Fakultät I – Geistes- und Bildungswissenschaften

Institut für Sprache und Kommunikation

Fachgebiet Audiokommunikation

Wahrnehmung von Phasenänderungen bei der Bewegung im Schallfeld von Lautsprechersystemen

Masterarbeit

vorgelegt von

Maximilian Bux

[redacted]

Erstgutachter

Prof. Dr. Stefan Weinzierl

FG Audiokommunikation

Technische Universität Berlin

Zweitgutachter

Dipl.-Ing. Florian Straube

FG Audiokommunikation

Technische Universität Berlin

15. Mai 2017

Danksagung

An dieser Stelle geht mein Dank an alle, die am Gelingen dieser Arbeit beteiligt waren. Neben allen Proband*innen der Hörversuche sind das Florian Straube, Fabian Brinkmann und Prof. Dr. Stefan Weinzierl.

Eidesstattliche Erklärung

Hiermit versichere ich gegenüber der Fakultät I der Technischen Universität Berlin, dass ich die vorliegende Arbeit selbstständig verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel benutzt habe. Alle Ausführungen, die anderen veröffentlichten oder nicht veröffentlichten Schriften wörtlich oder sinngemäß entnommen wurden, habe ich kenntlich gemacht. Die Arbeit hat in gleicher oder ähnlicher Fassung noch keiner anderen Prüfungsbehörde vorgelegen.

Berlin, den 15. Mai 2017 _____

Maximilian Bux

Zusammenfassung

Die Homogenität des Schallfeldes der für die Beschallung von größeren Zuhörerbereichen verwendeten modernen Lautsprecher-Systeme, die meist als Linienschallquellen (Line-Array-Systeme) ausgeführt werden, kann mit geeigneten Ansteuerfunktionen aus numerischen Optimierungsverfahren weiter verbessert werden. Die dazu in Straube et al., 2015 vorgeschlagenen Position-Index-Plots (PIPs) stellen eine grafische Repräsentation der orts- und frequenzabhängigen Amplitudenverteilung über den zu beschallenden Bereich dar, liefern jedoch keine Aussage über Phasenlage und -sprünge. Versierte Zuhörer berichten jedoch von flanger- und phaserartigen Artefakten bei Bewegung durch entsprechende Schallfelder. Ziel dieser Arbeit ist es, systematisch mit Hilfe geeigneter Hörversuche zu untersuchen, ob eine perzeptiv empfundene Minderung der Qualität durch Phasenänderungen bei der Bewegung von Zuhörern im Schallfeld auftritt und von welchen Faktoren (im Wesentlichen Bewegungsgeschwindigkeit des Hörers und Frequenzumfang der Anregungssignale) diese abhängen. Dazu werden mit Hilfe der Binauralsynthese entsprechende Schallfelder simuliert, in denen der Hörer die zu untersuchenden Bewegungen virtuell ausführt und perzeptiv bewertet. Aus diesen Erkenntnissen werden mögliche Prädiktionsmethoden vorgeschlagen, welche den Effekt im Voraus abschätzen und beispielsweise dazu dienen können, existierende numerische Optimierungsverfahren weiter zu verbessern.

Inhaltsverzeichnis

1 Einleitung ... 1

1.1 Motivation ... 1

1.2 Stand der Forschung ... 1

2 Methodik ... 3

2.1 Grundszenerie ... 3

2.1.1 Umgebung ... 3

2.1.2 Bewegungspfad ... 3

2.1.3 Bewegungsgeschwindigkeit ... 4

2.1.4 Art und Kontur der Phasenänderung ... 4

2.1.5 Anregungssignale ... 5

2.2 Binaurale Synthese ... 9

2.2.1 Modifikation der *AKshAura*()-Funktion ... 9

2.2.2 Sphärische Trajektorie ... 10

2.2.3 Amplituden-Hüllkurve ... 10

2.2.4 Kopfhörer-Kompensation ... 11

2.2.5 Loudness-Matching ... 14

3 Perzeptuelle Evaluation ... 15

3.1 Vorbetrachtungen ... 15

3.1.1 Verwendete Hard- und Software ... 15

3.1.2 Modifikation der Audio-Playout-Routine von WhisPER ... 15

3.2 Informeller Hörversuch ... 16

3.2.1 Methodik ... 16

3.2.2 Ergebnisse ... 18

3.3 Haupt-Hörversuch ... 19

3.3.1 Vorbemerkungen ... 19

3.3.2 Teil 1: 3-AFC-Test ... 21

3.3.3 Teil 2: ABCHR-Test ... 22

4 Ergebnisse ... 24

4.1 Vorbemerkungen ... 24

4.2 Teil 1: 3-AFC-Test ... 24

4.3 Teil 2: ABCHR-Test ... 25

4.3.1 Deskriptiv ... 25

4.3.2 Lilliefors-Test ... 30

4.3.3 Friedman-Test ... 31
4.3.4 Wilcoxon-Signed-Rank-Test ... 32
4.4 Fragebogen ... 33
5 Diskussion ... 35
5.1 Zusammenfassung ... 35
5.2 Prädiktionsindikatoren ... 36
5.2.1 Spektraler Flux ... 36
5.2.2 Autokorrelation ... 37
5.3 Fehlerbetrachtung ... 38
5.4 Ausblick ... 39
6 Literatur ... 40
7 Abbildungsverzeichnis ... i
8 Tabellenverzeichnis ... ii
9 Anhang ... iii

1 Einleitung

1.1 Motivation

Eines der wichtigsten Ziele moderner Beschallungssysteme ist es, den zu beschallenden Zuhörerbereich mit einem möglichst homogenen Schallfeld abzudecken, damit entsprechend viele Hörer ein gleichermaßen zufriedenstellendes akustisches Erlebnis erfahren können. Um dies zu erreichen, verwenden die meisten modernen Beschallungssysteme sogenannte Zeilenlautsprecher (Line-Arrays). Diese strahlen Schall im Nahfeld, ab einer von der Zeilenlänge abhängigen Grenzfrequenz gerichtet nach den Gesetzmäßigkeiten einer Linienschallquelle, ab (Möser, 2012, S. 93f). Um dieses Schallfeld homogen auf den Zuhörerbereich auszurichten, werden die einzelnen Segmente des Line-Arrays in der Regel mit unterschiedlichen Neigungswinkeln versehen. Hierbei ist es mittlerweile möglich und gängige Praxis, mit Hilfe numerischer Berechnungsverfahren auf eine noch größere Homogenität des Schallfeldes hin optimierte Ziel-Schallfelder zu finden. Damit lassen sich geeignete Ansteuerfunktionen der einzelnen Segmente der Linienschallquelle finden, die üblicherweise mittels FIR- oder minimalphasigen Filtern durch leistungsfähige digitale Signalprozessoren (DSPs) umgesetzt werden (Scheirman, 2015). In Straube et al., 2015 werden Strategien zur Bewertung solcher optimierten Ziel-Schallfelder vorgestellt. Hierbei werden unter anderem sogenannte Position-Index-Plots (PIPs) verwendet, die im Wesentlichen eine grafische Repräsentation der optimierten Ziel-Schallfelder sowohl für die homogen zu beschallenden, als auch für die zu vermeidenden Bereiche (beispielsweise Decken und Wände) sind. Diese Plots beinhalten die relevanten ortsabhängigen Betragsfrequenzgänge der optimierten Ziel-Schallfelder, liefern jedoch keine Informationen über ortsabhängige Phasenlagen und -sprünge, die möglicherweise zu Qualitätsbeeinträchtigungen der wahrgenommenen Schallfelder führen können. Ziel dieser Arbeit ist es, systematisch zu untersuchen, in welchem Maße stetige Phasenveränderungen bei der Bewegung im Schallfeld perzeptiv wahrgenommen werden und welche Einflussfaktoren, beispielsweise spektrale und temporale Eigenschaften der Anregungssignale und die Bewegungsgeschwindigkeit des Hörers, dabei maßgeblich eine Rolle spielen.

1.2 Stand der Forschung

Zur Wahrnehmung von Phasenänderungen in harmonisch komplexen Signalen gibt es bereits einige wenige Studien. Die wohl älteste und fundamentalste Untersuchung zur Wahrnehmbarkeit von Phasenvariationen akustischer Signale führte Herrmann von Helmholtz bereits in der zweiten Hälfte des 19. Jahrhunderts durch, indem er mehrere Stimmgabeln mit Elektromagneten kontinuierlich anregte und durch kontrolliertes Verstimmen (partielles Verschließen) von an den Stimmgabeln angekoppelten Resonanzröhren die Phasen der einzelnen Schwingungen einstellen konnte. Seine Schlussfolgerung aus diesem Experiment lautete:

> „Ich habe in dieser Weise mannigfaltige Combinationen der Töne mit verschiedenen Phasenunterschieden versucht, aber niemals gefunden, dass sich die Klangfarbe im geringsten dabei veränderte. Es war für den Klang immer vollständig gleichgültig, ob ich einzelne Partialtöne durch unvollständige Oeffnung der Röhren, oder durch deren Entfernung von den Stimmgabeln abschwächte, wodurch also die von uns aufgestellte Frage dahin entschieden wird, dass *die Klangfarbe des musikalischen Theils eines Klanges nur abhängt von der Zahl und Stärke der Theiltöne, nicht von ihren Phasenunterschieden.*“ (Helmholtz, 1865, S. 193).

Seine Erkenntnis aus diesem Experiment ist also, dass das menschliche Gehör nicht dazu in der Lage ist, Differenzen in der Phasenwinkellage von Amplitudenänderungen der einzelnen Schwingungen wahrzunehmen.

Aus Laitinen et al., 2013 lässt sich entnehmen, dass ein harmonisch komplexes Signal mit tiefer Grundfrequenz (um 100 Hz) spektral verfärbter und dünner wahrgenommen wird, wenn die Phasenlage in den Obertönen beeinflusst wird. Auch wird hierbei ermittelt, dass das menschliche Gehör Phasendifferenzen bei statischen Signalen spektral lokal wahrnimmt und umliegende Frequenzbereiche perzeptiv hiervon nur bedingt beeinflusst werden. Außerdem wird dargestellt, dass Phasenänderungen bei harmonisch komplexen Signalen, deren Grundfrequenzen über 800 Hz liegen, nicht mehr wahrgenommen werden. Dies bestätigen auch die Untersuchungen von Tachi und Isobe, 1973. Eine Erkenntnis ihrer Studie ist, dass Timbre-Veränderungen bei harmonisch komplexen Zweiklang-Tongemischen bei Variation der Phasenlage mit Frequenzen über 1,2 kHz nicht mehr wahrgenommen werden. Umgekehrt führen Phasenänderungen in höheren Frequenzen bei Signalen mit spektral tiefer Grundfrequenz zu einer Änderung des wahrgenommenen Bass-Fundamentes. Dieser Umstand ist ebenso als interessanter Aspekt im Hinblick auf die Thematik dieser Arbeit zu sehen, da Lautsprechersysteme für Beschallungszwecke auf Konzerten, Festivals und ähnlichen Veranstaltungen in der Regel hohe Schalldrücke im Bassbereich reproduzieren müssen.

Alle vorigen Betrachtungen beziehen sich ausschließlich auf die Wahrnehmung von Schallereignissen mit statischen Phasenänderungen. In einer realen, natürlichen Hörsituation kann sich die Phasenlage des vom Hörer wahrgenommenen Schallereignisses allein schon mit jeder Änderung der Position zur Schallquelle und im Besonderen durch Einflüsse der Raumakustik und dem Vorhandensein eines diffusen Schallfelds ändern. Damit kann das Schallereignis charakterisiert und unterscheidbar von anderen Schallereignissen werden (Griesinger, 2010). Der menschliche Hörsinn besitzt zwar Mechanismen, die dabei entstehenden Artefakte zu kompensieren, indem er beispielsweise Hörereignisse mit gelernten Mustern vergleicht (Bregman, 1994). Diese Kompensationsmechanismen funktionieren jedoch nur bis zu einem gewissen Grad und können bei einem sich ständig ändernden Schallfeld, beispielsweise bei kontinuierlicher Bewegung des Hörers in Relation zur Schallquelle, an die Grenzen getrieben werden. Hierzu existieren bisher noch keine systematischen Untersuchungen, die als Anregungssignale solche mit sich stetig verändernden Phasenlagen verwenden. Genau diese Situationen sind Gegenstand dieser Arbeit.

2 Methodik

Die für die Studien dieser Arbeit notwendigen Schallfelder werden mit Hilfe binauraler Synthese generiert und über Kopfhörer dargeboten. Die Realisierung erfolgt mit Hilfe numerischer Berechnungen mit der matrixbasierten Software MATLAB® auf Grundlage der AKTools (Brinkmann und Weinzierl, 2016). Als HRIR-Datensatz wird die zum Zeitpunkt des Verfassens dieser Arbeit aktuelle Version des an der TU Berlin im Fachgebiet Audiokommunikation gewonnenen FABIAN-Datensatzes verwendet (Brinkmann, 2017).

2.1 Grundszenerie

Die gesamte Arbeit baut auf Auralisationen synthetisch erzeugter binauraler Schallfelder, die den zu untersuchenden Effekt beinhalten, auf. Zunächst einmal gilt es zu klären, unter welchen Randbedingungen diese Schallfelder auralisiert und welche Merkmale und Eigenschaften dabei untersucht und gegenübergestellt werden sollen.

2.1.1 Umgebung

Die erste Betrachtung fällt auf die zu auralisierende Umgebung. Hier wird der einfachste Fall angenommen: Es existieren keine Reflexionen, der Hörer befindet sich in einer reflexionsfreien Umgebung, in dem ideale Freifeld-Schallausbreitungsbedingungen herrschen. Damit soll der zu untersuchende Effekt von allen anderen spektralen und temporalen Kolorationen seitens der Raumakustik frei gemacht werden, damit der Hörer davon nicht irritiert, abgelenkt oder gar überfordert wird.

2.1.2 Bewegungspfad

Der nächste zu berücksichtigende Aspekt ist der virtuelle Bewegungspfad des Hörers. Dieser wird für alle auralisierten Schallfelder als immer gleich definiert. Er soll einer tatsächlich real denkbaren Situation, beispielsweise der Demonstration eines Lautsprechersystems im kleineren Personenkreis unter nahezu Freifeld-Bedingungen, beispielsweise auf einer Wiese im Freien, entsprechen.

Auch hier wird der Reproduzierbarkeit wegen der einfachste Fall angenommen: Der Bewegungspfad soll geradlinig sein, kontinuierlich lotrecht zur Schallquelle stehen und an ihr mittig vorbeiführen. Es wird ein Bewegungspfad definiert, der in $d = 3$ Metern Entfernung zu einer gedachten (akustisch transparenten) Bühne ist, auf dessen Vorderkante die Test-Schallquelle auf Ohrhöhe des Hörers steht und omnidirektional abstrahlt. Der halbe Bewegungspfad ist mit $l = 2d = 6$ Meter lang, der gesamte Bewegungspfad entsprechend $2l = 12$ Meter. Der Mittelpunkt des gesamten Bewegungspfades liegt genau lotrecht zur gedachten Mittelachse der Schallquelle. Die Blickrichtung des Hörers ist dabei stets senkrecht zum Bewegungspfad, die Ohrenachsen des Hörers verlaufen also direkt durch die Bewegungsachse hindurch. Abbildung 1 skizziert die entsprechende Umgebung und den Bewegungspfad des Hörers.

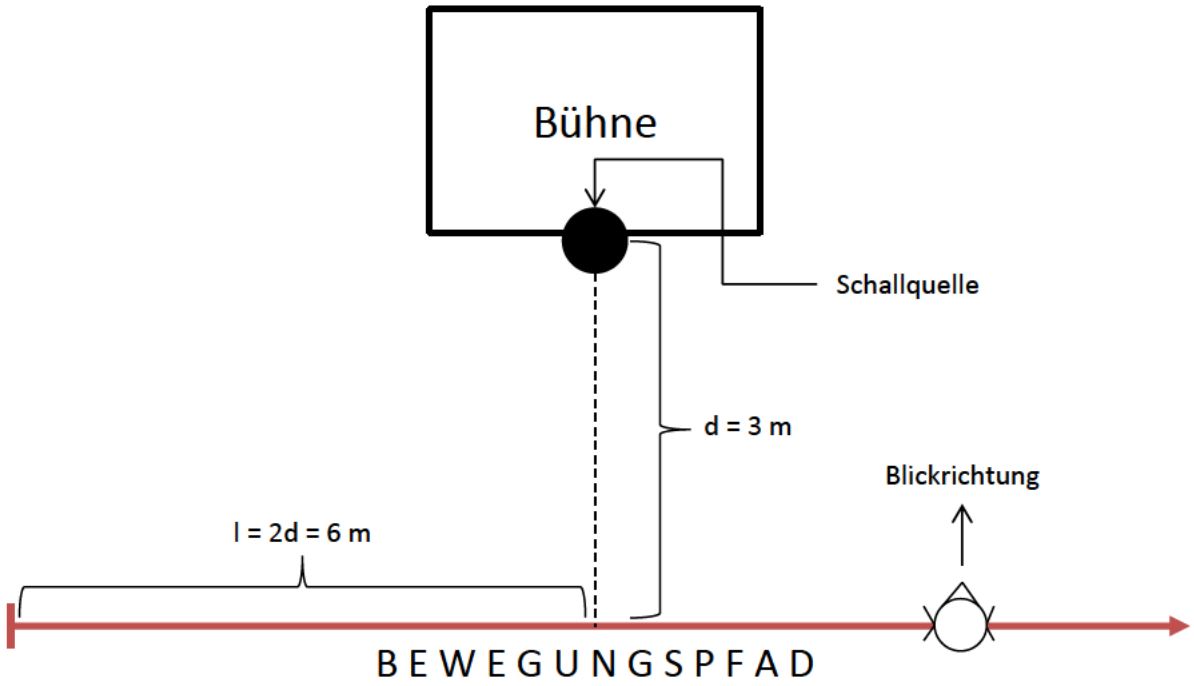

Abbildung 1: Skizze der Umgebung und des Bewegungspfades der Auralisation

2.1.3 Bewegungsgeschwindigkeit

Die Frage nach der veränderten Wahrnehmung bei Variation der Bewegungsgeschwindigkeit ist eine in dieser Arbeit essenziell zu untersuchende Größe. In Betracht kommen hier Bewegungsgeschwindigkeiten, die ein Mensch unter normalen Bedingungen in einer oben genannten Situation, nämlich der Demonstration eines Lautsprechersystems im kleineren Personenkreis, durchführen würde. In Betracht gezogen werden hierbei hauptsächlich langsame Gehgeschwindigkeiten von v = 2 km/h (entsprechend etwa 0,5 m/s) bis zur üblichen Schrittgeschwindigkeit von etwa v = 6 km/h (entsprechend 1,5 m/s). Um auch vergleichsweise schnelles Laufen mit zu berücksichtigen, wird auch eine normale schnellere Laufgeschwindigkeit eines Menschen von v = 12 km/h (entsprechend etwa 3 m/s) angenommen.

2.1.4 Art und Kontur der Phasenänderung

Bei der Art der Phasenänderungen des auralisierten Signals wird auch auf den einfachsten Fall zurückgegriffen: Die Änderung des Phasenwinkels wird nicht frequenzselektiv vorgenommen, sondern gilt für das gesamte Phasenspektrum. Dies entspricht also einer Phasenmodulation des gesamten Signals mit variierendem Modulationsgrad, abhängig vom momentanen Phasenverschiebungswinkel $\Delta\varphi_k$. Der Maximalbetrag des Phasenverschiebungswinkels $\Delta\varphi_{max}$ wird, ähnlich Helmholtz, 1865, S. 190f, auf eine Viertelperiode, entsprechend 90 Winkelgraden, festgelegt. Eine Phasenverschiebung von mehr als 90 Grad würde den Effekt während der virtuellen Bewegung lediglich doppeln, daher werden die Grenzen auf den Bereich zwischen 0 Grad und 90 Grad beschränkt.

Der von der momentanen Hörerposition in Bezug zur Mittelachse Δl_p abhängige Betrag des Phasenverschiebungswinkels $\Delta\varphi_k$ wird für den einfachsten Fall als lineare Funktion (n = 1) angenommen. Darüber hinaus soll untersucht werden, ob andere Kurvenformen (quadratisch mit n = 4, 2 und die entsprechenden Umkehrfunktionen mit n = ¼, ½) zu veränderten Wahrnehmungen führen. Somit wird der Phasenverschiebungswinkel $\Delta\varphi_k$ nach (1) verändert.

$$\Delta\varphi_k = \left(\frac{\Delta l_\mathrm{p}}{\Delta l_\mathrm{pmax}}\right)^n \cdot \Delta\varphi_\mathrm{max} \qquad n = \{4, 2, 1, 0.5, 0.25\} \tag{1}$$

Abbildung 2 veranschaulicht die damit resultierenden Konturen der Phasenverschiebungswinkel $\Delta\varphi_k$ in Abhängigkeit zum im 2.1.2 definierten Hörer-Bewegungspfad.

