

Diplomarbeit

Perzeptiver Vergleich von Schallfeldreproduktionsverfahren unterschiedlicher räumlicher Bandbreite

Matthias Frank

eingereicht an der Technischen Universität Graz im Juni 2009

Betreuer:

DI Jens Ahrens (T-Labs)
Dr.-Ing. Sascha Spors (T-Labs)
DI Franz Zotter (IEM)
Univ.-Ass. DI Dr. Alois Sontacchi (IEM)

Beurteiler:

O.Univ.-Prof. Mag. DI Dr. Robert Höldrich

Abstract

Perzeptiver Vergleich von Schallfeldreproduktionsverfahren unterschiedlicher räumlicher Bandbreite

Nearfield Compensated Higher-Order Ambisonics (HOA) und Wellenfeldsynthese (WFS) stellen die beiden bekanntesten Vertreter räumlich hochauflösender Schallfeldreproduktionsverfahren dar. Bei beiden Verfahren besteht die Wiedergabeordnung in der Praxis aus diskreten Lautsprechern, die im betrachteten zweidimensionalen Fall regelmäßig auf einem Kreis verteilt sind. Da die Verfahren allerdings zunächst auf der Formulierung einer kontinuierlichen Anordnung von Wiedergabequellen und deren Ansteuerungsfunktion beruhen, muss diese Ansteuerungsfunktion diskretisiert werden. Darin zeigt sich der wesentliche Unterschied zwischen den beiden untersuchten Verfahren: HOA begrenzt die winkelbezogene räumliche Bandbreite der Ansteuerungsfunktion vor der Diskretisierung in Abhängigkeit von der verwendeten Lautsprecheranzahl, während WFS keine explizite Bandbegrenzung verwendet. Dadurch kommt es bei den beiden Verfahren zu einer unterschiedlichen räumlichen Verteilung und Ausprägung von Artefakten.

Ziel dieser Arbeit ist der Vergleich von WFS und HOA durch Hörversuche. Im Vorfeld vermessene binaurale Impulsantworten einer 56-kanaligen Lautsprecheranordnung für unterschiedliche Kopfpositionen und -orientierungen wurden als Grundlage für die Hörversuche herangezogen. Mithilfe dieser Daten und eines dynamischen Wiedergabesystems kann das Lautsprechersystem präzise über Kopfhörer simuliert und so nahtlos zwischen verschiedenen Wiedergabeverfahren und Hörpositionen umgeschaltet werden.

Die Diplomarbeit findet in Zusammenarbeit mit dem Quality & Usability Lab der Deutschen Telekom Laboratories in Berlin statt, wo der praktische Teil der Arbeit durchgeführt wurde.

Perceptual comparison of sound field reproduction methods with different spatial bandwidth

Nearfield Compensated Higher-Order Ambisonics (HOA) and Wave Field Synthesis (WFS) are the most prominent sound field reproduction methods with high spatial resolution. In practice, both methods use a circular (for the considered two dimensional case) array with equally spaced discrete loudspeakers. As these methods are based on a formulation with continuous reproduction sources and their driving function, this driving function has to be discretized. The basic difference between HOA and WFS is due to the handling with this discretization: Whereas HOA limits the circular spatial bandwidth of the driving function before the discretization according to the number of used loudspeakers, WFS uses no explicit limitation. Thus it comes to a different spatial distribution and shape of artefacts.

The goal of this work is the comparison of WFS and HOA using listening tests. Previously measured binaural impulse responses of a circular 56-channel loudspeaker array for different head positions and orientations have been used as the basis of the listening tests. Given these data and a dynamic reproduction system, the loudspeaker system can be simulated accurately via headphones and thereby it can be seamlessly switched between different reproduction methods and listening positions.

This diploma thesis proceeds in cooperation with the Quality & Usability Lab of Deutschen Telekom Laboratories in Berlin, where the practical part of this work has been carried out.

Inhaltsverzeichnis

1	Einleitung	1
2	Theorie zu den Schallfeldreproduktionsverfahren	2
2.1	Gauß'scher Integralsatz und Green'sche Identität	2
2.2	Kirchhoff-Helmholtz-Integral	3
2.3	Wellenfeldsynthese (WFS)	4
2.3.1	Rekonstruktion von ebenen Wellen mit WFS	5
2.3.2	Rekonstruktion von Punktquellen mit WFS	6
2.3.3	2.5D Korrektur	7
2.4	Higher Order Ambisonics (HOA)	8
2.4.1	2.5D Korrektur	8
2.4.2	Rekonstruktion von ebenen Wellen mit HOA	9
2.4.3	Rekonstruktion von Punktquellen mit HOA	9
2.5	Diskretisierung	10
2.6	Analytische Simulation der Schallfeldreproduktionsverfahren	13
2.6.1	Reproduzierte Schallfelder HOA	13
2.6.2	Reproduzierte Schallfelder WFS	14
2.6.3	Zeitliche Spektren HOA	15
2.6.4	Zeitliche Spektren WFS	17
2.6.5	Lokale Variation der zeitlichen Spektren	18
3	Statistische Auswertungsverfahren für den Hörversuch	19
3.1	MATLAB: ANOVA1	19
3.1.1	Varianzanalyse	19
3.1.2	Boxplot	20
3.1.3	Beispiel	21
3.1.4	Signifikanzmatrix	23
3.2	Principle Component Analysis	24
3.2.1	Einführung in die PCA	24
3.2.2	Beispiel zur PCA	24
4	Entwurf und Durchführung des Hörversuchs	27
4.1	Generierung der Stimuli	27
4.1.1	Messung der BRIRs des Lautsprecherarrays	27

