenzen bei geringer Lautstärke (Käsbach, 2016). Dabei hat das verbreitern der Quelle nur geringe Auswirkungen auf Ort, Halligkeit und Klangfarbe der Quelle. Der ursprünglich aus der Raumakustik stammende Parameter „Interaural Cross Correlation" (IACC) lässt es zu, näherungsweise die Quellbreite von Phantomschallquellen in der Stereofonie zu beschreiben (Zotter und Frank, 2013). Unter Einbeziehung der interauralen Zeit- und Pegeldifferenzen gelingt dies genauer (Johnson und Lee, 2017).

Ableitend aus der Recherche ergeben sich folgende Fragestellungen:

Wo liegen die Wahrnehmungsschwellen von Differenzen im Frequenzgang und unterscheiden sich diese in Abhängigkeit von Frequenz, Signalart und akustischer Umgebung signifikant voneinander? Welche Zusammenhänge zwischen einer spektralen Differenz und der Frequenz, Signalart sowie akustischen Umgebung gibt es in Bezug auf klangliche und insbesondere geometrische Wahrnehmungsqualitäten? Bestätigen sich diesbezüglich die Ergebnisse vorheriger Untersuchungen zur ASW von Phantomschallquellen?

<u>Hypothesen zu den Wahrnehmungsschwellen:</u>

H_1 Die Wahrnehmungsschwellen spektraler Differenzen zwischen zwei Lautsprechern in der Stereofonie sind frequenzabhängig.

H_2 Räumliche Einflüsse verringern die Wahrnehmungsschwellen von spektralen Differenzen in der Stereofonie.

H_3 Zeitliche und spektrale Eigenschaften von Audiosignalen haben Einfluss auf die Wahrnehmungsschwellen von spektralen Differenzen in der Stereofonie.

Hypothesen zu den Einflüssen auf die stereofone Abbildung:

H_1 Differenzen im Frequenzgang in der Stereofonie erschweren frequenzabhängig die Lokalisation von Phantomschallquellen.

H_2 Differenzen im Frequenzgang in der Stereofonie vermindern abhängig von der Signalart die Lokalisierbarkeit der Hörereignisse.

H_3 Differenzen im Frequenzgang in der Stereofonie haben unter Wiedergabe mit Raumeinflüssen geringere Auswirkungen auf die Verminderung der Abbildungsschärfe als ohne.

3 THEORETISCHE GRUNDLAGEN

3.1 Psychoakustik

3.1.1 Lokalisation natürlicher Schallquellen

Theile (1980, S.7) definiert den Begriff Lokalisation als das ”(...) Zuordnungsgesetz zwischen dem Ort eines Hörereignisses außerhalb des Kopfes und bestimmten Merkmalen eines oder mehrerer Schallereignisse.“ Das Hörereignis ist das durch das Gehör verarbeitete Schallereignis. Dabei muss es in seinen spektralen, zeitlichen Eigenschaften und seines Ortes nicht mit dem des Schallereignisses übereinstimmen.

Für die Lokalisation von Schallquellen stehen dem Gehör binaurale und monaurale Merkmale zur Verfügung. Nach der 1907 von Lord Rayleigh eingeführten Duplex-Theorie, gelingt dies dem Gehör in der Horizontalebene primär über die Auswertung binauraler Unterschiede zwischen den Ohrsignalen in Form von interauralen Pegel- und Zeitdifferenzen (Interaural Level Differences (ILD) und Interaural Time Differences (ITD)). Ist die empfangene Frequenz tief und damit die Wellenlänge groß, nutzt das Gehör die ITDs für die räumliche Zuordnung des empfangenen Schalls. Frequenzanteile dessen Wellenlängen so kurz sind, dass der Schall sich nicht mehr um den Kopf beugen kann, sondern durch ihn Abgeschattet wird, lokalisiert das Gehör über die resultierenden ILDs (Blauert, 1974). Dabei unterstützen monaurale Ohrsignalmerkmale, wie die Architektur der Pinna, die Richtungseinordnung -vor allem bei seitlichem Schalleinfall- und der Vorne-/Hinten Erkennung (Langendijk und Bronkhorst, 2002). Die Kopfgröße und der Torso beeinflussen ebenfalls die am Trommelfell eintreffenden Signale. Je nach Einfallsrichtung erfährt der Schall so eine individuelle Filterung. Durch Messung der kopfbezogenen Übertragungsfunktion (HRTF) lassen sich die Richtcharakteristiken der Ohren festhalten. So unterschiedlich die Anatomie der Menschen, so verschieden sind auch ihre HRTFs (Møller et al., 1995). Die untere Grenze für die Auswertung der Hörereignisrichtung mittels ILDs liegt bei etwa 300Hz, die obere Grenze für ITDs durch die interaurale Phasenlaufzeit bei etwa 1600Hz. Darüber wird nicht mehr die Feinstruktur, sondern die Hüllkurve des Signals ausgewertet. Aufgrund dessen lassen sich diskrete Sinussignale oberhalb von 1600Hz nur durch die Auswertung von ITDs nicht mehr scharf lokalisieren, breitbandige jedoch schon. Der Schwellwert für gerade noch wahrnehmbare Richtungsänderungen, beträgt bei seitlichem Schalleinfall breitbandiger Signale ±4° (Blauert, 1974).

3.1.2 Lokalisation im überlagerten Schallfeld

Befindet sich ein Zuhörer im gleichen Abstand und Winkel zu zwei Schallquellen, die zeitgleich ein identisches, ausreichend breitbandiges Signal wiedergeben, findet dieser sich in einer unnatürlichen akustischen Umgebung wieder. Durch den symmetrischen Aufbau der Schallquellen zum Zuhörer, nimmt dieser auf halben Wege zwischen den beiden Schallquellen einen fiktiven Sender, die sogenannte Phantomschallquelle war. Der Ort der Phantomschallquelle wird über die Theorien der Summenlokalisation beschrieben. Es gibt diesbezüglich Ausarbeitungen von Warncke (1941), Franssen (1960), Wendt (1964) und weiteren Autoren. Bei der Summenlokalisation stimmen Hörereignisort und Schallquellenorte nicht mehr überein. In der Stereofonie würden idealerweise die beiden Lautsprecher als Realschallquelle nicht mehr wahrgenommen werden, stattdessen nur die Phantomschallquelle. Jedoch trüben Raumreflexionen, Abstrahlverhalten, Übersprechen der Lautsprechersignale sowie Pegel- und Zeitdifferenzen diese akustische Illusion. Die wahrgenommene Breite der Phantomschallquellen bleibt stets größer als die von realen Quellen. Beeinflusst wird sie auch durch den Abstand der Lautsprecher zueinander und dem Winkel der Lautsprecher zum Zuhörer (Frank, 2013).

Nach Theile (1980) sind die Theorien der Summenlokalisation für die Beschreibung der Bildung von Phantomschallquellen breitbandiger Signale jedoch unzureichend. Er schlägt hierfür ein Assoziationsmodell vor, was davon ausgeht, dass bei der Lokalisation von Schallquellen Assoziationsvorgänge bestimmend wirken. Die durch Hörerfahrung gesammelten Muster helfen dabei nur hinreichend ähnliche Ohrsignale zu einer Phantomschallquelle verschmelzen zu lassen (Theile, 1980). Durch Mechanismen aus der Gestaltpsychologie gelingt es dem Gehör aus einem komplexen Klangbild, wie es bei der stereofonen Wiedergabe von Musik vorliegt, einzelne Klangelemente zu lokalisieren (Moore, 2007). Diese Theorie erklärt auch, warum es trotz des Übersprechens der Lautsprecherkanäle und den damit verbundenen starken kammfilterartigen Verfärbungen zwischen den Ohrsignalen gelingt, die Phantomschallquellen dennoch getrennt zu lokalisieren und die wahrgenommenen Klangfarbenveränderungen gering zu halten (Theile, 1980; Martin, 2002).

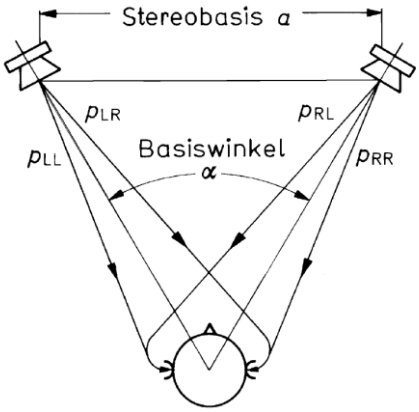


Abbildung 1: Stereo-Standardaufstellung mit einem Basiswinkel α = 60° (Blauert und Braasch, 2008, S.101)

Pegel- und Zeitdifferenzen zwischen den Lautsprechern lassen eine laterale Verschiebung der Phantomschallquelle auf der, in Abbildung 1 dargestellten Stereo-Standardaufstellung, *Stereobasis a* zu. Diese sollte zwischen zwei und vier Meter breit sein (EBU Tech 3276). Abbildung 2 zeigt auf der Ordinate die Verschiebung der Hörereignisrichtung in Abhängigkeit der Pegel- oder Zeitdifferenz zwischen den Lautsprechern. Es bedarf 30dB Pegeldifferenz für die vollständige Verschiebung der Hörereignisrichtung auf 30°. Eine Auslenkung von 30° allein durch Zeitdifferenzen lässt sich nicht erreichen. Dabei sind die erzeugten Schallquellen nicht ortsstabil. Bewegt sich der Zuhörer im überlagerten Schallfeld, hat dies unmittelbare Auswirkungen auf die geometrischen und klanglichen Eigenschaften der Phantomschallquellen. Die Abbildungsschärfe nimmt mit steigender Pegeldifferenz zwischen den Kanälen ab.

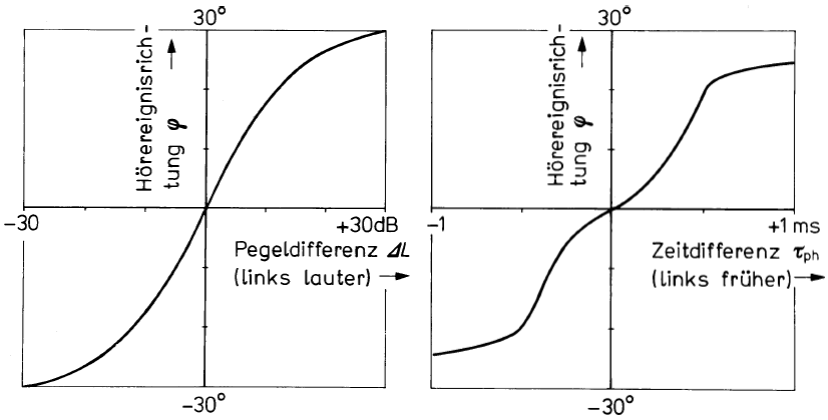


Abbildung 2: Auswirkungen von Pegel- und Zeitdifferenzen auf die Hörereignisrichtung. (Blauert und Braasch, 2008, S. 101)

3.1.3 Klangfarbenwahrnehmung

Die Klangfarbe wird in der Regel als das Attribut der auditorischen Wahrnehmung beschrieben, bei dem ein Zuhörer bei der Präsentation zwei gleichlauter Klänge in gleicher Tonhöhe unter identischen Umständen dennoch die Aussage treffen kann, dass sie unterschiedlich sind.

Klangfarbenunterschiede ergeben sich über die Zusammensetzung des Frequenzspektrums und zeitlichen Schwankungen der dargebotenen Schalle (Moore, 2007). Die psychoakustischen Empfindungsgrößen „Rauhigkeit“ und „Schärfe“ dienen als wichtige Begriffe zur Beschreibung der Klangfarbe. Änderungen der Klangfarbe nimmt das Gehör über seine auditorischen Filter war. Das bei der Cochlea anliegende Signal wird durch die mit steigender Mittenfrequenz angesiedelten Bandpassfilter gefiltert. Das resultierende interne Leistungsspektrum liegt in einer, durch die Bänder, in Frequenzgruppen geteilten Form an. Die Wahrnehmungsschwelle einer Klangfarbenänderung eines Signals ist abhängig von der Frequenz, Ausprägung und Bandbreite der Modifikation.

3.1.4 Verdeckung und Lautheit

Bei der Klangfarbenwahrnehmung spielen auch Maskierungseffekte eine Rolle. Durch spektrale Anhebungen werden benachbarte Frequenzen verdeckt. Insbesondere tiefe Frequenzen, welche erst am Ende der Cochlea die Basilarmembran in Resonanz versetzen, durchwandern die gesamte Länge der Membran und maskieren dadurch benachbarte, vor allem höhere Frequenzen. Die Bandbreite der Verdeckung hängt von den spektralen Eigenschaften des Maskers ab und wird als kritische Bandbreite bezeichnet. Sie wird mit der Erhöhung der betrachteten Frequenz und zunehmender Amplitude größer (Zwicker und Fastl, 1990). Abbildung 3 zeigt den Verdeckungseffekt eines schmalbandigen Rauschens bei einer Mittenfrequenz von 410Hz.

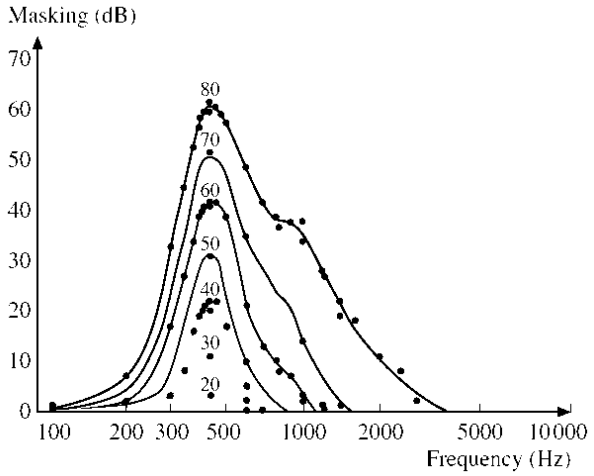


Abbildung 3: Verdeckung benachbarter Frequenzen durch ein schmalbandiges Rauschen in Abhängigkeit des Pegels und der Frequenz. (Moore, 2007)

Betrachtet man den Verlauf der 80dB Kurve, wird deutlich, dass ein Signal dieser Art und Stärke mehr Frequenzgruppen anregt, als ein schwach Ausgeprägtes. Die Lautheit eines Signals ergibt sich aus der Addition aller Teillautheiten der Frequenzgruppen. Auf Grund dessen werden

Anhebungen im Spektrum bei geringerer Ausprägung als Absenkungen entdeckt und haben größeren Einfluss auf die Lautheit (Zwicker und Fastl, 1990).

Vergleicht man zwei Signale unterschiedlicher Amplitude aber gleichen Energiegehaltes (das heißt, eins der Signale wird in seiner Amplitude reduziert aber in seiner Bandbreite so angepasst, dass der Energiegehalt unverändert bleibt), behalten sie die gleiche wahrgenommene Lautheit (Moore, 2007). Folglich beeinflussen Verzerrungen von Signalen durch Filter hoher Bandbreiten die Lautheit eines Signals stärker, als die von Filtern mit geringer Bandbreite (Toole und Olive, 1987).

3.2 Raumakustik

Die Qualität der stereofonen Abbildung hängt neben den verwendeten Lautsprechern maßgeblich von der Umgebung ab in der sie betrieben werden.

Bei Luftschall handelt es sich um mechanische Schwingungen von Teilchen in Fluiden. Werden elastische Kontinua deformiert, so pflanzt sich diese Bewegung in Gestalt von Druckdifferenzen fort. Eine Schallquelle animiert die Teilchen des umgebenden Fluids zu kleinen Schwingungen, welche aufgrund von Massenträgheit und Kompressibilität des Mediums als Kettenleiter, in Form von Longitudinalwellen, übertragen werden. Befindet sich die Schallquelle in einem Raum und der von ihr emittierte Schall trifft auf eine glatte Fläche, spiegelt diese den Schallstrahl nach dem Reflexionsgesetz in dem Winkel in dem er die Fläche erreicht hat. In Abhängigkeit des Reflexionsgrads des Materials der Fläche und der Dissipation der Luft, treffen die Schallstrahlen dieser sogenannten Spiegelschallquellen in geringerer Intensität und später als der Direktschall beim Empfänger ein. Spiegelschallquellen der ersten Ordnungen tragen den Namen Erstreflexionen. Die Dichte der am Empfängerort eintreffenden Reflexionen nimmt mit der Zeit zu und ihre Intensität ab. Diese Vorgänge sind der geometrischen Raumakustik zugeordnet. Ist die abgestrahlte Wellenlänge gegenüber den Raumbegrenzungsmaßen groß, kommt es zwischen den Flächen zur Bildung stehender Wellen mit lokalen Druckminima und -maxima. Diese wirken bis zur Schröderfrequenz, welche sich aus dem Volumen und der Nachhallzeit des Raumes berechnet, klangverfärbend. Ist die abgestrahlte Wellenlänge größer als die größte Raumausdehnung kommt es zum Druckkammereffekt, der einen Pegelanstieg der Frequenzen unterhalb der ersten Raummode zur Folge hat.

Trifft Schall auf Strukturen welche im Bereich seiner Wellenlänge oder größer liegen, wird dieser nicht mehr entsprechend des Reflexionsgesetzes gerichtet reflektiert, sondern diffus in alle Richtungen gestreut. Dabei teilt sich die Intensität auf und die Schallenergie verteilt sich gleichmäßiger im Raum. Diffuse Strukturen beschleunigen in einem Raum das Erreichen eines stationären Zustandes, in dem die Energiedichte örtlich konstant ist. Dieser Zustand stationärer

Energiedichte wird auch als diffuses Schallfeld bezeichnet und daher der statistischen Raumakustik zugeordnet. Wird ein Raum mit einem Impuls angeregt, so setzt sich das Schallfeld aus dem geometrischen und statistischen Anteil zusammen. In der Impulsantwort des Raumes sind in den ersten Millisekunden nach dem Eintreffen des Direktschalles die Erstreflexionen abgebildet und im weiteren Verlauf der statistische Anteil des Raumes, der als Nachhall bezeichnet wird. Abbildung 4 zeigt schematisch den Abklingvorgang der Schallenergie über die Zeit. Die Zeit welcher dieser Abklingvorgang benötigt wird als Nachhallzeit bezeichnet. Gängig ist die Angabe der Dauer, die der Schalldruck braucht um 60dB (T_{60}) zum Ausgangspegel abzufallen.

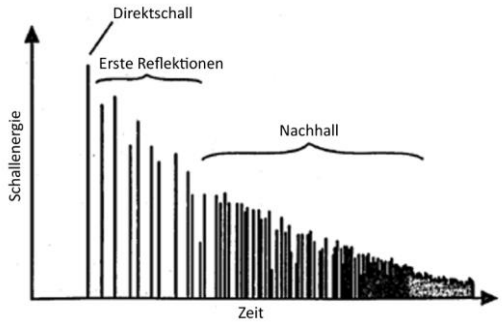


Abbildung 4: Schematische Darstellung für den Abklingvorgang eines Schallfeldes, nach Vorländer (2008)

Da die Nachhallzeit wesentliche Auswirkungen auf den Höreindruck hat, gibt es verschiedene Empfehlungen für Grenzwerte.[3] Die Grenze ab welcher der Direktschallanteil und der des Raumschalls gleich groß sind, bezeichnet man als Hallradius, bezieht man den Bündelungsgrad der Schallquelle mit ein, als Hallabstand. Der Hallabstand r_H welcher sich aus den Raumvolumen V, der Nachhallzeit T_{60} und dem Bündelungsgrad γ nach

$$r_H = \sqrt{\frac{\gamma\,\mathrm{V}}{100\pi T_{60}}} \qquad \text{Gl. (1)}$$

berechnet, beschreibt die kritische Abhördistanz (Critical Listening Distance, CLD). Mit zunehmender Nachhallzeit verringert sich die CLD. Befindet sich der Empfänger zu weit von der Grenze der CLD entfernt, sodass der Nachhall klar dominiert, nimmt die Klarheit und Deutlichkeit der Wiedergabe ab. Während in der Studiotechnik geraten wird sich nicht zu weit von der CLD zu entfernen (EBU Tech 3276-2), eignet sich in einer Wohnraumsituation der Hallabstand nicht für die Einordnung einer Empfohlenen Abhörposition (Toole, 2008).

Frühe Reflektionen können, sofern sie in einer gewissen Zeitdauer nach Eintreffen des Direktschalls den Zuhörer erreichen, maßgeblich zur wahrgenommenen Lautheit, Deutlichkeit

[3] Zum Beispiel DIN EN ISO 18041 (2016) für die Hörsamkeit in Räumen oder ITU-R BS.1116-1 und EBU Tech 3276-1998 für Hörversuche und die Beurteilung von Programmmaterial

und (Sprach-) Verständlichkeit eines Schallereignisses in einem Raum beitragen (Vorländer, 2008). Es konnte gezeigt werden, dass diese in ihrer Wirkung gleichbedeutend mit einer Verstärkung des Direktschalls sind (Bradley, 2003). Die Zeitdauer, bis zu der frühe Reflektionen konstruktiv zur Wahrnehmung einer Schallquelle beitragen, sind nicht festgelegt und variieren je nach Art des Schallereignisses. Typischerweise werden Werte zwischen 15 – 80ms nach Eintreffen des Direktschalls genannt (Vorländer, 2008; Warzybok et al., 2011). Die Wellenfronten der frühen Reflexionen welche das Ohr aus verschiedensten Richtungen erreichen, werden modelliert durch die Beschaffenheit der Flächen und Hindernisse auf die sie treffen. In Konsequenz verringert sich der Grad des IACC. In der Stereofonie hat dies ein Vergrößern der wahrgenommenen Räumlichkeit und der Größe der Phantomschallquellen sowie die Änderung der wahrgenommenen Entfernung der Quellen zur Folge (Toole, 1990). Das Absorbieren erster Reflexionen resultiert in einem analytischeren Klangbild und ist bei der Evaluierung von Audioprodukten zu bevorzugen. Für ein musikalisches Klangerlebnis sollten die ersten Reflexionen den Zuhörer ungetrübt erreichen (Kuhl und Plantz, 1978; Hartmann, 1983).

3.3 Akustische Messtechnik

In diesem Abschnitt werden die für die Interpretation der Messergebnisse von Lautsprechern benötigten Grundlagen behandelt.

Ein Lautsprecher lässt sich näherungsweise, sofern man nicht längere Zeit an seine Belastungsgrenze stößt, als ein lineares, zeitinvariantes System (LZI-System) betrachten. Regt man dieses an seinem Eingang mit einem Impuls an und nimmt die Antwort an seinem Ausgang auf, erhält man die Impulsantwort, welche das Übertragungsverhalten des Lautsprechers vollständig beschreibt. Die Änderung welche das Signal durch das System erfährt, wird mathematisch durch die Übertragungsfunktion beschrieben. Durch die Fouriertransformation der Zeitsignale der Ein- und Ausgangssignale des Systems $x(t)$ und $y(t)$, erhält man deren Spektren $Y(\omega)$ und $X(\omega)$ und die komplexe Übertragungsfunktion $H(\omega)$ ergibt sich aus

$$H(\omega) = \frac{Y(\omega)}{X(\omega)} \qquad \text{(Gl. 2)}$$

Bei LZI-Systemen werden die übertragenen Frequenzen lediglich in Betrag und Phase beeinflusst. Der Betragsfrequenzgang beschreibt die Abhängigkeit der Amplitude und der Phasenfrequenzgang die Abhängigkeit der Phase von der Frequenz. Für eine aussagekräftige Übertragungsfunktion, braucht es einen hohen Signal-/ Störabstand (S/N Ratio). Bestimmt wird dieser durch die Messumgebung, das Anregungssignal und das Messsystem. Die Umgebung sollte einen möglichst geringen Störgeräuschpegel aufweisen, das Anregungssignal einen hohen Energiegehalt und das Messsystem ein geringes Eigenrauschen aufweisen sowie verzerrungsarm

sein. Durch die Mittelung mehrerer Messungen und Erhöhung der Dauer des Anregungssignals, lässt sich das Verhältnis des Signal-/ Störabstands verbessern.[4] Für die Bestimmung der Übertragungsfunktion stehen verschiedene Anregungssignale mit unterschiedlichen Vor- und Nachteilen zur Verfügung. In der akustischen Lautsprechermesstechnik werden hauptsächlich Rausch- oder Gleitsinussignale (Sweep) verwendet (Ahnert und Feistel, 2009). Sämtliche Messungen die im Rahmen dieser Arbeit erstellt wurden, sind mit einem Gleitsinus, dessen Momentanfrequenz exponentiell von der Zeit abhängt (Log-Sweep), durchgeführt worden. Um Messungen vergleichbar und reproduzierbar zu halten, sollten sie unter gleichen akustischen und klimatischen Bedingungen durchgeführt, dass verwendete Messsystem kalibriert und etwaige nichtlinearitäten des Systems, wie beispielsweise die des Messmikrofons, kompensiert werden. Für die vollständige Erfassung des Abstrahlverhaltens eines Lautsprechers, ist das Messen im Fernfeld, unter der Einhaltung von definierten Fernfeldkriterien, notwendig. Ist dies aufgrund der akustischen Umgebung, bzw. der Ausmaße des Lautsprechers nicht möglich, werden Nahfeld- und gefensterte Fernfeldmessungen kombiniert. Hinweise zum Vorgehen beim Messen von Lautsprechern finden sich unter anderem in der DIN EN 60268-5, DIN EN 61305-5 und AES2-2012.