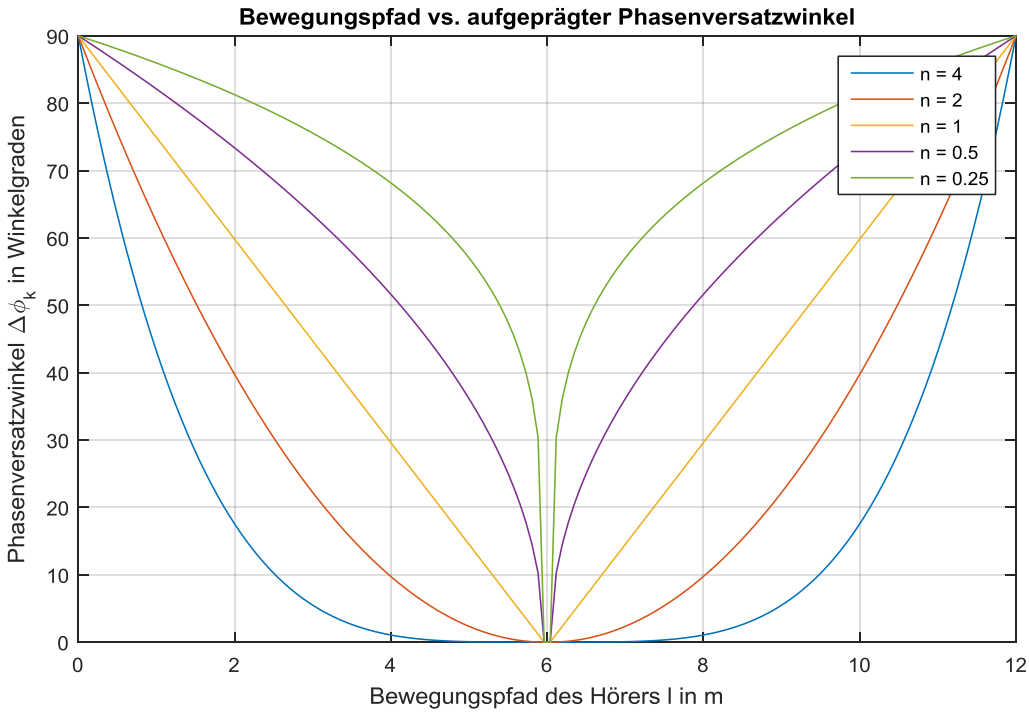

Abbildung 2: Konturen der Phasenversatzwinkel in Abhängigkeit des Hörer-Bewegungspfades

2.1.5 Anregungssignale

Generell sollen als Anregungssignale sowohl natürliches, als auch synthetisch erzeugtes Material zum Einsatz kommen. In Anbetracht der Tatsache, dass die Auralisation angelehnt sein soll an eine Demonstration von Beschallungs-Lautsprechern, mit denen in der Praxis in der Regel Audiosignale mit hoher dynamischer Auflösung wiedergegeben werden, wird sowohl ein in der Dynamik wenig komprimiertes Audiosample eines Drum-Sets mit hinzugemischter Bass- und Rhythmusgitarre verwendet als auch eine ohne Dynamikreduktion produzierte englischsprachige männliche Sprecherstimme. Um die Anregungssignale für den Leser dieser Arbeit zu visualisieren, wird das Spektrogramm jedes Signals in MATLAB® mit Hilfe der *mirspectrum*()-Funktion aus der MIR-Toolbox (Lartillot et al., 2008) berechnet und im Folgenden näher erläutert.

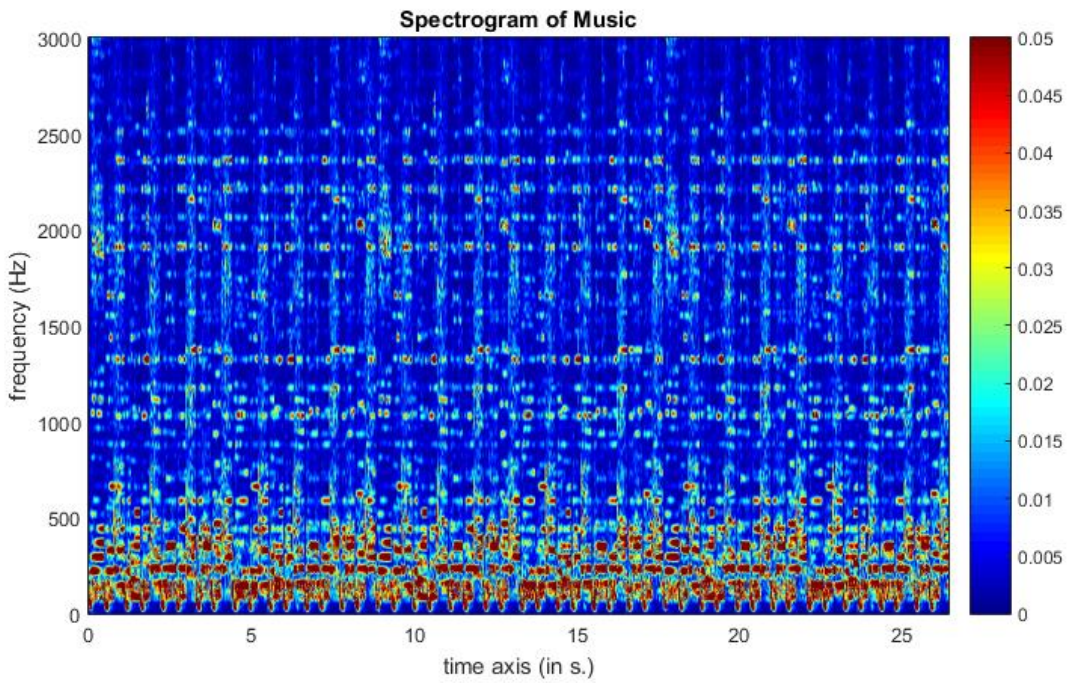

Abbildung 3: Spektrogramm des Musik-Signals

Das Spektrogramm des Musiksignals ist in Abbildung 3 dargestellt. Wie man erkennen kann, konzentriert sich bei dem Musiksignal die meiste Energie im Frequenzbereich unter 250 Hz. Daher stellt es ein gutes Test-Signal für den in 1.2 angesprochenen Effekt der möglicherweise veränderten Bass-Intensitäts-Wahrnehmung bei Veränderung der Phasenlage dar. Um den Effekt noch präsenter darzustellen, wurde der Bassanteil im Musiksignal mit Hilfe eines Constant-Q-PEQs (f_c = 60 Hz, Q = 1, Gain = +6 dB), berechnet in MATLAB® mit der *AKfilter*()-Funktion aus den AKTools, angehoben, bevor das Musiksignal auralisiert wird. Der resultierende Filter-Betragsfrequenzgang ist in Abbildung 4 ersichtlich.

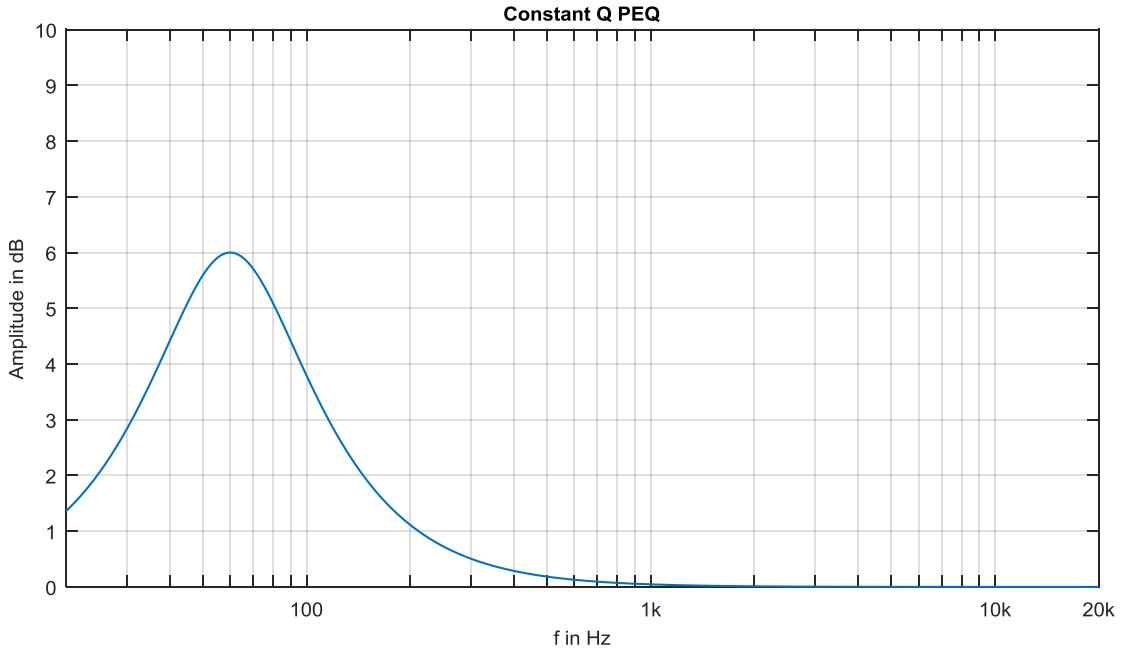

Abbildung 4: Betragsfrequenzgang des Bass-Boost-PEQs für das Musiksignal

Außerdem ist auffällig, dass sich das Spektrum oft und stetig wiederholt, sodass davon auszugehen ist, dass der Hörer bei der virtuellen Bewegung durch das Schallfeld zu jeder Zeit ein spektral ähnliches Schallereignis präsentiert bekommt.

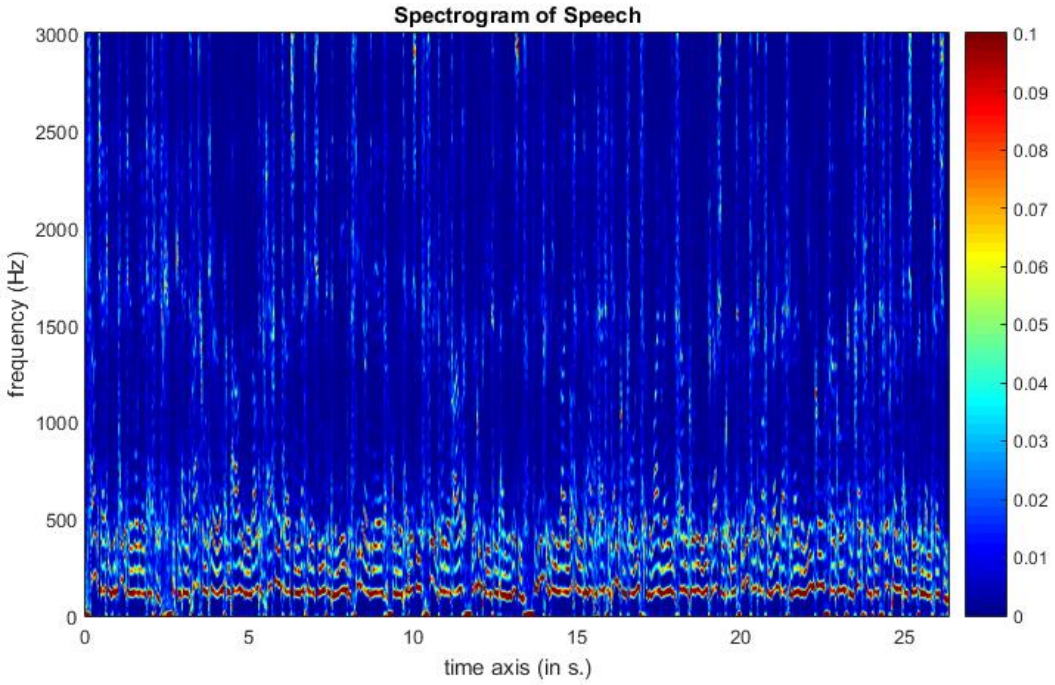

Abbildung 5: Spektrogramm des Sprach-Signals

Das Spektrogramm des Sprachsignals ist in Abbildung 5 dargestellt und lässt erahnen, dass der Sprecher mit relativ konstanter Prosodie gesprochen hat, da sowohl Grundfrequenz als auch Formanten niedriger Ordnung kaum schwanken. Damit ist auch hier sichergestellt, dass der Hörer im virtuellen Raum möglichst zu jeder Zeit ein ähnliches Schallereignis dargeboten bekommt.

Das synthetisch erzeugte Multitone-Signal, in MATLAB® erzeugt mit der AKTools-Funktion *AKsine*(), besteht aus 4 reinen Sinus-Teiltönen gleicher Amplitude mit den Frequenzen 55, 110, 220 und 440 Hz. Es ist also ein auf den Kammerton A gestimmter Klang mit vergleichsweise geringer spektraler Bandbreite, der die Hauptenergie im Bassbereich unter 250 Hz hat. Damit soll im konkreten Fall untersucht werden, ob der Effekt womöglich bei statischen Basstönen, die auch im realen Fall bei tief gestimmten natürlichen Instrumenten oder elektronisch erzeugter Musik durchaus vorkommen können, besonders gut wahrgenommen wird. Das Spektrogramm des Multitone-Signals ist in Abbildung 6 dargestellt.

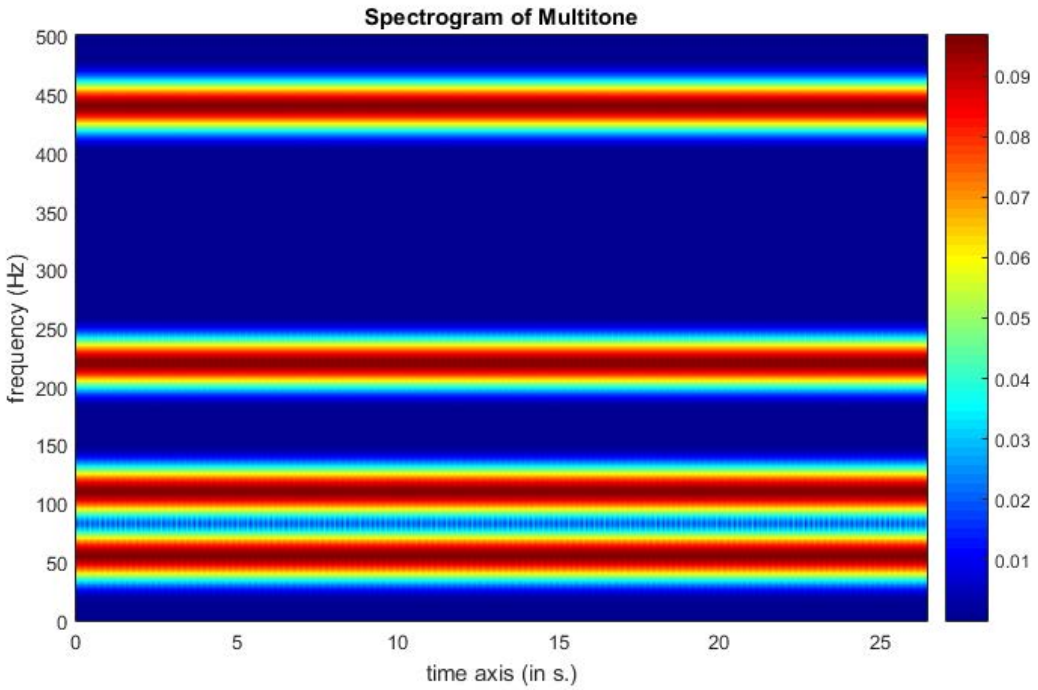

Abbildung 6: Spektrogramm des Multitone-Signals

Zu dem Pink-Noise-Signal, in MATLAB® mit der *AKnoise*()-Funktion aus den AKTools erzeugt, gibt es weitaus weniger zu erklären. Die Rauschleistung fällt hierbei per Definition pro Oktave mit etwa 3 dB ab. Damit entspricht es einem vom menschlichen Gehör in allen Frequenzbereichen in etwa gleich laut empfundenen Rauschen. Dieses Signal stellt spektral das Gegenteil zum Multitone-Signal dar, da es spektral sehr breit über die komplette Hörfläche ausgeweitet ist. Es soll als Ankerpunkt bei den Hörversuchen dienen, da es spektrale Verfärbungen im gesamten hörbaren Bereich aufdecken kann und sich dynamisch nahezu konstant über die gesamte Länge des Bewegungspfades verhält. Abbildung 7 veranschaulicht das Spektrogramm des Pink-Noise-Signals, beschränkt auf den spektralen Bereich von 0 Hz bis 3 kHz. Da hier eine lineare Darstellung der Frequenzachse vorliegt, kann man den spektralen Pegelabfall zu hohen Frequenzen hin erkennen.

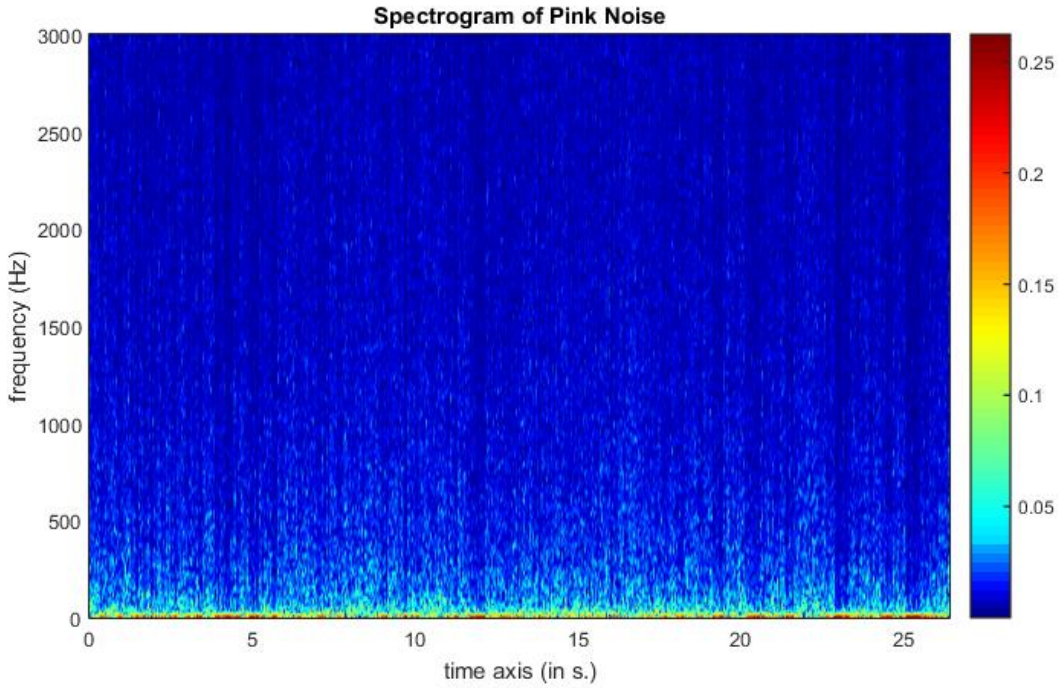

Abbildung 7: Spektrogramm des Pink-Noise-Signals

Die zeitliche Dauer der jeweiligen Signale T (in Sekunden) bestimmt die Dauer der auralisierten Bewegung und damit die virtuelle Bewegungsgeschwindigkeit des Hörers. Somit ist die Dauer der Signale T abhängig von der jeweiligen Ziel-Bewegungsgeschwindigkeit v (in km/h) nach 2.1.3 und berechnet sich mit (2)

$$T = \frac{2l \cdot 3{,}6}{v} \tag{2}$$

2.2 Binaurale Synthese

2.2.1 Modifikation der *AKshAura*()-Funktion

Als erster Schritt wird die *AKshAura*()-Funktion modifiziert, die Bestandteil der AKTools ist, um die Möglichkeit, den gesamten Phasengang der HRIRs aus dem FABIAN-Datensatz abhängig von der momentanen Position auf der Trajektorie nach den in Abschnitt 2.1.4 beschriebenen Funktionen zu verändern.

Die Funktion baut auf der *AKisht*()-Funktion auf, die mittels sphärisch-harmonischer Synthese nach Rafaely, 2015 aus den im SOFA-Format vorliegenden sphärisch-harmonischen Koeffizienten des gemessenen FABIAN-Datensatzes die komplexen Übertragungsfunktionen der einzelnen HRTFs pro Winkelabschnitt der Ziel-Trajektorie berechnet. Die jeweiligen zusätzlich aufzuprägenden Phasenwinkelverschiebungen $\Delta\varphi_k$ für die einzelnen komplexen Übertragungsfunktionen H_k werden nach diesem Schritt mit (3) abhängig ihrer momentanen Winkelposition gemäß der jeweiligen Zielfunktion nach (1) verschoben.

$$H_{\Delta\varphi k} = H_k \cdot \mathrm{e}^{\left(\mathrm{j}\frac{2\pi}{360°}\cdot\Delta\varphi_k\right)} \tag{3}$$

Der aufgeprägte zusätzliche Phasenversatz erreicht im konkreten Fall an den Randpositionen der Trajektorie seinen Maximalwert (entspricht einem Phasenversatz des originalen Phasenfrequenzgangs um den jeweils definierten Maximalwert), während er in der Mitte stets 0 Grad aufweist (entspricht dem originalen Phasenfrequenzgang, kein Phasenversatz). Somit ist der implementierte Phasenversatz achsensymmetrisch auf einer gedachten Mittelachse des halben Bewegungspfades, im konkreten Fall bei 6 Metern Lauflänge, und erreicht an der Position dieser Mittelachse immer 0 Grad Phasenversatz (siehe hierzu auch Abbildung 2).