4.1.2	Binaural Room Scanning	29
4.2	Versuchsaufbau	29
4.3	Auswahl der Stimuli/Klangbeispiele	31
4.4	Vorversuche	33
4.4.1	Erster Vorversuch	33
4.4.2	Weitere Vorversuche	36
4.5	Finaler Versuch	37
5	Auswertung und Interpretation des Hörversuchs	39
5.1	Eingangsvoraussetzungen	39
5.1.1	ProbandInnen	39
5.1.2	Normalverteilung	41
5.2	Abhängigkeit von Klangbeispiel und BRIR-Länge	42
5.3	Auswertung der einzelnen Attributpaare	44
5.3.1	Attributpaar 1: dünner - voller	44
5.3.2	Attributpaar 2: weicher - härter	45
5.3.3	Attributpaar 3: natürlicher - künstlicher	46
5.3.4	Attributpaar 4: dunkler - heller	47
5.3.5	Attributpaar 5: weniger/stärker räumlich	48
5.3.6	Attributpaar 6: näher - ferner	49
5.3.7	Attributpaar 7: schmaler - breiter	50
5.3.8	Attributpaar 8: weiter links - weiter rechts	51
5.3.9	Attributpaar 9: schlechter/besser lokalisierbar	52
5.3.10	Attributsignifikanz	53
5.4	Zusammenfassung der Attributpaare durch die PCA	54
5.5	Interpretation der Ergebnisse und Vergleich mit der Theorie und Simulation	59
5.5.1	Klangfarbe	59
5.5.2	Räumliche Attribute	59
6	Zusammenfassung	60
7	Danksagung	61

1 Einleitung

Die meisten heutzutage gebräuchlichen Wiedergabeverfahren mit Lautsprechern (Stereo, Surround [The90]) arbeiten nach dem Prinzip der Phantomschallquelle [Bla83]. Dieses Prinzip beruht auf einem psychoakustischen Effekt, der sich aus der Wahrnehmung einer Überlagerung von mehreren Lautsprechersignalen ergibt. Eine Verallgemeinerung der stereophonischen Technik stellt Vector-Base Amplitude-Panning dar [Pul97]. Dabei sind im zweidimensionalen Fall immer zwei Lautsprecher, im dreidimensionalen immer drei Lautsprecher an der Bildung einer einzelnen Phantomschallquelle beteiligt. Um den gewünschten Effekt zu erzielen, sollte der Abstand zu den beteiligten Lautsprechern annähernd gleich groß sein. Dadurch reduziert sich jedoch der optimale Abhörpunkt (auch Sweetspot genannt) bei kreis- oder kugelförmigen Lautsprecheranordnungen auf einen kleinen Bereich im Zentrum der Anordnung.

Es existieren auch Verfahren, die eine größere Hörzone bieten. Dies wird dadurch möglich, dass die virtuellen Quellen nicht durch psychoakustische Effekte wahrgenommen werden, sondern das Schallfeld der virtuellen Quellen möglichst exakt reproduziert wird. Zu diesen Schallfeldreproduktionsverfahren zählen Wellenfeldsynthese (WFS) [Ber88, BdVV93] und Nearfield Compensated Higher-Order Ambisonics (HOA) [Dan00, Dan03] (im Gegensatz dazu unterscheidet sich traditionelles Ambisonics [Ger72] durch reines Amplitudenpanning kaum von stereophonischen Verfahren). Zu jedem der beiden vorgenannten Verfahren, WFS und HOA, existieren Untersuchungen [WA01, SRA08, BDPW08, BDG⁺07, Hub02, FZS08, Aug04, Wit07] mit analytischen Berechnungen und Hörversuchen. Vergleichende Untersuchungen gibt es allerdings auf theoretischer Basis nur wenige [DNM03, SA08], vergleichende Untersuchungen anhand von Hörversuchen sind nicht bekannt.

Diese Arbeit stellt die ersten Schritte eines direkten Vergleichs von WFS und HOA durch Hörversuche dar.

2 Theorie zu den Schallfeldreproduktionsverfahren

Dieses Kapitel befasst sich mit der Theorie der verwendeten Schallfeldreproduktionsverfahren. Im ersten Schritt wird dabei das Kirchoff-Helmholtz-Integral allgemein hergeleitet. Dieses wird dann speziell im Fall eines Schallfeldes betrachtet und dient als Basis für die beiden Verfahren. Am Ende des Kapitels werden analytische Simulationen der reproduzierten Schallfelder, sowie der zeitlichen Spektren durchgeführt, um die Unterschiede zwischen den beiden Schallfeldreproduktionsverfahren aufzuzeigen.

2.1 Gauß'scher Integralsatz und Green'sche Identität

Der Gauß'sche Integralsatz stellt einen Zusammenhang zwischen der Divergenz eines Vektorfeldes $\nabla \mathbf{F} = \left(\frac{\partial}{\partial x} \frac{\partial}{\partial y} \frac{\partial}{\partial z} \right)^T \cdot \mathbf{F}$ und dem durch das Feld \mathbf{F} vorgegebenen Fluss Q durch eine geschlossene Oberfläche S her [Wik].