Vergleichende Messungen, um beispielsweise Abweichungen innerhalb einer Messreihe eines Lautsprechertyps festzustellen, sind auch ohne die Einhaltung von Messkriterien wie Nah- und Fernfeld sowie ohne Berücksichtigung von Raumeinflüssen möglich. Dennoch gilt es auf einen möglichst hohen Signal-Rausch-Abstand und konstante Messbedingungen zu achten. Für die Überprüfung und Selektion von Chassis bieten sich Messboxen an, bei denen sich das Messmikrofon und der Speaker Under Test (SUT) durch entsprechende Vorrichtungen am immer exakt selben Ort befinden. Der Lautsprecher strahlt hier in das Boxeninnere, indem sich auch das Messmikrofon, entweder als Grenzfläche eingelassen im Boden der Box oder in direkter Nähe zum Lautsprecher, auf der 0° Achse des Lautsprechers befindet.

3.4 Lautsprecher

Es gibt verschiedene Schallwandlerprinzipien unterschiedlicher Bauformen. Zu den am häufigsten verwendeten Prinzipien im HiFi, Studio und Beschallungsbereich, zählen die elektrodynamischen Direktstrahler und elektrodynamischen Kompressionstreiber. Lautsprecher werden als Ein- oder Mehr-Wege-System aufgebaut.

[4] Pro Verdoppelung der Anzahl der Messungen und Mittelung dieser, erhöht sich der Signal-Rausch-Abstand um 3dB. Pro Verdopplung der Messdauer ebenfalls um 3dB.

In Ein-Wege-Systemen, wie Mikro- oder Kleinstlautsprechern (Smartphones, Mobile-Lautsprecher), ELA-Anlagen oder Selbstbaulautsprechern werden Breitbandlautsprecher eingesetzt. Ziel bei der Auslegung des Wandlers ist die verzerrungsarme Wiedergabe über den größtmöglichen Frequenzbereich bei einer hohen Empfindlichkeit. In einem Mehr-Wege-System wird der Wiedergabebereich auf teilfrequenzoptimierte Lautsprecher aufgeteilt. Dadurch ergeben sich gegenüber Ein-Wege-Systemen einerseits Nachteile, wie Abweichungen in der komplexen Übertragungsfunktion in den Übergangsbereichen und Richtwirkungssprünge, andererseits Vorteile wie die Verringerung nichtlinearer Verzerrungen und der Reduzierung der Richtwirkung zu hohen Frequenzen.

Durch die Größe der Membran besitzt diese ein gewisses Maß an Trägheit, was dazu führt, dass mit steigender Frequenz nur noch die in der Nähe der Schwingspule befindlichen Bereiche bewegt werden. Es kommt zu konzentrischen Biegeschwingungen und einer daraus resultierenden, komplizierten Schwingungsverteilung, welche sich auf den Frequenzgang und das Abstrahlverhalten auswirken. Das Abstrahlverhalten eines realen Schallwandlers ergibt sich aus seinem Volumenfluss, bzw. aus seiner Schalldruckamplitude p_Q. Dabei folgt sie der Spaltfunktion, woraus sich die Richtwirkung berechnen lässt:

$$p = p_Q \frac{\sin\left(\pi \frac{l}{\lambda} sin\Theta\right)}{\pi \frac{l}{\lambda} sin\Theta} \qquad \text{(Gl. 3)}$$

Θ gibt den Winkel der 0° Hauptachse zum Aufpunkt an und läuft von $\pm\frac{\pi}{2}$. Bei dem vielfachen von π ergeben sich Nullstellen und zwischen diesen bilden sich Nebenkeulen aus. Je höher die Frequenz, desto schmaler wird die Hauptkeule und die Anzahl der Nebenkeulen nimmt zu. Ist die Wellenlänge größer als die Ausmaße l des Strahlers, ergeben sich keine Nebenkeulen und die Abstrahlung erfolgt in radialer Form ohne Richtwirkung. Die Richtwirkung des Lautsprechers bestimmt das Leistungsdichtespektrum des Diffusfeldes, welches das Direktfeld in einem Raum überlagert, und beeinflusst somit frequenzabhängig die Schallenergieverteilung im Raum. Tiefe Frequenzen sind stärker vertreten als hohe und verursachen zudem die Ausbildung eines modalen Schallfeldes. Darüber hinaus beeinflusst die Architektur des Gehäuses die abgestrahlte Wellenfront des Lautsprechers. Schallwandeffekte wie der „Baffle Step“, der den Übergang von kugelförmiger zu halbkugelförmiger Abstrahlung im Bereich der Gehäuseabmessungen beschreibt, und Kantenreflexionen, die zur Ausbildung von Spiegelschallquellen an eben diesen führen und Interferenzen verursachen. Aus einem gemessenen Amplituden- und Phasenfrequenzgang auf der 0° Hauptachse allein, lässt sich nicht feststellen, ob der Lautsprecher einen besonders neutralen oder unverfälschten Klang hat. All die genannten Faktoren beeinflussen die räumlichen Wiedergabeeigenschaften eines Lautsprechers.

Bei elektrodynamischen Lautsprechern werden Nichtlinearitäten in der Übertragungsfunktion im Wesentlichen durch drei Faktoren bestimmt. Die *Antriebskraft* der Membran, definiert über die Feldstärke F, nimmt mit zunehmender Auslenkung ab. Grund dafür ist, dass die Spule mit zunehmender Auslenkung zum Teil das Magnetfeld verlässt. Die Effizienz des Lautsprechers nimmt zu hohen Auslenkungen hin ab. Analog zur Antriebskraft verringert sich auch die *Schwingspuleninduktivität*, da die Induzierte Spannung mit dem Verlassen des Magnetfeldes geringer wird. Der Lautsprecher besitzt Tiefpasscharakter. Dagegen steigt die *Federsteife der Aufhängung* mit der Auslenkung. Wird durch hohe Schalldrücke der lineare Arbeitsbereich verlassen, beginnt die Zentrierspinne als „Bremse“ zu wirken und begrenzt die Auslenkung. Die resultierenden nichtlinearen Verzerrungen äußern sich im Frequenzspektrum zum einen als ganzzahlige Vielfache der anregenden Grundschwingungen (beschrieben durch den Klirrfaktor) und zum anderen in Form von Intermodulationsverzerrungen (beschrieben durch Differenztonfaktoren).

In der Herstellung von Chassis kommt es, trotz technischer Hilfsmittel wie Klebemaschinen oder Zentrierhilfen, durch Ungenauigkeiten in der Montage und Toleranzen der Rohmaterialien (Membran, Zentrierspinne, Sicke, Mischverhältnisse des Klebers, Dämpfungsmaterial, …) zu Exemplarstreuungen. Die Abweichungen wirken sich auf den komplexen Frequenzgang, den Impedanzgang und auf das Verzerrungsverhalten des Lautsprechers aus. Bei Tieftonlautsprechern ist es leichter geringere Fertigungstoleranzen einzuhalten als bei Hochtonlautsprechern. Beispielsweise ungleichmäßig oder zu viel aufgebrachter Kleber hat bei Hochtonlautsprechern, aufgrund ihrer geringeren Masse, größere Auswirkungen als auf die wesentlich schwereren Tieftonlautsprecher.

4 METHODIK

Der Hörversuch bestand aus zwei Teilen: Im ersten Teil wurden die Wahrnehmungsschwellen ermittelt und im zweiten Teil der Einfluss auf Wahrnehmungsqualitäten erhoben. Die Methodik dieser Arbeit wird gestützt von den Erkenntnissen vorheriger Studien. Die Wahrnehmungsschwellen von Überhöhungen im Frequenzgang bei Lautsprechern hängen von den folgenden Faktoren ab: Frequenzpunkt, Bandbreite der Frequenzüberhöhung, Hörfähigkeiten der Versuchsperson, Hörumgebung, Abstrahlverhalten des Lautsprechers und der Wahl der Stimuli (Toole und Olive, 1987). Es wurde angenommen, dass die Einflussfaktoren auch für spektrale Differenzen (Überhöhungen und Absenkungen im Frequenzgang) in der Stereofonie gelten.

4.1 Auswahl und Generierung der Stimuli

4.1.1 Auswahl der Frequenzpunkte der Differenzen im Frequenzgang

Um die verschiedenen Arbeitsbereiche der Lautsprecher im tief-, mittel- und hochfrequenten Bereich abzudecken, wurden die Differenzen an drei Frequenzpunkten erzeugt. Die Wahl der Frequenzpunkte an denen die Audiosignale manipuliert worden sind, erfolgte auf Grundlage von Messdaten welche im Rahmen der Chassis-Selektion erhoben wurden. Es sind Daten von 580 Tiefton-, 622 Tiefmittelton-, 596 Mittelton- und 733 Hochtonlautsprechern, also von insgesamt 2531 Lautsprechern verschiedener Modelle und Hersteller betrachtet worden. Dabei standen für jeden Chassistyp fünf Frequenzpunkte in dem für sie vorgesehenen Arbeitsbereich für die Analyse zur Verfügung. An jedem der Frequenzpunkte wurde die Standardabweichung der gemessenen Schalldruckpegel der verschiedenen Chassistypen bestimmt. Denn desto größer die Varianz der Messdaten an den Frequenzpunkten ist, desto größer sind die Differenzen zwischen den einzelnen Chassis. Bei jeden Chassistyp wurde der Frequenzpunkt an dem die Standardabweichung am größten war in eine Häufigkeitstabelle übertragen. Tabelle 1 zeigt, dass es insbesondere in den Randbereichen des Übertragungsbereichs zu den größten Pegeldifferenzen kommt. Weiterhin wurde deutlich, dass sich Abweichungen herstellerspezifisch, aber modellunabhängig in ähnlichen Bereichen bewegen.

Tabelle 1: Häufigkeiten der größten Standardabweichungen an den betrachteten Frequenzpunkten.

TT						
Frequenzpunkt	39Hz	59Hz	90Hz	150Hz	249Hz	350Hz
Häufigkeit	2	1	0	3	**9**	3
TMT						
Frequenzpunkt	100Hz	199Hz	399Hz	806Hz	1606Hz	3018Hz
Häufigkeit	2	3	0	3	2	**8**
MT						
Frequenzpunkt	150Hz	300Hz	500Hz	1000Hz	2000Hz	4000Hz
Häufigkeit	3	**4**	0	0	3	2
HT						
Frequenzpunkt	2000Hz	3000Hz	5000Hz	7000Hz	10000Hz	13000Hz
Häufigkeit	**7**	2	3	0	**6**	**6**

Auf Grundlage der Datenanalyse wurden die Frequenzpunkte 300, 2500 und 10000Hz ausgewählt. An den betrachteten Punkten lag die geringste Standardabweichung bei 0,11dB (bei einem Tiefmitteltonlautsprecher) und die höchste bei 1,69dB (bei einem Mitteltonlautsprecher). Da es sich bei den Datensätzen um Firmeninterna handelt, bleiben die vollständigen Daten unter Verschluss.

Die Frequenzpunkte wurden anschließend einer Note (gleichstufige Stimmung) in der Nähe zugeordnet, welche laut den Spektren aller Stimuli über ausreichend Energie verfügte. Die Differenzen wurden an den folgenden Frequenzpunkten f_0 erzeugt:

- D4, 293Hz
- E7, 2637Hz
- D9, 9397Hz

4.1.2 Bandbreite der Differenz im Frequenzgang

Die Bandbreite des manipulierenden Filters hat maßgeblichen Einfluss auf die Wahrnehmungsschwellen von Resonanzen: Pro Verdoppelung des Gütefaktors Q, sinkt die Wahrnehmungsschwelle um 3dB (Olive et al., 1997; Spiess 1990). Die Anwendung eines Filters mit einer größeren Filtergüte Q, als die Breite des kritischen Bandes in dem die Mittenfrequenz liegt, hat zur Folge, dass zwei (oder mehr) Frequenzgruppen angeregt werden. Bei gleichem Energiegehalt des Signals, wird die Frequenzüberhöhung folglich als lauter wahrgenommen, bzw.

früher entdeckt (Moore, 2007; Zwicker und Fastl, 1990). Je nach Frequenz beträgt die Bandbreite einer Gruppe 1/6 bis 1/4 Oktav (Moore, 2007). Bei der Analyse von 160 Hochtonlautsprecher-Messungen wurde festgestellt, dass auch Abweichungen größerer Bandbreite als die der Frequenzgruppen auftreten. Aufgrund dessen erfolgte die Erzeugung der Differenz durch das Filter mit Terz-Bandbreite (1/3 Oktav), was einer Filtergüte von

$$Q = \frac{f_0}{B} \approx 4.3 \qquad \text{(Gl. 4)}$$

mit der Mittenfrequenz f_0 und Bandbreite $B = f_2 - f_1$ entspricht.

4.1.3 Verwendete Signalarten

Die für die Hörversuche verwendeten Audiosignale sind auf Grundlage von Ergebnissen vorheriger Untersuchungen zu ähnlichen Fragestellungen ausgewählt worden. Es wurde Rosa Rauschen, Sprache, flächige Musik und impulshaltige Musik, geringerer spektraler Dichte genutzt. Rauschsignale gelten als das kritischste Testsignal für Abweichungen im Amplitudengang (Karjalainen et al., 1999, S.13). Die Ausdehnung von Phantomschallquellen, welche mit Lokalisationsungenauigkeiten korreliert, lässt sich am geeignetsten mit der menschlichen Stimme bewerten. Diese weist von sich aus eine sehr geringe ASW auf (Medina Victoria und Görne, 2009). Um den Praxisbezug zu vergrößern, wurden im zweiten Teil des Hörversuchs zusätzlich zwei Musikstücke bewertet. Ein flächiges, klassisches Musikstück und ein impulshaltiges, pop Musikstück. Bei Stimuli mit impulsiven Inhalten lassen sich Auslenkungen der Phantomschallquellen eher erkennen als bei solchen mit flächigen Inhalten (Olive et al., 1997; Toole 2008). Durch Frequenzüberhöhungen herbeigeführte Verfärbungen lassen sich besser mit stationären, breitbandigem Programmmaterial, mit einer hohen spektralen Dichte erkennen (Klippel, 1988; Toole und Olive, 1987). Die Wahrscheinlichkeit die Frequenzüberhöhung oder Absenkung zu treffen, ist bei Programmmaterial mit einer niedrigeren spektralen Dichte geringer als bei solchen mit einer hohen Dichte. Fällt eine Differenz innerhalb eines Frequenzbereiches des Stimulus in dem kaum Energie vorhanden ist, so wird die Wahrnehmungsschwelle entsprechend hoch ausfallen. Nachdem die Bereiche der Frequenzpunkte der zu erzeugenden Differenzen feststanden, wurden von den ausgewählten Stimuli die Zeitsignale fouriertransformiert und anhand der Spektren sichergestellt, dass diese in dem Bereich der Manipulationen über ausreichend Energie verfügen. Nach Sichtung mehrerer klassischer Musikstücke, wurde deutlich, dass diese generell über einen wesentlich geringen Anteil sehr hoher Frequenzen verfügen als Popmusikstücke. Dies zeigt sich auch in den in Abbildung 5 gezeigten Leistungsdichtespektren der beiden Musikstücke. Neben dem Rosa Rauschen (Leistungsdichte 1/f, 1,5 Sekunden) und der männlichen Sprecherstimme (4 Sekunden) wurden folgende Musiktitel verwendet:

- Berliner Philharmoniker - Serenade In D, K.320 Posthorn 6. Minuetto (Ausschnitt, Länge 12 Sekunden)
- Steely Dan - Cousin Dupree (Ausschnitt, Länge 20 Sekunden)

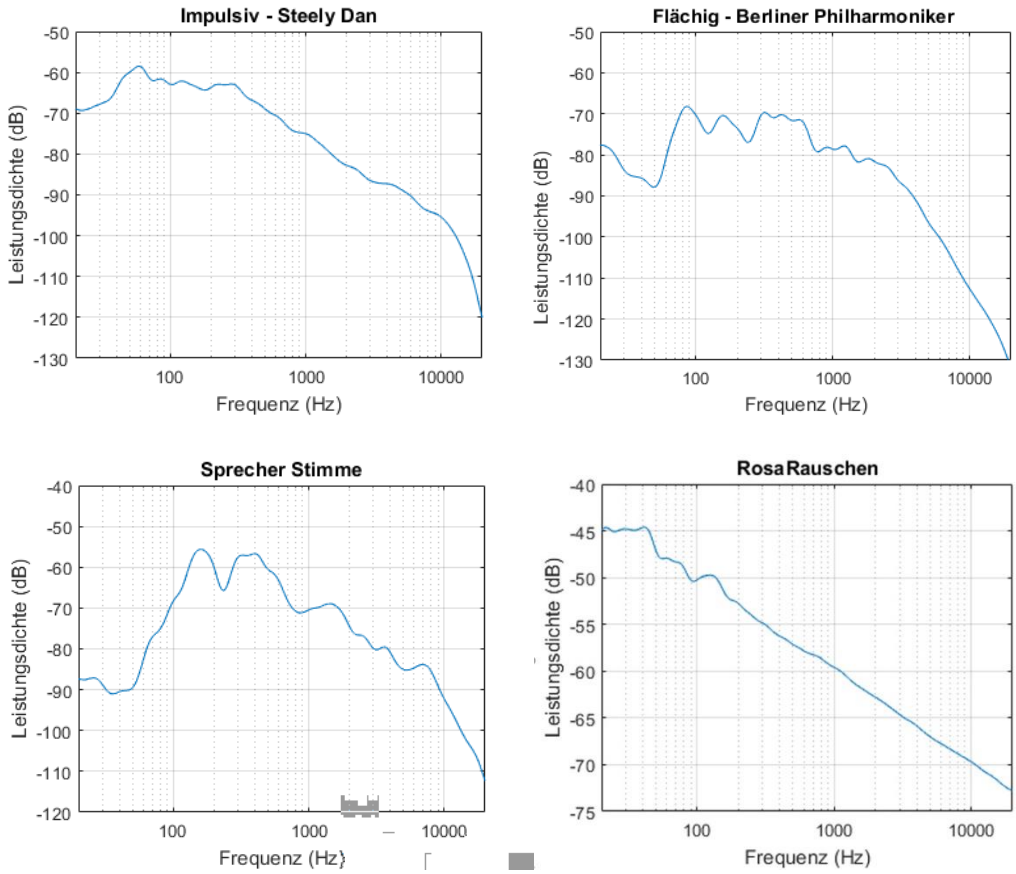


Abbildung 5: Abbildung der Leistungsdichtespektren der verwendeten Testsignale. Es wurde eine Terzbandglättung der Daten vorgenommen.

Die Lautheitsunterschiede der vier normalisierten Stimuli wurden über ein entsprechendes Gaining kompensiert.

4.1.4 Erzeugung der Differenzen im Frequenzgang

Die vier Signalarten wurden normalisiert, in Matlab eingelesen und die Differenzen über ein Filter in den Audiodaten erzeugt. Es wurde an jedem der Frequenzpunkte eine Differenz unterschiedlicher Amplitude erzeugt. Bei drei Frequenzpunkten ergaben sich so 12 verschiedene Darbietungen unterschiedlicher Abstufungen. Über eine Matlab-Funktion wurde am entsprechenden Frequenzpunkt der linke Kanal angehoben und invers der rechte Kanal am selben Frequenzpunkt um den gleichen Wert abgesenkt. Allerdings sind Absenkungen für das Gehör schwerer zu erkennen als Anhebungen (Bücklein 1962; Olive et al., 1997). Bei der Art der spektralen Differenz muss zwischen einseitiger (nur ein Kanal wird manipuliert) oder zweiseitiger Differenz (links und rechts werden manipuliert) unterschieden werden. Ein Eigenversuch

bestätigte die Relevanz der Unterscheidung der zwei Arten: Eine einseitige Differenz von 2dB wurde erkannt, eine zweiseitige nicht. Bei der einseitigen Differenz beträgt die Frequenzüberhöhung nur 1dB. Bei einem Delta von 4dB, also einer Anhebung und Absenkung von jeweils 2dB, wurde die Differenz hingegen erkannt. Einseitig erzeugte Differenzen führen demzufolge vermutlich zu niedrigeren Schwellwerten. Dennoch wurde sich für die Erzeugung zweiseitiger Differenzen entschieden, da zum einen die durch die Manipulation hervorgerufenen Lautheitsunterschiede kompensiert werden und zum anderen der Praxisbezug größer ist: In der Qualitätssicherung und im Selektionsprozess wird stets ein Toleranzschlauch, wie beispielhaft in Abbildung 6 zu sehen, welcher Anhebungen und Absenkungen zulässt, über die Referenz gelegt. Absenkungen sollten folglich in den Wahrnehmungsschwellen mitberücksichtigt werden.

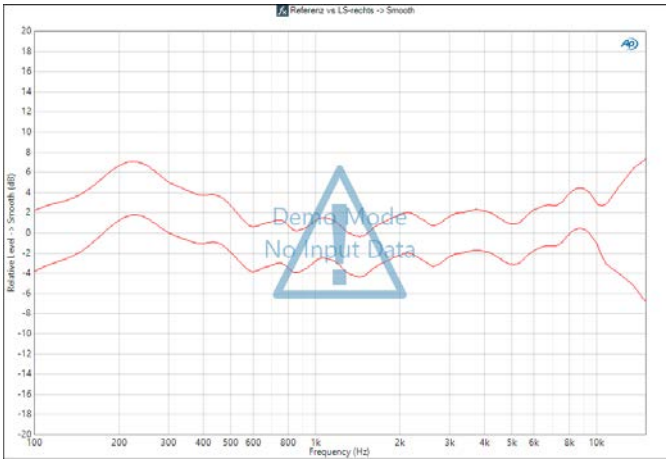


Abbildung 6: Toleranzschlauch in der Qualitätssicherung. Eigene Darstellung

Im Eigenversuch lagen die Stimuli mit Anhebungen im linken als auch rechten Kanal in randomisierter Reihenfolge vor und wurden anschließend getrennt ausgewertet. Ob die Anhebung im linken oder rechten Kanal stattfand, hatte dabei keinen Einfluss auf die Wahrnehmungsschwellen.