2.2.2 Sphärische Trajektorie

Die *AKshAura*()-Funktion platziert die Schallquelle bei einem Azimut-Wert von 0/360 Grad aus Sicht des Hörers hart vorne, bei 90 Grad hart links, bei 180 Grad hart hinten und bei 270 Grad hart rechts. Die horizontalen Winkel der sphärischen Ziel-Trajektorie auf der Azimut-Ebene berechnen sich, ausgehend von der in Abschnitt 2.1.2 definierten Grundgeometire der Auralisation und der Hörer-Position l_p, nach (4) für die erste Hälfte der Trajektorie, die die Schallquelle aus Sicht des Hörers von Richtung rechts nach vorne bewegt.

$$az_r = 360 - \left(90 - arctan\frac{d}{l_\mathrm{p}}\right) \tag{4}$$

Der zweite Teil der Trajektorie berechnet sich nach (5) für eine Bewegung der Schallquelle von vorne Richtung links.

$$az_l = 90 - arctan\frac{d}{l_\mathrm{p}} \tag{5}$$

Da eine rein horizontale Bewegung ausgeführt werden soll, betragen die Winkel der Ziel-Trajektorie auf der Elevations-Ebene stets 0 Grad für alle Winkelwerte korrespondierend zur Azimuth-Ebene.

2.2.3 Amplituden-Hüllkurve

Wie schon im vorigen Abschnitt diskutiert, bewegen sich die mit der *AKshAura*()-Funktion gerenderten Schallfelder auf rein sphärischen Bewegungspfaden auf der horizontalen Hörebene, im konkreten Fall also aus Sicht des Hörers halbkugelförmig von rechts über vorne nach links. Um diesen halbkugelförmigen Verlauf auf einen geraden Bewegungspfad aufzubrechen, muss der Pegel der generierten Schallfelder positionsabhängig abgeschwächt werden. Es wird hierbei von einer omnidirektional abstrahlenden Punktschallquelle ausgegangen, deren Schalldruck unter den idealerweise angenommenen Freifeldbedingungen nach dem linearen Zusammenhang (6) proportional mit dem Abstand r zur Schallquelle sinkt (Möser, 2012, S. 76).

$$p \sim \frac{1}{r} \tag{6}$$

Der auf 1 normierte Amplituden-Skalar A für alle Positionen auf dem halben Bewegungspfad berechnet sich hiernach relativ zum maximalen Abstand der Mittelposition auf dem halben Bewegungspfad Δl_{pmax} je nach Abstand des Hörers von ebendieser Mittelposition Δl_p nach (7).

$$A = \sin\left(arctan\frac{\Delta l_{\mathrm{pmax}}}{\Delta l_{\mathrm{p}}}\right) = \frac{\frac{\Delta l_{\mathrm{pmax}}}{\Delta l_{\mathrm{p}}}}{\sqrt{\left(\frac{\Delta l_{\mathrm{pmax}}}{\Delta l_{\mathrm{p}}}\right)^2 + 1}} \tag{7}$$

Die damit resultierende Amplituden-Hüllkurve für den halben Bewegungspfad wird an der Mitte des Bewegungspfads gespiegelt und verhält sich damit für den gesamten Bewegungspfad wie in Abbildung 8 dargestellt. Die skalaren Abschwächungswerte für jede der Positionen auf dem Bewegungspfad sind hier zur besseren Darstellbarkeit logarithmieret als dBFS-Werte dargestellt.

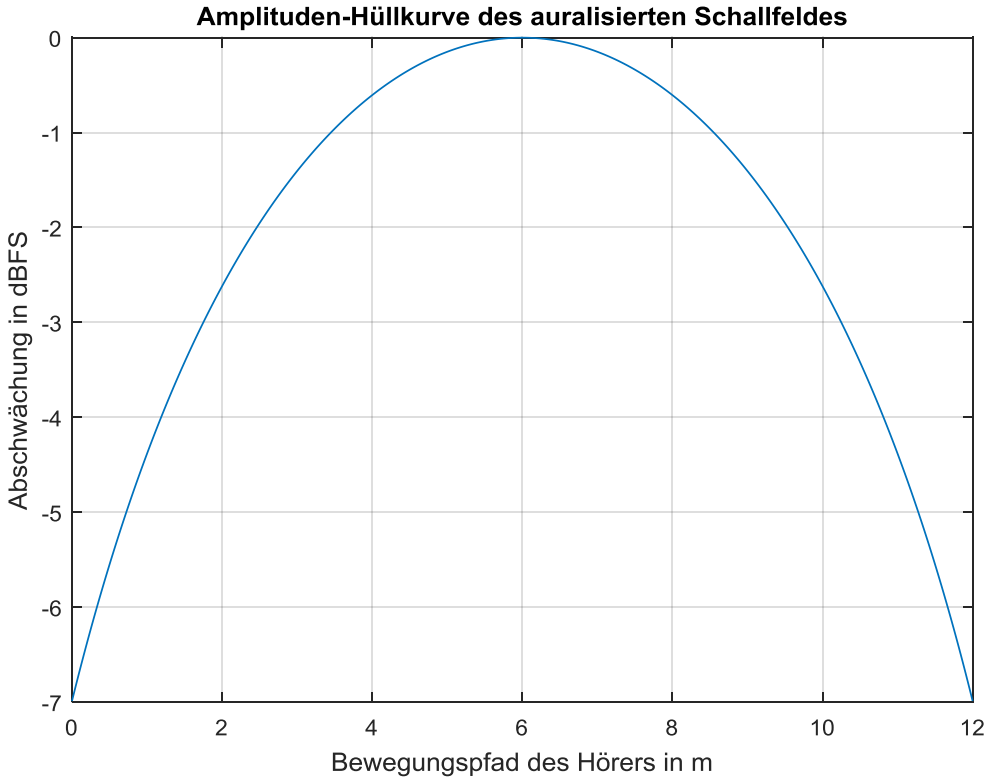

Abbildung 8: Abschwächung des auralisierten Schallfeldes vs. Bewegungspfad des Hörers

2.2.4 Kopfhörer-Kompensation

Wie bereits in Møller, 1992 erwähnt, verhalten sich übliche Kopfhörer (stellenweise hochgradig) nicht linear und müssen zur Reproduktion binauraler Signale auf ein möglichst lineares Übertragungsverhalten hin entzerrt (kompensiert) werden, damit sich beim Hörer ein natürliches und glaubhaftes Wahrnehmen der generierten binauralen Schallfelder ergibt. Brinkmann und Lindau, 2010 haben hierzu festgestellt, dass die Erfassung und Kompensation individueller Kopfhörer-Übertragungsfunktionen (HpTFs), die direkt mit dem Ziel-Kopfhörer und auf dem Kopf des Hörers durchgeführt werden, für jeden einzelnen Hörer die Natürlichkeit der binauralen Synthese steigern können. Dies wäre im Rahmen dieser Arbeit jedoch mit großem Aufwand verbunden, da jede einzelne Versuchsperson zunächst mit einem Messmikrofon im Ohrkanal ausgestattet werden müsste, die individuelle HpTF damit gemessen müsste, daraus der entsprechende Kompensationsfilter errechnet und in die zu darbietenden binauralen Signale gefaltet werden. Um

den Aufwand an dieser Stelle geringer zu halten, wurde als Kompromiss aus der Messung einer einzigen (nicht-individuellen) HpTF mit Hilfe des FABIAN-Kunstkopfs ein nicht-individueller Kompensationsfilter errechnet. Lindau und Brinkmann, 2012 sowie Schärer und Lindau, 2009 haben gezeigt, dass sich nach dieser Methodik keine nennenswert signifikanten Abweichungen gegenüber individuellen Kopfhörer-Kompensationsfiltern ergeben, die meisten Probanden haben sogar den von FABIAN stammenden nicht-individuellen Kopfhörer-Kompensationsfilter als natürlicher empfunden als ihre individuelle HpTF.

Der hierfür mit FABIAN gemessene und für alle Hörversuche dieser Arbeit verwendete Kopfhörer war ein Sennheiser HD800 (Abbildung 9). Die Bauform dieses Kopfhörers ist offen und ohrumschließend (zirkumaural). Dadurch wird die akustische Strahlungsimpedanz aus Sicht des Ohrkanals nicht beeinflusst und entspricht nahezu jener, die auch ohne aufgesetzten Kopfhörer vorhanden ist. Nach Møller, 1992 ist dies für eine natürliche Reproduktion binauraler Ohrsignale ein unabdingbarer Faktor.

Abbildung 9: Sennheiser HD800 auf FABIAN-Kunstkopf (Brinkmann, 2017)

Der Kompensationsfilter-Impuls für diesen Kopfhörer ist bereits in der zum Zeitpunkt dieser Arbeit aktuellen Version der FABIAN-Database vorhanden (Brinkmann, 2017). Er wurde aus dem komplexen Frequenzgang der FABIAN-Messung nach der Methodik der regulierten Inversion (Schärer und Lindau, 2009) mit Hilfe der in den AKTools enthaltenen Funktion *AKregulatedInversion*() gewonnen. Hierbei wurde eine manuelle Regulierungsfunktion erstellt, die aus zwei parametrischen EQs und einem High- und Low-Shelf-Filter besteht. In Abbildung 10 sind die Vorgänge und das Kompensationsergebnis für das linke Ohr dargestellt. Die beiden obere Grafik veranschaulichen den über 12 An- und Absetzvorgängen komplex gemittelten Betragsfrequenzgang des Sennheiser HD800, gemessen mit dem FABIAN-Kunstkopf (blaue Linie) und den jeweils dazu gültigen regulierten Bereich (schwarze Linie). Die mittlere Grafik zeigt den Betragsfrequenzgang des daraus errechneten Kompensationsfilters. Die untere Grafik stellt schließlich das Ergebnis des Kompensationsvorgangs dar, was dem errechneten Kopfhörerfilter, angewendet auf den gemessenen komplexen Kopfhörerfrequenzgang, entspricht. Das Ergebnis ist ein nahezu linearer Frequenzgang zwischen 50 Hz und 8 kHz mit aufgrund der Regulierung nur leicht kompensierten sehr schmalbandigen Resonanzen im oberen Hochtonbereich, die mutmaßlich aufgrund destruktiver Interferenzen an der Cavum Conchae entstehen (Schärer und Lindau, 2009). Sie lassen sich aufgrund der zur Kompensation nötigen sehr steilflankigen und schmalbandigen Anhebungen nicht sinnvoll entzerren, da sie als Ringing-Artefakte hörbar werden und das Klangbild insgesamt verschlechtern würden.

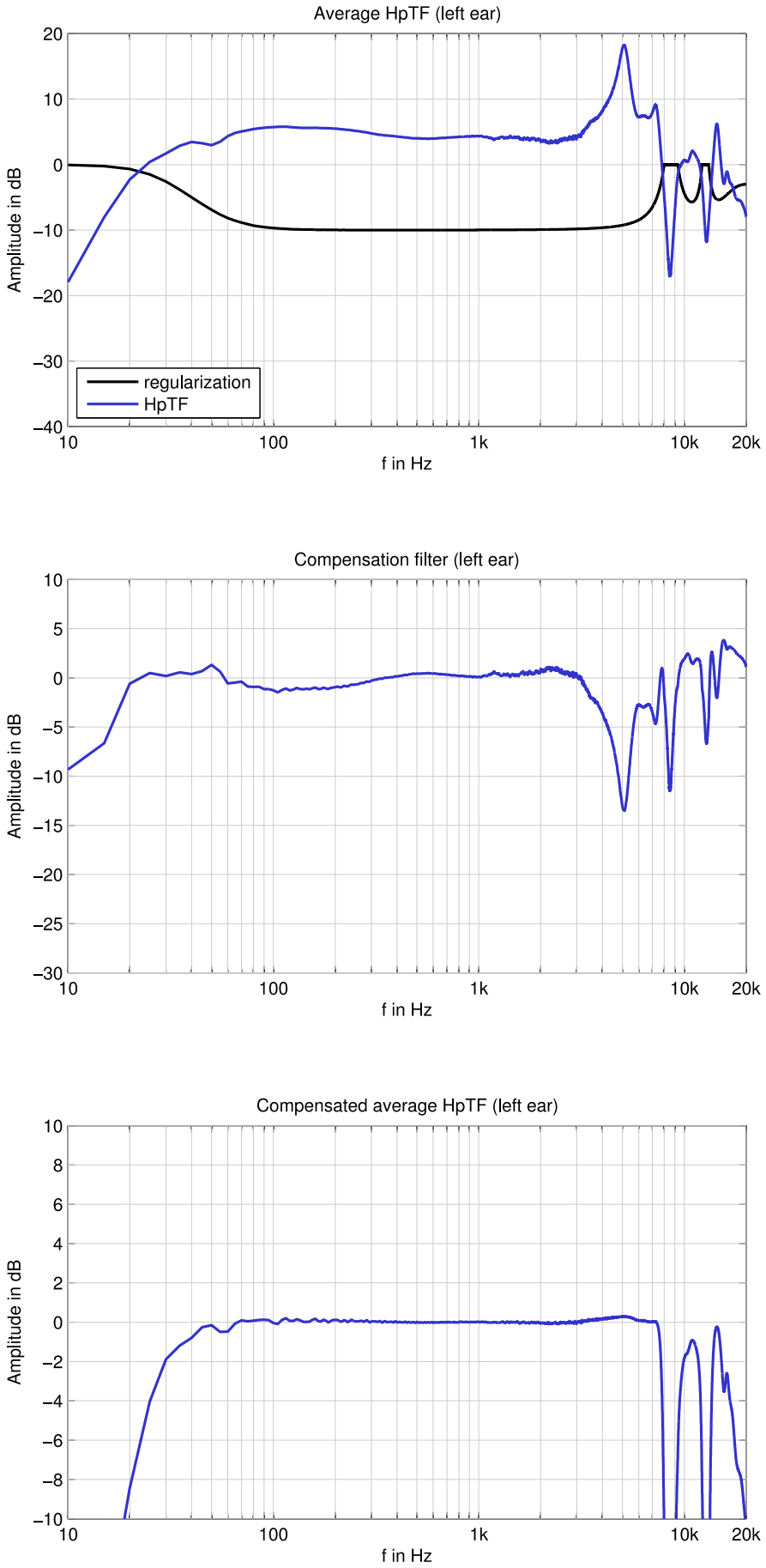

Abbildung 10: Ermittlungsvorgang des Kompensationsfilters für den HD800 (Brinkmann, 2017)

2.2.5 Loudness-Matching

Nachdem der Kopfhörer-Kompensationsfilter auf die gerenderten binauralen Stimuli gefaltet wurde, wird als letzter Schritt die Lautheit aller vier verschiedenen Anregungssignale aneinander angeglichen. Die unterschiedlichen Energiedichten der einzelnen Signale führen zu einer nicht einheitlichen Lautheitswahrnehmung, die es unbedingt zu vermeiden gilt. Der Grund hierfür ist, dass aufgrund der Linien gleicher Lautstärke-Wahrnehmung (Möser, 2012, S. 10) das lautere Signal als spektral voller und angenehmer empfunden wird, da das Gehör mit steigender Lautheit empfindlicher für tiefe und hohe Frequenzen wird. Somit würde nicht auszuschließen sein, dass empfundene spektrale Änderungen auf die Varianz in der Lautheit zurückzuführen sind und nicht auf die zu untersuchenden Effekte der Phasenvariationen.

Die Lautheits-Anpassung erfolgt hierbei mit dem gleichen Kopfhörer, der auch für die Hörversuche verwendet wurde. Ermittelt wurde die Lautheit folgendermaßen: Ein Expertenhörer verglich durch direktes A/B-Umschalten alle Stimuli miteinander und stellte iterativ dabei Abschwächungs-Skalare ein, bis alle Stimuli beim schnellen hintereinander Hören gleich laut wirkten. Die hiermit ermittelten Abschwächungs-Werte sind in Tabelle 1 aufgelistet. Der nach ITU-Empfehlung BS.1770 berechnete Lautheitswert ist jeweils zum Vergleich mit in der Tabelle enthalten (ITU, 2015).

Tabelle 1: Ermittelte Loudness-Werte aller Anregungssignale

Anregungssignal	Abschwächung [dB]	LU-Wert [LUFS]
Pink Noise	-3	-15,59
Multitone	-11,5	-6,44
Musik	0	-15,79
Sprache	-1,5	-16,96

Nicht verwunderlich hierbei ist, dass die nach Gehör eingestellten Abschwächungswerte in ihren Tendenzen den algorithmisch ermittelten Werten nach ITU BS.1770 ähneln. Das Multitonsignal wurde von beiden Methoden als das lauteste Signal wahrgenommen. Dies ist auch zu erwarten gewesen, da es zeitlich und spektral konstant ist und keine kleinen dynamischen Pegelsprünge aufweist, die die empfundene Lautheit senken würden.

3 Perzeptuelle Evaluation

3.1 Vorbetrachtungen

3.1.1 Verwendete Hard- und Software

Die perzeptive Evaluation des zu untersuchenden Effektes wird mit Hilfe der am Fachgebiet Audiokommunikation der TU Berlin entwickelten MATLAB®-Toolbox WhisPER durchgeführt (Ciba et al., 2014). Diese lief innerhalb der MATLAB®-Version R2015a auf einem Lenovo ThinkPad W550s-Laptop mit Windows 7 Professional x64 als Betriebssystem. Als Audio-Interface und Kopfhörerverstärker für den Sennheiser HD800-Kopfhörer kam ein RME ADI-2 PRO zum Einsatz. Der gesamte Aufbau ist schematisch in Abbildung 11 skizziert.

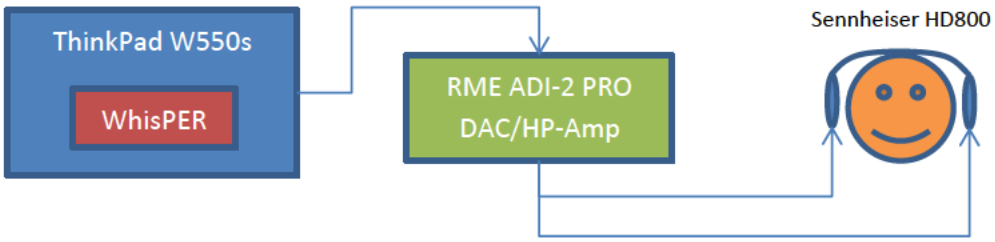

Abbildung 11: Schematischer Aufbau der Hard- und Software für die Hörversuche

3.1.2 Modifikation der Audio-Playout-Routine von WhisPER

Da das in WhisPER integrierte Audio-Playout noch auf der veralteten MATLAB®-eigenen *audioplayer*()-Funktion basiert und prinzipiell keine Low-Latency Audio-Playout-Funktionalität über ASIO-Treiber oder Ähnlichem bietet, wird die in den AKTools enthaltene Playrec-Routine als alternative Sound-Playout-Routine in WhisPER implementiert. Im Wesentlichen wird hier wie folgt vorgegangen: Zunächst wird beim Starten von WhisPER über die whisper.m überprüft, ob die AKTools mit Playrec vorhanden sind. Wenn nicht, wird über einen globalen True/False-Wert auf die ursprüngliche MATLAB®-eigene *audioplayer*()-Routine zurückgegriffen. Anschließend wird beim Öffnen des mainmenu.m-GUIs überprüft, ob ein Konfigurations-File (PR.mat) für Playrec mit den Einstellungen von Samplerate, Audiogerät, Kanalkonfiguration und Ähnlichem vorhanden ist. Falls ja, wird es geladen, falls nicht, erscheint ein zusätzliches GUI, um diese Einstellungen zu konfigurieren und in ein entsprechendes Konfigurations-File zu speichern. Außerdem wird ein Link zu dem Konfigurationsfenster im GUI erstellt, welches auch über den GUI-Reiter „settings/audio“ des Hauptmenüs aufgerufen werden kann. Somit ist eine Änderung der Audio-Playout-Konfiguration jederzeit auch vom WhisPER-Hauptmenü aus möglich.

Während der jeweiligen Hörversuchs-Routine wird immer dann, wenn ein Stimulus abgespielt werden soll, die *present_stimulus*()-Funktion aufgerufen, die zunächst über eine weitere implementierte *prStop*()-Funktion den Audio-Buffer von Playrec leert (und damit die Wiedergabe jeglicher gerade laufender Stimuli stoppt), den jeweiligen neuen Stimulus in den Audio-Buffer von Playrec lädt und ihn schließlich über das in dem Konfigurations-File angegebene Audiogerät ausgibt.