Man kann einerseits die Quellenstärke eines Vektorfeldes $\mathbf{F} = (F_x \ F_y \ F_z)^T$ als Hüllenintegral über den differentiellen Fluss dQ darstellen:

$$\begin{aligned} dQ &= \mathbf{F} \cdot \mathbf{n} \, dS \\ \Rightarrow Q &= \oint_S \mathbf{F} \cdot \mathbf{n} \, dS. \end{aligned} \quad (1)$$

Andererseits lässt sich die Quellenstärke auch als Volumensintegral über differentielle Hüllenelemente dV darstellen:

$$\begin{aligned} dQ &= \left(\frac{\partial F_x}{\partial x} \frac{\partial F_y}{\partial y} \frac{\partial F_z}{\partial z} \right) dx \, dy \, dz = \nabla \mathbf{F} \, dV \\ \Rightarrow Q &= \int_V \nabla \mathbf{F} \, dV. \end{aligned} \quad (2)$$

Setzt man die beiden Ansätze (1) und (2) für die Quellenstärke Q gleich, so ergibt sich der Gauß'sche Integralsatz [Weia]:

$$\oint_S \mathbf{F} \cdot \mathbf{n} \, dS = \int_V \nabla \mathbf{F} \, dV. \quad (3)$$

2. Green'sche Identität

Zur Herleitung der zweiten Green'sche Identität wird der Term

$$\mathbf{F} = G \nabla P - P \nabla G \quad (4)$$

für das Vektorfeld \mathbf{F} eingesetzt. G und P sind dabei jeweils zweifach stetig differenzierbare Funktionen. Das Einsetzen dieses Terms in den Gauß'schen Integralsatz ergibt [Weib]:

$$\begin{aligned} \oint_S (G \nabla P - P \nabla G) \cdot \mathbf{n} \, dS &= \int_V \nabla (G \nabla P - P \nabla G) \, dV \\ &= \int_V (\nabla G \nabla P + G \nabla^2 P - \nabla P \nabla G + P \nabla^2 G) \, dV \\ &= \int_V (G \nabla^2 P - P \nabla^2 G) \, dV. \end{aligned} \quad (5)$$

Für die Funktion P als homogene Lösung der Helmholtz-Gleichung gilt:

$$(\nabla^2 + k^2)P = 0. \quad (6)$$

Für die Green'sche Funktion G als inhomogene Lösung der Helmholtz-Gleichung gilt:

$$(\nabla^2 + k^2)G = -\delta(\mathbf{x} - \mathbf{x}_0). \quad (7)$$

Setzt man die homogene Lösung (6) und die Green'sche Funktion (7) in die rechte Seite der 2. Green'schen Identität (5) ein, so erhält man:

$$\oint_V (G\nabla^2 P - P\nabla^2 G) dV = \oint_V (-Gk^2 P + \delta(\mathbf{x} - \mathbf{x}_0)P + k^2 GP) dV = P. \quad (8)$$

Die linke Seite der 2. Green'schen Identität lässt sich vereinfachen zu:

$$\begin{aligned} \oint_S (G\nabla P - P\nabla G) \cdot \mathbf{n} dS &= \oint_S (G\nabla P \cdot \mathbf{n} - P\nabla G \cdot \mathbf{n}) dS \\ &= \oint_S \left(G \frac{\partial}{\partial \mathbf{n}} P - P \frac{\partial}{\partial \mathbf{n}} G \right) dS. \end{aligned} \quad (9)$$

2.2 Kirchhoff-Helmholtz-Integral

Aus (8) und (9) lässt sich die **3. Greensche Identität**, auch bekannt als das **Kirchhoff-Helmholtz-Integral**, anschreiben [SRA08]:

$$P(\mathbf{x}, \omega) = - \oint_{\partial V} \left(G(\mathbf{x}|\mathbf{x}_0, \omega) \frac{\partial}{\partial \mathbf{n}} S(\mathbf{x}_0, \omega) - S(\mathbf{x}_0, \omega) \frac{\partial}{\partial \mathbf{n}} G(\mathbf{x}|\mathbf{x}_0, \omega) \right) dS_0. \quad (10)$$

Das Kirchhoff-Helmholtz-Integral besagt für ein quellenfreies Volumen V , dass das Schallfeld einer Quelle S innerhalb der geschlossenen Hülle δV bzw. S_0 allein durch den Schalldruck $S(\mathbf{x}_0, \omega)$ und die Normalschnelle $\frac{\partial}{\partial \mathbf{n}} S(\mathbf{x}_0, \omega)$ auf der Hülle eindeutig bestimmt wird.