Verwendetes Filter

Für die Erzeugung der Differenzen wurde ein symmetrisches infinite impulse response (IIR) Peak Filter zweiter Ordnung genutzt. Die Koeffizienten wurden in der Matlab-Funktion *peak_eq()* nach Zölzer (2002, S. 55) erzeugt. Der Funktion werden Mittenfrequenz, Gain, Filtergüte und Abtastrate übergeben und sie gibt die b und a Koeffizienten aus. Diese werden, neben dem Signal, der *filter()* Funktion von Matlab in dem Skript *eqing.m* übergeben, wodurch die Signale entsprechend bearbeitet werden. Die erstellte Funktion *peak_eq()* und das Skript *eqing.m* finden sich auf dem beigelegten Datenträger. Die Abstufungen der Differenzen wurden zunächst auf Grundlage der Ergebnisse eines Eigenversuchs festgelegt, nach einem Vorversuch (siehe 4.7.1.5) aber in folgender Abstufung:

- Spektrale Differenz zwischen den Lautsprechersignalen in dB:
  0,5 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 28
- Die Differenz wurde bei den folgenden Mittenfrequenzen erzeugt: 293Hz, 2637Hz, 9397Hz
- Bei einer Filtergüte: Q = 4,3 entspricht N = 0,33 Oktaven (Terzbandbreite)
- Abtastrate: 44100Hz

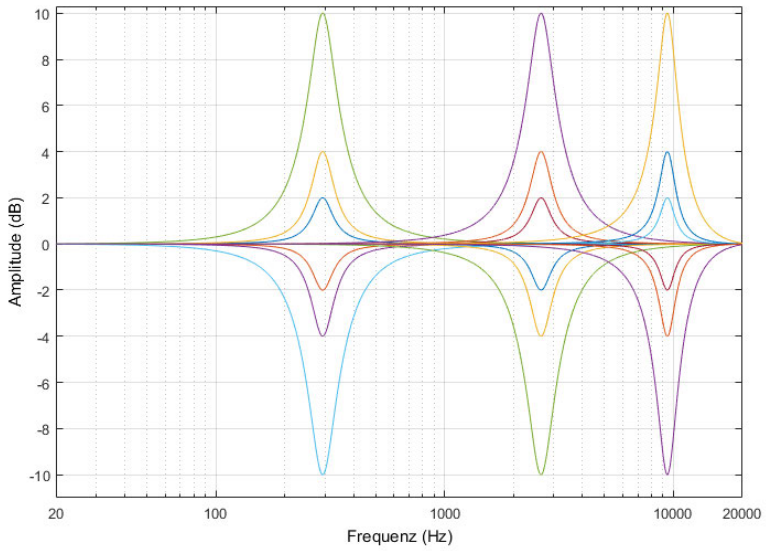


Abbildung 7: Auszug der Amplitudengänge der verwendeten Filter zur Erzeugung der Differenzen. 20dB, 8dB und 4dB Differenz an den drei Frequenzpunkten 293Hz, 2637Hz und 9397Hz.

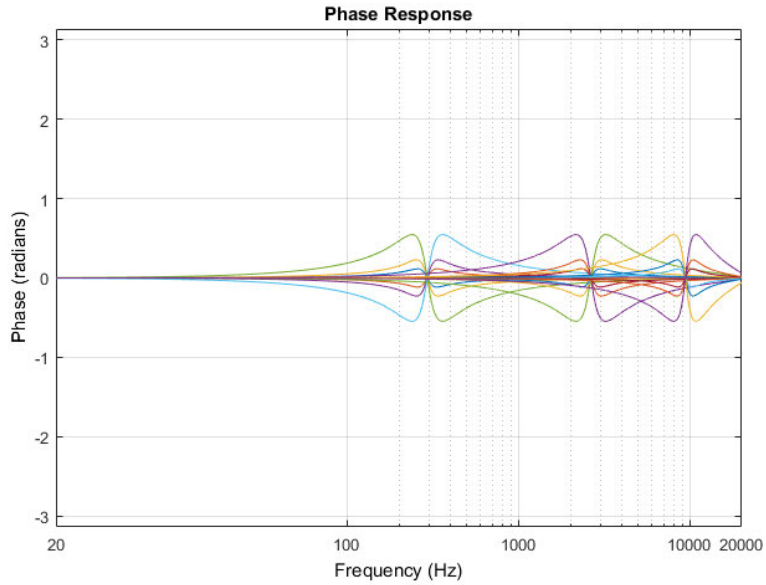


Abbildung 8: Auszug der Phasengänge der verwendeten Filter zur Erzeugung der Differenzen. 20dB, 8dB und 4dB Differenz an den drei Frequenzpunkten 293Hz, 2637Hz und 9397Hz.

Es wurden Koeffizienten für 168 Filter berechnet. Abbildung 7 zeigt exemplarisch die Amplitudenfrequenzgänge, Abbildung 8 die Phasenfrequenzgänge einiger der generierten IIR

Peak-EQ Filter 2. Ordnung. Die durch die Filter hervorgerufenen minimalphasigen Verschiebungen, lagen zu sehr hohen Differenzen oberhalb der Wahrnehmungsschwellen für Phasenverschiebungen.[5] Das heißt in den extremen Fällen wurde möglicherweise, je nach Person, auch die Phasenverschiebung mitbewertet. In dem Bereich der zu erwartenden Schwellwerte lag die Phasenverschiebung allerdings unterhalb der Wahrnehmungsschwelle. Insgesamt standen neben den vier Referenzen 672 Varianten für den Versuch der Schwellwertschätzung zur Verfügung.

4.2 Auswahl der Versuchslautsprecher

Ziel bei der Auswahl des Lautsprechers war es, ein Paar zu finden, dass nur minimale Differenzen in Amplituden- und Phasenfrequenzgang und ein möglichst homogenes Abstrahlverhalten aufweist. Das Abstrahlverhalten eines Lautsprechers nimmt, in Abhängigkeit vom Raum, wesentlichen Einfluss auf die wahrgenommene Qualität der stereofonen Wiedergabe (Müller, 1999; Toole, 2009). Ferner resultieren Phasendifferenzen zwischen den Lautsprechern und innerhalb eines Lautsprecherkanals zu tiefen Frequenzen in der Änderung räumlicher Informationen und zu hohen Frequenzen in Klangfarbenänderungen (Choisel und Martin, 2008).

Vor der Auswahl des Versuchslautsprechers wurde in einem Eigenversuch ein Drei-Wege-Lautsprecher mit einem Breitbandlautsprecher verglichen. Beide Lautsprecher wurden über eine digitale Entzerrung linearisiert. Der Breitbänder stand auf dem Gehäuse des Drei-Wege-Standlautsprechers, der Hochtöner des Drei-Wege-Systems und das Chassis des Breitbänders lagen etwa 18cm auseinander. Die Ohren befanden sich auf halber Höhe zwischen dem Hochton-Chassis und dem Breitbänder. Der Hörversuch wurde in einem akustisch optimierten Raum, mit einer Nachhallzeit von Rt_{60} = 0,25s (> 300Hz) durchgeführt. Erste Reflexionen werden in dem Raum durch poröse Absorber an den Wänden und der Decke stark vermindert. Zuerst sind die Schwellwerte mittels der Breitbänder, dann über das Drei-Wege-System ermittelt worden.

Bei Verwendung der Breitbandlautsprecher, lagen die Schwellwerte bei Rauschen und Sprache im Mittel etwa 1,5dB niedriger als bei dem Drei-Wege-System. Über das verwendete Audiointerface, konnte die Wiedergabe nahtlos zwischen den beiden Lautsprecherpaaren gewechselt werden. Die Abbildungsschärfe wurde beim Breitbandlautsprecher als höher wahrgenommen. Die Lokalisierung der Phantomschallquellen gelang leichter. Das Drei-Wege-System hat zwar eine höhere wahrgenommene Wiedergabequalität, diese ist jedoch für den Versuch nebensächlich. Der hauptsächliche Grund für die inkongruenten Schwellwerte und den

[5] Nach Empfehlung ITU-R BS.644-1. Es gibt zahlreiche Diskussionen über die Wahrnehmungsschwellen von Phasenunterschieden. Insbesondere die Signalart und akustische Umgebung beeinflussen in hohem Maße die Wahrnehmbarkeit von Phasenverschiebungen. Eine Abhandlung verschiedener Studien findet sich unter http://www.audioholics.com/room-acoustics/human-hearing-phase-distortion-audibility-part-2

Abweichungen in der Abbildungsschärfe, liegt vermutlich in dem differenten Abstrahlverhalten der beiden Lautsprechertypen. Der Breitbandlautsprecher weist zu hohen Frequenzen eine höhere Bündelung auf. Das Verhältnis von Direkt- zu Diffusschall ist größer, Einflüsse welche die Lokalisation verschlechtern fallen geringer aus (Bech und Zacharov, 2006). Eine weitere, mögliche Ursache, sind die durch den baulichen Versatz der Lautsprecher hervorgerufenen, unterschiedlichen Schallentstehungsorte, Richtwirkungssprünge und das passive Filternetzwerk. Filter ungleicher Ordnungen führen in den Übergangsbereichen zu ungünstigen Phasenbeziehungen. Zudem können durch Fertigungstoleranzen der Bauteile, die Phasengänge der Lautsprecher voneinander abweichen. Allerdings liegen gerade in den Übergangsbereichen die häufigsten Abweichungen der Lautsprecher zu ihrer Referenz.[6] Der Breitbandlautsprecher, ohne passives Filternetzwerk und mit nur einem Chassis, stellt in diesem Fall das kritischere Messinstrument als der Drei-Wege-Lautsprecher dar. Aufgrund der theoretischen Erkenntnisse und denen aus dem Eigenversuch, wurde als Versuchslautsprecher ein Breitbandlautsprecher verwendet.

Der eigens für die Untersuchung konstruierte Breitbandlautsprecher (10cm, Tang Band w4-2142) befand sich in einem 5,5L Bassreflexgehäuse. Die Gehäuseabstimmung wurde mit der Simulationssoftware *Boxsim* berechnet. Die Betragsfrequenzgänge der Lautsprecher wurden vor den Versuchen im RAR im Fernfeld in 2,65m Abstand gemessen. Abbildung 9 zeigt den Amplitudenfrequenzgang eines Versuchslautsprechers auf seiner 0° Achse.

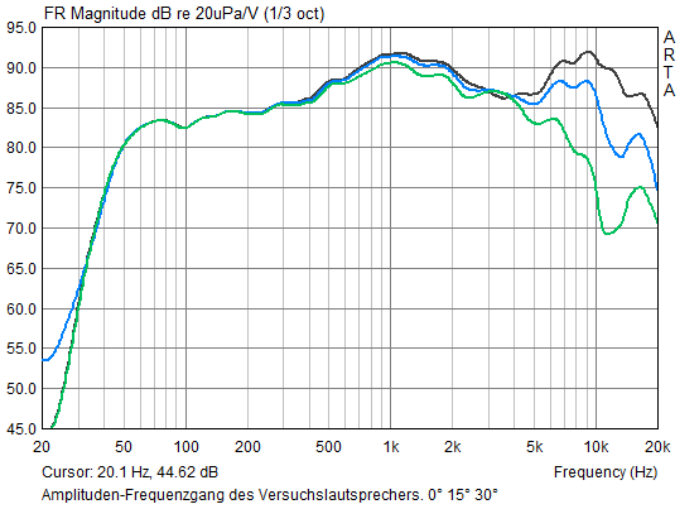


Abbildung 9: Amplitudenfrequenzgang des Versuchslautsprechers bei 0°, 15° und 30°.

[6] Siehe dazu **Fehler! Verweisquelle konnte nicht gefunden werden.**.

Anschließend wurden über die Matlab-basierte Software *Audio Weaver* digitale IIR-Filter im Signalfluss angelegt, um die Lautsprecher in Echtzeit zu linearisieren. **Fehler! Verweisquelle k onnte nicht gefunden werden.** zeigt die Parametrisierung der Filter.

Tabelle 2: Parameter der Filter für Entzerrung der Versuchslautsprecher

Filter Typ	Mittenfrequenz [Hz]	Güte	Amplitude [dB]
Peak EQ	100Hz	2	1,8dB
Peak EQ	110Hz	0,5	1,2dB
Peak EQ	1050Hz	0,7	-6,5dB
Peak EQ	1950Hz	3	-1,4dB
Peak EQ	6500Hz	3	-1,8dB
Peak EQ	8600Hz	2	-6,8dB
High Shelf	12000Hz	1	-2,9dB

Der Einsatz von IIR-Filtern zur Änderung des Betragsfrequenzganges geht mit minimalphasigen Änderungen einher. Damit die Veränderungen des Phasengangs in der Summe der Differenz- und Entzerrungsfilter nicht zu groß werden und um im Bereich der zu erwartenden Wahrnehmungsschwellen zeitliche Einflüsse auszuschließen, wurden für die Entzerrung linearphasige FIR-Filter eingesetzt. Dazu wurden die in **Fehler! Verweisquelle konnte nicht g efunden werden.** genannten Parameter an das Programm *rePhase* übergeben und dort die Koeffizienten (1000 Taps) des FIR-Filters generiert. Diese wurden anschließend in die Software *Audio Weaver* in ein FIR-Modul importiert. Die durch die Länge des Filters definierte Gruppenlaufzeit von 11,3ms ist bei diesem Versuch irrelevant und daher vernachlässigbar.

Da die Differenzen der Lautsprecher ohne Entzerrung zueinander in einem marginalen Bereich liegen (±0,3dB), wurde nur ein Lautsprecher gemessen und das Filter auf beide Kanäle dupliziert. Abbildung 10 zeigt den entzerrten Amplitudenfrequenzgang des Versuchslautsprechers bei 0°, 15° und 30°. In Abbildung 11 ist der Amplitudenfrequenzgang des linken und rechten Lautsprechers aufgetragen. Die Unterschiede der Frequenzgänge liegen im Frequenzbereich von 200Hz bis 10kHz bei ±0,3dB.

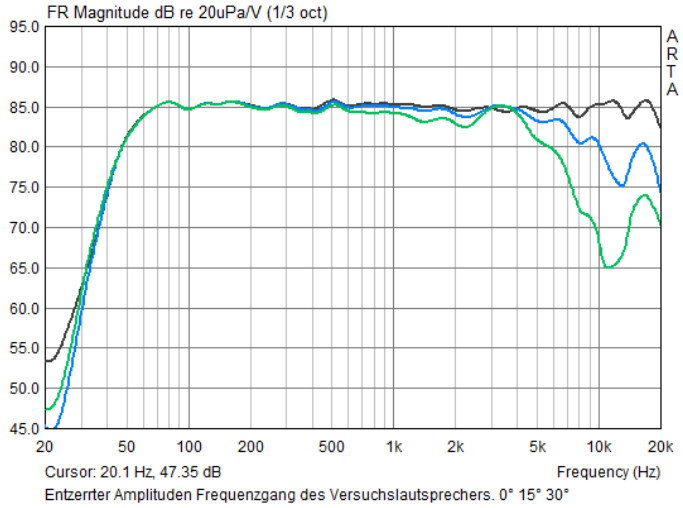


Abbildung 10: Entzerrter Amplitudenfrequenzgang des Versuchslautsprechers bei 0°, 15° und 30°.

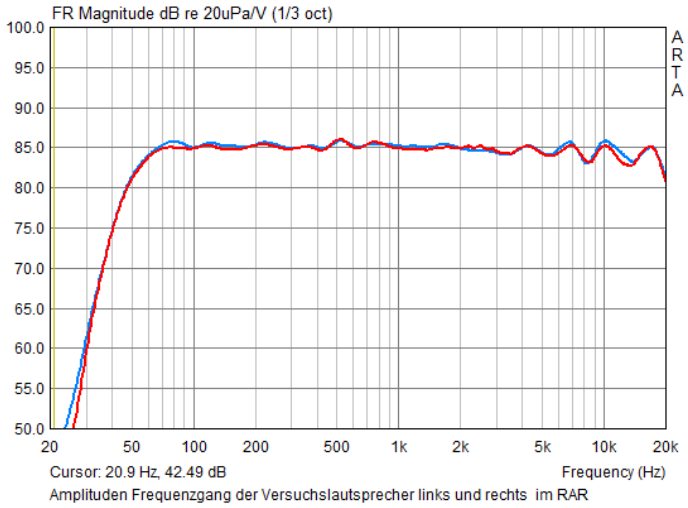


Abbildung 11: Entzerrter Amplitudenfrequenzgang der Versuchslautsprecher links und rechts.

4.3 Beschreibung der Versuchsräume

Da die ersten Reflexionen und der Nachhall eines Raumes wesentliche Einflüsse auf die Klangfarbe von Phantomschallquellen und die Abbildungsschärfe nehmen, werden die Hörversuche neben denen im reflexionsarmen Raum (RAR) in einem weiteren, akustisch weniger behandelten, Raum durchgeführt (Toole und Olive, 1987; Klippel, 1988). Einerseits sind in halligen Räumen spektrale Änderungen und Zeitfehler in der stereofonen Wiedergabe schwieriger zu hören als in trockenen (Spiess, 1990), was jedoch auch von der Art der verwendeten Stimuli abhängt. Uprichard und Choisel (2008) stellten fest, dass die akustische Umgebung nur bei klassischer Musik einen signifikanten Effekt auf die Wahrnehmungsschwelle von

Frequenzüberhöhungen hat, Popmusik nicht. Also solche mit hoher gegen welche mit geringer spektraler Dichte. Andererseits ist der Einfluss von Crosstalk umso größer, desto geringer die Nachhallzeit ist. Crosstalk bezeichnet das Übersprechen der Lautsprecherkanäle an den Ohren und verringert, umso größer sein Einfluss, die Natürlichkeit und Präzision der erzeugten Phantomschallquellen (Bock und Keele, 1986).

4.3.1 Reflexionsarmer Raum

Für kontrollierte Messungen unter Freifeldbedingungen wurden sog. reflexionsarme Räume konstruiert. Ein Freifeld ist eine theoretische akustische Umgebung, in dem die Wellenfront einer Punktschallquelle sich bis ins unendliche ungehindert ausdehnen kann. Sie trifft also weder auf reflektierende, noch streuende Flächen oder Hindernisse. In der Realität kann diese Umgebung nur näherungsweise geschaffen werden, weshalb es auch keine reflexionsfreien Räume gibt.

Der reflexionsarme Raum des Instituts für technische Akustik an der TU-Berlin umfasst ein Volumen von 1850 m³ mit 126 m² Grundfläche und einer unteren Grenzfrequenz von 63Hz. Es handelt sich um einen Vollraum ohne Schallharten Boden. Oberhalb der Grenzfrequenz herrschen Freifeldbedingungen und es ist ein nachhallfreies Messen und Hören möglich. Da der Hörversuch oberhalb der Grenzfrequenz stattfindet, ist der Hörversuch im RAR bereinigt von Raumeinflüssen.

4.3.2 Hörstudio

Um näherungsweise einen Bezug zur Realität des Hifi-Hörens herzustellen, wurde als zweiter Raum ein akustisch optimierter Hörraum nach ITU-R BS.1116-1 ausgewählt. Der Raum ist symmetrisch aufgebaut und es befinden sich nur wenige Einrichtungsgegenstände in ihm. Erste Reflexionen werden durch die modulare Bauweise des Raumes teilweise absorbiert, gestreut oder ungehindert zugelassen. Der Boden hat eine schallharte Oberfläche. Die ITU-R BS.1116-1 Empfehlung empfiehlt im Frequenzbereich von 200Hz bis 4000Hz eine Nachhallzeit T_m von

$$T_m = 0{,}25 * \sqrt[3]{\frac{V}{V_0}} \qquad \text{Gl. (4)}$$

mit V: Volumen des Raumes und V_0: Referenz Volumen von 100m³ und einer Toleranz von ±0,05s. Der Raum hat bei 9,55m x 5,21m x 3,12m ein V = 155,2m³, dies entspricht einer Nachhalltzeit $T_m = 0{,}29$s.

Um die Ergebnisse des Hörversuchs in Zusammenhang mit den akustischen Eigenschaften des Versuchsraums zu bringen, wurde seine Nachhallzeit erfasst. Dazu wurde zunächst mittels eines Dodekaeder-Lautsprecher an 14 unterschiedlichen Positionen die Nachhallzeit gemessen. Dabei

gab es pro Quellposition zwei Empfängerpositionen. Empfänger und Quellhöhe wurden ebenfalls variiert. Anschließend wurden die Messergebnisse exportiert und pro Frequenzpunkt über alle Nachhallwerte das arithmetische Mittel gebildet.[7] Der Verlauf der Nachhallzeit über die Frequenz bewegt sich, wie in Abbildung 12 zu erkennen, innerhalb des Toleranzbands der ITU-R BS.1116-1 (rote Linie), weißt aber eine Erhöhung von bis zu 0,045s im Frequenzbereich von 1000Hz bis etwa 6300Hz auf.

Abbildung 12: Nachhallzeit über die Frequenz des Hörstudios (blau) und Toleranzband nach ITU-R BS.1116-3 (rot).

Die Versuchsräume hatten dieselbe Temperatur und waren mit warmem Halogenlicht beleuchtet.

4.3.3 Messtechnischer Vergleich der Versuchsräume

Für die Hörversuche im Hörstudio wurden dieselben Filterparameter genutzt wie im RAR. Bei identischem Messaufbau lässt sich so allein der Einfluss des Raumes auf den Frequenzgang am Messpunkt, bzw. Hörplatz bewerten. In Abbildung 13 finden sich die Amplituden Frequenzgänge der Lautsprecher im Hörstudio (linker Lautsprecher, rechter Lautsprecher) und im RAR (schwarze Kurve).

[7] Nach DIN EN ISO 3382-1: Akustik – Messung von Parametern der Raumakustik.

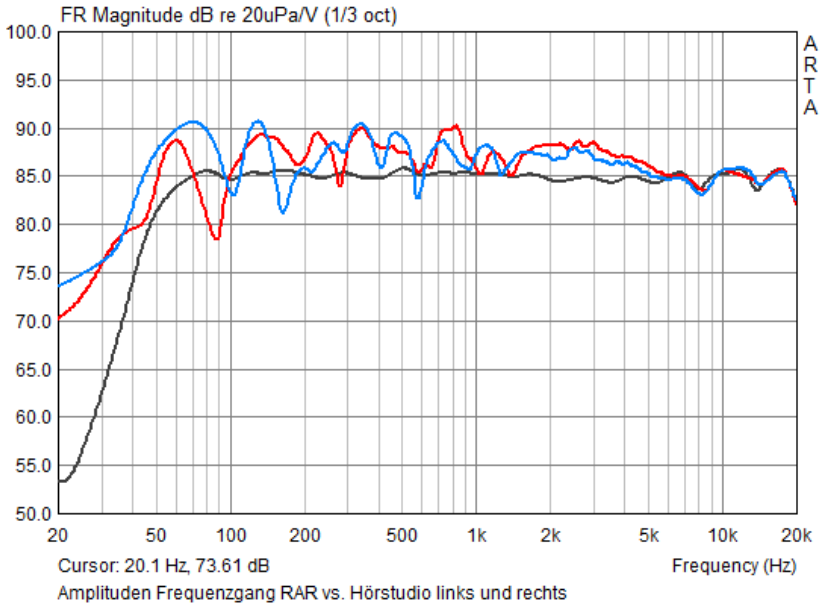


Abbildung 13: Amplitudenfrequenzgang der Lautsprecher **links** und **rechts** im Hörstudio und im **RAR**.