Als letzter Schritt wird die *prStop*()-Funktion auch an entsprechender Stelle in den Code der einzelnen Hörversuchs-Routinen integriert, die eine Stop-Taste im GUI haben und/oder an

bestimmten Stellen sicherstellen müssen, dass jegliche Stimuli gestoppt werden, beispielsweise beim Schließen eines Fensters.

Insgesamt lässt sich nun auch reaktionsschnell und ohne hörbare Latenz in der ABX- oder ABCHR-Versuchsumgebung zwischen zwei Stimuli hin- und herschalten, was einen angenehmen und schnellen A/B-Vergleich der Stimuli ermöglicht, wie er zuvor mit der latenzbehafteten *audioplayer*()-Routine nicht möglich gewesen ist.

3.2 Informeller Hörversuch

3.2.1 Methodik

Ein informell durchgeführter Hörversuch soll zunächst eine grobe Orientierung liefern, wie stark der Effekt der ortsabhängigen Phasenwinkelverschiebung $\Delta\varphi$ von Expertenhörern wahrgenommen wird. Hierzu wird mittels den in den vorigen Abschnitten dargestellten Methodiken die Grundszenerie mit maximalen Phasenverschiebungswinkeln von $\Delta\varphi_{max}$ = 0 … 90 Grad in 2-Grad-Schritten, also in insgesamt 45 Ausprägungsstufen, auralisiert, wobei die Kurvenform des ortsabhängigen Phasenverschiebungswinkels $\Delta\varphi$ immer n = 1, also einem rein linearen Anstieg, entspricht (siehe auch Abbildung 2). Dies wird für zwei unterschiedliche Bewegungsgeschwindigkeiten (v = 6 km/h und 12 km/h) und Anregungssignale durchgeführt, nämlich dem Pink-Noise als künstliches und dem Musik-Sample als natürliches Signal. Hiermit ergeben sich also insgesamt 2*2*45 = 180 unterschiedliche Stimuli.

Die Stimuli werden nach dem 3-Alternative-Forced-Choice-Paradigma (3-AFC, 2-Down-1-Up) mit WhisPER präsentiert (Abbildung 12).

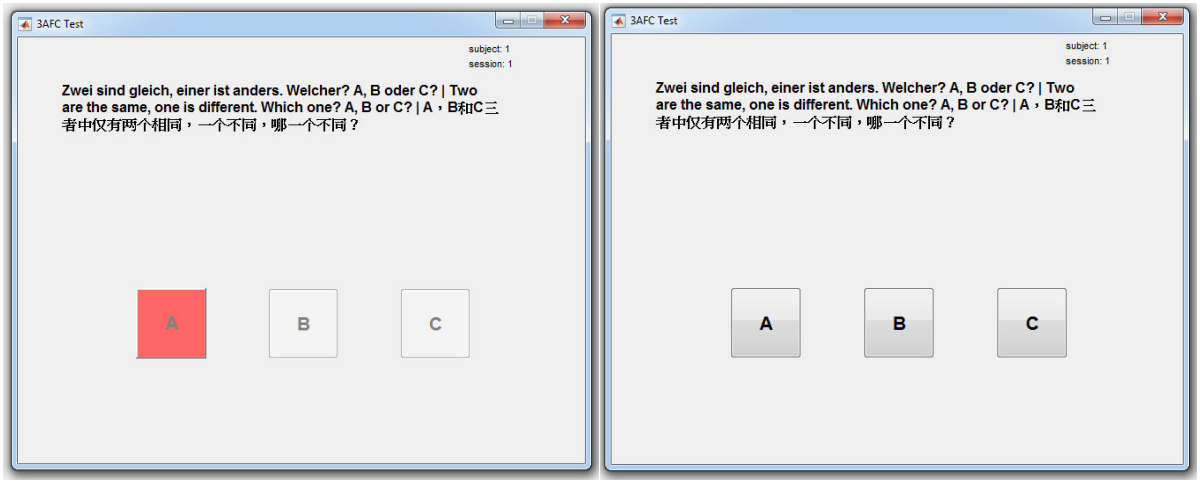

Abbildung 12: GUI des informellen 3AFC-Hörversuchs

Der Proband bekommt also pro Versuch insgesamt drei Stimuli (A, B und C) präsentiert (der rote Button bedeutet, dass der Stimulus gerade wiedergegeben wird), wovon zwei identisch und ohne Effekt ($\Delta\varphi_{max}$ = 0 Grad) sind und einer den Effekt der Phasenverschiebung in unterschiedlicher Ausprägung beinhaltet. Anschließend muss der Proband sich entscheiden, in welchem der drei Stimuli er den Effekt gehört hat, indem er im GUI einen von den drei Buttons (A, B oder C) anklickt. Unmittelbar danach startet der nächste Versuch. Hat der Proband den Effekt erkannt und richtig zugeordnet, wird die Effektstärke ($\Delta\varphi_{max}$) verringert, andernfalls erhöht. So nähert sich der Proband von Versuch zu Versuch an einen für ihn individuellen Schwellwert der wahrnehmbaren Effektstärke ($\Delta\varphi_{max}$) an, bei dem er keinen Unterschied mehr feststellen kann.

Hierbei wird jedoch nicht, wie gerade skizziert, ein einfaches Staircase-Verfahren, sondern ein adaptives psychometrisches Maximum-Likelihood-Verfahren nach der ZEST-Methode (King-Smith et al., 1994) angewendet, da dies durch die Einbeziehung einer vorher abgeschätzten und parametrisierten psychometrischen Zielfunktion sowie durch die Zwischenergebnisse des Versuchs während des Versuchsablaufs die gesamte Dauer des Verfahrens erheblich verkürzen kann, ohne dass die Genauigkeit des Schwellwert-Ergebnisses allzu groß darunter leidet (Otto und Weinzierl, 2009). Die Erkennbarkeit des Effektes wird als eher gering vermutet, daher hilft diese Vorgehensweise auch, Ermüdungs- und Konzentrationsschwierigkeiten der Versuchsperson mit zunehmender Versuchsdauer vorzubeugen.

Für den informellen Hörversuch wird das ZEST-Verfahren folgendermaßen parametrisiert: Grundlage ist eine logistische psychometrische Funktion mit einer (relativ großen) Spannweite von $\beta = 6$ und einer Fall-Rate von $\lambda = 1$ % (Abbildung 13). Hiermit wurde eine relativ flache psychometrische Funktion eingestellt, weil noch nicht genau eingeschätzt werden kann, in welchen Bereichen der Effektstärke ($\Delta\varphi_{max}$) sich der Schwellwert voraussichtlich befinden wird.

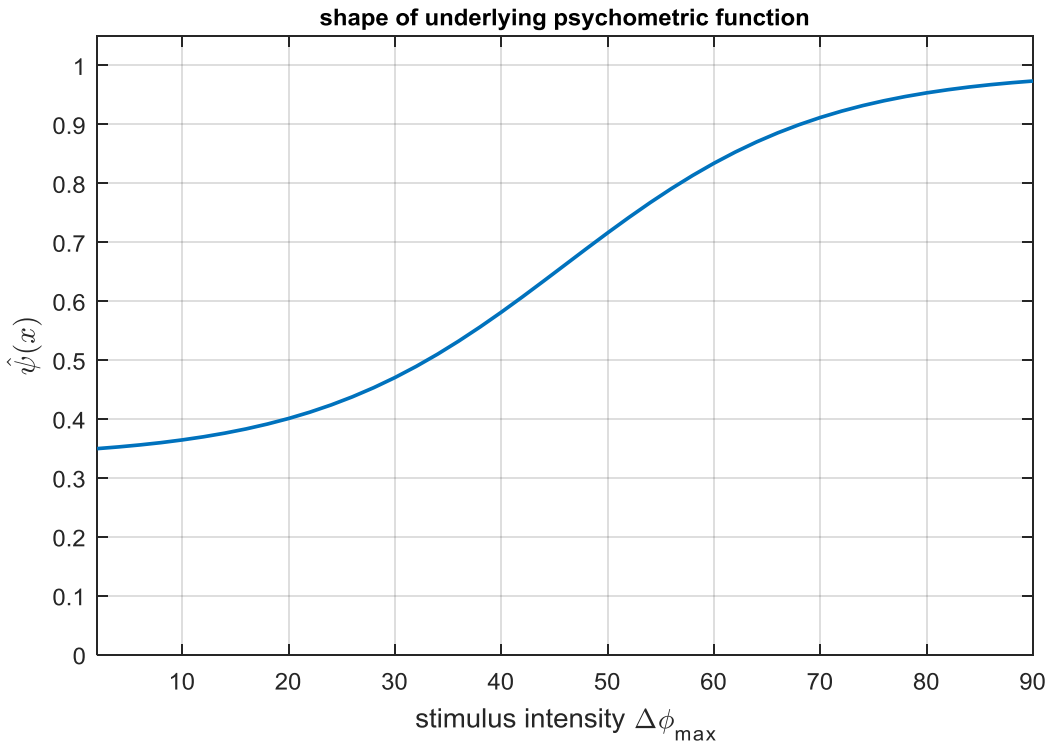

Abbildung 13: Psychometrische Funktion des informellen Hörversuchs

Die a-priori-Wahrscheinlichkeitsdichtefunktion wird nach Gauss-Verteilung angenommen, wobei der Mittelwert auf die Mitte der beiden Extremwerte ($\Delta\varphi_{max}$ = 0 Grad und 90 Grad), also auf $\Delta\varphi_{max}$ = 45 Grad und einer Standardabweichung von $\Delta\varphi_{max}$ = 30 Grad eingestellt wird (Abbildung 14). Außerdem wird festgelegt, dass der jeweilige Durchgang dann beendet wird, wenn ein Konfidenzlevel von 95 % (mit einer Breite n = 7) bezüglich der Schwellwertfindung erreicht ist.

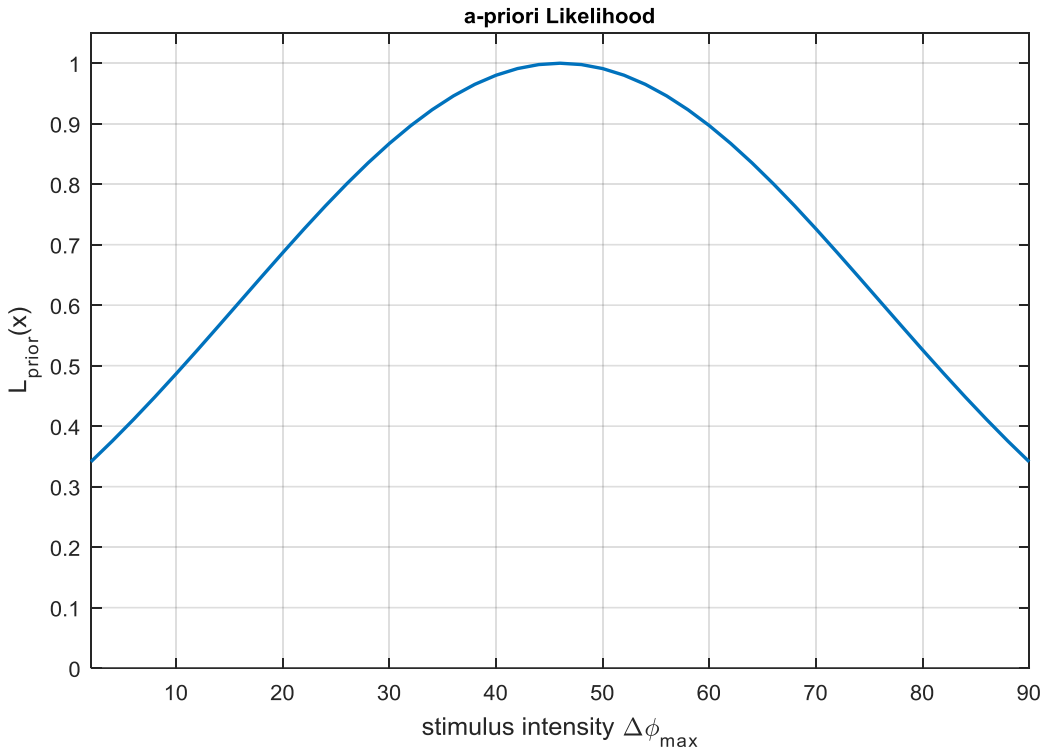

Abbildung 14: A-priori-Funktion des informellen Hörversuchs

3.2.2 Ergebnisse

Der informelle Hörversuch wurde mit insgesamt 6 Expertenhörern aus der Audiokommunikation mit einem geschulten Gehör und guten bis sehr guten Kenntnissen von Hörversuchen durchgeführt. Als Hörraum diente dabei jeweils immer ein ruhiger Büroraum, in dem die im Abschnitt 3.1.1 beschriebene Hard- und Software aufgebaut und betrieben wurde. Der Abhörpegel wurde von jedem Expertenhörer vor Versuchsbeginn auf die individuellen Präferenzen eingestellt. Das deskriptive Ergebnis der Schwellwerte von $\Delta\varphi_{max}$ in Abhängigkeit des Anregungssignals und der Bewegungsgeschwindigkeit ist in Abbildung 15 als Boxplot dargestellt.

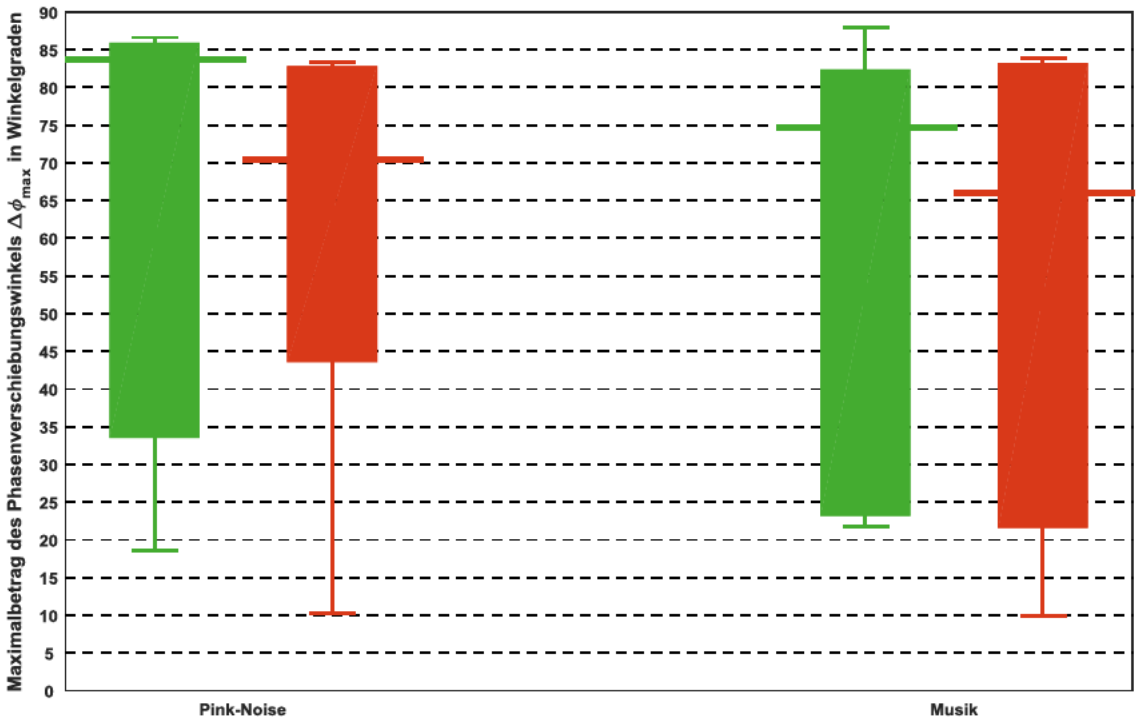

Abbildung 15: Ergebnis des informellen Hörversuchs (grün v=5 km/h, rot v=12 km/h)

Da die gewonnenen Daten aufgrund der kleinen Stichprobengröße nicht normalverteilt sind, werden hier nur Spannweiten anstatt von Konfidenzintervallen dargestellt. Auffällig ist, dass die kleinsten Messwerte nicht unter $\Delta\varphi_{max}$ = 10 Grad reichen, daher wird als erste Erkenntnis für den Haupt-Hörversuch mitgenommen, dass es unnötig ist, Phasenverschiebungswinkel kleiner 10 Grad zu untersuchen. Außerdem kann man erkennen, dass die Median-Werte aller Stimuli um ein $\Delta\varphi_{max}$ von 70 Grad schwanken, sodass als zweite Erkenntnis festgestellt wird, dass die Schwelle der Wahrnehmung des Effekts von erfahrenen Hörern um die 70 Grad liegt und damit vermutlich für unerfahrene Hörer noch höher sein wird. Generell kann damit festgestellt werden, dass die zu untersuchende Effektstärke nicht besonders groß ist, was auch durch die beiden unterschiedlichen Anregungssignale keine wesentliche Veränderung erfährt. Die Tendenz einer etwas stärkeren Streuung zu geringeren Werten hin ist bei den beiden Anregungssignalen mit der schnelleren Bewegungsgeschwindigkeit zu erkennen, was vermuten lässt, dass es bei schnelleren Bewegungen einfacher für den Hörer war, überhaupt einen Unterschied zu erkennen, was als dritte Erkenntnis aus dem informellen Hörversuch mitgenommen werden kann.

3.3 Haupt-Hörversuch

Auf Grundlage der im vorigen Abschnitt gewonnenen Erkenntnisse wird der Haupt-Hörversuch wie folgt implementiert und parametrisiert.

3.3.1 Vorbemerkungen

Der Hörversuch wird im MiniLAB der TU Berlin, einer kleinen Hörkammer (LxBxH ca. 2 m x 1 m x 2.6 m, V = ca. 5 m³) mit guter Außengeräuschdämmung (L_{ruhe} < 30 dB(A)) durchgeführt. Da der verwendete Sennheiser HD800 ein zirkumauraler Kopfhörer offener Bauart ist, bei dem Außengeräusche nahezu ungedämpft zu den Ohren des Hörers gelangen, ist die hohe Außengeräuschdämmung besonders wichtig, damit der Hörer nicht von störenden Geräuschen gestört und/oder abgelenkt wird. Der absolute Schalldruckpegel der dargebotenen Stimuli wird mit

einem RTA/SPL-Meter (t-bone MPAA1), an dem Pink-Noise-Signal orientiert, auf $L_{stimuli}$ = 65 dB(A) eingestellt. Da der Kopfhörer nicht dicht mit dem Ohr abschließen muss, wird der Pegel durch einfaches Einbringen des Messmikrofons zwischen Treiber des Kopfhörers und Pinna bei aufgesetztem Kopfhörer gemessen. Der Absolutpegel von 65 dB(A) wird gewählt, da er ein auch über einen längeren Zeitraum angenehm und ermüdungsfrei anhörbarer Pegel ist. Er hat jedoch den Nachteil, noch in dem Bereich zu liegen, in dem das Gehör nach den Linien gleicher Lautstärke-Wahrnehmung (Möser, 2012, S. 10) im tief- und hochfrequenten Bereich etwas unempfindlich ist. Dies wird jedoch zugunsten eines angenehmeren Hörens und demzufolge einer später einsetzenden Hörer-Ermüdung in Kauf genommen. Abbildung 16 stellt den Hörraum mit dem entsprechend verwendeten Equipment dar.

Abbildung 16: Hörversuchsumgebung im MiniLAB

Jede Versuchsperson wird vor Versuchsbeginn mündlich in den Versuchsablauf eingewiesen. Dabei wird zunächst die Grundszenerie anhand Abbildung 1 erklärt, damit sich die Versuchsperson ein gedankliches Bild von der auralisierten Situation machen kann. Anschließend wird auf die Testverfahren und deren Ablauf an sich eingegangen. Außerdem wird die Versuchsperson entsprechend angewiesen, sich Zeit zu lassen, im Zweifel die erste Intuition walten zu lassen und sich jederzeit bei Bedarf kurze Pausen zwischen den einzelnen Durchgängen zu nehmen.

3.3.2 Teil 1: 3-AFC-Test

Der Hörversuch wird in zwei Teile gesplittet. Zunächst soll in einem Schwellwert-Detektionsverfahren generell herausgefunden werden, in welchen Bereichen der Schwellwert von $\Delta\varphi_{max}$ bei allen Probanden liegt. Der Versuchsablauf ist prinzipiell identisch mit der Methodik, die in Abschnitt 3.2.1 dargestellt wird.

Aufgrund der zu erwartenden eher gering wahrnehmbaren Unterschiede und der daraus resultierenden potenziell schnellen Ermüdung der Probanden wird allerdings nur ein Durchgang durchgeführt, der als Anregungssignal das Pink-Noise-Signal bei einer mittleren Bewegungsgeschwindigkeit von v = 6 km/h verwendet. Sowohl das Anregungssignal als auch die Bewegungsgeschwindigkeit befinden sich nach den Erkenntnissen aus dem informellen Hörversuch im Wahrnehmungskontinuum voraussichtlich in etwa der Mitte des zu untersuchenden Bereichs, deshalb wird sich hier darauf beschränkt. Zudem ist das Pink-Noise-Signal ein beliebtes Test-Signal bei der praktischen Demonstration von Beschallungs-Lautsprechersystemen, daher wird es in dem Fall auch als eine Art Ankersignal angesehen.