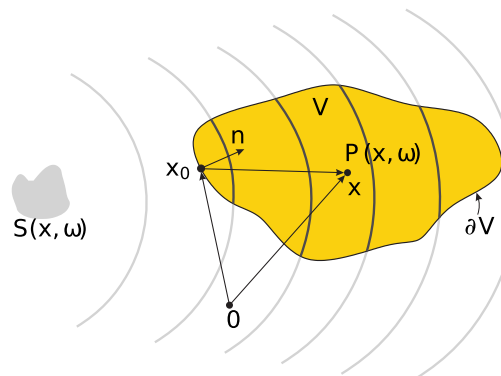


Abbildung 1: Illustration zum Kirchhoff-Helmholtz-Integral aus [SA08]

Diese Formel ist der Ausgangspunkt der Reproduktion des Schallfeldes $S(\mathbf{x}_0, \omega)$ mittels Lautsprecher als Sekundärquellen. Das Abstrahlverhalten dieser Lautsprecher wird durch die Green'sche Funktion $G(\mathbf{x}|\mathbf{x}_0, \omega)$ und deren Ableitung in Normalrichtung repräsentiert. Für eine physikalisch perfekte Rekon-

struktion bräuchte man also sowohl Schalldruck-(Monopol-) als auch Schallschnellequellen (Dipolquellen) als Sekundärquellen. Lautsprecher lassen sich am ehesten durch Monopolquellen approximieren, deshalb wird versucht den Dipolanteil zu eliminieren. Als Konsequenz dieser Eliminierung ist das Schallfeld außerhalb des Volumens V (außerhalb der Lautsprecheranordnung) nicht gleich 0. Die beiden in dieser Arbeit untersuchten Verfahren führen diese Eliminierung auf verschiedene Art durch. Nach der Eliminierung ergibt sich allgemein für das resultierende Schallfeld $P(\mathbf{x}, \omega)$ [SA08]:

$$P(\mathbf{x}, \omega) = - \oint_{\partial V} D(\mathbf{x}_0, \omega) G(\mathbf{x}|\mathbf{x}_0, \omega) dS_0. \quad (11)$$

Der Term $D(\mathbf{x}_0, \omega)$ ist die Ansteuerungsfunktion (engl. driving function). Er bestimmt, was die Lautsprecher, modelliert durch die Green'sche Funktion $G(\mathbf{x}|\mathbf{x}_0, \omega)$, wiedergeben.

In der weiteren Herleitung wird nur der untersuchte Fall einer kreisförmigen Wiedergabeanordnung behandelt. Die Positionen \mathbf{x}_0 der Sekundärquellen werden deshalb durch den Radius R und die Winkel α_0 ausgedrückt:

$$\mathbf{x}_0 = R \begin{pmatrix} \cos(\alpha_0) \\ \sin(\alpha_0) \end{pmatrix}. \quad (12)$$

2.3 Wellenfeldsynthese (WFS)

Bei der Wellenfeldsynthese (WFS) wird der Neumann'sche Ansatz für die Green'sche Funktion gewählt, da dort die Ableitung in Normalrichtung und damit der Dipolanteil in (10) verschwindet [SA08]:

$$\frac{\partial}{\partial \mathbf{n}} G_N(\mathbf{x}|\mathbf{x}_0, \omega)|_{\mathbf{x}_0 \in \partial V} \stackrel{!}{=} 0, \quad (13)$$

$$G_N(\mathbf{x}|\mathbf{x}_0, \omega)|_{\mathbf{x}_0 \in \partial V} = 2G_0(\mathbf{x}|\mathbf{x}_0, \omega). \quad (14)$$

Die Neumann Randbedingung modelliert die Sekundärquellenkontur ∂V als schallhart, da die Schallschnelle auf der Kontur auf 0 gesetzt wird. Dadurch wird das erzeugte Schallfeld dort virtuell reflektiert. Man unterdrückt diese Reflexionen, indem man die rückwertigen Sekundärquellen, deren Normalvektor keinen Anteil in Ausbreitungsrichtung hat, abschaltet. Dazu wird die Fensterfunktion $a(\mathbf{x}_0)$ bzw. $a(\alpha_0)$ verwendet [SA08]. Zusätzlich muss die Sekundärquellenkontur konvex sein, damit sich das Schallfeld außerhalb des Hörbereichs nicht in den Hörbereich ausbreitet. Diese Voraussetzung wird im betrachteten Fall ohnehin erfüllt. Damit ergibt sich (10) für eine kreisförmige Anordnung zu:

$$P(\mathbf{x}, \omega) = - \oint_{\partial V} 2a(\mathbf{x}_0) \frac{\partial}{\partial \mathbf{n}} S(\mathbf{x}_0, \omega) G_0(\mathbf{x}|\mathbf{x}_0, \omega) dS_0 \quad (15)$$

$$= - \int_0^{2\pi} \underbrace{2a(\alpha_0) \frac{\partial}{\partial \mathbf{n}} S(\mathbf{x}_0, \omega)}_{D(\mathbf{x}_0, \omega)} G_0(\mathbf{x}|\mathbf{x}_0, \omega) d\alpha_0. \quad (16)$$

Für Wellenfeldsynthese (WFS) ergibt sich die Ansteuerungsfunktion allgemein aus der Ableitung des virtuellen Schallfeldes in Normalrichtung und der Fensterfunktion für die Auswahl der Sekundärquellen. Im Folgenden wird die Ansteuerungsfunktion für zwei Arten von virtuellen Quellen, ebene Welle und Punktquelle, hergeleitet.