Der Einfluss des Raumes auf den Amplitudenfrequenzgang an der Hörposition wird sichtbar wenn, wie in Abbildung 13 gezeigt, die in beiden Räumen gemessenen Amplitudenfrequenzgänge des Versuchslautsprechers übereinander gelegt werden. In beiden Fällen wurde der Frequenzgang auf der 0° Achse des Lautsprechers und in 2,65m Abstand gemessen. An der Hörposition im Hörstudio, ist der Schalldruckpegel des Hochtonbereichs gegenüber dem Tief- und Mitteltonbereich um etwa 3dB verringert. Das modale Verhalten des Raumes spiegelt sich im Frequenzgang mit Welligkeiten von bis zu 10dB wieder. Von den ersten Reflexionen hervorgerufene Interferenzen sind zwar in ihrer Ausprägung und Frequenz in einem hohen Maße von der Messposition abhängig, resultieren aber an der Hörposition in Welligkeiten von bis zu 5dB. Möglichweise unterstützt die in Abbildung 12 erkennbare Erhöhung der Nachhallzeit des Raumes im mittleren bis hohen Frequenzbereich, den Anstieg von bis zu 3,5dB im selben Frequenzbereich an der Hörposition, wenn auch, durch die gegenüber dem Dodekaeder-Lautsprecher erhöhte Bündelung zu hohen Frequenzen, nicht so breitbandig. Ab etwa 7kHz überwiegt der Anteil des Direktschalls und räumliche Einflüsse haben kaum noch Auswirkungen auf den Amplitudenfrequenzgang an der Hörposition. Darüber hinaus beeinflussen die akustischen Eigenschaften des Raumes den Amplitudenfrequenzgang des linken und rechten Lautsprechers an der Hörposition im unterschiedlichen Maße und führen dort zu spektralen Differenzen. So existiert durch die Positionierung der Lautsprecher im Raum bei 290Hz bereits eine Differenz von etwa 4dB zwischen linken und rechtem Kanal.

4.4 Stichprobenumfang

Expertenhörer lassen einen kleineren Stichprobenumfang zu, da sie zu ähnlichen Ergebnissen gelangen wie deutlich größere Gruppen mit unerfahrenen Hörern (Toole, 2005). Für die Hörversuche wurden nur Probanden eingeladen, welche über Hörerfahrung verfügen. Das heißt sie beschäftigen sich beruflich oder als Student aktiv mit Audiotechnik, Akustik oder Musik. Die Versuchsgruppe setzte sich so aus Mitarbeitern eines Audiounternehmens und Studenten der TU-Berlin, aus dem Studiengang *Audiotechnologie und Kommunikation*, der UdK-Berlin, aus dem Studiengang *Tonmeister* und der SAE, aus dem Studiengang *Audio Production*, zusammen. Im Durchschnitt befassten sich die Probanden rund 2 Stunden pro Tag mit einem der genannten Tätigkeitsfelder. Die Versuchsteilnehmer waren zwischen 18 und 66 Jahren alt, das Durchschnittalter lag bei 34 Jahren. Drei der 17 Probanden waren weiblich. Acht der 17 Teilnehmer hatten bereits bei mindestens einem Hörexperiment teilgenommen.

4.5 Beschreibung der Versuchsverfahren

Standardisierte oder ausführlich untersuchte perzeptive Messverfahren ermöglichen es, hoch valide und objektive probabilistische Zusammenhänge zwischen physikalischen Reizen und der auditiven Wahrnehmung herzustellen. Die Wahrnehmungsschwellen von spektralen Differenzen in der Stereofonie, wurden mit einem adaptiven, psychometrischen Verfahren geschätzt. Für die Untersuchung von Einflüssen der Differenzen auf ausgewählte Wahrnehmungsqualitäten, wurde als Bewertungsinstrument eine Skala mit uni- oder bipolaren Adjektivpaaren verwendet.

4.5.1 Verfahren für die Schätzung der Wahrnehmungsschwellen

Eine Schwelle stellt den Zusammenhang zwischen einem festgelegten, zum Beispiel akustischen, Parameter und der merklichen Änderung eines Reizes dar. Der Schwellwert liegt in dem Punkt, wo Änderungen an einem Signal durch den Parameter gerade bemerkbar werden. Durch äußere Einflüsse, der Tagesform der Versuchsperson oder die von Mensch zu Mensch unterschiedlichen Empfindungsstärken, liegt dieser Schätzwert im statistischen Mittel mehrerer Messungen. Ziel des angewandten Versuchsverfahrens ist es, aus den ermittelten Daten, bei möglichst geringen Zeitaufwand, valide Schätzwerte der Wahrnehmungsschwellen von Differenzen im Frequenzgang zu gewinnen. Die Schätzwerte werden über das erweiterte, klassische, psychometrische Konstanzverfahren, dem 3-AFC (Alternative-Forced-Choice), mit dynamischer Anpassung der Reizstärke (mittels ZEST-Verfahren (Zippy Estimation by Sequential Testing, King-Smith et al., 1994)) ermittelt.

In einem Forced-Choice-Verfahren ist die Versuchsperson gezwungen eine Entscheidung zwischen einer gewissen Anzahl an dargebotenen Reizen zu treffen. Einfache Forced-Choice-Verfahren mit zwei Reizen, wobei einer der beiden die Referenz und einer manipuliert ist, haben den Nachteil der möglichen willkürlichen Bewertung durch die Versuchsperson. Um die Ratewahrscheinlichkeit auf 33% zu verringern werden bei der 3-AFC-Methode der Versuchsperson drei Reize angeboten. Nur einer der drei Reize entspricht nicht der Referenz. Aufgabe des Probanden ist es, aus den drei Reizen den auszuwählen, bei dem es sich nicht um eine Blindprobe handelt.

Eine Methode um Wahrnehmungsschwellen zu bestimmen ist das adaptive, psychometrische ZEST Verfahren nach King-Smith (King-Smith et al. 1994). Neben weiteren adaptiven Verfahren ist dies für die Ermittlung von Wahrnehmungsschwellen das effizienteste (Otto, 2008). Adaptiven Verfahren wird eine bestimmte Verteilungsfunktion als psychometrisches Modell der zu untersuchenden Wahrnehmung zu Grunde gelegt. Dadurch können die Schrittweiten kontrolliert werden und die Reizstärken in hoher Auflösung vorliegen, ohne dass das Finden eines validen Schätzwertes der Wahrnehmungsschwelle eine hohe Anzahl an Durchgängen (Trails) erfordert. Das ZEST Verfahren nutzt die Weibull-Funktion, welche eine logarithmische Skalierung der physikalischen Reize vorsieht. Im Wendepunkt der psychometrischen Funktion, wo ihre Steigung maximal ist, liegt der Schätzwert der Wahrnehmungsschwelle. Um die Anzahl der Trails möglichst gering zu halten, berechnet ZEST nach jeder Bewertung den Schätzwert der aktuellen Wahrnehmungsschwelle und legt die entsprechende Reizstärke für die nächste Darbietung vor. Gegenüber den Empfehlungen der hohen Anzahl an Trails anderer adaptiver Verfahren, gelten bei ZEST bereits 20 als ausreichend (Otto, 2008). Dies verkürzt die Versuchsdauer erheblich. Die dem Verfahren zu Grunde liegenden Bayes-Schätzung, bestimmt aus der berechneten posterior probability density function (posterior pdf) den Schätzwert. Für den Hörversuch wurde die psychometrische Funktion nach den Empfehlungen aus der WhisPER Dokumentation, welche sich auf verschiedene Untersuchungen stützen, parametrisiert (Spannweite β=2.8, Fallrate λ=0.01). Für die a-priori-Wahrscheinlichkeitsdichtefunktion wurde eine Gauss-Verteilung angenommen. Der Wendepunkt der psychometrischen Funktion lag bei 12,5dB spektraler Differenz zwischen den Lautsprechern.

Die für das Verfahren verwendeten Reizstärken finden sich in Abschnitt 4.1.3.

4.5.2 Verfahren für das Messen des Einflusses auf Wahrnehmungsqualitäten

Im zweiten Teil des Hörversuchs wurde der Einfluss von spektralen Differenzen in der Stereofonie auf komplexe Klangmerkmale hin untersucht. Dafür wurden Wahrnehmungs-qualitäten aus den Kategorien *Klangfarbe*, *Geometrie* und *Allgemein* des *Spatial Audio Quality*

Inventory (SAQI) (Lindau, 2014) verwendet. Bei der Auswahl wurden nur Attribute berücksichtigt, bei denen eine Beeinflussung durch spektrale Differenzen, auf Grundlage vorheriger Untersuchungen, zu erwarten war. In einem Eigenversuch wurden zunächst 14 Wahrnehmungsqualitäten bewertet, wovon anschließend sechs verworfen worden sind. Ziel war es, durch das Reduzieren der Anzahl an Attributen den Zeitbedarf des Hörversuchs zu verringern, ohne solche zu vernachlässigen, die auf spektrale Differenzen feststellbare Auswirkungen haben. Die Wahrnehmungsqualität *Klangfarbe* wurde nicht weiter differenziert, in der Kategorie *Geometrie* wurde das Attribut *Ausdehnung,* durch das Zusammenfassen der drei Wahrnehmungsqualitäten *Tiefen-, Breiten-* und *Höhenausdehnung* gebildet. Durch die Kompensation des Lautheitsanstieges, wurden eventuelle Lautheitsunterschiede nicht mit bewertet. Folgende Attribute wurden im Hörversuch bewertet:

- **Unterschied**
  Skalenpole: kein Unterschied - sehr großer Unterschied
  Wahrgenommener *Unterschied* des verzerrten Stimulus zu seiner Referenz.
- **Klangfarbe**
  Skalenpole: kein Unterschied - sehr großer Unterschied
  Wahrgenommener Unterschied hinsichtlich der *Klangfarbe* des verzerrten Stimulus zu seiner Referenz.
- **Lokalisierbarkeit**
  Skalenpole: schwieriger lokalisierbar - einfacher lokalisierbar
  Wahrgenommener Unterschied hinsichtlich der *Lokalisierbarkeit* des verzerrten Stimulus zu seiner Referenz.
- **Ausdehnung**
  Skalenpole: ausgedehnter - kompakter
  Wahrgenommener Unterschied hinsichtlich des Attributes *Ausdehnung* des verzerrten Stimulus zu seiner Referenz. Die Ausdehnung beschreibt die wahrgenommene Größe einer Schallquelle in alle Richtungen.
- **Räumliches Zerfallen**
  Skalenpole: fusionierter - zerfallener
  Wahrgenommener Unterschied hinsichtlich des Attributes *räumliches Zerfallen* des verzerrten Stimulus zu seiner Referenz. „Schallquellen, die erfahrungsgemäß eine einheitliche räumliche Gestalt haben sollten, erscheinen räumlich separiert. […]
  Beispiel: Griffgeräusche und Töne einer Instrumenten-Phantomschallquelle kommen nicht vom selben Ort.“ (Lindau, 2014, S.34)

- **Entfernung**
  Skalenpole: näher - ferner
  Wahrgenommener Unterschied hinsichtlich des Attributes *Entfernung* des verzerrten Stimulus zu seiner Referenz. Die Entfernung beschreibt die „wahrgenommene Distanz einer Schallquelle“ (Lindau, 2014, S.34) oder der Phantomschallquellen.
- **Natürlichkeit**
  Skalenpole: unnatürlicher - natürlicher
  Wahrgenommener Unterschied hinsichtlich des Attributes *Natürlichkeit* des verzerrten Stimulus zu seiner Referenz. Skalenpole bilden hier die Adjektive *unnatürlicher - natürlicher*.
- **Gefallen**
  Skalenpole: gefällt weniger - gefällt mehr
  Bewertung des Attributes *Gefallen* des verzerrten Stimulus zu seiner Referenz.

Bei der Referenz handelt es sich um das ursprüngliche, nicht verzerrte Signal. Außer bei den Bewertungen der Wahrnehmungsqualitäten *Klangfarbe* und *Ausdehnung* wurden als Skalenpole die im SAQI festgelegten bipolaren Adjektive verwendet.

4.6 Versuchsaufbau

Die Lautsprecher wurden in Standard-Stereo-Aufstellung mit einem Öffnungswinkel von 30° aufgestellt (ITU-R BS.1116-3). Bei einer Basisbreite vom 2,65m betrug der Hörabstand zu den Lautsprechern ebenfalls 2,65 m. Damit die Einflüsse des Raumes auf die stereofone Wiedergabe nicht zu gering waren, wurden die Lautsprecher zumindest für die unteren beiden Frequenzpunkte oberhalb der Critical Listening Distance (Abschnitt 3.2) aufgestellt. Basierend auf den Raummaßen und der ermittelten Nachhallzeit, ergibt sich für punktförmige Schallquellen ein Mindestabstand von 1,3 m. Da der Bündelungsgrad des Lautsprechers unbekannt war, wurde zu mittleren bis hohen Frequenzen ein Wert von $\gamma = 4$ (entspricht Hyperniere) angenommen.[8] Mit diesem Wert ergibt sich im Hörstudio ein Hörabstand von mindesten 2,5 m. Zu höheren Frequenzen bündelt der Lautsprecher recht stark (bei 10kHz unter 30 ° etwa -15dB) und sein Abstrahlverhalten würde einer keulenförmigen Richtcharakteristik gerechter werden. Dadurch würde sich der Abstand aber auf 4,29m erhöhen, was zu Platzproblemen im Versuchsaufbau geführt hätte. Zudem ergibt sich daraus die Fragestellung, in wie fern sich bei der hochfrequenten Differenz ein Unterschied in den Wahrnehmungsschwellen zwischen dem RAR und dem

[8] Den Messungen aus Abbildung 10 zur Folge, fällt der Frequenzgang unter 30° ab etwa 4,5kHz (3dB).

Hörstudio ergibt. Die Differenz der Ankunftszeit des Direktschalls zwischen den Lautsprechern lag unter 0,02ms, weshalb der Einfluss interauraler Zeitdifferenzen vernachlässigbar ist.

Die 0° Lautsprecherachsen wurden auf den Gehöreingang der Versuchspersonen gerichtet, was einem Winkel von etwa 60° zur Stereobasis entsprach. Die Probanden saßen auf einem höhenverstellbaren Stuhl der, um ihn gegen versehentliches Verrutschen zu sichern in einer Bodenplatte fixiert war. Der Kopf der Probanden wurde nicht fixiert, denn unbewusste Kopfbewegungen erleichtern das Lokalisieren von Schallquellen sowie die Bewertung räumlicher Attribute und führen zu konsistenteren Ergebnissen (Blauert, 1974; Choisel und Zimmer, 2003; Frank, 2013). Es wurde jedoch der Hinweis gegeben, dass der Kopf, insbesondere im Hörversuchsteil der Wahrnehmungsschwellen, möglichst stillgehalten werden sollte.

Die beiden Versuchsräume unterscheiden sich optisch deutlich voneinander. Die Diskrepanz zwischen dem was gesehen und was gehört wird beeinflusst die Bewertung (Toole, 2005; Linkwitz, 2009). Um die Einflüsse visueller Reize zu reduzieren und die Versuchsräume in ihrer visuellen Wahrnehmung in Blickrichtung der Versuchspersonen ähnlicher zu gestalten, wurde vor den Lautsprechern ein akustisch transparenter, dunkelgrauer Vorhang, mit einer Größe von 300 cm mal 400 cm installiert. So konnten visuelle Asymmetrien des Aufbaus im Hörstudio, durch zum Beispiel Einrichtungsgegenstände, verdeckt werden. Um die Versuchspersonen bei der Bewertung der geometrisch räumlichen Attribute zu unterstützen, wurde auf dem Vorhang die Mitte zwischen den Lautsprechern durch eine weiße Linie markiert. Abbildung 14 zeigt den Aufbau im RAR, mit der Zentrierplatte des Stuhls und dem Vorhang mit Markierung der Phantommitte.

Abbildung 14: Versuchsaufbau im reflexionsarmen Raum.

Die Hörversuche wurden mit Hilfe der an der TU-Berlin programmierten Software *WhisPER 1.9* (Ciba, Brinkmann und Lindau, 2015) durchgeführt. Die auf Matlab-Skripten basierte Software stellt neben den implementierten Versuchsverfahren die Eingabeoberfläche für die Probanden

bereit und fasst die gesammelten Daten in Tabellen und Grafiken zusammen. Für den Hauptversuch wurde die AFC- und die MUSHRA-Oberfläche verwendet. Die MUSHRA-Oberfläche wurde so angepasst, dass auch zweiseitige Bewertungen möglich sind und sich der Nullpunkt in der Mitte der Skala befindet. Abbildung 15 und Abbildung 16 zeigen die Bedienoberfläche für den ersten und den zweiten Teil des Hörversuchs.

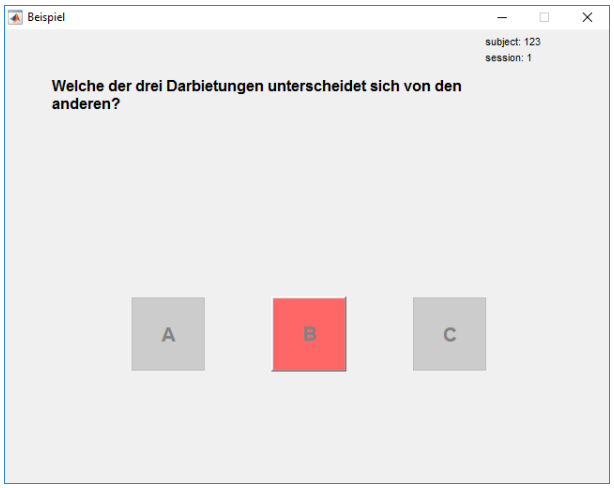


Abbildung 15: Bewertungsinstrument WhisPER. AFC-Oberfläche für die Schätzung der Wahrnehmungsschwellen.

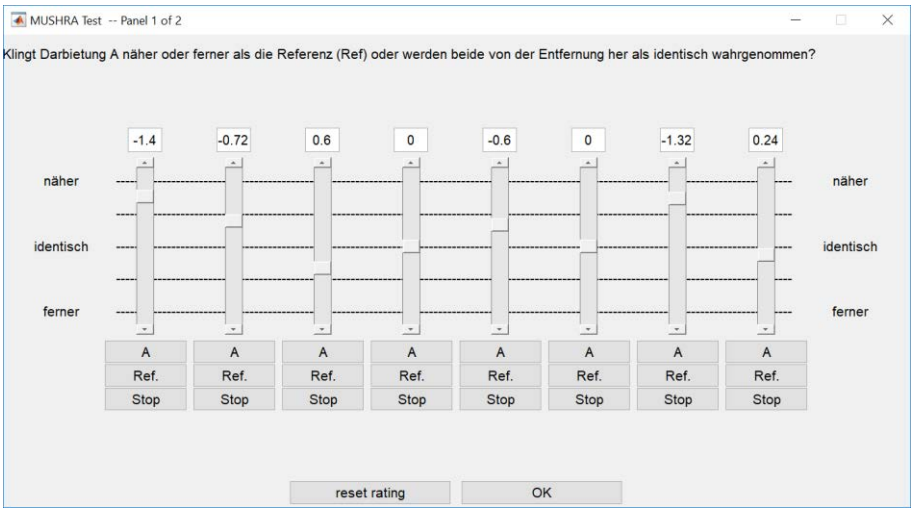


Abbildung 16: Bewertungsinstrument WhisPER. Angepasste MUSHRA-Oberfläche. Erst nach dem Anklicken der Schaltflächen jeder Darbietung wurde die OK-Schaltfläche freigegeben.

Die Eingaben erfolgten über eine Funkmaus und einen LCD-TV, welcher mittig zwischen den Lautsprechern positioniert wurde. Die akustischen Einflüsse des Bildschirms wurden im RAR gemessen und bewegten sich, wie auch die des nahezu akustisch transparenten Vorhanges, in einem vernachlässigbaren Bereich (< 0,2dB). Durch den symmetrischen Aufbau wurde das Schallfeld zudem gleichermaßen beeinflusst.

Die von *WhisPER* abgespielten Signale durchflossen das FIR-Filter in *AudioWeaver* und gelangten über ein USB-Audiointerface (RME Fireface UC) in die Endstufe zu den

Lautsprechern. Zu Verständigungs- und Messzwecken war in jedem Raum ein Mikrofon aufgebaut. Über einen kleinen aktiven Lautsprecher konnte so auf eventuelle Zwischenfragen der Versuchsteilnehmer eingegangen werden.

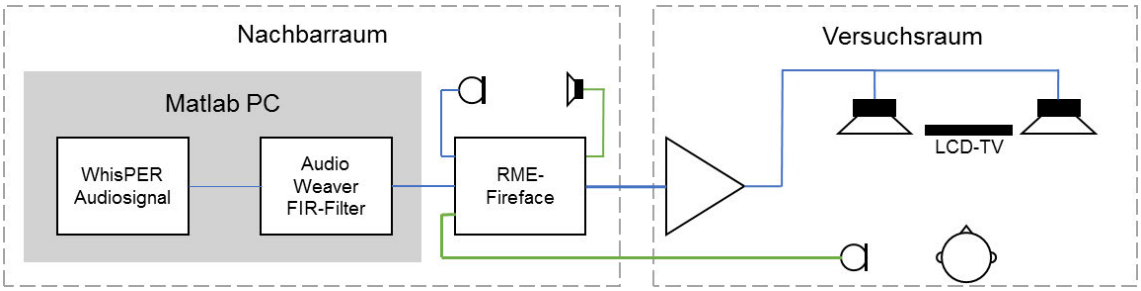


Abbildung 17: Schematischer Versuchsaufbau und Signalfluss.

Abbildung 17 zeigt schematisch den Versuchsaufbau und Signalfluss. In beiden Räumen wurde der Ausgangspegel so gewählt, dass am Hörplatz ein Schalldruckpegel von 75dB(A) (L_{eq} Rosa Rauschen) erreicht wurde.

4.7 Versuchsdurchführung

Über die Hörversuche soll einerseits die Schwelle der Wahrnehmung von physikalisch messbaren Differenzen geschätzt und andererseits ein Zusammenhang zwischen den Differenzen und den subjektiven Hörempfindungen gefunden werden.

4.7.1 Vorversuch

In einem Vorversuch wurde die Methodik zum Schätzen der Wahrnehmungsschwellen und Eigenschaften der qualitativen Erhebung überprüft. Folgende Punkte waren dabei von besonderen Interesse:

- Anzahl der notwendigen Trials für die Schätzung der Wahrnehmungsschwelle über das adaptive QUEST-Verfahren bestimmen.
- Die ermittelten Wahrnehmungsschwellen dienen als Grundlage für die Wahl der psychometrischen Funktion im Hauptversuch.
- Die im Vorversuch ermittelte Wahrnehmungsschwelle, dient als Grundlage für die Höhe der Differenz zwischen den Lautsprechern in dem zweiten Teil des Hauptversuchs.
- Überprüfung, ob Differenzen bei 9397Hz in der qualitativen Erhebung auch bewertet werden sollen.
- Überprüfung, ob die Bewertung aller vier Stimuli in der qualitativen Erhebung erkenntnisgewinnend ist.
- Überprüfung, ob die Attribute und die Beschriftung der Skalenpole sinnvoll gewählt wurde.

4.7.1.1 Versuchsaufbau

Der Aufbau entspricht dem des Hauptversuchs. Allerdings fand der Vorversuch in einem Raum mit einer Grundfläche von 27m² und einem Volumen von 102m³ bei einer Nachhallzeit T_{60} = 0,2s. Dies entspricht im Mittel in etwa den beiden Räumen des Hauptversuches.

4.7.1.2 Stichprobenumfang

Bei den vier Versuchspersonen handelte es sich um männliche Mitarbeiter eines Audiotechnikunternehmens zwischen 26 und 32 Jahren. Drei der vier Teilnehmer waren Hobbymusiker und haben schon mal an einem Hörversuch teilgenommen. Keiner der Teilnehmer hatte zuvor an einem Versuch zu Resonanzschwellen oder ähnlichen Untersuchungen teilgenommen. Alle Personen versicherten keine wissentliche Hörschädigung zu haben. Die Versuche wurden innerhalb von zwei Tagen durchgeführt.