Um die Versuchsperson mit dem Verfahren des 3-AFC-Tests und der auralisierten Hörsituation vertraut zu machen, wird dem eigentlichen Testverfahren eine Trainingsphase vorangestellt. Diese besteht nur aus den zwei extremsten Stimuli-Ausprägungen: $\Delta\varphi_{max}$ = 0 Grad und 90 Grad und wird nach 10 Trials abgebrochen. Die Versuchsperson wird auch mit einem akustischen Marker auf den Stimulus mit $\Delta\varphi_{max}$ = 90 Grad hingewiesen. Ihm wird ein kurzer (200 ms dauernder) 1 kHz-Sinuston an den Anfang und auf dem linken Kanal hinzugemischt, um der Versuchsperson das Erkennen des bearbeiteten Stimulus zu zeigen und sie gleichzeitig mit dem Verfahren (drei Stimuli, davon zwei gleiche und einer abweichend) besser vertraut zu machen.

Ausgelegt wird das 3-AFC-Verfahren erneut nach der 2-Down-1-Up-Regel. Die Schwellwertfindung wird auch hier wieder nach dem adaptiven psychometrischen Maximum-Likelihood-Verfahren nach der ZEST-Methode (King-Smith et al., 1994) ausgelegt. Dieses wird nach den Erkenntnissen aus dem informellen Hörversuch diesmal folgendermaßen parametrisiert: Die Spanne bleibt bei $\beta = 6$, ebenso die Fall-Rate $\lambda = 1$ %, was eine psychometrische Funktion identisch mit der Konfiguration des informellen Hörversuchs (nach Abbildung 13) ergibt. Die ebenfalls Gauss'sche a-priori-Wahrscheinlichkeitsdichtefunktion wird jedoch an die Erkenntnisse aus dem informellen Hörversuch angepasst: Der wahrnehmbare Unterschied ist demnach als sehr gering anzunehmen und wird wahrscheinlich im obersten Bereich von $\Delta\varphi_{max}$ liegen (Median bei $\Delta\varphi_{max}$ = 70 Grad), daher wird sie, im Unterschied zum informellen Hörversuch, nach rechts verschoben. Der Mittelwert wird also auf einen $\Delta\varphi_{max}$ von 70 Grad konfiguriert und auch die Standardabweichung wird auf einen $\Delta\varphi$ von 40 Grad erhöht. Die a-priori-Funktion fällt also flacher aus und ist im oberen Bereich des Wahrnehmungskontinuums angesiedelt. Außerdem wird keine Terminierung nach Erreichen eines Konfidenzintervalls gewählt, sondern eine fixe Terminierung nach 20 Versuchen, da sich gezeigt hat, dass die Versuchsdauer bei Versuchspersonen, die den Effekt gut wahrnehmen können und damit von der a-priori-Funktion stärker abweichen, sonst unnötig in die Länge gezogen wird. Die damit resultierende a-priori-Funktion ist in Abbildung 17 dargestellt.

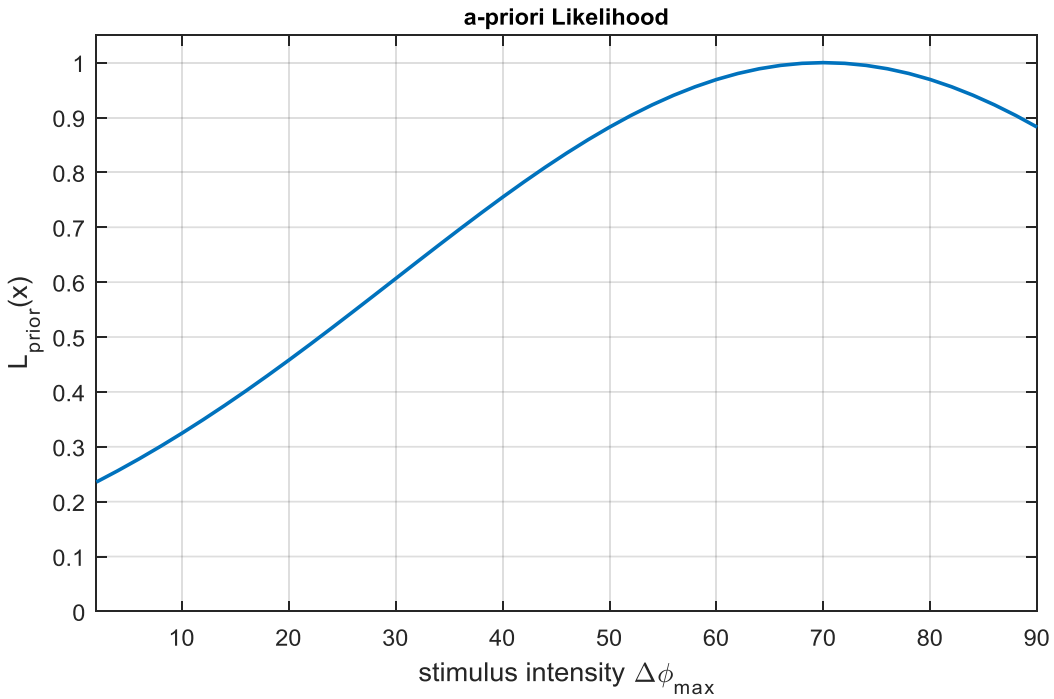

Abbildung 17: A-priori-Funktion des ZEST-3-AFC-Tests

3.3.3 Teil 2: ABCHR-Test

Der zweite Teil des Haupt-Hörversuchs bedient sich des ABCHR-Paradigmas. In WhisPER wird dieses Verfahren prinzipiell mit Hilfe eines GUIs nach Abbildung 18 durchgeführt.

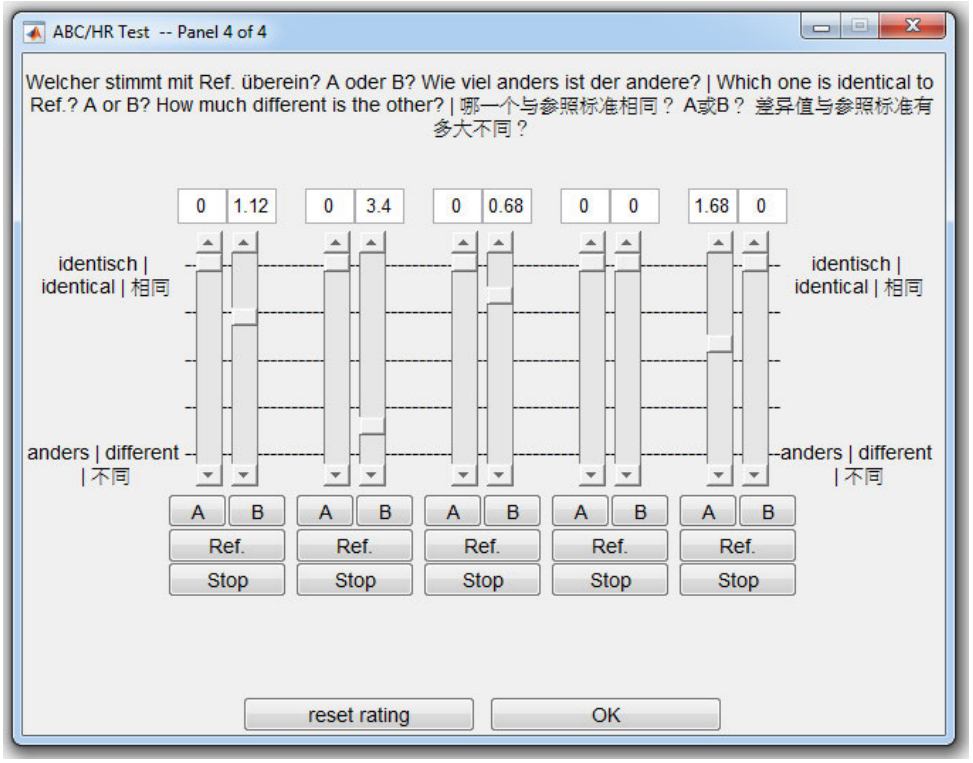

Abbildung 18: Das ABCHR-Test-GUI von WhisPER

Die Versuchsperson erhält jeweils ein Set von Stimulus-Paaren, repräsentiert als Buttons A und B mit jeweils einem zugehörigen Schieberegler. Einer davon bekommt per Zufall immer die Referenz ohne Effekt ($\Delta\varphi_{max}$ = 0 Grad) zugewiesen, der jeweils andere beinhaltet den Stimulus mit Effekt in jeweils unterschiedlichen Ausprägungen. Aus den Erkenntnissen des informellen Hörversuchs geht hervor, dass ein $\Delta\varphi_{max}$ < 10 Grad selbst von Expertenhörern nicht mehr wahrgenommen wird. Daher markiert dieser Wert die unterste Stufe der Effekt-Ausprägung. Ab diesem Wert wird der Effekt äquidistant in 20-Grad-Schritten bis hinauf zu 90 Grad stufenweise erhöht ($\Delta\varphi_{max}$ = 10, 30, 50, 70, 90 Grad). Insgesamt ergeben sich pro Audiosignal und Bewegungsgeschwindigkeit so 5 Abstufungen des Effekts, die mit dem ABCHR-Hörexperiment untersucht werden.

Die Versuchsperson muss im Verlauf des Hörversuchs mit Hilfe der Schieberegler angeben, wie sehr der wahrgenommene Stimulus jeweils von dem für alle Stimulus-Paare gleichen Referenz-Stimulus, der den Effekt nicht beinhaltet, abweicht. Der Schieberegler stellt eine Skala von 0 (identisch mit Referenz) bis 5 (komplett anders als die Referenz) dar und ist in seinem Regelbereich kontinuierlich. Die Versuchsperson muss damit gleich zwei Aufgaben pro Stimulus-Paar erledigen: Zunächst muss herausgefunden werden, welcher der beiden Stimuli mit der Referenz übereinstimmt. Daraus ergibt sich der abweichende Stimulus als der jeweils andere. Er ist von der Versuchsperson mittels Schieberegler in seiner individuell wahrgenommenen Abweichung mit der oben beschriebenen Skala zu bewerten und einzustellen.

Da insgesamt vier verschiedene Audiosignale in fünf verschiedenen Effekt-Ausprägungen mit drei unterschiedlichen virtuellen Bewegungsgeschwindigkeiten untersucht werden, ergeben sich insgesamt 5*4*3 = 60 A/B-Paare (Tabelle 2). Die Darbietungsreihenfolge der unterschiedlichen Anregungssignale und Effekt-Stärken des Phasenversatzes werden randomisiert, während die dreistufige Variation der Bewegungsgeschwindigkeit immer bei v = 3 km/h startet und bei 12 km/h endet.

Tabelle 2: Übersicht aller Stimuli des ABCHR-Hörvergleichs

	v [km/h]			
$\Delta\varphi_{max}$	Pink-Noise	Multitone	Musik	Sprache
10 Grad	3, 6, 12	3, 6, 12	3, 6, 12	3, 6, 12
30 Grad	3, 6, 12	3, 6, 12	3, 6, 12	3, 6, 12
50 Grad	3, 6, 12	3, 6, 12	3, 6, 12	3, 6, 12
70 Grad	3, 6, 12	3, 6, 12	3, 6, 12	3, 6, 12
90 Grad	3, 6, 12	3, 6, 12	3, 6, 12	3, 6, 12

Wie aus Tabelle 2 zu erkennen ist, stellt der gesamte Versuchsablauf eine große Menge an zu untersuchenden Stimuli dar, daher wird dieser Teil mit insgesamt 60 Minuten Hörzeit angenommen. Diese Zeit wird auch als Maximum an Konzentrationszeit einer durchschnittlichen Versuchsperson angenommen. Jede Versuchsperson wird neben der Möglichkeit, jederzeit kurze Pausen einzulegen, zudem angewiesen, bei diesem Teil ganz besonders die erste Grund-Intuition walten zu lassen und die 5-stufige Bewertung der Unterschiedlichkeit nicht allzu präzise vornehmen zu müssen, da sonst die Versuchsdauer in unzumutbare Längen gezogen werden würde.

4 Ergebnisse

4.1 Vorbemerkungen

Insgesamt haben an dem Hörversuch 24 Personen teilgenommen. Jede Versuchsperson hat freiwillig und anonymisiert einen Fragebogen (siehe Anhang) ausgefüllt. Aus der Summe dieser Fragebögen gehen die folgenden Kenntnisse über den soziodemografischen Hintergrund der Menge an Versuchspersonen hervor.

16 der teilnehmenden Personen ordneten sich des männlichen Geschlechts zu, acht Personen haben sich dem weiblichen Geschlecht zugeordnet. Der Alters-Median der gesamten Menge an Versuchspersonen liegt bei 27,5 Jahren. 20 von ihnen haben angegeben, musikalisch und/oder technisch kundig zu sein. Die übrigen vier Probanden stuften sich als nicht musikalisch-technisch kundig ein.

Die Teilnehmer an dem Hörexperiement sind also im Wesentlichen als eher junge, mit dem Fachgebiet kundige Menschen einzuordnen. Deswegen kann auch erwartet werden, dass die Hörfähigkeit der meisten Versuchspersonen über dem Durchschnitt liegt und insgesamt als gut zu bezeichnen ist.

4.2 Teil 1: 3-AFC-Test

Der eher als informell zu betrachtende 3-AFC-Test mit dem Pink-Noise-Signal liefert einen über alle Versuchspersonen gemittelten Median-Schwellwert-Verlauf nach Abbildung 19.

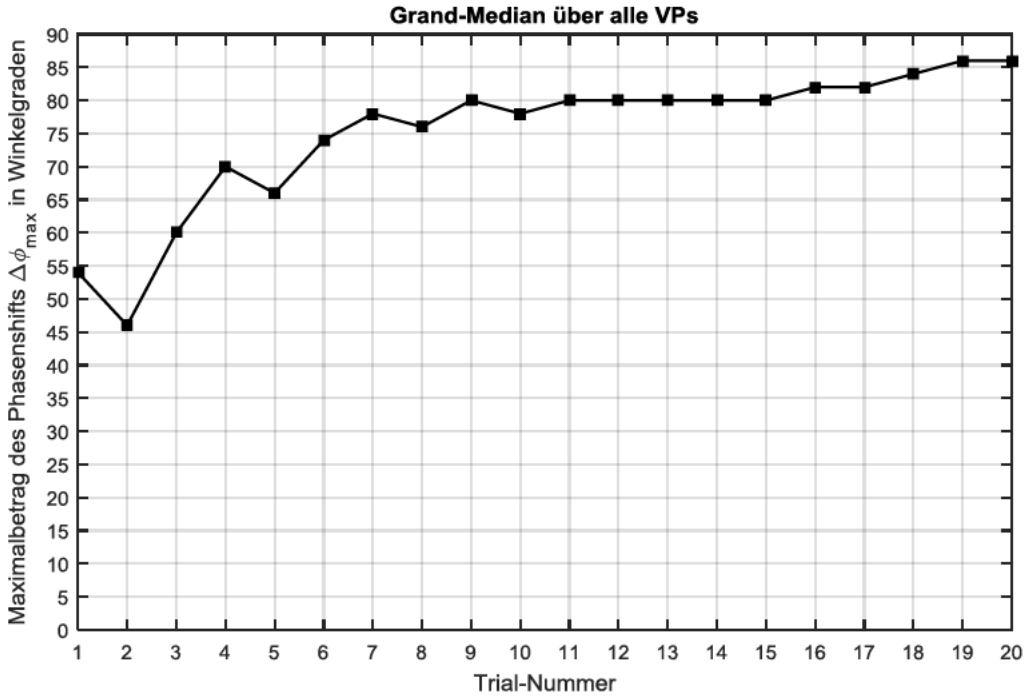

Abbildung 19: Grand-Median-Schwellwert-Verlauf des 3-AFC-Tests

Wie bereits in Abschnitt 3.3.2 erklärt, wird der Einstiegspunkt von $\Delta\varphi_{max}$ (bei Trial-Nummer 1) in die Nähe des im informellen Hörversuch von Expertenhörern ermittelten Median-Wertes gelegt. Wie nun zu erkennen ist, ist dieser Punkt vermutlich zu niedrig gewählt worden, da die meisten

Versuchspersonen schon an diesem Punkt keine Unterschiede feststellen konnten. Resultierend ist daraus zu erkennen, dass sich die Schwellwertkurve zügig zum oberen Bereich der stärksten Ausprägung des Effekts hin bewegt und zum Ende hin nahezu den Maximalwert annimmt. Der Grand-Median-Schwellwert über alle Versuchspersonen und Trials liegt bei einem $\Delta\varphi_{max}$ von 85,75 Grad. Dieser Wert liegt also sehr nahe an der maximalen Effekt-Ausprägung. Abbildung 20 veranschaulicht die Schwellwert-Ergebnisse der einzelnen Versuchspersonen als Boxplot. Da hier mehr die Spanne und der Median-Wert über alle Schwellwerte interessant sind, wird anstatt Konfidenzintervallen die Spanne der einzelnen Schwellwerte dargestellt (schwarze dünne Linien). Die dicke schwarze Linie in der Box symbolisiert den Grand-Median-Schwellwert über alle Versuchspersonen.

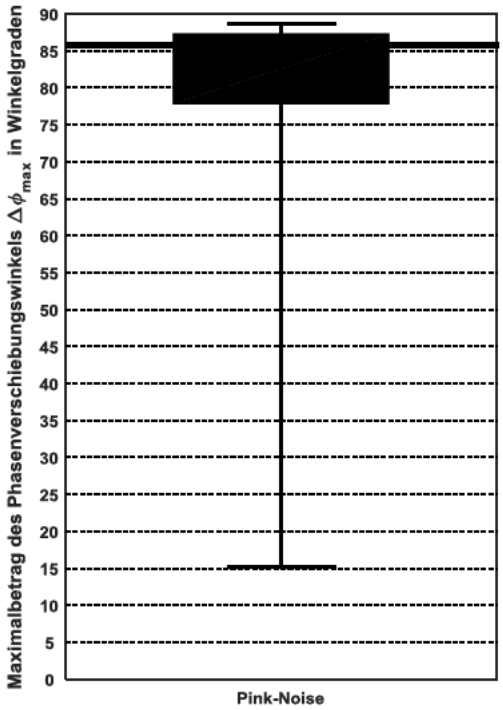

Abbildung 20: Boxplot der Schwellwerte des 3-AFC-Tests

Wie hier gut zu erkennen ist, bewegen sich die Schwellwerte der meisten Versuchspersonen im obersten Bereich zwischen einem $\Delta\varphi_{max}$ von 77 … 87 Grad. Dies bestätigt nur die oben angeführte Beobachtung, dass die Detektion des Effekts, trotz der vorangestellten Trainingsphase, den Versuchspersonen durchweg schwer gefallen ist. Wie sich dies mit unterschiedlichen Anregungssignalen und unter verschiedenen virtuellen Bewegungsgeschwindigkeiten verhält, wird nun im folgenden Abschnitt mit dem ABCHR-Test ermittelt und dargestellt.

4.3 Teil 2: ABCHR-Test

4.3.1 Deskriptiv

Die Ratings des ABCHR-Tests $R_{mit/ohnePhasenversatz}$ werden aus der kontinuierlichen fünfstufigen Skala der einzelnen Schieberegler, die die Versuchspersonen eingestellt haben, zunächst auf den Wert 1 normiert. Anschließend werden sie als normierte Differenzgrade D von -1 bis +1 abgebildet und weiterverarbeitet (8).

$$D = \frac{R_{\text{mitPhasenversatz}}}{5} - \frac{R_{\text{ohnePhasenversatz}}}{5} \tag{8}$$

Ein D von -1 bedeutet hierbei zum einen, dass die Versuchsperson die Referenz und das Signal mit Phasenversatz richtig zugeordnet hat. Zum anderen hat die Versuchsperson das Signal mit Phasenversatz als sehr ungleich im Vergleich zur Referenz bewertet. D = 0 bedeutet, dass die Versuchsperson keinen Unterschied zwischen Referenz und dem Signal mit Effekt feststellen konnte und angegeben hat, dass beide Signale identisch wahrgenommen wurden (beide Schieberegler in der obersten Stellung). D = +1 bedeutet im anderen Extremfall, dass die Versuchsperson das Signal mit Phasenversatz für das Referenzsignal gehalten und das Signal ohne Phasenversatz als das Signal mit Phasenversatz wahrgenommen hat, und zwar in voller Ausprägung. Referenzsignal und Signal mit Phasenversatz wurden in diesem Fall also komplett vertauscht, was darauf hindeuten würde, dass die Versuchsperson große Schwierigkeiten hatte, den Effekt überhaupt wahrzunehmen und bei der Antwort möglicherweise geraten hat.