2.3.1 Rekonstruktion von ebenen Wellen mit WFS

In diesem Abschnitt wird die Ansteuerungsfunktion für eine ebene Welle hergeleitet:

$$D_{pw}(\mathbf{x}_0, \omega) = -2a_{pw}(\alpha_0) \frac{\partial}{\partial \mathbf{n}} S(\mathbf{x}_0, \omega).$$

Für eine ebene Welle mit Eintrittswinkel Φ_{pw} gilt im zweidimensionalen Fall [Wil99]:

$$S_{pw}(\mathbf{x}_0, \omega) = e^{-i\mathbf{k}_{pw}^T \mathbf{x}} = e^{-ik_{pw,x}x} e^{-ik_{pw,y}y}. \quad (17)$$

Die Fensterfunktion (Auswahl der Sekundärquellen am Ort \mathbf{x}_0) für eine ebene Welle mit Eintrittswinkel Φ_{pw} lautet allgemein:

$$a_{pw}(\mathbf{x}_0) = \begin{cases} 1, & \text{für } \mathbf{n}_{pw}^T \mathbf{n}_{x_0} > 0, \\ 0, & \text{sonst.} \end{cases}$$

Für eine runde Sekundärquellenverteilung ergibt sich die Fensterfunktion in Abhängigkeit des Winkels α_0 der einzelnen Sekundärquellen zu:

$$a_{pw}(\alpha_0) = \begin{cases} 1, & \text{für } \cos(\alpha_0 - \Phi_{pw}) > 0, \\ 0, & \text{sonst.} \end{cases} \quad (18)$$

Der Normalvektor für den untersuchten Fall einer kreisförmigen Wiedergabeordnung berechnet sich als:

$$\mathbf{n} = \begin{pmatrix} \cos(\alpha + \pi) \\ \sin(\alpha + \pi) \end{pmatrix} = \begin{pmatrix} -\cos(\alpha) \\ -\sin(\alpha) \end{pmatrix} = - \begin{pmatrix} \cos(\alpha) \\ \sin(\alpha) \end{pmatrix}. \quad (19)$$

Für die Ableitung $\frac{\partial}{\partial \mathbf{n}}$ in Richtung dieses Normalvektors gilt:

$$\frac{\partial}{\partial \mathbf{n}} = - \left(\cos(\alpha) \frac{\partial}{\partial x} + \sin(\alpha) \frac{\partial}{\partial y} \right). \quad (20)$$

Im Speziellen ergeben sich die partiellen Ableitungen für den Fall einer ebenen Welle zu:

$$\frac{\partial}{\partial x} S_{pw}(\mathbf{x}_0, \omega) = -ik_{pw,x} e^{-ik_{pw,x}x}, \quad (21)$$

$$\frac{\partial}{\partial y} S_{pw}(\mathbf{x}_0, \omega) = -ik_{pw,y} e^{-ik_{pw,y}y}. \quad (22)$$

Für die Ansteuerungsfunktion $D_{pw}(\alpha_0, \omega)$ ergibt sich im Falle einer ebenen Welle aus (21), (22) und (20) also [SR06]:

$$\begin{aligned} D_{pw}(\alpha_0, \omega) &= -2a_{pw}(\alpha_0) \frac{\partial}{\partial \mathbf{n}} S_{pw}(\mathbf{x}_0, \omega) \\ &= -2a_{pw}(\alpha_0) ik \cos(\alpha_0 - \Phi_{pw}) e^{-i\mathbf{k}_{pw}^T \mathbf{x}}. \end{aligned} \quad (23)$$

2.3.2 Rekonstruktion von Punktquellen mit WFS

Im Folgenden wird die Ansteuerungsfunktion für eine Punktquelle berechnet. Die Vorgehensweise ist dabei analog zum vorherigen Abschnitt. Die Ansteuerungsfunktion lautet allgemein:

$$D_{sw}(\mathbf{x}_0, \omega) = -2a_{sw}(\alpha_0) \frac{\partial}{\partial \mathbf{n}} S_{sw}(\mathbf{x}_0, \omega).$$

Für eine Punktquelle am Ort \mathbf{x}_s gilt mit $r = |\mathbf{x}_0 - \mathbf{x}_s|$:

$$S_{sw}(\mathbf{x}_0, \omega) = \frac{e^{-ikr}}{r}. \quad (24)$$

Die Fensterfunktion (Auswahl der Sekundärquellen \mathbf{x}_0) für eine Punktquelle am Ort \mathbf{x}_s lautet allgemein:

$$a_{sw}(\mathbf{x}_0) = \begin{cases} 1, & \text{für } (\mathbf{x}_0 - \mathbf{x}_s)^T \mathbf{n}_{\mathbf{x}_0} > 0, \\ 0, & \text{sonst.} \end{cases}$$

Für eine kreisförmige Sekundärquellenverteilung ergibt sich Fensterfunktion in Abhängigkeit des Winkels α_0 der einzelnen Sekundärquellen zu:

$$a_{sw}(\alpha_0) = \begin{cases} 1, & \text{für } -(\cos(\alpha_0)(x_0 - x_s) + \sin(\alpha_0)(y_0 - y_s)) > 0, \\ 0, & \text{sonst.} \end{cases} \quad (25)$$

Die partielle Ableitung in x-Richtung lautet:

$$\begin{aligned} \frac{\partial}{\partial x} S_{sw}(\mathbf{x}_0, \omega) &= \frac{\left[\frac{\partial(e^{-ikr})}{\partial x} \cdot r \right] - [e^{-ikr} \cdot \frac{\partial r}{\partial x}]}{r^2} \\ &= -\frac{e^{-ikr}}{r} \cdot (x - x_s) \left[\frac{1}{r^2} + ik \right]. \end{aligned} \quad (26)$$