4.7.1.3 Stimuli

Für die Bestimmung der Wahrnehmungsschwellen lagen das Rauschen und die Sprachdarbietung in den in Tabelle 3 gezeigten Reizstärken vor:

Tabelle 3: Abstufung der Reizstärken der Darbietungen im Vorversuch

Reiznummer	Differenz	Reiznummer	Differenz
1	0,5dB	10	5dB
2	1dB	11	6dB
3	1,5dB	12	7dB
4	2dB	13	9dB
5	2,5dB	14	10dB
6	3dB	15	12dB
7	3,5dB	16	16dB
8	4dB	17	24dB
9	4,5dB	18	30dB

Die Abstufungen wurden auf der Grundlage von Ergebnissen eines Eigenversuchsmit mit dem adaptiven Staircase Verfahren gewählt. In dem Bereich der durch den Versuchsleiter erwarteten Schwelle betrug die Schrittweite der Differenzen 0,5dB.[9]

In der qualitativen Erhebung wurde die Differenz aller in 4.1 genannten Stimuli zwischen linkem und rechtem Kanal auf 16dB festgelegt.

4.7.1.4 Versuchsdurchführung

Nachdem die Probanden durch den Versuchsleiter über den Versuchsablauf instruiert worden sind, wurde zunächst der Schwellwertversuch durchgeführt. Die Schätzwerte der Schwellen wurden über das psychometrische adaptive QUEST-Verfahren ermittelt. Nach 40 Trails oder wenn das Konfidenzintervall der Schwellwertschätzung unterhalb die gesetzte Grenze fiel (Konfidenzniveau 95 %), wurde der Messvorgang gestoppt. Die Werte der Parameter für das QUEST-Verfahren richteten sich nach den Empfehlungen der *WhisPER* Dokumentation, welche sich auf Ergebnisse vorheriger Studien stützen (Ciba, Brinkmann und Lindau, 2015).

Nach einer kurzen Pause von ein paar Minuten wurde der zweite Teil des Hörversuches durchgeführt. Dazu wurde das in *WhisPER* implementierte MUSHRA Verfahren genutzt (ITU-R BS.1534-3, 2015). Allerdings wurde aufgrund der gewählten Differenz von 16dB und den damit großen klanglichen Unterschieden, auf einen Anker und eine verstecke Referenz verzichtet.

[9] Dieser wurde unter identischen Bedingungen vor dem Vorversuch durchgeführt. Die Schwelle lag hier im Mittel bei einer Differenz von 4dB.

Die Versuchspersonen haben auf einer Skala von 0-100 Punkten bewerten müssen, wie sehr sich die dargebotene Referenz von dem Signal mit der erzeugten Differenz unterscheidet. Die Skalenpole waren beschriftet mit „identisch“ (100) und „unterschiedlich“ (0). Die Attribute entsprachen den in 4.5.2 genannten. Um Kontexteffekte bezüglich der Reihenfolge der Stimuli auszuschließen, wurden diese vor jedem Durchlauf randomisiert.

4.7.1.5 Ergebnisse und Diskussion: Wahrnehmungsschwellen der Differenzen

Über die durch das QUEST-Verfahren ermittelten Schätzwerte der Wahrnehmungsschwellen der einzelnen Versuchspersonen, wurde für jede Darbietung das arithmetische Mittel gebildet.

Tabelle 4: Arithmetisches Mittel und Standardabweichung der Wahrnehmungsschwellen des Vorversuchs

Stimulus	$\bar{x}$	**s**
Rauschen 293Hz	6,75dB	3,77
Rauschen 2900Hz	6,25dB	4,34
Rauschen 9600Hz	5,75dB	5,56
Sprache 293Hz	11dB	2,6
Sprache 2960Hz	11,5dB	1,7
Sprache 9600Hz	11dB	3,87

Wie aus Tabelle 4 ersichtlich wird, streuen die Antworten beim Rauschen mehr als bei Sprache, allerdings liegen hier auch die Schwellwerte deutlich niedriger. Eine Erklärung hierfür sind die wesentlich gröberen Abstufungen zu hohen Differenzen (siehe Tabelle 3). Die erwarteten Schwellwerte wurden von dem Versuchsleiter zu niedrig angesetzt. Auffällig ist, dass die Unterschiede zwischen den Frequenzpunkten der erzeugten Differenzen recht gering sind, bei Sprache sogar beinahe gleich. Allerdings ist der Grad der statistischen Validität, aufgrund des geringen Stichprobenumfangs und der gewählten Abstufung zu gering, um an dieser Stelle Schlussfolgerungen für den Hauptversuch ziehen zu können.

Durchschnittlich kam es durch das gewählte Konfidenzniveau von 95 % zu einem Abbruch der Sektionen nach 25,13 Trails. In 6 von 24 Fällen wurde erst nach 40 Trails abgebrochen. Dies lag vermutlich vor allem an Zufällen, bei denen zu Beginn der Sektion mehrmals richtig geraten wurde. Abbildung 18 zeigt exemplarisch den Verlauf Schätzfunktion eines Probanden und verdeutlicht den Sachverhalt.

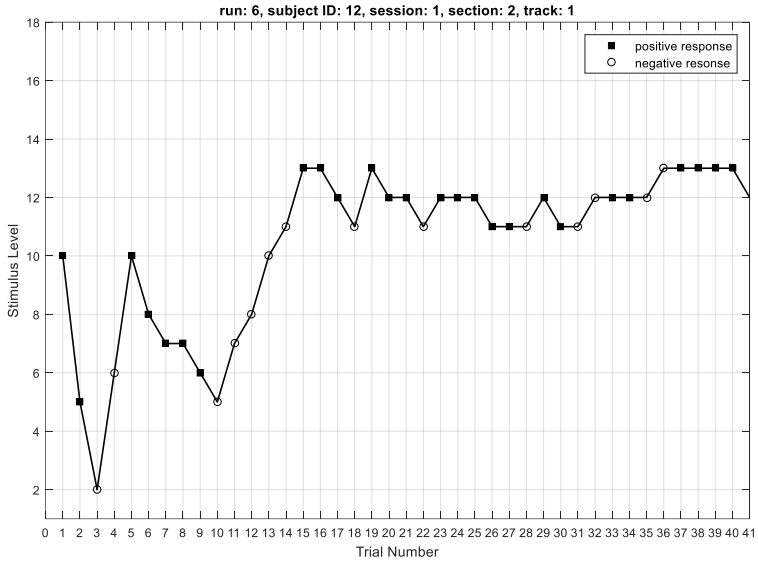


Abbildung 18: Abbruch nach 40 Trails. In den ersten Trails hatte die Versuchsperson offensichtlich mehrmals zufällige Treffer.

In mehreren Durchgängen der Versuchspersonen, bei denen erst nach 40 Trails abgebrochen wurde, ist zudem zu erkennen, dass sich der Schwellwert in den letzten Wiederholungen etwas erhöht (wie auch in Abbildung 18). Die Versuchspersonen beklagten sich am Ende des Versuchs, dass die Konzentration bei der Schwellwertermittlung bei zu vielen Trails ohne Pause nachgelassen habe. Aufgrund der gewonnenen Erkenntnisse wurden für den Hauptversuch folgende Änderungsmaßnahmen ergriffen:

- Für die Schätzung der Wahrnehmungsschwellen wurde auf das effizientere ZEST-Verfahren umgestellt.[10]
- Die Abstufungen der Stimuli wurden auch zu hohen Differenzen hin verringert und gleichmäßige dB-Schrittgrößen gewählt, sodass hier die Vorteile des ZEST-Verfahrens genutzt werden konnten.[11]

4.7.1.6 Ergebnisse und Diskussion: Erhebung des Einflusses auf Wahrnehmungsqualitäten

Aufgrund der geringen Anzahl an Probanden und deren unterschiedlichen Antwortverhalten streuen die Werte um den Median teilweise sehr stark. Dennoch lassen sich Tendenzen aus den Ergebnissen abschätzen, die als ausreichend angenommen wurden, um die in diesem Kapitel Eingehens genannten Fragen zu beantworten.

[10] Zu den Versuchsverfahren im Hauptversuch siehe 4.5.1
[11] Zu den Abstufungen im Hauptversuch siehe 4.1.3

Die Untersuchung von Differenzen bei 9397Hz ist aufgrund der differenten Bewertung gegenüber den anderen Frequenzpunkten sinnvoll. In den meisten Fällen, wurde die erzeugte Differenz bei dem höchsten der drei Frequenzpunkte, als am wenigsten beeinflussend bewertet.

Bei den meisten Attributen wurde der Unterschied zwischen dem verzerrten Rauschsignal und der unverzerrten Referenz am größten bewertet. Demzufolge gilt das Rauschen, wie schon bei der Schwellwertmessung, hier als das kritischste Testsignal. Das Bewertungsverhalten vom Rauschen deckt sich nicht mit dem des flächigen Musikstücks und das der Sprache nicht mit dem der impulsiven Musik. Dies verdeutlicht die Abhängigkeit der Bewertung vom Stimulus, weshalb alle vier Darbietungen im Hauptversuch getestet wurden.

Die verwendeten Skalenpole *identisch* bis *unterschiedlich* haben bei den Attributen *Natürlichkeit* und *Gefallen* für Irritation bei manchen Versuchspersonen gesorgt. Der Vergleich der Stimuli bei der Bewertung der *Natürlichkeit* mittels der genannten Skalenpole wurde als zu abstrakt empfunden. Mit den gewählten Skalenpolen ließ sich nicht herausfinden ob Stimuli mit Differenzen besser gefallen als die Referenz selbst, sondern nur ob sie höchstens gleich gut gefallen. Auch konnte bei den Bewertungen des Attributs *Entfernung* nicht differenziert werden, ob die Phantomschallquellen näher oder ferner wirken. Gäbe es bezüglich der geometrischen Attribute gegensätzliche Ergebnisse zu anderen Studien, wie z.B. dass sich die Lokalisierbarkeit durch spektrale Differenzen vereinfacht, bliebe dies in dem Design des Vorversuchs unentdeckt. Daher fiel die Entscheidung die Skalen aller Attribute, bis auf *Unterschied* und *Klangfarbe* für den Hauptversuch bipolar auszulegen und den Nullpunkt in die Mitte der Skala zu legen.

Im zweiten Versuchsteil lag die spektrale Differenz an allen Frequenzpunkten bei 16dB, rund 10dB über dem Schwellwert des Rauschens und 5dB über dem der Sprache und wurde von den Probanden als ausreichend groß bewertet. Dennoch wurde, um Fehler im Versuchsaufbau sowie Fehleinschätzungen der Versuchspersonen zu erkennen und ausschließen, für den Hauptversuch im Bewertungsverfahren eine verstecke Referenz (*Hidden Reference* - HR) hinterlegt.

4.7.2 Hauptversuch

Die Hörversuche im Hörstudio und dem reflexionsarmen Raum fanden an unterschiedlichen Tagen statt, sodass für wiederkehrende Versuchspersonen mehrere Tage zwischen den Versuchen in den Räumen lagen. Um mögliche Reihenfolge- oder Trainingseffekte zu berücksichtigen, war die eine Hälfte der Teilnehmer zuerst im RAR und die andere Hälfte zuerst im Hörstudio.

Nach der Begrüßung haben die Versuchsteilnehmer schriftliche Informationen zum Versuchsablauf, den Definitionen der Wahrnehmungsqualitäten und einen Probandenfragebogen

vorgelegt bekommen.[12] Anschließend sind offene Fragen, insbesondere zu den Wahrnehmungsqualitäten, beantwortet worden. Die Instruktionen wurden für alle Teilnehmer möglichst gleich ausgeführt und nur so viele Informationen zu dem Versuch weitergegeben wie notwendig. Den Teilnehmern ist nicht mitgeteilt worden an welchen Frequenzpunkten genau die Differenzen erzeugt wurden und welche Ausprägung sie hatten. Danach wurden Teilnehmer in den Versuchsraum geführt und der Stuhl auf die korrekte Höhe gebracht. Dazu wurde ein Laserdistanzmessgerät an den Phase-Plug des Lautsprechers gehalten und die sitzende Person hat die Stuhlhöhe auf Zuruf so angepasst, dass der Laserpunkt sich am Gehöreingang befand. Um den Versuchsteilnehmern die Bedienoberfläche zu erklären, wurde vorab ein Probedurchlauf des 3-AFC Tests mit drei Wiederholungen gemacht. Um die Funktionsweise zu verdeutlichen wurden, neben dem unbearbeiteten Signal, Stimuli mit der extremsten Differenz von 28dB abgespielt. Nach Abschluss des ersten Versuchsteils wurden die Teilnehmer gefragt, ob sie eine kurze Pause machen möchten, die meisten stimmten hier zu. Anschließend wurde die Funktionsweise der Bedienoberfläche des zweiten Versuchsteils erklärt und nach der Beantwortung möglicher offener Fragen begann der Versuch. Hier wurde auch der ausdrückliche Hinweis gegeben auf die Beschriftung der Skala zu achten. Der Versuchsleiter befand sich während der Versuche im Nachbarraum. Bei Fragen oder Problemen konnten die Teilnehmer über ein Mikrofon mit dem Versuchsleiter kommunizieren. Im Versuchsraum lag eine Liste mit den Begriffsdefinitionen aus, sodass die Probanden diese bei Unklarheiten jederzeit einsehen konnten. Die Versuchszeit hing stark vom zweiten Teil ab, in dem die Teilnehmer die Stimuli so oft anhören konnten wie sie wollten. Im Durchschnitt lag die Versuchszeit bei etwa 65 Minuten.

[12] Die vorgelegten Informationen und der Fragebogen finden sich im Anhang.

5 ERGEBNISSE

5.1 Wahrnehmungsschwellen

Zunächst wurden die Daten visuell über ein Quantil-Quantil-Diagramm, wie in Abbildung 19 gezeigt, und anschließend mittels des Shapiro-Wilk-Tests auf Normalverteilung geprüft. Bei nur geringen Abweichungen der Punkte von der Geraden, kann eine Normalverteilung der Daten angenommen werden. Allerdings konnte über dieses optische Verfahren bei den betrachteten 12 Datensätzen nicht eindeutig festgestellt werden, ob die Wahrscheinlichkeit einer Normalverteilung in allen Fällen ausreichend groß ist.

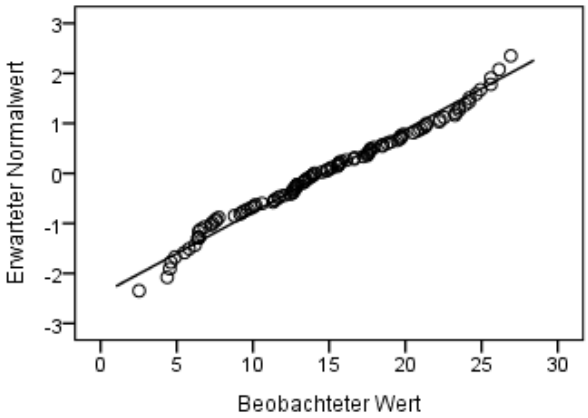


Abbildung 19: Beispiel eines Quantil-Quantil-Diagramm zum Vergleichen der Quantile der Daten aller Wahrnehmungsschwellen bei 293Hz gegen die theoretischen Quantile einer Normalverteilung.

Daher wurde anschließend mit dem Shapiro-Wilk-Test ein strengerer Test auf Normalverteilung durchgeführt. Bei einem Signifikanzniveau von $\alpha = 5\%$ musste, bis auf einem Fall, die Nullhypothese, dass die Daten normalverteilt sind, beibehalten werden. Tabelle 5 zeigt die Ergebnisse des Tests. Die rot markierte Zahl kennzeichnet den Fall bei dem keine Normalverteilung der Daten angenommen werden kann.

Tabelle 5: Ergebnisse des Shapiro-Wilk-Tests auf Normalverteilung.

Raum	Stimulus	Signifikanz	Raum	Stimulus	Signifikanz
RAR	Rauschen 293Hz	0.472	HS	Rauschen 293Hz	0.09
RAR	Rauschen 2637Hz	0.059	HS	Rauschen 2637Hz	0.186
RAR	Rauschen 9397Hz	0.767	HS	Rauschen 9397Hz	0.985
RAR	Sprache 293Hz	0.304	HS	Sprache 293Hz	0.639
RAR	Sprache 2637Hz	0.707	HS	Sprache 2637Hz	0.67
RAR	Sprache 9397Hz	0.035	HS	Sprache 9397Hz	0.071

Bei Durchführung des Tests nach Kolmogorov-Smirnov mit einer Signifikanzkorrektur nach Lilliefors, musste die Nullhypothese für den Datensatz *RAR Sprache 9397Hz* hingegen ebenfalls beibehalten werden (p=0.053). Da die ANOVA mit Messwiederholung zudem relativ robust gegenüber Verletzungen der Normalverteilungsannahme ist (Pagano, 2013, S.398), wurde angenommen, dass alle Daten normalverteilt sind.

Mit den verbundenen Gruppen und normalverteilten Daten, wurde eine dreifaktorielle Varianzanalyse mit Messwiederholung auf allen Faktoren, eine rmANOVA (von **r**epeated **m**easures **an**alysis **of va**riance), durchgeführt. Als Faktoren wurden der *Raum* (zwei Stufen), die *Signalart* (zwei Stufen) und die *Frequenz* (drei Stufen) festgelegt. Ausreißer waren nicht vorhanden. Der Mauchly-Tests auf Sphärizität zeigte kein signifikantes Ergebnis zwischen den Gruppen ($p<0.05$), die Sphärizitätsannahme war somit nicht verletzt und Homoskedastizität zwischen allen Stufen gegeben.

5.1.1 Haupteffekte

Die Nullhypothese der rmANOVA nimmt an, dass es keinen Unterschied zwischen den Mittelwerten der Gruppen gibt ($\alpha = 5\ \%$). Das Ergebnis der für die Anpassung des α-Niveaus mit Bonferronikorrektur durchgeführten rmANOVA, zeigt keinen statistisch signifikanten Unterschied der Wahrnehmungsschwellen von spektralen Differenzen zwischen den Lautsprechern im RAR oder Hörstudio ($F(1,16)=1.5$, $p=0.239$). Ein Überblick über das Streuungsmaß, die Minima und Maxima sowie die Lage der Mittelwerte liefert der in Abbildung 20 gezeigte Boxplot.

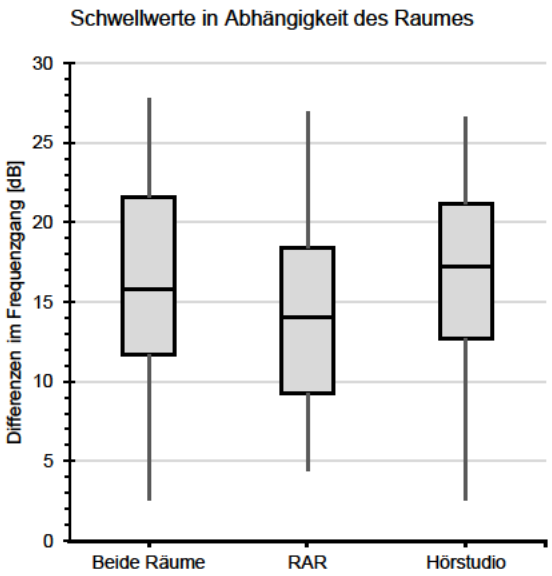


Abbildung 20: Boxplot der Wahrnehmungsschwellen in den beiden Räumen unter Einbeziehung der Schwellen beider Darbietungen und der drei Frequenzpunkte.

Dieser veranschaulicht, dass die Quantile der Wahrnehmungsschwellen im Hörstudio höher als die im RAR liegen. Die Spanne des 25%-Quantil zum 75%-Quantil ist in beiden Fällen ähnlich, das Delta der Standardabweichung beider Räume liegt bei 0,64dB. Im Hörstudio liegt das Minimum der Wahrnehmungsschwelle von Differenzen im Frequenzgang mit 2,52dB rund 1,86dB niedriger als im RAR. Der grand Median aller Messungen (*Beide Räume*) liegt erwartungsgemäß dazwischen.

Die verwendete Signalart, ob Rauschen oder Sprache, hat einen statistisch signifikanten Einfluss auf die Wahrnehmungsschwelle von spektralen Differenzen ($F(1,16)=112$, $p<0.001$, partielles $\eta^2=0.88$). Die Standardabweichung und die Spannweite zwischen dem ersten und dritten Quantil, sind unter Verwendung eines Rauschens ($s = 6{,}07\text{dB}\ x_{0.75} - x_{0.25} = 9{,}4\text{dB}$) größer als unter Verwendung von Sprache ($s = 5{,}41\text{dB}\ x_{0.75} - x_{0.25} = 8{,}23\text{dB}$). Der Median liegt mit 14,05dB um 1,62dB niedriger als bei Sprache. Die minimale Wahrnehmungsschwelle liegt bei Rauschen mit 2,52dB um 1,85dB niedriger als bei Sprache.

Wie bei der Signalart ist der Haupteffekt im Faktor *Frequenz* wirksam: Der Frequenzbereich in dem die Differenzen erzeugt werden, hat einen signifikanten Einfluss auf die Wahrnehmungsschwelle ($F(2,32)=22.9$, $p<0.001$, partielles $\eta^2=0.59$). Es besteht somit ein Unterschied der durchschnittlichen Wahrnehmungsschwellen zwischen mindestens zwei der Frequenzpunkte. Um herauszufinden welche der Punkte sich signifikant voneinander unterscheiden, wurde abschließend, zur Vermeidung eines Typ-1-Fehlers, ein Bonferroni-korrigierter Post-Hoc-Mehrfachvergleichs-Test durchgeführt. Dieser zeigte, dass es zwischen den Frequenzpunkten 293Hz und 2637Hz (Δ 5.62, 95 % - KI[Δ 3.46, 7.8]) sowie 9397Hz und 2637Hz (4.9, 95 % - KI[2.56, 7.21]), einen signifikanten Unterschied der Wahrnehmungs-schwellen gibt.

5.1.2 Interaktionseffekte

Zwischen den Faktoren gibt es nur in einem Fall einen Wechselwirkungseffekt: Die Wahrnehmungsschwelle spektraler Differenzen in der Lautsprecherstereofonie ist abhängig von der Faktorstufenkombination *Signalart* und *Frequenz* ($F(2,32)=15.3$, $p<0.001$, partielles $\eta^2=0.49$), aber nicht vom Raum (Raum*Frequenz: $p=0.18$, Raum*Signalart: $p=0.233$).

In den Abbildung 21, Abbildung 22 und Abbildung 23 sind Boxplots zu den frequenzabhängigen Wahrnehmungsschwellen abgebildet. Der Median der Schwellwerte lag bei den Differenzen an den Frequenzpunkten von 2637Hz mit 12,67dB am niedrigsten und mit 18,27dB bei 293Hz am höchsten. Knapp darunter lag, mit einem Median von 17,59dB, die durchschnittliche Wahrnehmungsschwelle von Differenzen bei 9397Hz. Die Streuung der Schwellwerte ist bei 9397Hz am größten. Ein Grund dafür ist der, gegenüber den anderen Frequenzpunkten, größere Unterschied der Schwellwerte zwischen den beiden Räumen.

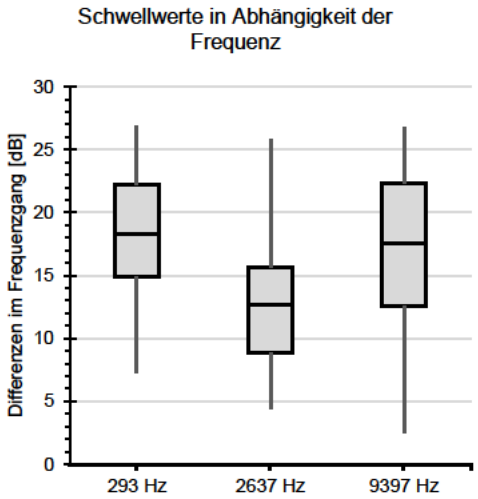


Abbildung 21: Boxplot der Wahrnehmungsschwellen von Differenzen an den drei verwendeten Frequenzpunkten unter Einbeziehung der Schwellen beider Räume und Stimuli.