4.3.1.1 Pink-Noise

Abbildung 21 stellt das Ergebnis der normierten Differenzgrade als Boxplots über die unterschiedlichen Effekt-Ausprägungen in Form von stufenweise erhöhten $\Delta\varphi_{max}$ mit den drei unterschiedlichen Bewegungsgeschwindigkeiten für das Pink-Noise-Signal dar. Auch hier sind Spannweiten (lange dünne vertikale Linien) und Mediane über alle Antworten (dicke kurze Linien durch die Boxen) dargestellt.

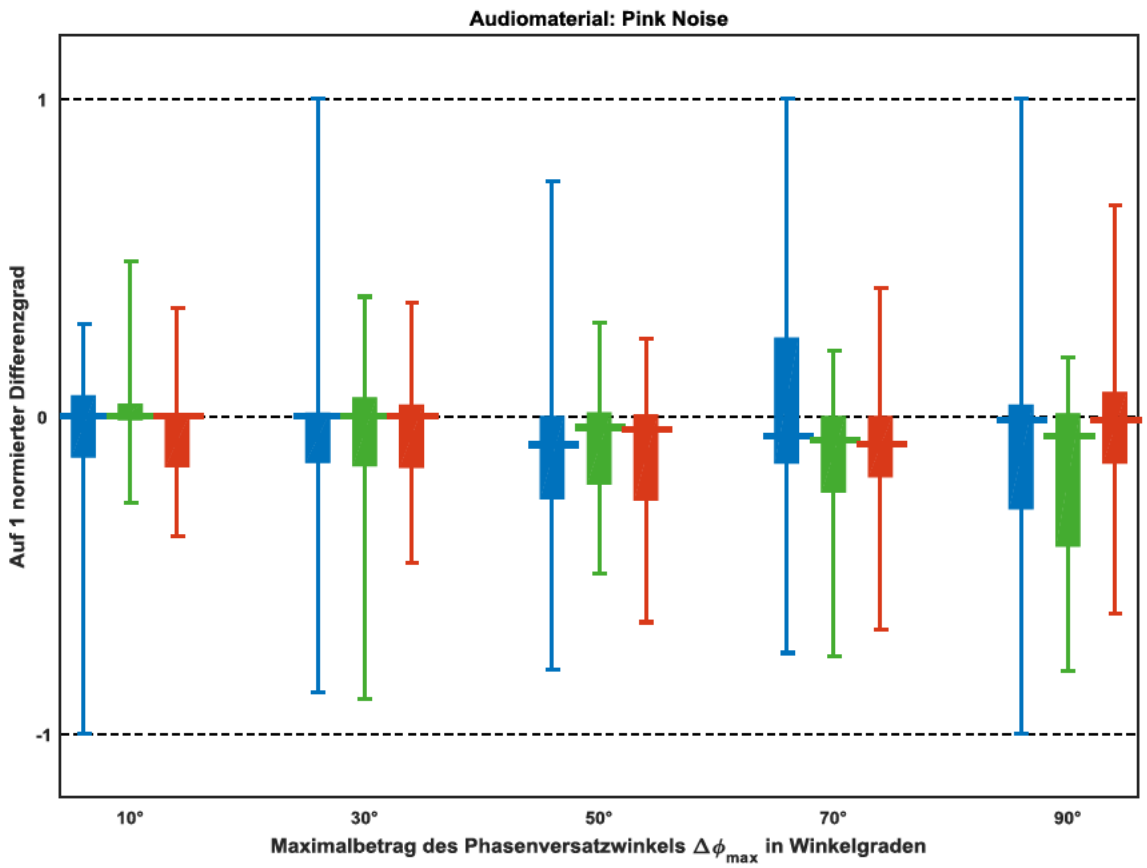

Abbildung 21: Deskriptives Ergebnis des ABCHR-Tests für das Pink-Noise-Signal

Blau steht hierbei für die langsame Bewegungsgeschwindigkeit von v = 3 km/h, grün für die mittlere Geschwindigkeit von v = 6 km/h und rot ist die schnellste Bewegungsgeschwindigkeit von v = 12 km/h. Wie deutlich zu erkennen ist, kommen die Median-Werte bei allen drei Bewegungsgeschwindigkeiten bei einem $\Delta\varphi_{max}$ von 10 und 30 Grad auf einen Differenzgrad von 0. Dies bedeutet, dass die meisten Versuchspersonen bei diesem Anregungssignal und den beiden geringsten Phasenversatzwinkeln keinen Unterschied zur Referenz feststellen konnten. Bei den höheren Stufen von $\Delta\varphi_{max}$ von 50 Grad bis 90 Grad sind hingegen die Median-Werte etwas unter der Nulllinie. Hier wurden also Unterschiede wahrgenommen, wenn auch nur in sehr geringer Ausprägung. Betrachtet man an dieser Stelle die unterschiedlichen Bewegungsgeschwindigkeiten, fällt auch hier keine besondere Tendenz der besseren Detektierfähigkeit zu höheren oder niedrigeren Bewegungsgeschwindigkeiten auf. Insgesamt ist an der Spanne der Antworten zu erkennen, dass es eine Tendenz zum richtigen Zuordnen von Referenzsignal und Signal mit Effekt gibt und extreme Ausreißer hiervon noch am ehesten bei der geringsten Bewegungsgeschwindigkeit vorhanden sind.

4.3.1.2 Multitone

Wie das Antwortverhalten der Versuchspersonen für das bassbetonte Multitone-Signal aussieht, zeigt Abbildung 22.

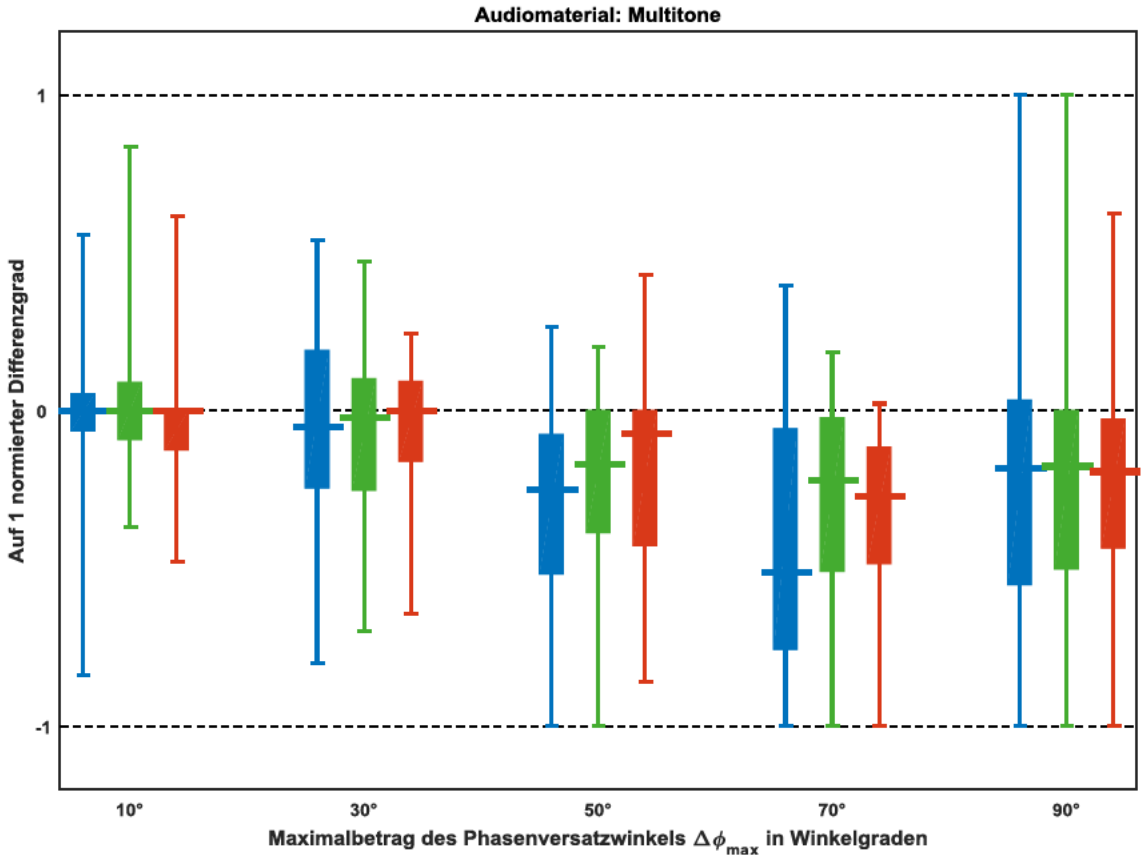

Abbildung 22: Deskriptives Ergebnis des ABCHR-Tests für das Multitone-Signal

Auch hier ist auffällig, dass die beiden geringsten Maximalbeträge des Phasenversatzwinkels von 10 und 30 Grad zu keiner nennenswerten Wahrnehmung des Effekts führen, bei 10 Grad sind die Median-Werte für alle drei Bewegungsgeschwindigkeiten direkt auf der Nulllinie. Ab einem $\Delta\varphi_{max}$ von 50 Grad ändert sich dies jedoch. Hier tendierten die Antworten eher zu negativen Differenzgraden und damit richtiger Zuordnung von Referenz- und Effektsignal. Am deutlichsten kann man eine

Abhängigkeit der Bewegungsgeschwindigkeit von der Wahrnehmbarkeit des Effekts bei einem $\Delta\varphi_{max}$ von 70 Grad und v = 3 km/h beobachten, hier tendiert der Median-Wert aller Antworten deutlich in Richtung -1. Für die anderen Fälle und einen $\Delta\varphi_{max}$ von 90 Grad spielt die Veränderung der Bewegungsgeschwindigkeit jedoch nur eine untergeordnete Rolle. Vergleicht man aber die gesamte Tendenz des Antwortverhaltens über die Spannweite aller Fälle, kann man bei dem Multitone-Signal eine deutliche Verschiebung in den negativen Bereich des Differenzgrades erkennen, was für eine gesteigerte Detektierbarkeit des Effekts steht, und zwar viel deutlicher als bei dem Pink-Noise-Signal. Dies liefert einen ersten Hinweis darauf, dass die Wahrnehmbarkeit des Effekts eher vom Audiomaterial als von der Bewegungsgeschwindigkeit abhängt.

4.3.1.3 Musik

Abbildung 23 stellt das Ergebnis des ABCHR-Tests für das Musik-Anregungssignal dar.

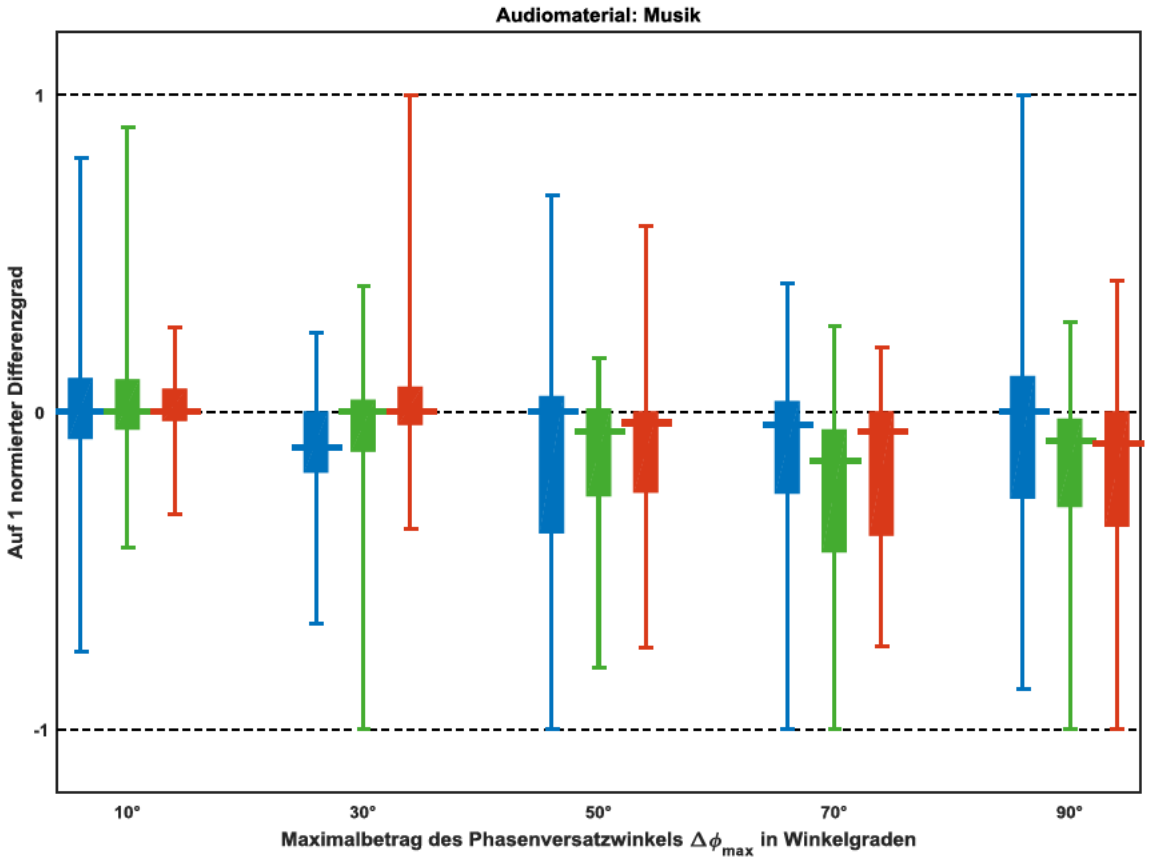

Abbildung 23: Deskriptives Ergebnis des ABCHR-Tests für das Musiksignal

Auch hier lässt sich erkennen, dass die kleinste Ausprägung von $\Delta\varphi_{max}$ (10 Grad) im Median nicht wahrgenommen wird und auch ein $\Delta\varphi_{max}$ von 30 Grad zu einer kaum nennenswerten Wahrnehmungssteigerung führt. Etwas stärker wird die Wahrnehmung bei höheren Phasenversatzwinkeln, jedoch ist auch hier die Neigung zum Wahrnehmen des Effekts nur gering und unwesentlich von der Bewegungsgeschwindigkeit abhängig. Im Vergleich zu dem übrigen Audiomaterial sind die Ausprägungen hier vergleichbar mit dem Pink-Noise-Signal, wenn auch ein unwesentlich mehr ausgeprägter bei höheren Phasenversatzwinkeln. Auch hier gibt es keine klare Neigung zu einem mit steigendem $\Delta\varphi_{max}$ stärker wahrnehmbaren Effekt, die Abweichungen der Mediane sind eher als gering zu betrachten.

4.3.1.4 Sprache

Das deskriptive Ergebnis für das ABCHR-Hörexperiment mit dem Sprachsignal veranschaulicht Abbildung 24.

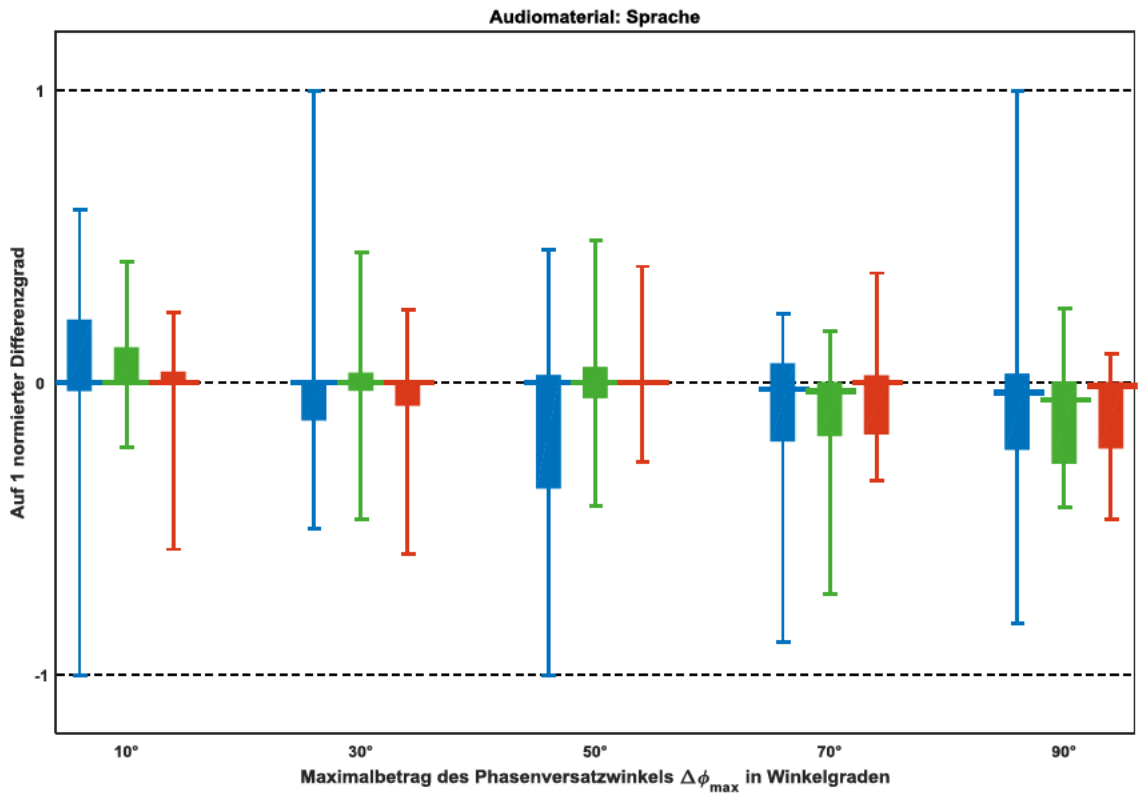

Abbildung 24: Deskriptives Ergebnis des ABCHR-Tests für das Sprachsignal

Hier fällt besonders die Neigung auf, keine Unterschiede gehört zu haben, die Werte schwanken alle mehr oder weniger um die Nulllinie. Auch bei großen Ausprägungen von $\Delta\varphi_{max}$ (70 und 90 Grad) ist die Abweichung der Mediane von der Nulllinie nur marginal feststellbar. Die Spanne der Antworten tendiert hierbei jedoch etwas zum negativen Bereich des normierten Differenzgrades hin, während sie bei den kleineren Ausprägungen von $\Delta\varphi_{max}$ eher in den positiven Bereich hineinreicht. Dies deutet darauf hin, dass es besonders für das Sprachsignal sehr schwer für die Versuchspersonen gewesen ist, den Effekt überhaupt herauszuhören.

Betrachtet man die Tendenzen aller deskriptiven Ergebnisse über alle drei Bewegungsgeschwindigkeiten, fällt für die langsame Bewegungsgeschwindigkeit (blaue Boxen) eine durchweg leicht größere Streuung/Spanne auf. Dies deutet daraufhin, dass es für die Versuchspersonen bei der langsamen Bewegungsgeschwindigkeit am uneindeutigsten war, welcher Stimulus den Effekt beinhaltet und wie groß die Ausprägung des Effektes ist. Für die beiden schnelleren Bewegungsgeschwindigkeiten ist die Spanne tendenziell ein bisschen kleiner, was auf eine gesteigerte Eindeutigkeit in der Wahrnehmung des Effekts der Phasenverschiebung hindeutet.

4.3.1.5 Kurvenform

Abbildung 25 stellt das deskriptive Ergebnis der normierten Differenzgrade für die Wahrnehmbarkeit des Effekts bei unterschiedlichen Kurvenformen von $\Delta\varphi$ über den Bewegungspfad dar.