Analog dazu gilt für die Ableitung in y-Richtung:

$$\frac{\partial}{\partial y} S_{sw}(\mathbf{x}_0, \omega) = -\frac{e^{-ikr}}{r} \cdot (y - y_s) \left[\frac{1}{r^2} + ik \right]. \quad (27)$$

Für die Ansteuerungsfunktion $D_{sw}(\mathbf{x}_0, \omega)$ ergibt sich für den Fall einer Punktquelle aus (26) und (27) also [SRA08]:

$$\begin{aligned} D_{sw}(\mathbf{x}_0, \omega) &= -2a_{sw}(\alpha_0) \frac{\partial}{\partial \mathbf{n}} S(\mathbf{x}_0, \omega) \\ &= -2a_{sw}(\alpha_0) \frac{e^{-ikr}}{r} \left[\frac{1}{r^2} + ik \right] (\mathbf{x}_0 - \mathbf{x}_s)^T \mathbf{n}_{\mathbf{x}_0}. \end{aligned} \quad (28)$$

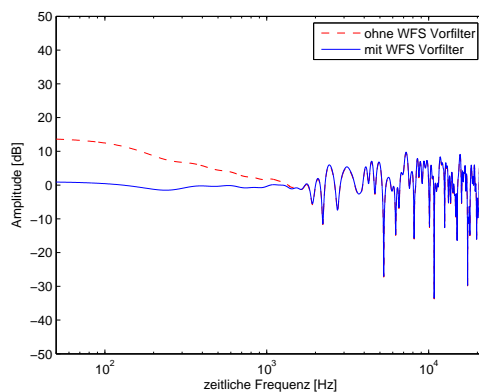
2.3.3 2.5D Korrektur [SRA08]

Bei den meisten WFS-Systemen sind die Sekundärquellen auf einer Ebene verteilt, was dazu führt, dass die reproduzierten Schallfelder nur zweidimensional in dieser Ebene kontrolliert werden können. Für eine rein zweidimensionale Wellenausbreitung werden aber sekundäre Linienquellen benötigt, die Zylinderwellen abstrahlen. Will man dazu Monopolstrahler verwenden, so muss die Ansteuerungsfunktion verändert werden. Die folgende Berechnung drückt die Green'sche Funktion einer Linienquelle [Wil99] durch die einer Punktquelle aus. Dabei wird die Fernfeldapproximation für Hankelfunktionen 0. ter Ordnung [AS72] verwendet:

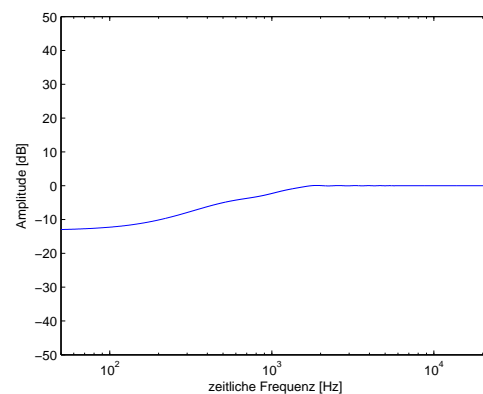
$$G_{2D}(\mathbf{x} - \mathbf{x}_0, \omega) = iH_0^{(2)}(kr) \quad (29)$$

$$\begin{aligned} &\approx i\sqrt{\frac{2i}{\pi kr}} e^{-ikr} \\ &\approx -\sqrt{\frac{2r}{\pi ik}} \underbrace{\frac{e^{-ikr}}{r}}_{G_{3D}(\mathbf{x}-\mathbf{x}_0, \omega)} \end{aligned} \quad (30)$$

Um eine rein zweidimensionale Wellenausbreitung mit Monopolstrahlern zu approximieren, muss eine vorherige Filterung der Ansteuerungsfunktion oder des Eingangssignals durchgeführt werden. Es ergibt sich ein Filter mit halber Ordnung (siehe Abbildung 2), welches aber nur bis zur Aliasingfrequenz (siehe 2.5) eingesetzt wird. Oberhalb dieser Frequenz kommt durch das räumliche Aliasing zusätzliche Energie hinzu [AS09]. Trotzdem kommt es noch zu Amplitudenfehlern, da der Pegel von Linienquellen nur 3dB pro Abstandsverdopplung abnimmt und der von Punktquellen 6dB.



(a) Wiedergabesystem mit/ohne WFS Vorfilter



(b) verwendetes WFS Vorfilter von [GAS08b]

Abbildung 2: Auswirkung des WFS Vorfilters: Betragsschwingungsgang für die zentrale Hörposition (für den Fall der Wiedergabe einer virtuellen ebenen Welle aus 0° Richtung mit 56 Lautsprechern)

2.4 Higher Order Ambisonics (HOA)

Im Ansatz des traditionellen Ambisonics [Ger72, BV95] werden die Sekundärquellen als ebene Wellen betrachtet. Diese Annäherung trifft allerdings nur zu, wenn der Radius der Sekundärquellenkontur sehr groß ist. Die Entstehung eines *Sweetspot* wird dadurch erklärt, dass das Feld der Punktquellen nur in diesem sehr kleinen Raumbereich den Charakter von ebenen Wellen haben. Der *Sweetspot* ist ein kleiner Hörbereich im Zentrum der Anordnung, in dem das reproduzierte Feld dem gewünschten sehr gut entspricht. Im Gegensatz zum traditionellen Ansatz werden beim hier verwendeten *Nearfield Compensated Higher-Order Ambisonics* Monopolquellen als Sekundärquellen angenommen.