Dies zeigen die Boxplots in Abbildung 22 und Abbildung 23, welche in aufgeschlüsselter Form die Abhängigkeit der Wahrnehmungsschwellen von der Signalart und der Frequenz veranschaulichen. Beim Betrachten der Mediane wird zudem deutlich, dass bei 9,3kHz und Rauschen als Signalart, die durchschnittlichen Wahrnehmungsschwellen auf gleichem Niveau liegen. Räumliche Einflüsse haben an dieser Stelle die geringsten Auswirkungen. Der Median des Schwellwertes bei 293Hz lag mit 19,7dB (*Rauschen*) und 20dB (*Sprache*) im Hörstudio am höchsten.

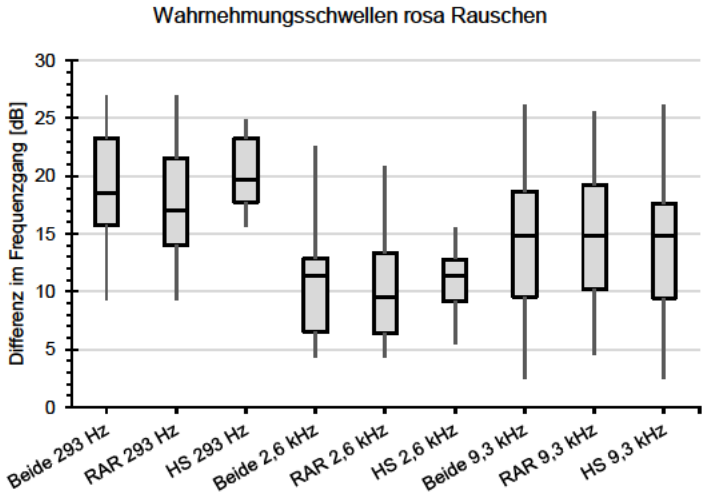


Abbildung 22: Boxplot der Wahrnehmungsschwellen in Abhängigkeit von *Frequenz* und *Raum* (HS=Hörstudio) bei Verwendung von „Rosa Rauschen“ als Stimulus.

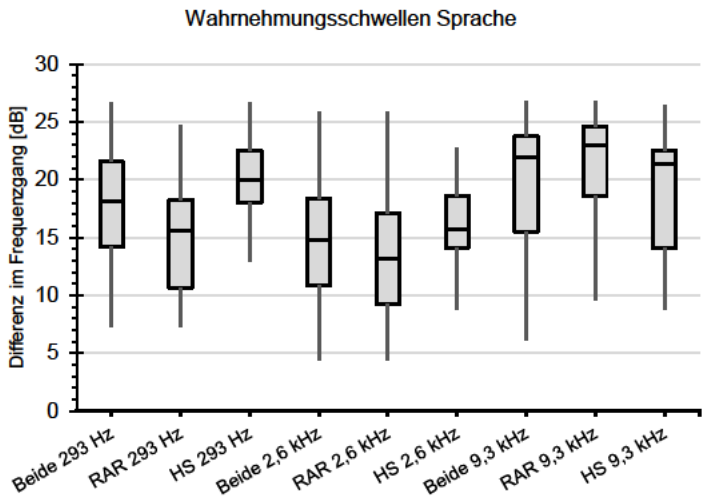


Abbildung 23: Boxplot der Wahrnehmungsschwellen in Abhängigkeit von Frequenz und Raum (HS=Hörstudio) bei Verwendung von Sprache als Stimulus.

Tabelle 6 zeigt die pro Frequenz niedrigsten Durchschnittswerte und Minima der Wahrnehmungsschwellen. Bis auf das Minimum bei 9397Hz, liegen die im Versuch festgestellten, niedrigsten Wahrnehmungsschwellen und der Median im RAR stets tiefer als im Hörstudio. Ein, aufgrund der Lage der Mediane, einzeln durchgeführter gepaarter t-Test zeigte, dass sich die Wahrnehmungsschwellen zwischen den Räumen nur bei 293Hz und der Signalart *Sprache* signifikant unterscheiden, $t(16)=-2{,}77$, $p=0{,}014$, $|d| = 0{,}67$ (mittlerer Effekt nach: Cohen, 1988).

Tabelle 6: Auflistung der niedrigsten Durchschnittswerte und Minima der Wahrnehmungsschwellen von Differenzen im Frequenzgang unter Angabe der Merkmalsausprägungen unter denen diese erhoben wurden.

Frequenz	**293Hz**	**2637Hz**	**9397Hz**
Minimum	7,24dB	4,4dB	2,5dB
Raum **Stimulus**	RAR Sprache	RAR Rauschen & Sprache	RAR Sprache
Median	15,6 dB	9,5dB	13,2dB
Raum **Stimulus**	RAR Sprache	RAR Rauschen	Studio Rauschen

5.1.3 Abhängigkeit der Wahrnehmungsschwellen vom Alter oder der Hörerfahrung

In den nach dem Hörversuch ausgeteilten Fragebögen wurden unter anderem das Alter und der aktive Umgang mit Audiotechnik, Akustik oder Musik in Stunden pro Tag abgefragt. Im Folgenden wird die Stärke des Zusammenhangs zwischen den beiden unabhängigen Variablen (UV) *Alter* und *aktiver Umgang* sowie der abhängigen Variable, der Wahrnehmungsschwelle, quantifiziert. Da es sich bei allen Variablen um metrisch skalierte Merkmale handelt, wurde der Korrelationskoeffizient nach Bravais/Pearson berechnet. Dieser geht von einem linearen Zusammenhang der Daten aus, also, dass die Wahrnehmungsschwellen mit steigendem Alter zu-, abnehmen oder beispielsweise mit wachsender Hörerfahrung sinken (Pagano, 2009). Die numerische Auswertung ergab einen Korrelationskoeffizienten von r = -0,05 beim *Alter*, was einem Determinationskoeffizienten von R^2 = 0,0025 entspricht. Bei der unabhängigen Variable (UV) *aktiver Umgang* mit Audio, Akustik oder Musik in Stunden pro Tag ergab sich ein r = 0,04, was einem Determinationskoeffizienten von R^2 = 0,0019, entspricht. In beiden Fällen liegt somit ein nur äußerst geringer Zusammenhang vor. Abbildung 24 verdeutlicht diesen Sachverhalt. In den beiden Graphen sind der grand Median und die Schwellwerte jeweils in Abhängigkeit vom Alter oder dem aktiven Umgang mit Audiotechnik, Akustik oder Musik in Stunden pro Tag abgebildet. Die gepunktete Linie stellt die Regressionsgerade der Schwellwerte dar. In Abhängigkeit vom Alter lässt sich kein Trend erkennen. In Abhängigkeit vom aktiven Umgang zeichnet sich ein schwacher Trend in Richtung Verringerung der Wahrnehmungsschwelle mit zunehmenden aktiven Umgang in Stunden pro Tag ab.

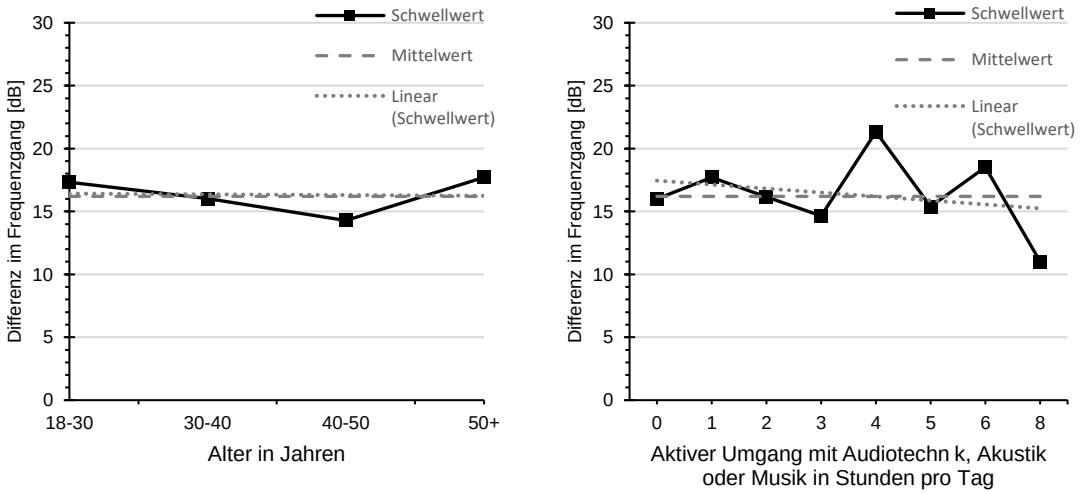


Abbildung 24: Zusammenhang der Wahrnehmungsschwellen der Differenzen und dem Alter (links) oder dem aktiven Umgang mit Audiotechnik, Akustik oder Musik in Stunden pro Tag (rechts).

5.2 Einfluss auf Wahrnehmungsqualitäten

Durch die Anzahl der abhängigen und unabhängigen Variablen ergaben sich, inklusive der versteckten Referenz (**H**idden **R**eference, HR), 128 zu bewertende Darbietungen. Über beide Räume sind 4096 Bewertungen abgegeben worden. Den ordinalskalierten Wahrnehmungsqualitäten wurden über die fünfstufige Ratingskala Zahlen als Ergebniswert zugewiesen.

5.2.1 Aufbereitung der Daten

Zunächst wurden die Bewertungen in Matlab, von der durch WhisPER vorgegebenen kontinuierlichen Skalierung der Werte von 0 bis 1, in den Wertebereich 0 bis 4 für die unipolaren Attribute und -2 bis 2 für die bipolaren Attribute transformiert. Über eine Matlab-Funktion wurden alle Bewertungen, in dem die HR mit mehr als 20% gewertet wurde (also >1 im unipolaren Fall und >±0,5 im bipolaren Fall), gekennzeichnet. Die dem Stimulus zugehörigen Darbietungen wurden bei Verletzung des Kriteriums innerhalb des Attributs ebenfalls gekennzeichnet. Ist beispielsweise beim Attribut *Unterschied*, die HR vom Stimulus *Rauschen* mit 1,3 bewertet worden, sind die drei weiteren Darbietungen des Rauschens, innerhalb des Attributs *Unterschied*, gekennzeichnet worden. Alle anderen Bewertungen wurden, sofern das Kriterium nicht verletzt wurde, nicht markiert. Nach der tabellarischen Aufbereitung der Daten wurden diese in das Statistikprogramm *SPSS* importiert. Durch die Markierung der gekennzeichneten Fälle als *fehlend*, sind diese in späteren Untersuchungen nicht berücksichtigt worden. Tabelle 7 zeigt die Summen der Häufigkeiten, in denen die HR auf der Skala mit mindestens 20%, bzw. ±10% bewertet wurde.

Tabelle 7: Prozentuale Häufigkeiten in denen die versteckte Referenz höher als das Kriterium es zulässt bewertet wurde.

	Rauschen	Sprache	Musik impulsiv	Musik flächig	Summe
Unterschied	28,1	9,4	25,0	37,5	**25,0**
Klangfarbe	28,1	3,1	9,4	31,3	**18,0**
Lokalisierbarkeit	6,3	3,1	15,6	31,3	**14,1**
Ausdehnung	9,4	6,3	12,5	34,4	**15,6**
Räum. Zerfallen	6,3	3,1	12,5	21,9	**10,9**
Entfernung	12,5	3,1	15,6	21,9	**13,3**
Natürlichkeit	15,6	6,3	9,4	18,8	**12,5**
Gefallen	6,3	6,3	6,3	12,5	**7,8**
Summe	**14,1**	**5,1**	**13,3**	**26,2**	

Auffällig ist, dass insbesondere bei den unipolaren Attributen *Unterschied* und *Klangfarbe*, bei denen der Nullpunkt nicht in der Mitte lag, die Häufigkeit der Verletzung des Kriteriums am höchsten ist. Trotz des Hinweises zu Beginn des zweiten Teils des Hörversuchs, stets auf die Skalierung und die Verschiebung des Nullpunktes bei den beiden Attributen zu achten, haben vermutlich vereinzelte Versuchsteilnehmer die Mitte der Skala als Nullpunkt gewählt (entspricht einer Wertung von 2). Dies geht einerseits aus Anmerkungen der Teilnehmer hervor, die den Fehler während des Versuchs selbst bemerkt haben und zum anderen aus dem Antwortverhalten darauffolgender Bewertungen, in dem die HR von manchen Probanden stets als solche erkannt wurde. Mit durchschnittlich 26,2% war die Unsicherheit einen Unterschied wahrgenommen zu haben beim Stimulus *Flächige Musik* am größten, am sichersten waren sich die Probanden mit 5,1% beim Stimulus *Sprache*.

5.2.2 Korrelationen der Wahrnehmungsqualitäten

Ein Shapiro-Wilk Test auf Normalverteilung zeigte, dass die Bewertungen der Attribute in allen Fällen nicht normalverteilt sind. Für die Korrelationsanalyse wurde daher die Methode nach Spearman verwendet, welche sich nicht auf normalverteilte Daten beschränkt und nichtlineare Zusammenhänge erkennt. Tabelle 8 zeigt die Wahrscheinlichkeiten für Korrelationen der bipolaren Attribute bei einer maximal zulässigen Irrtumswahrscheinlichkeit von $\alpha = 5\%$[13].

Tabelle 8: Korrelationskoeffizienten der bipolaren Attribute ρ (rho) nach Spearman. Signifikante Korrelationen sind grün markiert.

		Lokalisierbk.	Ausdehnung	R.Zerfallen	Entfernung	Natürlichkeit	Gefallen
Lokalisation	r_s	1.000	.299	.350	-.018	.438	.423
	p		.000	.000	.765	.000	.000
Ausdehnung	r_s	.299	1.000	.328	.005	.169	.098
	p	.000		.000	.934	.005	.096
R.Zerfallen	r_s	.350	.328	1.000	-.008	.318	.389
	p	.000	.000		.890	.000	.000
Entfernung	r_s	-.018	.005	-.008	1.000	-.065	-.125
	p	.765	.934	.890		.273	.034
Natürlichkeit	r_s	.438	.169	.318	-.065	1.000	.581
	p	.000	.005	.000	.273		.000
Gefallen	r_s	.423	.098	.389	-.125	.581	1.000
	p	.000	.096	.000	.034	.000	

[13] Für alle folgenden statistischen Analysen dieser Arbeit wurde ein Signifikanzniveau $\alpha = 5\%$ angenommen.

Die unipolaren Attribute *Unterschied* und *Klangfarbe* korrelieren ebenfalls miteinander (r_s=0.598, p<0.001). Wie in der vorliegenden Arbeit erläutert, geht eine verminderte Abbildungsschärfe mit einer geringeren Lokalisierbarkeit, erhöhter Ausdehnung und räumlichen Zerfallen von Phantomschallquellen einher. Die Skalenpole der geometrischen Wahrnehmungsqualitäten wurden so beschriftet, dass mit einer Bewertung im positiven Zahlenbereich (0 bis 2) Abbildungsschärfe mindernde Eindrücke der Probanden abgebildet werden:

- Lokalisation: schwieriger lokalisierbar
- Ausdehnung: ausgedehnter
- Räumliches Zerfallen: zerfallener
- Entfernung: näher
- Natürlichkeit: unnatürlicher
- Gefallen: gefällt weniger

Die signifikante Korrelation aller räumlichen Wahrnehmungsqualitäten zeigt, dass das Vorhandensein einer spektralen Differenz, unter Berücksichtigung beider Räume und aller Stimuli, zu einer verminderten Abbildungsschärfe führt. Zwischen *Lokalisierbarkeit*, *Natürlichkeit* und *Gefallen* besteht der größte, positive lineare Zusammenhang. Dies lässt die Schlussfolgerung zu, dass eine stereofone Wiedergabe mit schwieriger zu lokalisierenden Phantomschallquellen, in einer unnatürlicheren und dadurch weniger gefallenden Wiedergabe resultiert. Bei der wahrgenommenen Entfernung von Phantomschallquellen kann kein oder ein nur sehr geringer Zusammenhang mit den anderen Attributen ausgemacht werden.

5.2.3 Haupt- und Interaktionseffekte

Für die statistische Analyse wird ein Mehrebenenmodell angewandt. Nur einzelne fehlende Fälle führen bei Messwiederholungs-Varianzanalysen zu einem listenweisen Fallausschluss, das heißt die Versuchsperson wird inklusive aller Messwiederholungen in der Analyse nicht berücksichtigt. Dies würde bedeuten, dass sich der Stichprobenumfang erheblich verringern würde, da nur eine geringe Anzahl an Versuchspersonen die versteckte Referenz in keinem von 32 Fällen bewertet haben. Die Bedingung der Sphärizität muss hier nicht erfüllt sein, die Varianz-Kovarianz-Struktur der Daten kann frei gewählt werden. Allerdings ist bei der unrestriktivsten Kovarianzstruktur (alle Varianzen und Kovarianzen verschieden) der Modellfit aufgrund der hohen Anzahl der zu schätzenden Parametern am geringsten (IBM, 2013). Eine ANOVA mit Messwiederholung geht von einer zusammengesetzt symmetrischen Kovarianzstruktur aus, diese Annahme wurde für die Analysen der Daten dieser Arbeit im Mehrebenenmodell ebenfalls getroffen. Das Akaike-Informationskriterium (AIC) erreicht bei dem Kovarianztyp zusammengesetzt symmetrisch einen

Wert von AIC = 757,65. Im explorativen Vergleich mit anderen Modellen, wurde der Modellfit unter der Verwendung dieses Kovarianztyps als sehr hoch eingeschätzt. Für die Untersuchung von Haupt- und Interaktionseffekten wurden die Daten in longitudinale Form gebracht.

Unter Einbeziehung der unipolar skalierten Attribute *Unterschied* und *Klangfarbe* zeigt der Typ I der festen Effekte, dass die Stufen der Faktoren *Signalart* ($F(3,528)=11.85$, $p<0.001$) und *Frequenz* ($F(2,528)=38.82$, $p<0.001$) einen signifikanten Effekt auf die Bewertung haben. Der *Raum* ($F(1,528)=1.19$, $p=0.275$) und das *Attribut* ($F(1,528)=2.5$, $p=0.114$) jedoch nicht. Die Differenz der Mittelwerte der Attribute *Unterschied* und *Klangfarbe* liegt bei Berücksichtigung aller Stimuli und unter Einbeziehung der Bewertung in beiden Räumen bei 0,1. Wie Tabelle 9 zeigt, hat unter Einbeziehung aller bipolar skalierten Attribute, mindestens ein Attribut einen signifikanten Effekt auf die Bewertung ($F(5,1755)=12.8$, $p<0001$). Der Faktor Frequenz hat den größten Einfluss. Außerdem zeigen sich Interaktionseffekte zwischen den Faktoren *Attribut* und *Frequenz* sowie *Signal* und *Frequenz.*

Tabelle 9: Tests auf feste Effekte aller bipolaren Attribute.

Quelle	Zähler-Freiheitsgrade	Nenner-Freiheitsgrade	F-Wert	Signifikanz
Attribut	5	1755	12,816	,000
Raum	1	1755	,274	,600
Signal	3	1755	7,680	,000
Frequenz	2	1755	33,667	,000
Attribut * Raum	5	1755	,937	,456
Attribut * Signal	15	1755	1,092	,358
Attribut * Frequenz	10	1755	7,659	,000
Raum * Signal	3	1755	,176	,912
Raum * Frequenz	2	1755	2,480	,084
Signal * Frequenz	6	1755	6,897	,000
Attribut * Raum * Signal	15	1755	,546	,916
Attribut * Raum * Frequenz	10	1755	,990	,450
Attribut * Signal * Frequenz	30	1755	1,136	,280
Raum * Signal * Frequenz	6	1755	,982	,436
Attribut * Raum * Signal * Frequenz	30	1755	,712	,874

5.2.3.1 Einfluss der Frequenz der spektralen Differenz auf die Wahrnehmungsqualitäten

Die Analyse der Daten durch das Mehrebenenmodell zeigt, dass, außer bei den Attributen *räumliches Zerfallen* und *Entfernung*, die Stufung des Faktors *Frequenz* stets signifikante

Unterschiede in der Bewertung der Qualitäten bewirkt. Am größten ist der F-Wert bei dem Attribut *Natürlichkeit* ($F(2,294)=34.2$; $p<0.001$), gefolgt von *Gefallen* ($F(2,309)= 24.47$; $p<0.001$) und *Unterschied* ($F(2,252)= 21.3$; $p<0.001$).[14] Bonferroni korrigierte Post-hoc Tests zeigten bei den Attributen *Natürlichkeit* und *Gefallen* signifikante Unterschiede zwischen allen Frequenzpunkten ($p<0.001$), bei dem Attribut *Unterschied* nur zwischen den Frequenzpunkten 2637Hz und 293Hz, bzw. 9397Hz ($p<0.001$). Demzufolge nimmt der Frequenzbereich, in dem die spektrale Differenz zwischen zwei Lautsprechern liegt, im unterschiedlichen Maße Einfluss auf die *Natürlichkeit*, das *Gefallen*, die Größe des wahrgenommenen *Unterschieds*, die *Klangfarbe*, die *Lokalisation* und *Ausdehnung* von Phantomschallquellen.

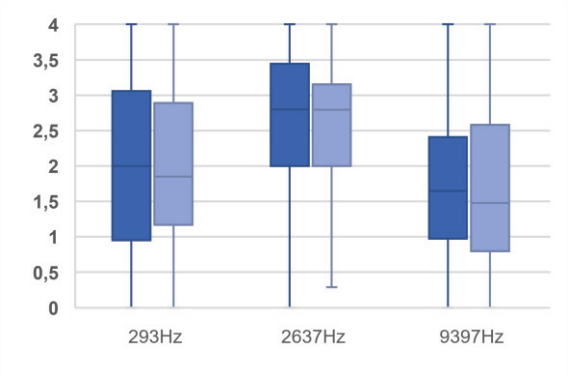
Die Boxplots in Abbildung 25 verdeutlichen den Zusammenhang der Frequenz der spektralen Differenz und der Bewertung der Wahrnehmungsqualitäten. Aus Gründen der Übersichtlichkeit finden sich die Beschriftungen der Ordinaten der beiden Boxplots in Tabelle 10. Die Mediane aller Attribute zeigen die Tendenz, dass der Einfluss spektraler Differenzen zu hoher und niedriger Frequenz geringer ist als zu der mittleren. Die Standardabweichung der arithmetischen Mittel der Attribute, ist bei 2637Hz mit $s = 0{,}39$ am höchsten, während sie bei 9397Hz mit $s = 0{,}1$ am niedrigsten ist. Dies liegt an dem geringen Einfluss der spektralen Differenz bei 9397Hz auf die abgefragten Attribute.

Bei 293Hz variieren die Mittelwerte der Bewertungen zwischen den Attributen kaum und liegen dabei, bis auf bei dem Attribut *Entfernung,* im positiven Bereich, was einer Verminderung der Abbildungsschärfe gleichkommt.

Bei 2637Hz ist der Einfluss einer spektralen Differenz auf die stereofone Abbildung am größten. Diese Erkenntnis deckt sich mit den Ergebnissen des ersten Teils, in dem festgestellt wurde, dass die Wahrnehmungsschwelle bei 2637Hz am geringsten ist. Die Wiedergabe gefällt hier gegenüber den anderen Frequenzpunkten am wenigsten und gilt als am unnatürlichsten.

Bei 9397Hz ist, trotz des beinahe mittleren Unterschieds durch die spektrale Differenz, der Einfluss auf die abgefragten Wahrnehmungsqualitäten gering (<0.5 auf der Bewertungsskala) oder im Mittel gar nicht vorhanden. Insgesamt ist der Einfluss von Differenzen im Frequenzgang zweier Lautsprecher bei 9397Hz am geringsten.

[14] Anmerkung: Die Höhe der Freiheitsgrade variiert aufgrund der unterschiedlichen Anzahl ausgeschlossener Fälle.