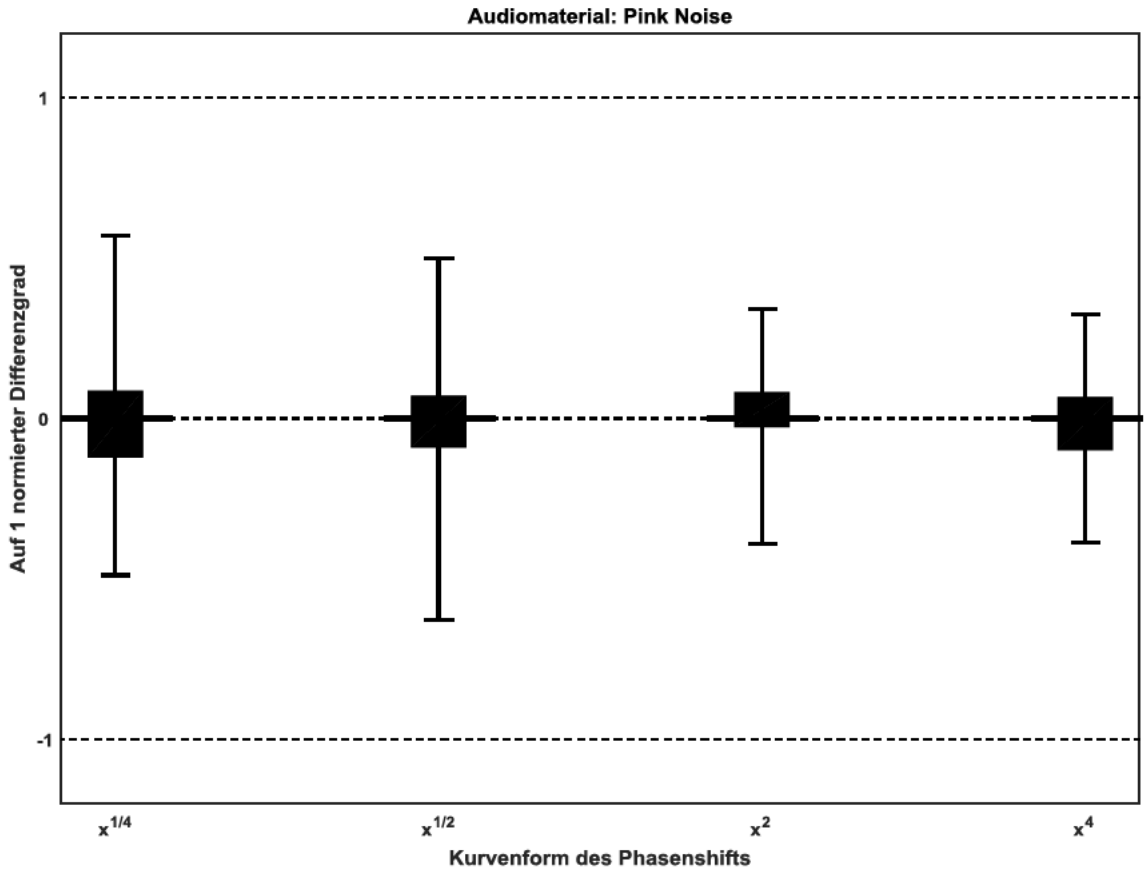

Abbildung 25: Deskriptives Ergebnis des ABCHR-Tests für Variationen der Kurvenformen

Hierzu gibt es vor allem eins festzustellen: Die wahrnehmbaren Unterschiede beschränken sich ausschließlich auf Streuungen um die Nulllinie herum, die Median-Werte liegen allesamt genau auf dem Nullwert. Somit konnten die Versuchspersonen keinen nennenswerten Unterschied bei veränderten Kurvenformen feststellen. Daher kann als Erkenntnis festgestellt werden, dass die Kurvenform der positionsabhängigen Phasenverschiebung die Wahrnehmbarkeit des Effekts nicht verändert. Im weiteren Verlauf der statistischen Analyse wird dieser Fall daher auch ausgeklammert.

4.3.2 Lilliefors-Test

Für das weitere Vorgehen der statistischen Analyse muss zunächst abgeschätzt werden, ob die Daten des ABCHR-Hörexperimentes normalverteilt sind. Dies geschieht in MATLAB® mit der Statistik-Toolbox und der *lillietest*()-Funktion, die ein Lilliefors-Test (Lilliefors, 1969) mit allen Kombinationen aus den fünfstufigen Ausprägungen von $\Delta\varphi_{max}$ und den vier unterschiedlichen Anregungssignalen mit jeweils drei verschiedenen Bewegungsgeschwindigkeiten durchführt. Das Testergebnis ist in Tabelle 3 aufgelistet.

Tabelle 3: p-Werte des Lilliefors-Tests

		$\Delta\varphi_{max}$				
Anregungssignal	v [km/h]	10°	30°	50°	70°	90°
Pink Noise	3	0,001	0,001	0,093	0,003	0,011
	6	0,001	0,077	0,072	0,022	0,003
	12	0,002	0,101	0,036	0,163	0,024
Multitone	3	0,003	0,362	0,304	0,277	0,500
	6	0,001	0,500	0,107	0,154	0,500
	12	0,001	0,089	0,035	0,500	0,080
Musik	3	0,046	0,093	0,007	0,024	0,054
	6	0,041	0,001	0,068	0,002	0,009
	12	0,011	0,001	0,010	0,002	0,252
Sprache	3	0,003	0,001	0,049	0,007	0,003
	6	0,004	0,001	0,012	0,001	0,014
	12	0,001	0,003	0,001	0,013	0,001

Als Signifikanzlevel wird hierbei 5 % angenommen. In der Tabelle sind alle Fälle in roter Schrift aufgeführt, bei denen dieses Kriterium verletzt wird (p > 5 %), die Nullhypothese (= Daten sind nicht normalverteilt) also angenommen werden muss. Wie zu erkennen ist, trifft dies für eine nicht zu vernachlässigende Menge aller Fälle zu, sodass nicht angenommen werden kann, dass die Daten normalverteilt sind.

4.3.3 Friedman-Test

Aus der Erkenntnis des vorigen Abschnitts, dass die Ratings des ABCHR-Hörexperiments über alle Versuchspersonen an einigen Stellen nicht normalverteilt sind, ergibt sich die Notwendigkeit, die weitere statistische Analyse mit einem nichtparametrischen Verfahren durchzuführen: Dem Friedman-Test (Corder und Foreman, 2009, S. 79). Dieser wird in MATLAB® mit der *friedman*()-Funktion aus der Statistik-Toolbox durchgeführt.

Mit dem Friedman-Test wird als erstes untersucht, ob ein statistisch signifikanter Effekt zwischen der Variation des Anregungssignals und der Variation der Bewegungsgeschwindigkeit zu beobachten ist. Der Friedman-Test wird über alle unterschiedlichen Anregungssignale insgesamt drei Mal durchgeführt, also für jede der drei Stufen der Bewegungsgeschwindigkeit. Für die drei Geschwindigkeitsstufen wird er entsprechend vier Mal, also über alle vier Anregungssignale, berechnet. Da der 3-AFC-Test eingangs gezeigt hat, dass der Effekt für die meisten Probanden erst ab einem $\Delta\varphi_{max}$ von nahezu 90 Grad hörbar wird, werden bei dem Friedman-Test alle Antworten jeweils für den Fall $\Delta\varphi_{max}$ = 90 Grad berücksichtigt, alle anderen werden ausgeklammert. Tabelle 4 listet das Ergebnis des Friedman-Tests für den Vergleich des Effekts der Anregungssignale auf, Tabelle 5 das entsprechende Ergebnis für den Effekt der unterschiedlichen Bewegungsgeschwindigkeiten.

Tabelle 4: Ergebnis des Friedman-Tests für die Variation der Anregungssignale

Bei v [km/h]	χ^2	p
3	6,649	0,091
6	7,55	0,056
12	13,88	0,003

Tabelle 5: Ergebnis des Friedman-Tests für die Variation der Bewegungsgeschwindigkeit

Bei Anregungssignal	χ^2	p
Pink Noise	1,08	0,581
Multitone	0,6	0,742
Musik	4,37	0,113
Sprache	0,35	0,841

Als Signifikanzniveau wird auch hier 5 % festgelegt. Wie aus Tabelle 4 zu erkennen ist, kann man bei der Variation der Anregungssignale nur für den Fall v = 12 km/h die Nullhypothese verwerfen, während für die anderen beiden Bewegungsgeschwindigkeiten das Signifikanzniveau nicht ganz erreicht wird. Daraus lässt sich schließen, dass die Probanden erst bei hohen Bewegungsgeschwindigkeiten den Effekt bei Variation der Anregungssignale sicher erkannt haben. Jedoch muss auch gesagt werden, dass die Signifikanzwerte der anderen beiden Fälle mit niedrigeren Bewegungsgeschwindigkeiten zahlenmäßig nicht allzu weit auseinander und in unmittelbarer Nähe des Signifikanzniveaus von 5 % liegen, sodass hieraus vermutet werden kann, dass die Detektion des Effekts auch bei niedrigeren Bewegungsgeschwindigkeiten möglich ist.

Tabelle 5 lässt erkennen, dass der Effekt der variierenden Bewegungsgeschwindigkeit bei keinem der Anregungssignale signifikant aufgefallen ist, da die p-Werte weit über dem auch hier gewählten Signifikanzniveau von 5 % liegen. Lediglich im Fall des Musiksignals kann eine Erhöhung der Wahrscheinlichkeit gegenüber den anderen Fällen beobachtet werden, sodass hier eine Tendenz festzustellen ist, dass der Effekt der variierten Bewegungsgeschwindigkeit noch am ehesten bei dem Musiksignal aufgefallen ist.

4.3.4 Wilcoxon-Signed-Rank-Test

Um herauszufinden, an welcher Stelle aller Paarkombinationen der Anregungssignale bei dem Fall v = 12 km/h die statistisch signifikanten Abweichungen auftreten, wird als Post-Hoc-Test der Wilcoxon-Signed-Rank-Test (Corder und Foreman, 2009, S. 55) mit ebendiesem Fall in MATLAB® mit der signrank()-Funktion aus der Statistik-Toolbox implementiert. Nach (9) errechnen sich mit n = 4 (Anregungssignale) und r = 2 (Paare) hierfür insgesamt C = 6 Paarkombinationen.

$$C_{n,r} = \frac{n!}{r!\,(n-r)!} \tag{9}$$

Der Test wird somit insgesamt sechs Mal für die Paarkombinationen aus Tabelle 6 durchgeführt. Die Ergebnisse werden in Form von p- und Z-Werten in Tabelle 6 dargestellt.

Tabelle 6: Ergebnis des Wilcoxon-Signed-Rank-Tests

Paarkombination	p	Z
Pink Noise <-> Multitone	0,015	+2,433
Pink Noise <-> Musik	0,175	+1,357
Pink Noise <-> Sprache	0,299	+1,039
Multitone <-> Musik	0,083	-1,733
Multitone <-> Sprache	0,012	-2,524
Musik <-> Sprache	0,056	-1,912

Das auch hier gewählte Signifikanzniveau von 5 % muss mittels Bonferroni-Korrektur nach (10) verändert werden, um einen Typ-1-Fehler zu vermeiden (Corder und Foreman, 2009, S. 59).

$$p_{\text{korr}} = \frac{p_{\text{sig}}}{C} \tag{10}$$

Damit verringert sich das Signifikanzniveau mit C = 6 Testdurchläufen auf 0,83 %. Dieses niedrige Signifikanzniveau erreicht kein p-Wert aus obiger Tabelle, weswegen nicht genau festgestellt werden kann, bei welchem der Paare der Effekt deutlich erkannt wird.

Trotzdem lassen sich Tendenzen feststellen. Paarkombinationen mit sowohl dem Musiksignal als auch ganz besonders dem Multitone-Signal zeigen im Vergleich zu den übrigen Paarkombinationen niedrige p-Werte. Daher kann angenommen werden, dass der Effekt besonders bei dem Musiksignal und mit noch etwas höherer Sicherheit bei dem Multitone-Signal von den Versuchspersonen erkannt wurde.

4.4 Fragebogen

In dem Fragebogen, den jede Versuchsperson nach erfolgreichem Absolvieren des Hörexperiments ausgefüllt hat, wird eine offene Frage gestellt, die die Versuchsperson dazu auffordert, in eigenen Worten zu beschreiben, an welchen Merkmalen sie etwaige Unterschiede wahrgenommen hat. Die Auswertung der Antworten erfolgt durch Einteilen der gesamten Menge an Antworten in Kategorien und anschließende Auszählung der Häufigkeiten der Antworten innerhalb einer Kategorie. Die Kategorien werden in Tabelle 7 aufgelistet und erklärt.

Tabelle 7: Kategorien der Antworten zur Wahrnehmung des Effektes

Kategorie	Beschreibung
Temporal	Zeitliche Effekte wie Verzögerung und Streckung
Spektral	Spektrale Effekte wie Veränderungen im Bass oder den Höhen (Über/Unterbetonung)
Amplitude	Veränderung der wahrgenommenen Lautstärke und des Absolut-Pegels
Tonhöhe	Veränderung der wahrgenommenen Tonhöhe, Verstimmung der Tonhöhe
Sonstige	Sonstige wahrgenommene Effekte wie Veränderung der Spielweise
Keine Antwort	Entweder keine Antwort gegeben oder angegeben, dass es nicht möglich ist, die empfundene Wahrnehmung in Worten zu beschreiben

Das Ergebnis der Häufigkeitsverteilung der Antworten aus der offenen Frage ist im Diagramm in Abbildung 26 grafisch veranschaulicht.

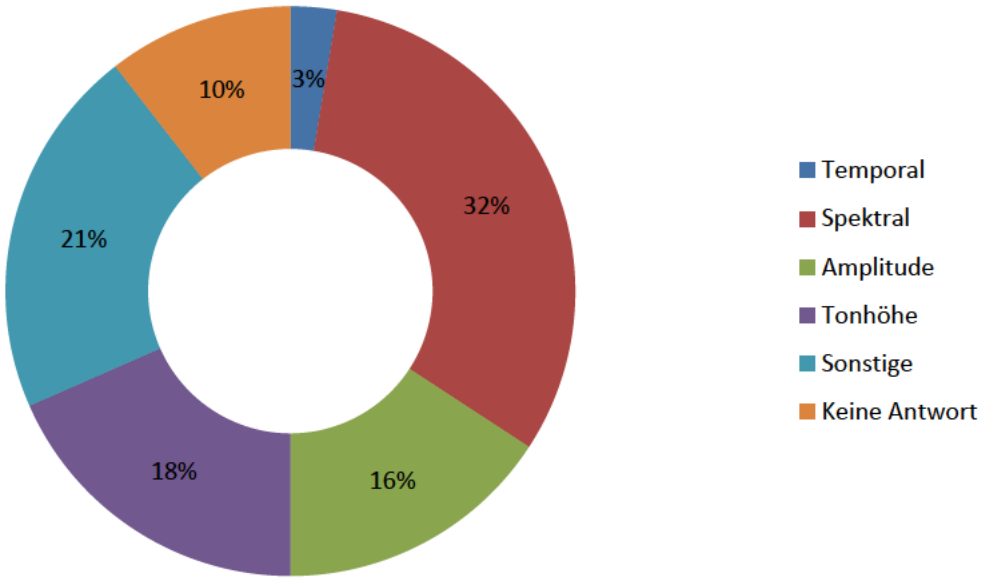

Abbildung 26: Auswertung der offenen Frage zur Wahrnehmung des Effekts

Wie hieraus zu erkennen ist, führt die Beeinflussung des Phasenversatzwinkels hauptsächlich zu Veränderungen der Wahrnehmung im spektralen Bereich. Häufig wurde hier der verändert wahrgenommene Bassbereich genannt, der meist als dünner und fundamentarmer beschrieben wurde. Auch wurde häufig eine Veränderung der empfundenen Tonhöhe angegeben sowie sonstiger Effekte, wie beispielsweise der wahrgenommenen veränderten Spielweise bei dem Musiksignal. Dies kann ebenso auf spektralen Variationen beruhen, wenn der Grundtonbereich verändert wahrgenommen wird. Einige Probanden empfanden auch eine leichte Variation der gesamten Lautstärke bei Darbietung der Auralisation mit Phasenversatz sowie den temporalen Strukturen. Generell war es aber auch für einzelne Versuchspersonen schwer, überhaupt Angaben zu ihren individuellen Wahrnehmungen zu machen, da sie den Effekt als nahezu unhörbar beschrieben haben, weswegen der Anteil an keinen Antworten nicht vernachlässigbar gering ausfällt.

5 Diskussion

5.1 Zusammenfassung

Wie schon die Ergebnisse des informellen 3-AFC-Hörversuchs vermuten ließen, ist die Wahrnehmbarkeit des untersuchten Effekts der ortsabhängigen Phasenmodulation als gering einzustufen. Aus dem ersten Teil des Haupt-Hörexperiments, dem nach den Erkenntnissen des informellen Hörversuchs parametrisierten 3-AFC-Schwellwertdetektionsverfahren, geht schließlich hervor, dass die meisten Probanden bereits bei der maximalen Ausprägung des Phasenmodulation Schwierigkeiten hatten, den Effekt überhaupt wahrzunehmen. Der mit diesem Verfahren gefundene Schwellwert liegt mit einem linearen Verlauf der ortsabhängigen Phasenmodulation bei einem $\Delta\varphi_{max}$ von 85,75 Grad, erreicht also nahezu den als maximal definierten Wert von 90 Grad.

Dies bestätigen auch die Ergebnisse aus dem ABCHR-Hörvergleich, dem zweiten Teil des Hörexperiments. Zunächst ist hieraus die Erkenntnis festzuhalten, dass es unerheblich ist, ob der ortsabhängige Phasenversatz nach einer linearen Funktion vorgenommen oder die Kurvenform dabei quadratisch oder entsprechend mit der Umkehrfunktion moduliert wird. Auch hat sich herausgestellt, dass die Variation der Lauf-Bewegungsgeschwindigkeit in von Menschen üblichen Geh- und Laufgeschwindigkeiten zu keiner signifikanten Beeinflussung der Wahrnehmung der Phasenmodulation führt.

Viel eher ist die Ausprägung der Wahrnehmbarkeit der ortsabhängigen Phasenmodulation jedoch abhängig vom Audiomaterial, welches in den auralisierten Szenerien dargeboten wird. Hier kann eine statistisch signifikante Detektion des Effekts bei der höchsten Bewegungsgeschwindigkeit von v = 12 km/h festgestellt werden. Wie weitere Post-Hoc-Tests gezeigt haben, kann anhand der Daten aus diesem Hörexperiment nicht sicher vorausgesagt werden, bei welchem Audiomaterial die Detektion am stärksten ist. Jedoch lassen sich Tendenzen feststellen, die besonders das Multitone-Signal wie auch das Musiksignal als Audiomaterial klassifizieren, bei welchem der Effekt der Phasenmodulation besonders gut hörbar wird.

Generell ist es interessant, sowohl an der Grand-Median-Schwellwertkurve des 3-AFC-Tests (Abbildung 19) als auch an den Boxplots des ABCHR-Tests zu erkennen, dass die Wahrnehmbarkeit bei großen $\Delta\varphi_{max}$ (ab zirka 70 Grad) in sich selbst nicht groß abweicht. Dies deutet zum einen darauf hin, dass der Effekt nur marginal wahrgenommen wird und es den Versuchspersonen schwer gefallen ist, größere $\Delta\varphi_{max}$ absolut einzuordnen. Zum anderen macht dies den Anschein, dass der Effekt zwar wahrgenommen wird („da ist etwas anders“), aber das Ausmaß der Manipulation schwer einzuschätzen ist. Zum anderen setzt möglicherweise auch ein Gewöhnungseffekt ein, der es den Versuchspersonen schwerer macht, den Effekt in seinem Ausmaße absolut einzuordnen.

Wie in Abschnitt 1.2 bereits diskutiert, geht sowohl aus den Erkenntnissen von Helmholtz, 1865 als auch aus den Forschungen von Laitinen et al., 2013 hervor, dass Änderungen in der Phasenwinkellage bei statischen Signalen vom menschlichen Gehör immer als Änderungen der Amplitude einzelner Frequenzen, also als eine Veränderung der spektralen Komponente des Hörereignisses, wahrgenommen werden. Dies konnte für den hier untersuchten Fall der bewegten Hörerposition bestätigt werden, da die Versuchspersonen bei der Beantwortung der offenen Frage nach dem Merkmal, an dem sie etwaige Unterschiede erkannt haben, wiederholt von einem veränderten Spektrum gesprochen haben. Häufig wurde dabei der abgeschwächt wahrgenommene

Bassbereich genannt, weswegen die Behauptungen von Laitinen et al., 2013, dass die Veränderung des Phasenspektrums zu einem dünner wahrgenommenen Bassfundament führt, auch für den Fall dynamischer Signale und Bewegung des Hörers und/oder der Schallquelle bestätigt werden kann.

Im konkreten Anwendungsfall, beispielsweise der Demonstration eines neuen Beschallungslautsprecher-Systems, kann aus den gewonnenen Erkenntnissen dieser Arbeit abgeleitet werden, dass positionsabhängige Abweichungen der Phasenwinkellage, wenn überhaupt, nur zu marginalen Beeinflussungen in der Wahrnehmung der Schallereignisse führen. Die dabei resultierenden Wahrnehmungen spielen sich hauptsächlich im Bassbereich ab. Die Hörsamkeit hängt damit im Wesentlichen von dem dargebotenen Audiosignal ab. Beinhaltet das Audiosignal wenig bis gar keinen Bassanteil, wie beispielsweise bei einem reinen Sprachsignal, kann die Phasenwinkellage bei der Optimierung von Schallfeldern nahezu vernachlässigt werden. Anders sieht die Situation aus bei langen, spektral wenig veränderlichen tiefen Tönen, wie sie beispielsweise in elektronischer Musik vorkommen können. Hier können etwaige ortsabhängige Phasenfehler in nicht vernachlässigbarem Maße wahrgenommen werden. Resultierend würde in diesem Fall der Bassbereich ortsabhängig als mehr oder weniger „voll“ empfunden.

5.2 Prädiktionsindikatoren

In diesem Abschnitt werden Prädiktionsindikatoren aus der Audio-Content-Analyse vorgeschlagen, die sich möglicherweise dazu eignen können, die Wahrnehmbarkeit des in den vorigen Abschnitten untersuchten Effekts vorauszusagen. Hierfür werden die in Abschnitt 2.1.5 vorgestellten Anregungssignale in MATLAB® mit Hilfe diverser Funktionen aus der MIR-Toolbox (Lartillot et al., 2008) analysiert.