Die Bestimmung der Ansteuerungsfunktion $D(\mathbf{x}_0, \omega)$ geschieht hier nicht auf direktem Wege, sondern im transformierten Bereich der räumlichen Frequenz ν entlang des Winkels α . Die Ansteuerungsfunktion ergibt sich erst aus der Rücktransformation von $\mathring{D}_\nu(\omega)$. Entscheidend ist dabei, dass die räumliche Bandbreite (auch Ordnung genannt) ν_{max} begrenzt wird [Dan00, SA08] (siehe 2.5).

$$D(\alpha_0, \omega) = \sum_{\nu=-\nu_{max}}^{\nu_{max}} \mathring{D}_\nu(\omega) e^{i\nu\alpha_0} \quad (31)$$

Für eine kreisförmige Verteilung (mit dem Radius R) von Sekundärquellen lässt sich das reproduzierte Schallfeld folgendermaßen anschreiben [AS08a]:

$$P(\mathbf{x}, \omega) = \oint_0^{2\pi} D(\alpha_0, \omega) \cdot G(\mathbf{x} - \mathbf{x}_0, \omega) R d\alpha_0. \quad (32)$$

Diese Gleichung kann als eine zirkuläre Faltung interpretiert werden und mithilfe des Faltungstheorems [Wil99]

$$\mathring{P}_\nu(r, \omega) = 2\pi R \mathring{D}_\nu(\omega) \mathring{G}_\nu(r, \omega) \quad (33)$$

erhält man für die Ansteuerungsfunktion für den Fall einer kreisförmigen Sekundärquellenverteilung [AS08a]:

$$\mathring{D}_\nu(\omega) = \frac{1}{2\pi R} \frac{\mathring{P}_\nu(r, \omega)}{\mathring{G}_\nu(r, \omega)}. \quad (34)$$

2.4.1 2.5D Korrektur

Die vorherige Berechnung der Ansteuerungsfunktion (34) gilt nur bei der Reproduktion zweidimensionaler Schallfelder. Da zur Reproduktion über ein kreisförmiges Wiedergabesystem meistens Punktquellen als Sekundärquellen benutzt werden, muss auch hier eine Anpassung der Ansteuerungsfunktion vorgenommen werden. Der wichtigste Schritt ist dabei die Referenzierung des wiedergegebenen Schallfeldes auf den Mittelpunkt des Wiedergabesystems. Durch weitere Umformung und Anwendung der Regel von de l'Hôpital erhält man die 2.5D Ansteuerungsfunktion [AS08a]:

$$\mathring{D}_\nu(\omega) = \frac{1}{2\pi R} \frac{\mathring{P}_{|\nu|}^\nu(\omega)}{\mathring{G}_{|\nu|}^\nu(\omega)}. \quad (35)$$

\mathring{P} und \mathring{G} sind dabei die Spherical Harmonic Expansions der virtuellen Quellen bzw. der Sekundärquellen. Die expliziten Terme für eine ebene Welle mit der Eintrittsrichtung (Φ_{pw}, Θ_{pw}) und eine Punktquelle

im Punkt \mathbf{x}_s bzw. (r_s, α_s, β_s) sind gegeben als [AS08a]:

- ebene Wellen:

$$\check{P}_n^m(\omega) = 4\pi(-i)^n Y_n^m(\Phi_{pw}, \Theta_{pw})^*$$

- Punktquellen:

$$\check{P}_n^m(\omega) = -ikh_n^{(2)}(kr_s) Y_n^m(\alpha_s, \beta_s)^*$$

mit

$$Y_n^m(\alpha, \beta) = \sqrt{\frac{(2n+1)(n-m)!}{4\pi(n+m)!}} \cdot P_n^m(\cos \beta) \cdot e^{im\alpha}$$

und P_n^m ist die zugeordnete Legendre Funktion m -ter Ordnung des Grades n [Wil99]. * kennzeichnet die konjugiert komplexe Funktion.

2.4.2 Rekonstruktion von ebenen Wellen mit HOA

Damit lässt sich nun die Ansteuerungsfunktion für eine ebene Welle im 2.5D Fall ($\Theta_{pw} = \beta_s = \frac{\pi}{2}$) berechnen [AS08b]:

$$\begin{aligned} \mathring{D}_{pw}(\nu, \omega) &= \frac{1}{2\pi R} \frac{\check{P}_{|\nu|}^\nu(\omega)}{\check{G}_{|\nu|}^\nu(\omega)} \\ &= -\frac{2}{ikR} \frac{(-i)^{|\nu|}}{h_{|\nu|}^{(2)}(kR)} e^{(-i\nu\Phi_{pw})}. \end{aligned} \quad (36)$$