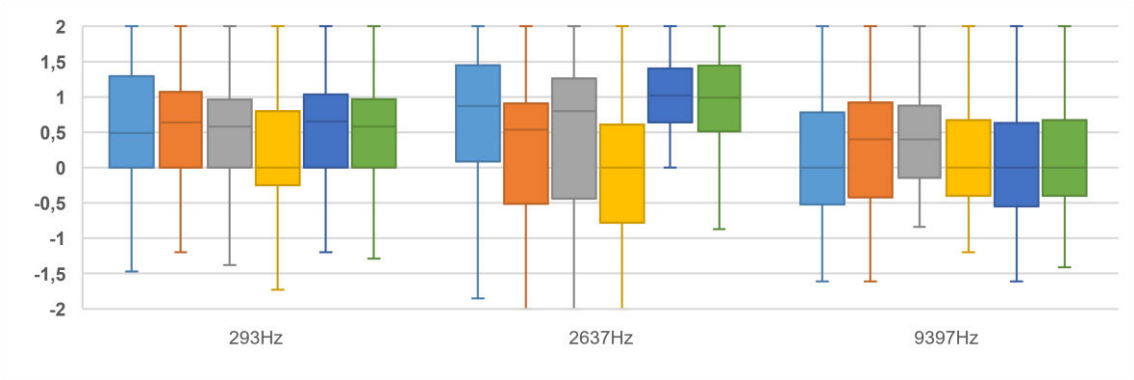


Abbildung 25: Boxplots der Bewertungen in Abhängigkeit des Attributes und der Frequenz. **Oben**: Ergebnisse der einseitigen Bewertungen. **Unten**: Ergebnisse der zweiseitigen Bewertungen.

Tabelle 10: Skalenpole der Wahrnehmungsqualitäten.

Einseitige Bewertung	**0**	**2**	**4**
Unterschied	gar keiner	mittlerer	sehr großer
Klangfarbe	gar keiner	mittlerer	sehr großer
Zweiseitige Bewertung	**- 2**	**0**	**2**
Lokalisierbarkeit	einfacher lokalisierbar	identisch	schwieriger lokalisierbar
Ausdehnung	kompakter	identisch	ausgedehnter
Räumliches Zerfallen	fusionierter	identisch	zerfallener
Entfernung	ferner	identisch	näher
Natürlichkeit	natürlicher	identisch	unnatürlicher
Gefallen	gefällt mehr	identisch	gefällt weniger

5.2.3.2 Einfluss der Signalart bei einer spektralen Differenz auf die Wahrnehmungsqualitäten

Die Bewertung der Attribute *Unterschied* ($F(3,252)=5.28$ $p=0.002$), *Klangfarbe* ($F(3,276)=6.89$ $p<0.001$), *Lokalisation* ($F(3,282)=2,7$ $p=0.044$) und *Räumliches Zerfallen* ($F(3,297)=2.95$ $p=0.033$) ist durch die Art des Signals in dem die spektralen Differenzen erzeugt wurden beeinflusst worden. Ein Bonferroni korrigierter Post-hoc Test zeigt, dass sich die Bewertungen des wahrgenommenen *Unterschieds* nur bei der Signalart *Flächige Musik* signifikant von den anderen Signalarten unterscheiden. Die Ergebnisse der Analyse sind in Tabelle 11 aufgeführt.

Tabelle 11: Paarweiser Vergleich der Stufungen der UV Signalart. Vermeidung der Alphafehlerkumulierung durch Bonferroni-Korrektur.

Signal		Diff.der Mittelwerte	Standard Fehler	Freiheitsgrade	Signifikanz	95% Konfidenzintervall für Differenz	
						Untergrenze	Obergrenze
Flächig	Impulsiv	-,503	,185	252	,043	-,996	-,009
	Rauschen	-,724	,187	252	,001	-1,222	-,226
	Sprache	-,545	,178	252	,014	-1,018	-,072

Ob als Signalart ein *Impulsives*-, ein *Rausch*- oder *Sprachsignal* verwendet wurde, hatte hingegen keine signifikante Auswirkung auf die Bewertung des Unterschiedes. Abbildung 26 zeigt die Boxplots der drei Attribute in denen Haupteffekte wirksam sind. Um die Unterschiede zu verdeutlichen, wurden ergänzend zu den Medianen zusätzlich die arithmetischen Mittel durch ein Kreuz gekennzeichnet. Bei der Wahrnehmungsqualität *Klangfarbe* unterscheiden sich die Bewertungen bei der Verwendung der Signalart *Flächig* zu *Rauschen* und *Impulshaltig* signifikant voneinander ($p=0.07$ und $p<0.001$).

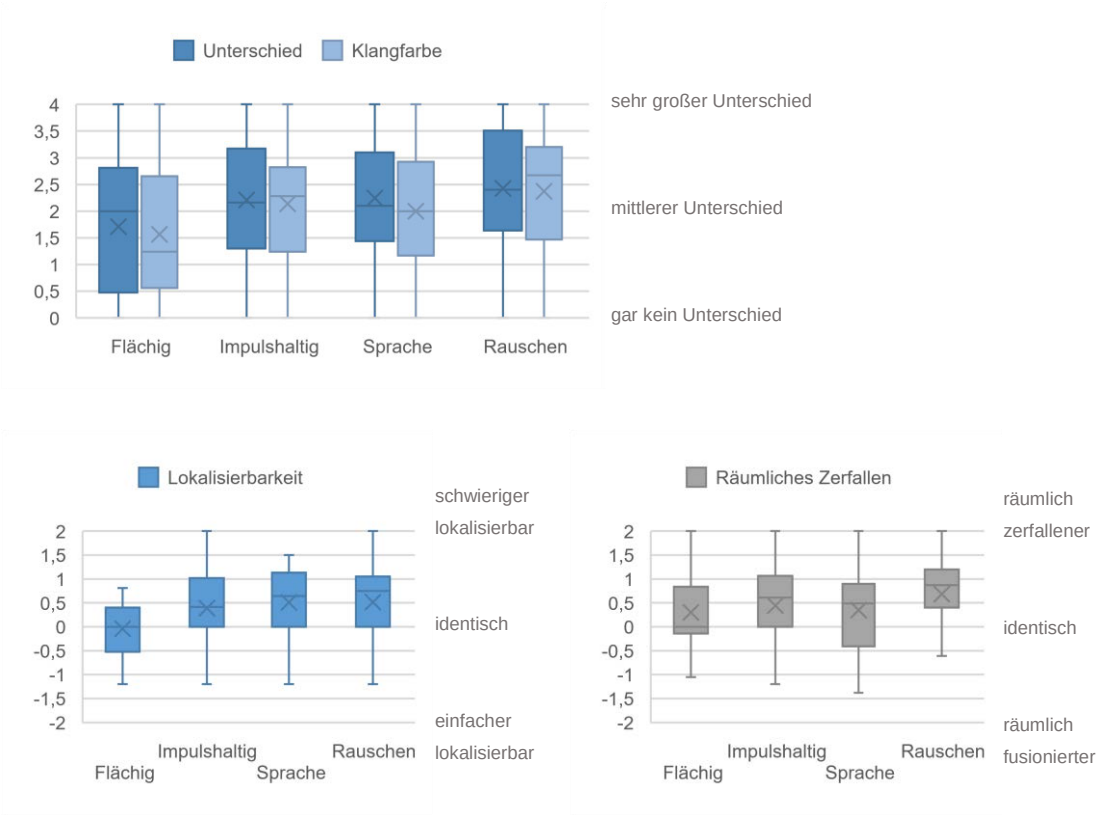


Abbildung 26: Boxplots der Bewertungen in Abhängigkeit des Attributes und der Signalart bei denen Haupteffekte wirksam sind. Die Kreuze markieren die arithmetischen Mittel. **Oben**: Ergebnisse der einseitigen Bewertungen. **Unten**: Ergebnisse der zweiseitigen Bewertungen.

Bei der Wahrnehmungsqualität *Lokalisierbarkeit* unterscheiden sich die Bewertungen *Flächig* und *Rauschen* signifikant voneinander (p=0.049). Der Boxplot zeigt zudem, dass eine Differenz bei dem klassischen Musikstück (*Flächig*) im Mittel keinen Unterschied in der Lokalisierbarkeit der Phantomschallquellen macht. *Sprache* und *Rauschen* als Monosignale liegen hingegen mit ihrer Einflussnahme auf die *Lokalisierbarkeit* gleich auf. Bei ihnen gelingt die Lokalisierung mit einer Differenz im Frequenzgang schwieriger.
Beim *räumliches Zerfallen* unterscheiden sich die Bewertungen bei der Verwendung der Signalart *Flächig* zu *Rauschen* signifikant voneinander (p=0.039). Insgesamt ist der durch eine spektrale Differenz hervorgerufene Unterschied bei dem flächigen Signal (klassisches Musikstück) niedriger als bei der Verwendung der anderen Signalarten. Am größten wurden der *Unterschied*, die *Klangverfärbung*, die *Lokalisierbarkeit* und das *räumliche Zerfallen* beim *Rauschen* als Signalart wahrgenommen.

5.2.3.3 Einfluss des Raumes bei einer spektralen Differenz auf die Wahrnehmungsqualitäten

Bei der UV *Raum* sind keine Haupteffekte wirksam, es existieren jedoch Interaktionseffekte: Bei der Bewertung der Wahrnehmungsqualität *Klangfarbe* ($F(2,276)=5.1$, $p<0.007$) und *Ausdehnung* ($F(2,282)=3.959$, $p<0.02$) beeinflusst der *Raum,* in Abhängigkeit von der *Frequenz* der Differenz das Antwortverhalten.

Die Boxplots in Abbildung 27 zeigen, dass sich die Bewertungen der beiden Attribute zwischen den Räumen insbesondere bei 9397Hz unterscheiden. Im *Hörstudio* wurden hier Unterschiede in der *Klangfarbe* größer und Phantomschallquellen als a*usgedehnter* wahrgenommen. Es ist auch zu erkennen, dass sich die Bewertungen bei 2693Hz zwischen den Räumen kaum unterscheiden. Im Hörstudio haben Differenzen bei 293Hz eine tendenziell geringere Auswirkung auf die *Klangfarbe* als im RAR.

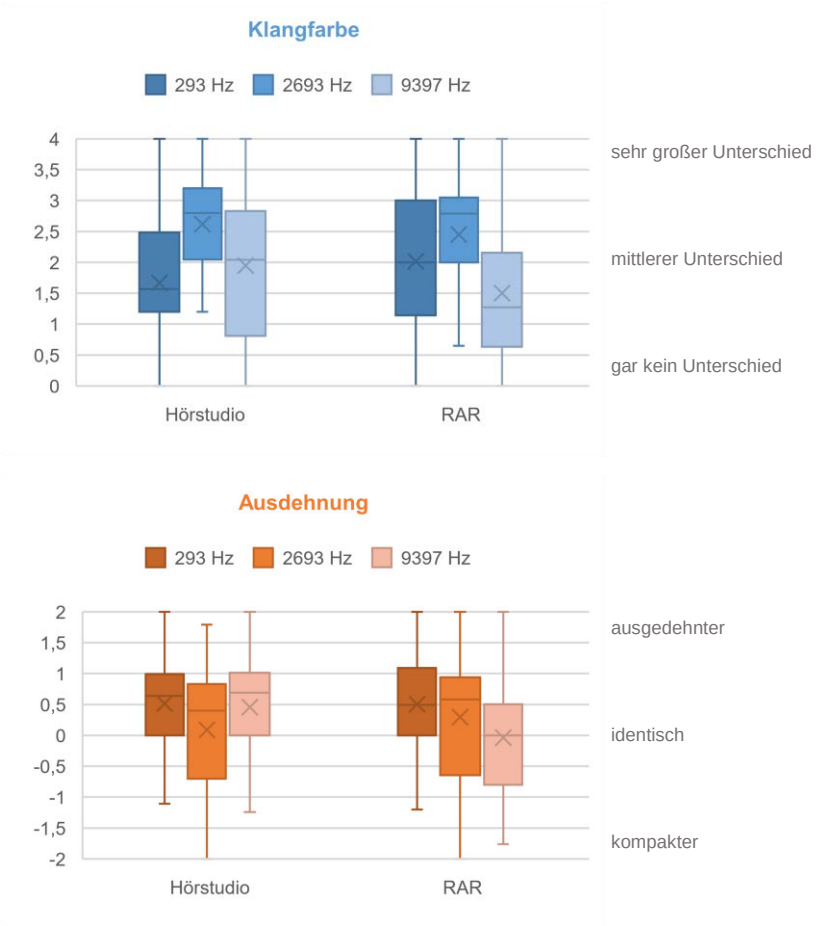


Abbildung 27: Boxplots der Bewertungen in Abhängigkeit des Attributes, des Raumes und der Frequenz. Die Kreuze markieren die arithmetischen Mittel.

6 DISKUSSION

6.1 Wahrnehmungsschwellen von Differenzen im Frequenzgang

Die einleitend erwähnte ITU Empfehlung BS.644-1 (1990), führt als Wahrnehmungsschwelle von Pegeldifferenzen zwischen zwei Lautsprechern im Frequenzbereich von 125Hz bis 10kHz einen Wert von 0,5dB auf. Dieser Wert deckt sich mit der gerade noch merklichen Änderung im Frequenzgang von Rosa Rauschen (Toole und Olive, 1987; Spiess, 1990). Unter- und oberhalb dieser Frequenzen liegt der minimale Schwellwert bei 2dB bzw. 1,5dB. In diesem Versuch lagen die Minima der Wahrnehmungsschwellen an den drei Frequenzpunkten wesentlich höher (293Hz 7,24dB, 2637Hz 4,4dB und 9397Hz 2,5dB). Dabei sind die Umstände unter denen die Schwellwerte ermittelt wurden unklar: An welchen Frequenzpunkten und mit was für einer Bandbreite wurden die Differenzen erzeugt, welche Signalart wurde verwendet, wie gestaltete sich der Versuchsaufbau und die akustische Umgebung? Wie in dieser Arbeit beschrieben, nimmt jeder der aufgeführten Punkte Einfluss auf die Erkennung von Abweichungen im Amplitudenfrequenzgang. In dieser Untersuchung konnten statistisch signifikante Unterschiede der Wahrnehmungsschwellen in Abhängigkeit der Frequenz und der Signalart festgestellt werden (H_1 und H_3 werden angenommen). Bei dem Frequenzpunkt 2637Hz liegt der Median der Wahrnehmungsschwelle am niedrigsten. Die erzeugte, terzbreite Differenz wirkt dabei in dem Frequenzbereich, in dem unser Gehör am empfindlichsten ist (Zwicker und Fastl, 1990).

Nach dem Abschluss des zweiten Hörversuchs wurden die Versuchsteilnehmer befragt, auf welche Eigenschaften der Wiedergabe sie achten, um die Differenzen zu erkennen. Dabei gab es solche, die vor allem die Auslenkung der Phantommitte und solche die die durch die Differenz hervorgerufene Verfärbung als Hinweis für eine spektrale Differenz nutzten. Eine mögliche Erklärung für den Unterschied der Minima der Wahrnehmungsschwellen zwischen 263Hz und 2637Hz, bzw. 9397Hz ist, dass unser Gehör zu hohen Frequenzen empfindlicher auf durch Pegeldifferenzen hervorgerufene Positionsänderungen der Phantomschallquellen, als zu tiefen reagiert (Blauert, 1974).

Die Feststellung, dass Abweichungen im Amplitudenfrequenzgang in stark reflektierenden Räumen besser zu hören sind (Spiess, 1990), lässt sich auf die Hörbarkeit von Differenzen im Frequenzgang zwischen zwei Lautsprechern unter den hier getesteten Bedingungen nicht übertragen. Zwar gilt das Hörstudio mit seiner geringen Nachhallzeit nicht als stark reflektierend, aber es sind gegensätzliche Tendenzen erkennbar, auch wenn der Einfluss des Raumes auf die Schwellwerte nicht signifikant ist (H_2 wird nicht angenommen). Lediglich bei 293Hz und der Verwendung von Sprache als Signalart, gibt es einen signifikanten Unterschied zugunsten des

reflexionsarmen Raums. Eine mögliche Erklärung für die Abweichungen zu Spiess (1990) ist, dass durch den räumlichen Einfluss, die Nutzung der Position, bzw. der Auslenkung der Phantomschallquelle, als Hilfsmittel für die Erkennung einer Differenz, erschwert wird.

Die nur sehr geringen Unterschiede der Mediane zwischen den Räumen bei 9397Hz, sind vermutlich auf den bereits hohen Bündelungsgrad der Lautsprecher in diesem Frequenzbereich zurückzuführen. Räumliche Einflüsse werden durch das Abstrahlverhalten in diesem Frequenzbereich minimiert. Dies verdeutlicht den Einfluss des verwendeten Lautsprechers auf die Wahrnehmungsschwelle von Differenzen im Frequenzgang.

Die Wahrnehmungsschwellen sind altersunabhängig. Aufgrund von altersbedingtem Hörverlust ist eine erhöhte Varianz der Schwellwerte zu sehr hohen Frequenzen, höher als der hier gemessenen, aber nicht ausgeschlossen. Die Hörpraxis (der aktive Umgang mit Musik, Tontechnik oder Akustik) hat einen nur sehr geringen Einfluss auf die Höhe der Wahrnehmungsschwellen. Allerdings wurden die niedrigsten Schwellwerte ausschließlich von Probanden erreicht, die sich, dem Autor wissentlich nach, intensiv mit Audiotechnik und dem aktiven Hören beschäftigen. Erfahrungswerte scheinen hier relevanter zu sein als die Hörpraxis pro Tag.

6.2 Einfluss von Differenzen im Frequenzgang auf die stereofone Abbildung

Die vorliegende Arbeit deckt sich insofern mit den Ergebnissen vorheriger Studien, als dass eine spektrale Dekorrelation der Lautsprechersignale und eine damit verbundene, sinkende Kohärenz der Schallsignale, zu einer erhöhten räumlichen Ausdehnung des Hörereignisses führt (Jefress, Blodgett und Deatherage, 1962; Chernyak und Dubrovsky, 1968). Die Ausdehnung der Phantomschallquellen ist dabei, wie von Zotter und Frank (2013) sowie Gribben und Lee (2016) gezeigt, frequenzabhängig. Dass die wahrgenommene Ausdehnung mit sinkender Frequenz zunimmt, deckt sich mit den Erkenntnissen von Käsbach (2016). Die von Käsbach (2016) beschriebene Frequenzabhängigkeit des räumlichen Zerfallens von Phantomschallquellen konnte in diesem Versuch nicht festgestellt werden. Möglicherweise war für einige Probanden die erzeugte Differenz zu gering, um überhaupt ein Zerfallen zu erkennen oder sie verfügten über ein zu geringes Hörverständnis, um sich der Bedeutung dieser Wahrnehmungsqualität im Klaren zu sein. Die Alternativhypothese H_1, dass Differenzen im Frequenzgang in der Stereofonie frequenzabhängig die Lokalisation von Phantomschallquellen erschweren, kann angenommen werden.

Bei 2637Hz ist der Einfluss einer Differenz im Frequenzgang am größten. Hier ist unser Gehör am empfindlichsten und lokalisiert die Phantomschallquellen vor allem über Pegeldifferenzen (ILD). Bei 9397Hz ist bei allen Signalarten der Energiegehalt geringer als zu den tieferen Frequenzen. Vermutlich ist aus diesem Grund der Einfluss einer Differenz bei dieser Frequenz so niedrig und der wahrgenommene Unterschied zu einer unveränderten Wiedergabe am geringsten.

Bei dem klassischen Musikstück waren die Unterschiede durch eine Differenz im Frequenzgang am niedrigsten und beim Rauschen am höchsten. Das Rauschsignal galt, wie in anderen Studien gezeigt, auch hier als die kritischste Signalart. Dass Hörer bei Sprache eine relativ direkte, lokalisierungsscharfe Wiedergabe bevorzugen und bei Musik eine räumlichere Darbietung (Klippel, 1988), hat sich in diesem Versuch insofern tendenziell bestätigt, als dass ein signifikanter Unterschied bei der Wahrnehmungsqualität *Lokalisierbarkeit* zwischen den Signalarten *klassische Musik* und *Rauschen* gefunden wurde (H_2 kann angenommen werden). *Sprache* liegt hier zumindest im arithmetischen Mittel mit dem *Rauschen* gleichauf, die Varianz vermeidet allerdings den signifikanten Unterschied zwischen *Sprache* und *klassischer Musik*.

Der räumliche Einfluss ist bei der Höhe der angewandten Differenz sehr gering. Im *Hörstudio* wurden bei 9397Hz Unterschiede in der *Klangfarbe* größer und Phantomschallquellen als a*usgedehnter* wahrgenommen. Auch wenn sich die Wahrnehmungsschwellen zwischen den Räumen bei 9397Hz nicht signifikant voneinander unterscheiden, unterstützt der Raum demzufolge den Einfluss spektraler Differenzen. Dies deckt sich mit Toole (1990), wonach Resonanzen durch räumliche Einflüsse verstärkt werden. Die Alternativhypothese H_3 kann nicht angenommen werden.

6.3 Limitationen und Fehlerbetrachtung

Akustische Umgebung, Stimulus, akustisches Verhalten der Lautsprecher und die Parameter der Differenz beeinflussen die Wahrnehmungsschwellen und die Einflüsse auf die stereofone Abbildung. Für möglichst reproduzierbare Ergebnisse wurden die Hörversuche in zwei Räumen, einer davon ein reflexionsarmer Raum, durchgeführt. Dennoch gibt es genügend Freiheitsgrade, die dazu führen, dass die Ergebnisse nur für die hier verwendeten Versuchsaufbauten gültig sind. Die kontrollierten Bedingungen unter denen die Versuche stattfanden, lassen nur näherungsweise Rückschlüsse auf die ökonomische Validität der Ergebnisse zu. Zu verschieden sind die akustischen Eigenschaften von Lautsprechern, der Umgebung und Signalen.

Der hohe Ausschluss von Fällen aufgrund der Bewertung der versteckten Referenz deutet an, dass der durch die spektrale Differenz erzeugte Unterschied, insbesondere zu hohen Frequenzen, zu klein war. Darüber hinaus hätte möglicherweise der Hinweis, dass die spektralen Differenzen an drei gleichen Frequenzpunkten mit immer der gleichen Ausprägung erzeugt werden und bei jeder

Signalart genau eine versteckte, unveränderte Darbietung existiert, die Anzahl der Bewertungen der HR erheblich verringert. So waren sich die Probanden unsicher und wollten möglicherweise Unterschiede hören wo keine vorhanden waren. Das Mischen von uni- und bipolaren Skalen und der damit verbundenen Verschiebung des Nullpunktes führte vermutlich, trotz der gegebenen Hinweise, ebenfalls zu vielen Ausfällen.

Die niedrigen Wahrnehmungsschwellen der Differenzen im Vorversuch, bestätigten die Vorgehensweise der Filterung der Signale mit IIR-Filtern. Jedoch lagen diese im Hauptversuch deutlich höher, sodass eine Beeinflussung der Messergebnisse durch die von dem IIR-Filter hervorgerufene minimalphasige Änderung, für die Versuchsteilnehmer mit sehr hohen Wahrnehmungsschwellen, nicht ausgeschlossen werden kann. Um eine Beeinflussung auch bei den höchsten Schwellen auszuschließen, ist die Verwendung von linearphasigen Filtern notwendig. Fraglich ist, ob Probanden deren Wahrnehmungsschwelle von spektralen Differenzen bei 20dB oder mehr liegen, in der Lage sind Phasenverschiebungen in der Größenordnung von 60° zu erkennen.

Möglicherweise war die Versuchszeit etwas zu Lang. Einige Versuchsteilnehmer äußerten im Nachhinein, dass bereits zum Ende des ersten Teils des Hörversuchs (Wahrnehmungsschwellen) die Konzentration nachgelassen hat, wodurch es zu Verpassern kam. Eine Verkürzung wäre nur durch Reduzierung der Frequenzpunkte oder die Untersuchung nur einer Signalart möglich gewesen. Möglicherweise wäre eine Pause zwischen den Signalarten hilfreich gewesen. Des Weiteren wurden manche Probanden im Hörstudio von Störgeräuschen aus dem Arbeitsbetrieb abgelenkt. Eine gänzliche Vermeidung der Außengeräusche wäre nur nach Betriebsschluss möglich gewesen, was aufgrund der Sicherheitsbestimmungen aber nicht möglich war.