5.2.1 Spektraler Flux

Wie im vorigen Abschnitt zusammenfassend erwähnt, ist die Wahrnehmbarkeit der ortsabhängigen Phasenmodulation vor allem bei spektral eher statischen Signalen auffällig. Die spektrale Varianz kann mit dem Spectral-Flux-Feature ermittelt werden. In Abbildung 27 ist der mit einem Moving-Average-Filter über 100 Audioframes mit der Länge von 512 Audiosamples geglättete spektrale Flux für alle vier Anregungssignale aus diesem Hörversuch dargestellt.

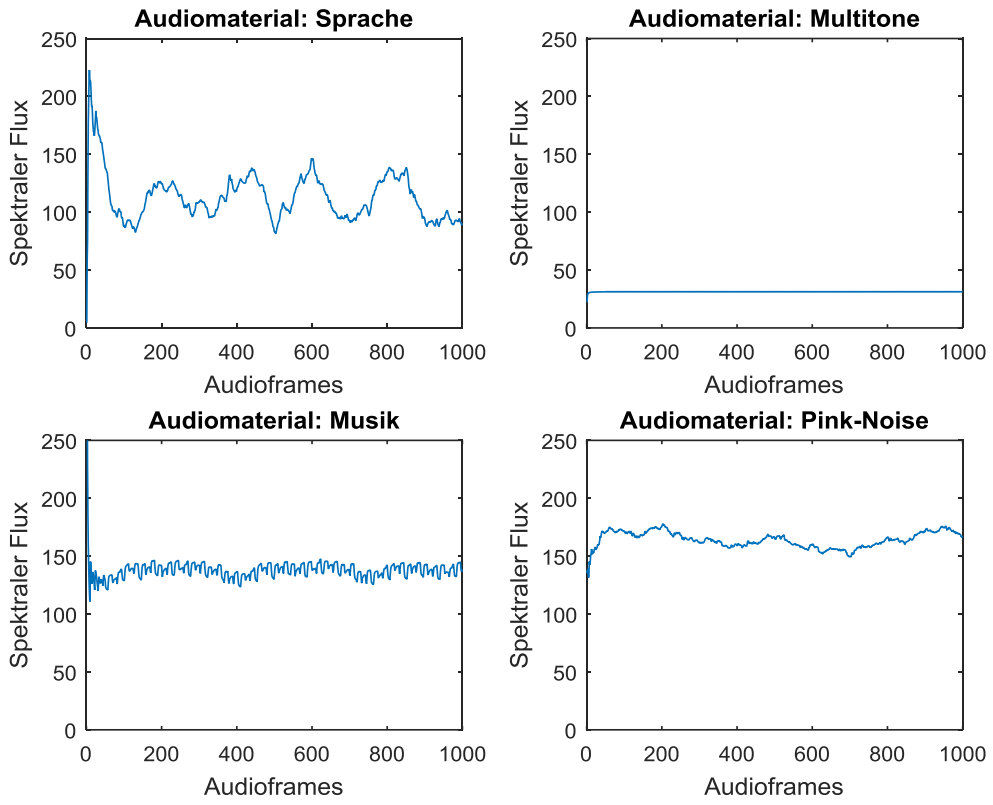

Abbildung 27: Spectral Flux aller vier Anregungssignale

Auffällig ist hierbei, dass der spektrale Flux bei dem Multitone-Signal konstant und niedrig über die gesamte Laufzeit des Signals ist. Auch der spektrale Flux des Musiksignals, bei dem der Effekt der ortsabhängigen Phasenmodulation ebenso erhöht wahrgenommen wurde, weist einen glatten Verlauf auf. In diesem Zusammenhang wird also die Aussage getroffen, dass je glatter und niedriger der spektrale Flux eines Audiosignals ausfällt, desto höher ist die Wahrscheinlichkeit, dass anschließende Phasenwinkelverschiebungen hörbar werden können.

5.2.2 Autokorrelation

Im Zeitbereich ist der Effekt möglicherweise mit Hilfe der Autokorrelationsfunktion voraussehbar. Diese Funktion stellt die Ähnlichkeit des Signals mit sich selbst zu einem früheren Zeitpunkt dar. Hiermit weisen reine Töne, also stark sich periodisch wiederholende Signale, eine hohe Autokorrelation auf, was für eine geringe Varianz des Frequenzinhaltes des Signals spricht. Abbildung 28 veranschaulicht die auf 1 normierten Autokorrelationskoeffizienten aller Anregungssignale des Hörversuchs.

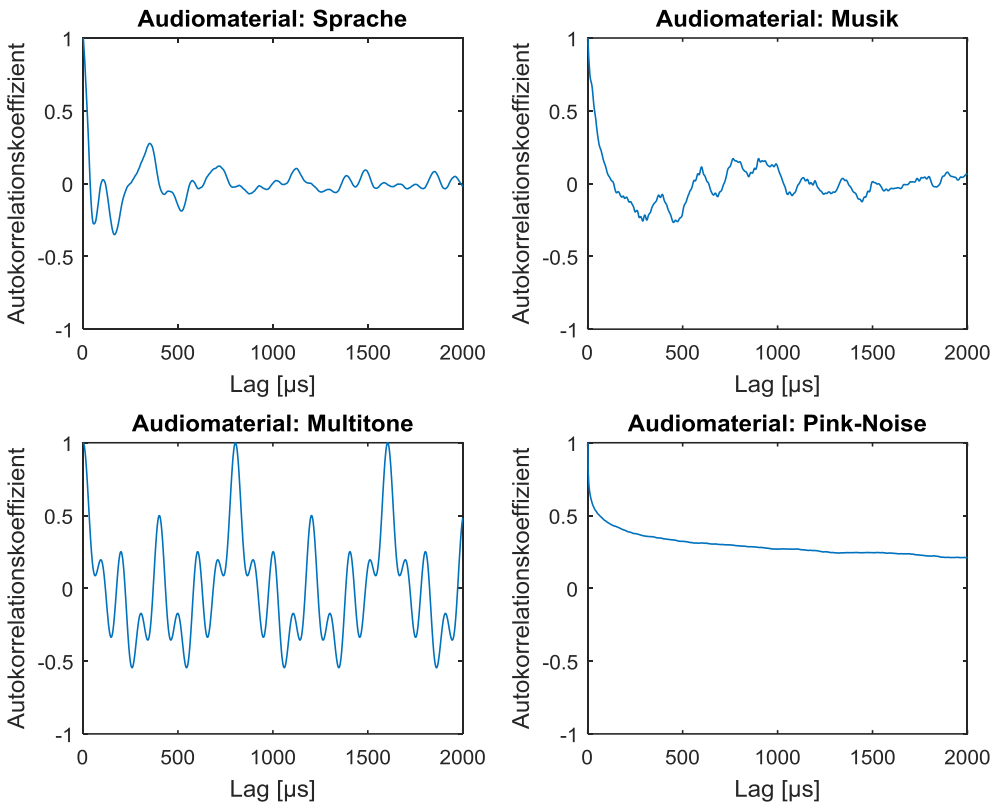

Abbildung 28: Autokorrelation aller vier Anregungssignale

Es wird deutlich, dass die Sprach- und Musiksignale mit sich selbst eher unähnlich sind, wobei das Sprachsignal aufgrund des Vorhandenseins nur einer einzigen klangerzeugenden Quelle, die periodische Impulse erzeugt (Stimmlippen), etwas periodischer korreliert. Das Pink-Noise-Signal korreliert nur zur Beginn sehr stark mit sich selbst, fällt dann aber aufgrund des rein stochastisch auftretenden Frequenzgehalts steilflankig ab. Das mit Abstand von allen Darstellungen am meisten abweichende Signal ist das Multitone-Signal, was aus mehreren reinen Sinustönen besteht und demnach stellenweise periodisch in sich sehr ähnlich, mit einem Phasenversatz jedoch periodisch stellenweise sehr verschieden ist. Erkennbar ist dies an der periodisch immer wieder auf den Maximalwert zurückkehrenden Autokorrelationskoeffizient, was für eine stellenweise maximale Ähnlichkeit mit sich selbst steht. Ein solcher Verlauf der Autokorrelation kann also im Hinblick auf die Hörsamkeit der Phasenwinkelverschiebungen auf eine hohe Wahrnehmbarkeit hindeuten.

5.3 Fehlerbetrachtung

In den auralisierten Hörszenarien wird jegliche Raumakustik ausgeklammert, um den Effekt isoliert von diesen Einflüssen zu betrachten. Dies stellt eine in der Praxis nur sehr selten anzutreffende Situation dar, da ausgereifte Beschallungssysteme häufig in geschlossenen Räumen betrieben und auf diese Situation hin optimiert werden.

Die in dieser Arbeit angewandten Phasenwinkelmanipulationen beeinflussen das gesamte Phasenspektrum und sind nicht frequenzselektiv. Häufig ist es jedoch in der Praxis so, dass der Phasenfrequenzgang frequenzselektive Veränderungen erfährt. Auch nimmt er in diesen Fällen in der Regel stellenweise beliebige Verläufe an. Dieses Verhalten hätte man mit randomisierten

Phasenwinkelbeeinflussungen simulieren können, wie es auch in Laitinen et al., 2013 vorgeschlagen und durchgeführt wird.

In Bezug auf das Hörversuchsverfahren und die Rücksprache mit den teilnehmenden Probanden kann festgestellt werden, dass die gesamte Prozedur für die meisten Probanden zu lange und zu ermüdend gewesen ist. Hier hätte eine Reduktion der zu untersuchenden Fälle auf weniger Variationen den Hörversuchsprozess beschleunigt. Auch ist gerade die Länge des ABCHR-Hörversuchsverfahrens von der Entscheidungsfreudigkeit der Versuchsperson abhängig, während das AFC-Paradigma mit seinem stringenten Verlauf den Proband zu schnelleren Entscheidungen zwingt. Möglicherweise hätte die Reduktion der Menge an Stimuli-Variationen auf einige wenige und die Durchführung nach dem 3-AFC-Verfahren zu einem angenehmeren Versuchsablauf für die Probanden geführt. Zudem hat der über alles betrachtet eher gering wahrnehmbare Effekt zu einer höheren Frustration bei einigen Probanden geführt, die durch eine kürzere Versuchszeit möglicherweise hätte vermieden werden können.

5.4 Ausblick

Im Hinblick auf die im vorigen Abschnitt erwähnten Aspekte wäre ein weiter zu untersuchender Fall die frequenzselektive und ortsabhängige Beeinflussung des Phasengangs und die damit zusammenhängende veränderte Wahrnehmung. Möglicherweise sind hierbei Phasenveränderungen in bestimmten schmalbandigen Frequenzbereichen noch weniger wahrnehmbar als die breitbandige Variation des Phasenverschiebungswinkels, wie sie in dieser Arbeit vorgenommen wurde.

Ein weiterer, auch nicht zu vernachlässigender genereller Aspekt bei der Darbietung von Audiosignalen über Kopfhörer ist die fehlende Körperschallübertragung. Gerade im Hinblick auf die veränderte Basswahrnehmung wäre es möglicherweise nützlich, den Effekt der veränderten Basswahrnehmung bei Beeinflussung des Phasengangs im Vergleich zur veränderten Basswahrnehmung bei Anregung des Probanden mit Körperschall gegenüberzustellen. In einer realen Konzert-Situation, bei der in der Regel ein hohes Maß an Körperschall durch hohe Schalldruckpegel im Bassbereich übertragen wird, ist die Beeinflussung der Wahrnehmbarkeit des Bassfundamentes möglicherweise viel eher von der Wahrnehmung tieffrequenten Schalldrucks über den eigenen Körper, als über einen nicht optimierten Phasenfrequenzgang abhängig.

Hierzu könnte in einem weiteren Schritt der Effekt der ortsveränderlichen Phasenmodulation bei Beschallungssystemen in Räumen im Hinblick auf die Interaktion mit unterschiedlich optimierten Raumakustiken untersucht werden. Gerade in schwierig zu beschallenden Räumen mit großen Nachhallzeiten wie Kirchen und Stadien ist der eher subtil wahrnehmbare Effekt der Phasenwinkelmanipulation möglicherweise bedeutungslos gegenüber den Kolorationen und anderen Effekten von großen Nachhallzeiten.

Im Hinblick auf numerische Optimierungsverfahren der von Beschallungslautsprechern erzeugten Schallfelder kann aus den in dieser Arbeit gewonnenen Erkenntnissen künftig der Aspekt mit einbezogen werden, welches Audiomaterial voraussichtlich über das Beschallungssystem dargeboten wird. Sind es spektral über den gesamten Audiofrequenzbereich ausgedehnte Signale mit sehr viel Bassanteil, sollte eine ortsabhängige Phasenlinearität sichergestellt werden. Handelt es sich jedoch um spektral gering ausgedehnte Signale mit wenig bis gar keinem Bassanteil, kann die Optimierung des Zielschallfeldes auf einen homogenen Betragsfrequenzgang hin beschränkt werden, während die Optimierung des Phasenfrequenzgangs hierbei keine fundamentale Beachtung erfahren muss.

6 Literatur

Bregman, Albert S. (1994): *Auditory Scene Analysis: The Perceptual Organization of Sound*. Cambridge, Massachusetts: MIT Press.

Brinkmann, Fabian (2017): „The FABIAN head-related transfer function data base.“.

Brinkmann, Fabian und Alexander Lindau (2010): „On the effect of individual headphone compensation in binaural synthesis.“ Vortrag auf der DAGA, Berlin.

Brinkmann, Fabian und Stefan Weinzierl (2016): „AKTools.“.

Ciba, Simon; Fabian Brinkmann und Alexander Lindau (2014): „WhisPER. A MATLAB toolbox for performing quantitative and qualitative listening tests.“.

Corder, Gregory W. und Dale I. Foreman (2009): *Nonparametric Statistics for Non-Statisticians: A Step-by-Step Approach*. John Wiley & Sons.

Griesinger, David (2010): „Phase coherence as a measure of acoustic quality, part one: the neural mechanism.“ Vortrag auf der Proceedings of the 20th International Congress on Acoustics (ICA), Sydney.

Helmholtz, Hermann von (1865): *Die Lehre von den Tonempfindungen als physiologische Grundlage für die Theorie der Musik*. Braunschweig: Vieweg.

ITU (2015): *Recommendation ITU-R BS.1770-4: Algorithms to measure audio programme loudness and true-peak audio level*.

King-Smith, P. E.; S. S. Grigsby; A. J. Vingrys; et al. (1994): „Efficient and unbiased modifications of the QUEST threshold method: theory, simulations, experimental evaluation and practical implementation.“ In: *Vision Research*, 34(7), S. 885–912.

Laitinen, Mikko-Ville; Sascha Disch und Ville Pulkki (2013): „Sensitivity of Human Hearing to Changes in Phase Spectrum.“ In: *J. Audio Eng. Soc*, 61(11), S. 860–877.

Lartillot, Olivier; Petri Toiviainen und Tuomas Eerola (2008): „A Matlab Toolbox for Music Information Retrieval.“ In: Christine Preisach; Professor Dr Hans Burkhardt; Professor Dr Lars Schmidt-Thieme; et al. (Hrsg.) *Data Analysis, Machine Learning and Applications*, Springer Berlin Heidelberg, (Studies in Classification, Data Analysis, and Knowledge Organization), S. 261–268.

Lilliefors, Hubert W. (1969): „On the Kolmogorov-Smirnov Test for the Exponential Distribution with Mean Unknown.“ In: *Journal of the American Statistical Association*, 64(325), S. 387–389.

Lindau, Alexander und Fabian Brinkmann (2012): „Perceptual Evaluation of Headphone Compensation in Binaural Synthesis Based on Non-Individual Recordings.“ In: *Journal of the Audio Engineering Society*, 60(1/2), S. 54–62.

Møller, Henrik (1992): „Fundamentals of binaural technology.“ In: *Applied Acoustics*, 36(3–4), S. 171–218.

Möser, Michael (2012): *Technische Akustik*. Berlin: Springer.

Otto, Stefanie und Stefan Weinzierl (2009): „Comparative simulations of adaptive psychometric procedures.“ In: *Jahrestagung der Deutschen Gesellschaft für Akustik*, S. 1276–1279.

Rafaely, Boaz (2015): *Fundamentals of Spherical Array Processing*. Berlin, Heidelberg: Springer.

Schärer, Zora und Alexander Lindau (2009): „Evaluation of Equalization Methods for Binaural Signals.“ Vortrag auf der Audio Engineering Society Convention 126 Audio Engineering Society.

Scheirman, David (2015): „Large-Scale Loudspeaker Arrays: Past, Present and Future (Part Two—Electroacoustic Considerations).“ *AES Proceedings*, Vortrag auf der 59. AES Conference on Sound Reinforcement Engineering and Technology, Montreal, Canada.

Straube, Florian; Frank Schultz; Michael Makarski; et al. (2015): „Evaluation Strategies for the Optimization of Line Source Arrays.“ *AES Proceedings*, Vortrag auf der 59. AES Conference on Sound Reinforcement Engineering and Technology, Montreal, Canada.

Tachi, Susumu und Takashi Isobe (1973): „Effect of Phase Relations on the Timbre of Harmonic Tones.“ In: *Japanese journal of medical electronics and biological engineering*, 11(2), S. 108–116.

7 Abbildungsverzeichnis

Abbildung 1: Skizze der Umgebung und des Bewegungspfades der Auralisation 4
Abbildung 2: Konturen der Phasenversatzwinkel in Abhängigkeit des Hörer-Bewegungspfades 5
Abbildung 3: Spektrogramm des Musik-Signals 6
Abbildung 4: Betragsfrequenzgang des Bass-Boost-PEQs für das Musiksignal 6
Abbildung 5: Spektrogramm des Sprach-Signals 7
Abbildung 6: Spektrogramm des Multitone-Signals 8
Abbildung 7: Spektrogramm des Pink-Noise-Signals 9
Abbildung 8: Abschwächung des auralisierten Schallfeldes vs. Bewegungspfad des Hörers 11
Abbildung 9: Sennheiser HD800 auf FABIAN-Kunstkopf (Brinkmann, 2017) 12
Abbildung 10: Ermittlungsvorgang des Kompensationsfilters für den HD800 (Brinkmann, 2017) 13
Abbildung 11: Schematischer Aufbau der Hard- und Software für die Hörversuche 15
Abbildung 12: GUI des informellen 3AFC-Hörversuchs 16
Abbildung 13: Psychometrische Funktion des informellen Hörversuchs 17
Abbildung 14: A-priori-Funktion des informellen Hörversuchs 18
Abbildung 15: Ergebnis des informellen Hörversuchs (grün v=5 km/h, rot v=12 km/h) 19
Abbildung 16: Hörversuchsumgebung im MiniLAB 20
Abbildung 17: A-priori-Funktion des ZEST-3-AFC-Tests 22
Abbildung 18: Das ABCHR-Test-GUI von WhisPER 22
Abbildung 19: Grand-Median-Schwellwert-Verlauf des 3-AFC-Tests 24
Abbildung 20: Boxplot der Schwellwerte des 3-AFC-Tests 25
Abbildung 21: Deskriptives Ergebnis des ABCHR-Tests für das Pink-Noise-Signal 26
Abbildung 22: Deskriptives Ergebnis des ABCHR-Tests für das Multitone-Signal 27
Abbildung 23: Deskriptives Ergebnis des ABCHR-Tests für das Musiksignal 28
Abbildung 24: Deskriptives Ergebnis des ABCHR-Tests für das Sprachsignal 29
Abbildung 25: Deskriptives Ergebnis des ABCHR-Tests für Variationen der Kurvenformen 30
Abbildung 26: Auswertung der offenen Frage zur Wahrnehmung des Effekts 34
Abbildung 27: Spectral Flux aller vier Anregungssignale 37
Abbildung 28: Autokorrelation aller vier Anregungssignale 38

8 Tabellenverzeichnis

Tabelle 1: Ermittelte Loudness-Werte aller Anregungssignale 14
Tabelle 2: Übersicht aller Stimuli des ABCHR-Hörvergleichs 23
Tabelle 3: p-Werte des Lilliefors-Tests 31
Tabelle 4: Ergebnis des Friedman-Tests für die Variation der Anregungssignale 31
Tabelle 5: Ergebnis des Friedman-Tests für die Variation der Bewegungsgeschwindigkeit 32
Tabelle 6: Ergebnis des Wilcoxon-Signed-Rank-Tests 32
Tabelle 7: Kategorien der Antworten zur Wahrnehmung des Effektes 33

9 Anhang

Die Anhänge dieser Arbeit sind in elektronischer Form auf der beiliegenden CD-ROM zu finden. Konkret handelt es sich um folgende Inhalte:

- Dieses Dokument im PDF-Format
- Matlab-Funktionen und Skripte im M-Format
- Fragebogen der Hörversuche im PDF-Format
- Datensätze der Hörversuche im XLSX-Format
- Sämtliche verwendete Audiodateien im WAV-Format
- Referenzierte Literatur im PDF-Format
- Zotero-RDF-Export des Literaturverzeichnisses dieser Arbeit.