6.4 Implikationen und Ausblick

Viele Hörer bevorzugen die durch den Raum entstehende Räumlichkeit einer trockenen Nahfeldwiedergabe und nehmen die dabei entstehende „Delokalisierung“ der Schallquellen hin. Ebenso bevorzugen viele Hörer eine, durch erste Reflexionen angereicherte Wiedergabe (Toole, 1990). In diesem Versuch wurden die Differenzen im Frequenzgang in ihrer Amplitude nicht verändert (16dB). Interessant wäre es daher zu erforschen, zu welchem Grad die Delokalisierung von Phantomschallquellen für das Gefallen einer stereofonen Wiedergabe förderlich wären.

Die Ergebnisse dieser Untersuchung zeigen, dass die akustische Umgebung keinen signifikanten Einfluss auf die Wahrnehmungsschwellen von spektralen Differenzen zwischen zwei Lautsprechern hat. Allerdings lassen sich Tendenzen erkennen, dass die Schwellwerte in einem reflexionsbehafteten Raum höher liegen. Um einen möglichen signifikanten Einfluss der

akustischen Umgebung auf die Schwellwerte auszumachen, wären Messungen in einem Raum mit einer wesentlich höheren Nachhallzeit notwendig.

Weiterhin wäre die Untersuchung des Zusammenhangs der Wahrnehmungsschwellen spektraler Differenzen in Abhängigkeit vom Abstrahlverhalten des Lautsprechers interessant. Wie in dem Eigenversuch festgestellt, lagen die Schwellwerte bei der Verwendung eines Ein-Wege-Systems etwas niedriger als bei dem getesteten Mehr-Wege-System, trotz linearer Entzerrung beider Systeme. Müller (1999, S. 226-229) führt Ergebnisse von Untersuchungen auf, die zeigen, dass insbesondere in reflexionsbehafteten Räumen hörbare Unterschiede zwischen verschiedenen, linearisierten Lautsprechern auftreten.

Zotter und Frank (2013) zeigen, dass sich der ursprünglich aus der Raumakustik stammende Parameter „Interaural Cross Correlation“ (IACC) eignet, um näherungsweise die Quellbreite von Phantomschallquellen in der Stereofonie zu beschreiben. Daher stellt sich außerdem die Frage, inwieweit sich daraus Rückschlüsse auf Abbildungsschärfe eines Lautsprecherpaares ziehen lassen und ob sich der Parameter eignet, um die diesbezügliche Wiedergabequalität von Lautsprecherpaaren zu messen und zu definieren.

Um Abweichungen zwischen Lautsprecherpaaren weiter zu reduzieren und die Beeinflussung auf die stereofone Abbildung weiter zu minimieren, empfiehlt es sich nicht nur die Chassis, sondern auch die Toleranzen der Bauteile der Frequenzweiche zu berücksichtigen. Frequenzweichen ließen sich, wie auch die Chassis, mittels des Paarungsalgorithmus selektieren. Interessant wäre eine Betrachtung der Dimensionen der durch Bauteiltoleranzen hervorgerufenen Differenzen in Amplituden- und Phasenfrequenzgang.

6.5 Einfluss der Erkenntnisse in die Qualitätssicherung der Lautsprecher-Fertigung

Im Rahmen dieser Arbeit wurde ein Messsystem für die Qualitätssicherung in der Lautsprecherfertigung eines High-End Audiotechnik-Herstellers evaluiert und eingerichtet. Die Erkenntnisse aus dieser Arbeit dienen der Festlegung der Toleranzgrenzen im Chassis-Selektionsprozess und sollen das Verständnis für die Auswirkungen von Differenzen im Frequenzgang zweier Lautsprecher in der Stereofonie schärfen.

Es wurde eine automatisierte Testprozedur eingerichtet, die von allen Chassis im der Wareneingangskontrolle durchlaufen wird. Je nach Chassistyp (Tief-, Mittel-, oder Hochtonlautsprecher), unterscheiden sich die Messparameter und die akustische Messumgebung. Um einen zu hohen Ausschuss und die damit verbundenen Kosten gering zu halten, liegt die Toleranz im Amplitudenfrequenzgang an dieser Stelle bei ±2dB. Darüber hinaus wird der

Impedanzgang, das Klirrverhalten und die Rub and Buzz Werte geprüft. Die Messdaten des Amplitudenfrequenzgangs der Chassis welche den Prüfvorgang bestehen werden terzbandgeglättet und anschließend von einem Paarungsalgorithmus verarbeitet. Dieser arbeitet mit frequenzabhängigen Gewichtungen. Dadurch können die gröberen Toleranzen aus der Eingangsprüfung, zwischen einem Lautsprecherpaar, minimiert werden. So sind nach der Paarung insbesondere im akustisch wirksamen Übertragungsbereichs des Chassis die Abweichungen eines Paares nur noch sehr gering. Die in dieser Arbeit gemessenen, minimalen Wahrnehmungsschwellen von ± 1,25dB Differenz werden so nicht überschritten. Als Toleranzgrenze wurde im Paarungsalgorithmus ein Wert von ± 1dB festgelegt. Das System gewährleistet so bei einem (zuliefererabhängig) geringem Ausschuss und geringer Abweichung vom Summenfrequenzgang der Lautsprecherreferenzbox, eine hohe Paargleichheit der Chassis, dessen spektrale Differenzen unterhalb der hier gemessenen Wahrnehmungsschwellen liegen. Eine verringerte Natürlichkeit und Abbildungsschärfe der Wiedergabe werden so vermieden.

ZUSAMMENFASSUNG

Die stereofone Lautsprecherwiedergabe ermöglicht durch eine akustische Täuschung des Gehörs die räumliche Anordnung sogenannter Phantomschallquellen zwischen den Lautsprechern. Voraussetzung für die geringe Ausdehnung und einfache Lokalisierbarkeit der entstehenden Hörereignisorte, ist eine möglichst hohe Kohärenz der Schallsignale (Jeffress, Blodgett und Deatherage 1962). Demzufolge sollten die in der Stereo-Signalübertragungskette verwendeten Komponenten im linken und rechten Kanal hinsichtlich ihres elektroakustischen Übertragungsverhaltens idealerweise identisch sein. Aufgrund von Ungenauigkeiten in der Herstellung der Komponenten lässt sich dieses Ziel aber nicht erreichen. In der Lautsprecherstereofonie führt die spektrale Ungleichheit der Einzelquellen zu Differenzen zwischen diesen. Um die Differenzen möglichst gering zu halten, werden durch einige Hersteller frequenzabhängige Toleranzbereiche festgelegt und die Lautsprecher paarweise selektiert. Die Ergebnisse dieser Arbeit geben Aufschluss über sinnvolle Toleranzgrenzen in der paarweisen Selektion von Lautsprechern und liefern einen Überblick des Einflusses spektraler Differenzen auf insbesondere akustisch, geometrische Wahrnehmungsqualitäten. Dazu wurden die Messdaten verschiedener Lautsprecher analysiert und basierend darauf Frequenzpunkte festgelegt, bei denen die Differenzen erzeugt wurden. In Hörversuchen wurden raum-, signal und frequenzabhängig die Wahrnehmungsschwellen von Differenzen im Frequenzgang gemessen und der Einfluss auf die stereofone Abbildung ermittelt. Für eine möglichst hohe akustische Passung zweier Lautsprecher, ist zunächst die messtechnische Erfassung der akustischen Eigenschaften der Lautsprecher und anschließend eine Paarung innerhalb definierter Toleranzen erforderlich. Im Rahmen dieser Arbeit wurde ein dazu geeignetes System eingerichtet und entsprechend der Erkenntnisse parametrisiert.

LITERATURVERZEICHNIS

Ahnert, W., & Feistel, S. (2009). Einmessung und Verifizierung raumakustischer Gegebenheiten und von Beschallungsanlagen. In M. Möser (Hrsg.), *Messtechnik der Akustik*. Springer-Verlag.

Bech, S., & Zacharov, N. (2006). *Perceptual Audio Evaluation - Theory, Method and Application* (1. Aufl.). Chichester, England ; Hoboken, NJ: John Wiley & Sons.

Blauert, J. (1974). *Räumliches Hören*. Stuttgart: Hirzel.

Blauert, J., & Braasch, J. (2008). Räumliches Hören. In S. Weinzierl (Hrsg.), *Handbuch der Audiotechnik*. Berlin: Springer Science & Business Media.

Bock, T. M., & Keele, J. (1986). The Effects of Interaural Crosstalk on Stereo Reproduction and Minimizing Interaural Crosstalk in Nearfield Monitoring by the Use of a Physical Barrier: Part 1. Gehalten auf der Audio Engineering Society Convention 81, Audio Engineering Society.

Bücklein, R. (1962). Hörbarkeit von Unregelmäßigkeiten in Frequenzgängen bei akustischer Übertragung. *Frequenz*, *16*(3), 103–108.

Chernyak, R. I., & Dubrovsky, N. A. (1968). Pattern if the noise images and the binaural summation of loudness for the differen interaural correlation of noise. *Reports of the 6th international congress on acoustics*, A-53 bis A-55.

Choisel, S., & Martin, G. (2008). Audibility of Phase Response Differences in a Stereo Playback System. Part 2: Narrow-Band Stimuli in Headphones and Loudspeakers. Gehalten auf der Audio Engineering Society Convention 125, Audio Engineering Society.

Choisel, S., & Zimmer, K. (2003). A pointing Technique with Visual Feedback for Sound Source Localization Experiments. Gehalten auf der Audio Engineering Society Convention 115, Audio Engineering Society.

Ciba, S., Brinkmann, F., & Lindau, A. (2015, Mai 2). WhisPER v1.9.0 User Documentation. TU-Berlin.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. Hillsdale, N.J.: L. Erlbaum Associates.

European Broadcasting Union (Hrsg.). (1998, Mai). EBU Tech. 3276-2 Listening conditions for the assessment of sound programme material: monophonic and two–channel stereophonic.

Frank, M. (2013). Source Width of Frontal Phantom Sources: Perception, Measurement, and Modeling. *Archives of Acoustics*, *38*(3), 311–319.

Gribben, C., & Lee, H. (2016). The Perception of Vertical Image Spread by Interchannel Decorrelation. Gehalten auf der Audio Engineering Society Convention 140, Audio Engineering Society.

Hartmann, W. M. (1983). Localization of sound in rooms. *The Journal of the Acoustical Society of America*, *74*(5), 1380–1391.

IBM (Hrsg.). (2013). IBM SPSS Advanced Statistics 22. Abgerufen von ftp://public.dhe.ibm.com/software/analytics/spss/documentation/statistics/22.0/de/client/Manuals/IBM_SPSS_Advanced_Statistics.pdf am 14.12.2017

International Telecommunication Union (Hrsg.). (1990, Juni). Recommendation ITU-R BS.644-1 - Audio quality parameters for the performance of a high-quality sound-programme transmission chain. ITU-R.

International Telecommunication Union (Hrsg.). (2015, Februar). Recommendation ITU-R BS.1116-3 - Methods for the subjective assessment of small impairments in audio systems. ITU-R.

International Telecommunication Union (Hrsg.). (2015, Oktober). Recommendation ITU-R BS.1534-3 - Method for the subjective assessment of intermediate quality level of audio systems. ITU-R.

Jeffress, L. A., Blodgett, H. C., & Deatherage, B. H. (1962). Effect of Interaural Correlation on the Precision of Centering a Noise. *The Journal of the Acoustical Society of America*, *34*(8), 1122–1123.

Johnson, D., & Lee, H. (2017). Just Noticeable Difference in Apparent Source Width Depending on the Direction of a Single Reflection. Gehalten auf der Audio Engineering Society Convention 142, Audio Engineering Society.

Karjalainen, M., Piirilä, E., Järvinen, A., & Huopaniemi, J. (1999). Comparison of Loudspeaker Equalization Methods Based on DSP Techniques. *Journal of the Audio Engineering Society*, *47*(1/2), 14–31.

Käsbach, J. (2016). *Characterizing apparent source width perception*. Technical University of Denmark. Abgerufen von http://www.hea.elektro.dtu.dk/-/media/Centre/HEA_Hearing_Systems/hea/english/research/PhD%20Thesis%20PDF/00_25-Ksbach-submitted-version.ashx?la=da am 14.12.2017

King-Smith, P. E., Grigsby, S. S., Vingrys, A. J., Benes, S. C., & Supowit, A. (1994). Efficient and unbiased modifications of the QUEST threshold method: Theory, simulations, experimental evaluation and practical implementation. *Vision Research*, *34*(7), 885–912.

Klippel, W. (1988). *Zusammenhang zwischen objektiven Lautsprecherparametern und subjektiver Qualitätsbeurteilung*. Technische Hochschule Dresden.

Kuhl, W., & Plantz, R. (1978). Die Bedeutung des von Lautsprechern abgestrahlten diffusen Schalls für das Hörereignis. *Acta Acustica united with Acustica*, *40*(3), 182–191.

Langendijk, E. H. A., & Bronkhorst, A. W. (2000). Contribution of spectral cues to human sound localization. *The Journal of the Acoustical Society of America*, *112*(4), 1583–1596.

Lindau, A. (2014). Spatial Audio Quality Inventory (SAQI). Test Manual.

Linkwitz, S. (2009). The Challenge to Find the Optimum Radiation Pattern and Placement of Stereo Loudspeakers in a Room for the Creation of Phantom Sources and Simultaneous Masking of Real Sources. Gehalten auf der Audio Engineering Society Convention 127, Audio Engineering Society.

Martin, G. (2002). Interchannel Interference at the Listening Position in a Five-channel Loudspeaker Configuration. Gehalten auf der Audio Engineering Society Convention 113, Audio Engineering Society.

Medina Victoria, J., & Görne, T. (2009). Apparent Source Width in ITU Surround. Gehalten auf der Audio Engineering Society Convention 126, Audio Engineering Society.

Møller, H., Sørensen, M. F., Hammershøi, D., & Jensen, C. B. (1995). Head-Related Transfer Functions of Human Subjects. *Journal of the Audio Engineering Society*, *43*(5), 300–321.

Moore, B. (2007). Psychoacoustics. In T. D. R. Prof (Hrsg.), *Springer Handbook of Acoustics* (S. 459–501). Springer New York.

Müller, S. (1999, Januar 22). *Digitale Signalverarbeitung für Lautsprecher*. RWTH Aachen, Aachen.

Olive, S. E., Schuck, P. L., Ryan, J. G., Sally, S. L., & Bonneville, M. E. (1997). The Detection Thresholds of Resonances at Low Frequencies. *Journal of the Audio Engineering Society*, *45*(3), 116–128.

Otto, S. (2008, Juli 8). *Vergleichende Simulation adaptiver, psychometrischer Verfahren zur Schätzung von Wahrnehmungsschwellen*. TU-Berlin, Berlin.

Pagano, R. R. (2013). *Understanding statistics in the behavioral sciences*. Belmont, CA: Wadsworth/Cengage Learning.

Spiess, K. (1990). *Zur Hörbarkeit linearer Verzerrungen in Räumen*. Technische Hochschule Aachen, Aachen.

Theile, G. (1980). *Über die Lokalisation im überlagerten Schallfeld*. Technische Universität Berlin, Berlin.

Toole, F. (2008). *Sound Reproduction: The Acoustics and Psychoacoustics of Loudspeakers and Rooms*. Amsterdam ; Boston: Focal Press.

Toole, F. E. (1990). Loudspeakers and Rooms for Stereophonic Sound Reproduction. Gehalten auf der Audio Engineering Society Conference: 8th International Conference: The Sound of Audio, Audio Engineering Society.

Toole, F. E. (2005). Audio-Science in the Service of Art. Harman. Abgerufen von http://www.harman.com/sites/default/files/white-paper/12/11/2015%20-%2005%3A54/files/AudioScience.pdf am 14.12.2017

Toole, F. E., & Olive, S. E. (1987). The Modification of Timbre by Resonances: Perception and Measurement. In *Audio Engineering Society Convention 83*.

Uprichard, S., & Choisel, S. (2008). The Influence of Acoustic Environment on the Threshold of Audibility of Loudspeaker Resonances. Gehalten auf der Audio Engineering Society Convention 125, Audio Engineering Society.

Vorländer, M. (2008). *Auralization - Fundamentals of Acoustics, Modelling, | Michael Vorländer | Springer*. Springer.

Warzybok, A., Rennis, J., Doclo, S. (2011). Influence of early reflections on speech intelligibility under different noise conditions. Gehalten auf dem Forum Acusticum Aalborg Denmark 2011.

Wendt, K. (1964). *Das Richtungshören bei der Überlagerung zweier Schallfelder bei Intensitäts- und Laufzeitstereophonie*. Aachen, Techn. Hochsch., Diss., 1964.

Zotter, F., & Frank, M. (2013). Efficient Phantom Source Widening. *Archives of Acoustics*, *38*(1), 27–37.

Zwicker, E., & Fastl, H. (1990). *Psychoacoustics: Facts and Models*. Springer-Verlag.

TABELLENVERZEICHNIS

Tabelle 1: Häufigkeiten der größten Standardabweichungen an den betrachteten Frequenzpunkten. 17

Tabelle 2: Parameter der Filter für Entzerrung der Versuchslautsprecher 24

Tabelle 3: Abstufung der Reizstärken der Darbietungen im Vorversuch 37

Tabelle 4: Arithmetisches Mittel und Standardabweichung der Wahrnehmungsschwellen des Vorversuchs 38

Tabelle 5: Ergebnisse des Shapiro-Wilk-Tests auf Normalverteilung. 42

Tabelle 6: Auflistung der niedrigsten Durchschnittswerte und Minima der Wahrnehmungsschwellen von Differenzen im Frequenzgang unter Angabe der Merkmalsausprägungen unter denen diese erhoben wurden. 46

Tabelle 7: Prozentuale Häufigkeiten in denen die versteckte Referenz höher als das Kriterium es zulässt bewertet wurde. 48

Tabelle 8: Korrelationskoeffizienten der bipolaren Attribute ρ (rho) nach Spearman. Signifikante Korrelationen sind grün markiert. 49

Tabelle 9: Tests auf feste Effekte aller bipolaren Attribute. 51

Tabelle 10: Skalenpole der Wahrnehmungsqualitäten. 53

Tabelle 11: Paarweiser Vergleich der Stufungen der UV Signalart. Vermeidung der Alphafehlerkumulierung durch Bonferroni-Korrektur. 54

ABBILDUNGSVERZEICHNIS

Abbildung 1: Stereo-Standardaufstellung mit einem Basiswinkel α = 60° (Blauert und Braasch, 2008, S.101) ... 8

Abbildung 2: Auswirkungen von Pegel- und Zeitdifferenzen auf die Hörereignisrichtung. (Blauert und Braasch, 2008, S. 101) ... 8

Abbildung 3: Verdeckung benachbarter Frequenzen durch ein schmalbandiges Rauschen in Abhängigkeit des Pegels und der Frequenz. (Moore, 2007) ... 9

Abbildung 4: Schematische Darstellung für den Abklingvorgang eines Schallfeldes, nach Vorländer (2008) ... 11

Abbildung 5: Abbildung der Leistungsdichtespektren der verwendeten Testsignale. Es wurde eine Terzbandglättung der Daten vorgenommen. ... 19

Abbildung 6: Toleranzschlauch in der Qualitätssicherung. Eigene Darstellung ... 20

Abbildung 7: Auszug der Amplitudengänge der verwendeten Filter zur Erzeugung der Differenzen. 20dB, 8dB und 4dB Differenz an den drei Frequenzpunkten 293Hz, 2637Hz und 9397Hz. ... 21

Abbildung 8: Auszug der Phasengänge der verwendeten Filter zur Erzeugung der Differenzen. 20dB, 8dB und 4dB Differenz an den drei Frequenzpunkten 293Hz, 2637Hz und 9397Hz. ... 21

Abbildung 9: Amplitudenfrequenzgang des Versuchslautsprechers bei 0°, 15° und 30°. ... 23

Abbildung 10: Entzerrter Amplitudenfrequenzgang des Versuchslautsprechers bei 0°, 15° und 30°. ... 25

Abbildung 11: Entzerrter Amplitudenfrequenzgang der Versuchslautsprecher links und rechts. ... 25

Abbildung 12: Nachhallzeit über die Frequenz des Hörstudios (blau) und Toleranzband nach ITU-R BS.1116-3 (rot). ... 27

Abbildung 14: Amplitudenfrequenzgang der Lautsprecher links und rechts im Hörstudio und im RAR. 28

Abbildung 15: Versuchsaufbau im reflexionsarmen Raum. ... 33

Abbildung 16: Bewertungsinstrument WhisPER. AFC-Oberfläche für die Schätzung der Wahrnehmungsschwellen. ... 34

Abbildung 17: Bewertungsinstrument WhisPER. Angepasste MUSHRA-Oberfläche. Erst nach dem Anklicken der Schaltflächen jeder Darbietung wurde die OK-Schaltfläche freigegeben. ... 34

Abbildung 18: Schematischer Versuchsaufbau und Signalfluss. ... 35

Abbildung 19: Abbruch nach 40 Trails. In den ersten Trails hatte die Versuchsperson offensichtlich mehrmals zufällige Treffer. ... 39

Abbildung 20: Beispiel eines Quantil-Quantil-Diagramm zum Vergleichen der Quantile der Daten aller Wahrnehmungsschwellen bei 293Hz gegen die theoretischen Quantile einer Normalverteilung. ... 42

Abbildung 21: Boxplot der Wahrnehmungsschwellen in den beiden Räumen unter Einbeziehung der Schwellen beider Darbietungen und der drei Frequenzpunkte. ... 43

Abbildung 22: Boxplot der Wahrnehmungsschwellen von Differenzen an den drei verwendeten Frequenzpunkten unter Einbeziehung der Schwellen beider Räume und Stimuli. ... 45

Abbildung 23: Boxplot der Wahrnehmungsschwellen in Abhängigkeit von *Frequenz* und *Raum* (HS=Hörstudio) bei Verwendung von „Rosa Rauschen“ als Stimulus. ... 45

Abbildung 24: Boxplot der Wahrnehmungsschwellen in Abhängigkeit von Frequenz und Raum (HS=Hörstudio) bei Verwendung von Sprache als Stimulus. ... 46

Abbildung 25: Zusammenhang der Wahrnehmungsschwellen der Differenzen und dem Alter (links) oder dem aktiven Umgang mit Audiotechnik, Akustik oder Musik in Stunden pro Tag (rechts). ... 47

Abbildung 26: Boxplots der Bewertungen in Abhängigkeit des Attributes und der Frequenz. Oben: Ergebnisse der einseitigen Bewertungen. Unten: Ergebnisse der zweiseitigen Bewertungen. ... 53

Abbildung 27: Boxplots der Bewertungen in Abhängigkeit des Attributes und der Signalart bei denen Haupteffekte wirksam sind. Die Kreuze markieren die arithmetischen Mittel. Oben: Ergebnisse der einseitigen Bewertungen. Unten: Ergebnisse der zweiseitigen Bewertungen. 54

Abbildung 28: Boxplots der Bewertungen in Abhängigkeit des Attributes, des Raumes und der Frequenz. Die Kreuze markieren die arithmetischen Mittel. 56

ANHANG

Die Anhänge dieser Arbeit finden sich auf dem der Arbeit beigelegten Datenträger. Dabei handelt es sich um die folgenden Daten:

- Masterarbeit in elektronischer Form *Stammverzeichnis*
- Informationen und Fragebogen zum Hörversuch *Stammverzeichnis*
- Matlab-Funktionen und Skripte *Matlab*
- Daten der Hörversuche als SAV *Daten der Hörversuche*
- Verwendete Stimuli als WAV *Stimuli*
- Referenzierte Literatur *Referenzierte Literatur*
- Referenzierte Literatur als RIS *Referenzierte Literatur*