2.4.3 Rekonstruktion von Punktquellen mit HOA

Ebenso ergibt sich die Ansteuerungsfunktion für eine Punktquelle zu [AS08b]:

$$\begin{aligned} \mathring{D}_{sw}(\nu, \omega) &= \frac{1}{2\pi R} \frac{\check{P}_{|\nu|}^\nu(\omega)}{\check{G}_{|\nu|}^\nu(\omega)} \\ &= \frac{1}{2\pi R} \frac{h_{|\nu|}^{(2)}(kr)}{h_{|\nu|}^{(2)}(kR)} e^{(-j\nu\Phi_{sw})}. \end{aligned} \quad (37)$$

2.5 Diskretisierung

Um die Ansteuerungsfunktion mit N diskreten Lautsprechern wiedergeben zu können, muss diese an den einzelnen Lautsprecherpositionen abgetastet werden. Bei dieser Diskretisierung wird das Winkelspektrum periodisch fortgesetzt [SR06]:

$$\mathring{D}_s(\nu, \omega) = \sum_{\eta=-\infty}^{\infty} \mathring{D}(\nu + \eta N, \omega). \quad (38)$$

Wie auch bei der zeitlichen Diskretisierung kann es dadurch zu Aliasing kommen. Streng genommen ist Aliasing erst dann gegeben, wenn die spektralen Wiederholungen ins Basisband hinein ragen. Die Wiedergabe der Wiederholungen außerhalb des Basisbandes ist eher ein Rekonstruktionsfehler, wenn die Sekundärquellen diese nicht ausblenden. Da die hier verwendeten Sekundärquellen ein unendliches Winkelspektrum besitzen, werden der Einfachheit halber immer dann von räumlichem Aliasing gesprochen, wenn spektrale Wiederholungen wiedergegeben werden. Räumliche Aliasing führt zu Artefakten im reproduzierten Schallfeld.

Dennoch gibt es bei HOA Raumbereiche, in denen kein räumliches Aliasing herrscht. Das reproduzierte Schallfeld kann als Summe der Produkte der Ansteuerungsfunktion und der Green'schen Funktion der Lautsprecher über alle räumlichen Frequenzen ν dargestellt werden. Über eine Fourier-Reihenentwicklung kann die Punktquelle der Green'schen Funktion in Besselfunktionen mit dem Argument kr zerlegt werden [SA08]. In Abbildung 3 sind Besselfunktionen für verschiedene Ordnungen dargestellt. Man erkennt, dass das Maximum der Besselfunktion für steigende Ordnungen erst bei einem größeren Wert für kr auftritt. Hält man die zeitliche Frequenz (und damit k) fest, so heißt das, dass höhere Ordnungen bzw. räumliche Frequenzen erst bei größerem Abstand r vom Mittelpunkt der Wiedergabeanordnung an der Schallfeldreproduktion beteiligt sind.

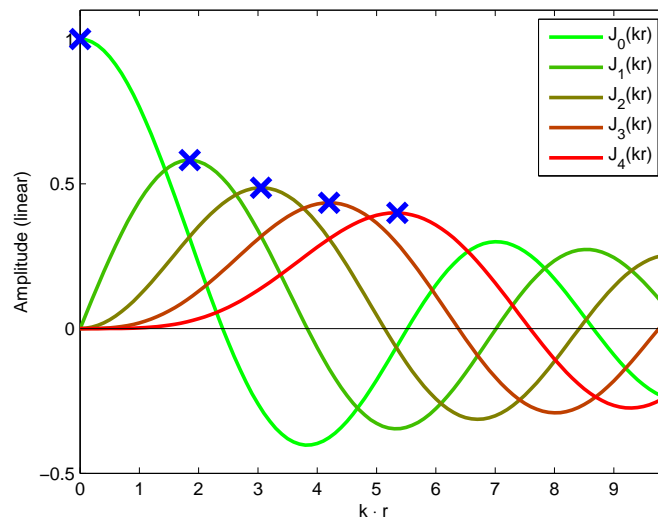


Abbildung 3: Besselfunktionen erster Art für verschiedene Ordnungen, blaue Kreuze kennzeichnen die Maxima

Wie bereits vorher erwähnt, wird bei HOA die räumliche Bandbreite (Ordnung) begrenzt. Die maximale Ordnung ν_{max} wird durch die Lautsprecheranzahl bestimmt. Für eine kreisförmige, zweidimensionale Wiedergabeanordnung gilt [Dan00, SA08]:

$$\nu_{max} = \lfloor \frac{N-1}{2} \rfloor, \quad (39)$$

wobei $\lfloor x \rfloor$ eine Abrundung auf die nächste ganze Zahl darstellt.

Dadurch überlappen sich die Spektren bei HOA nicht (siehe Abbildung 4) und es ergibt sich abhängig von der zeitlichen Frequenz ein Radius, innerhalb dem eine aliasingfreie und damit auch klanglich perfekte Schallfeldreproduktion möglich ist. Dieser Raumbereich wird als *Sweetspot* bezeichnet. Für größere Radien und höhere zeitliche Frequenzen nimmt das räumliche Aliasing zu. Bei Verringerung der Lautsprecheranzahl und damit der Ordnung verkleinert sich der Sweetspot (siehe Abbildung 6), da die Überlappung der Spektren schon bei kleineren räumlichen und zeitlichen Frequenzen eintritt (siehe Abbildung 5).

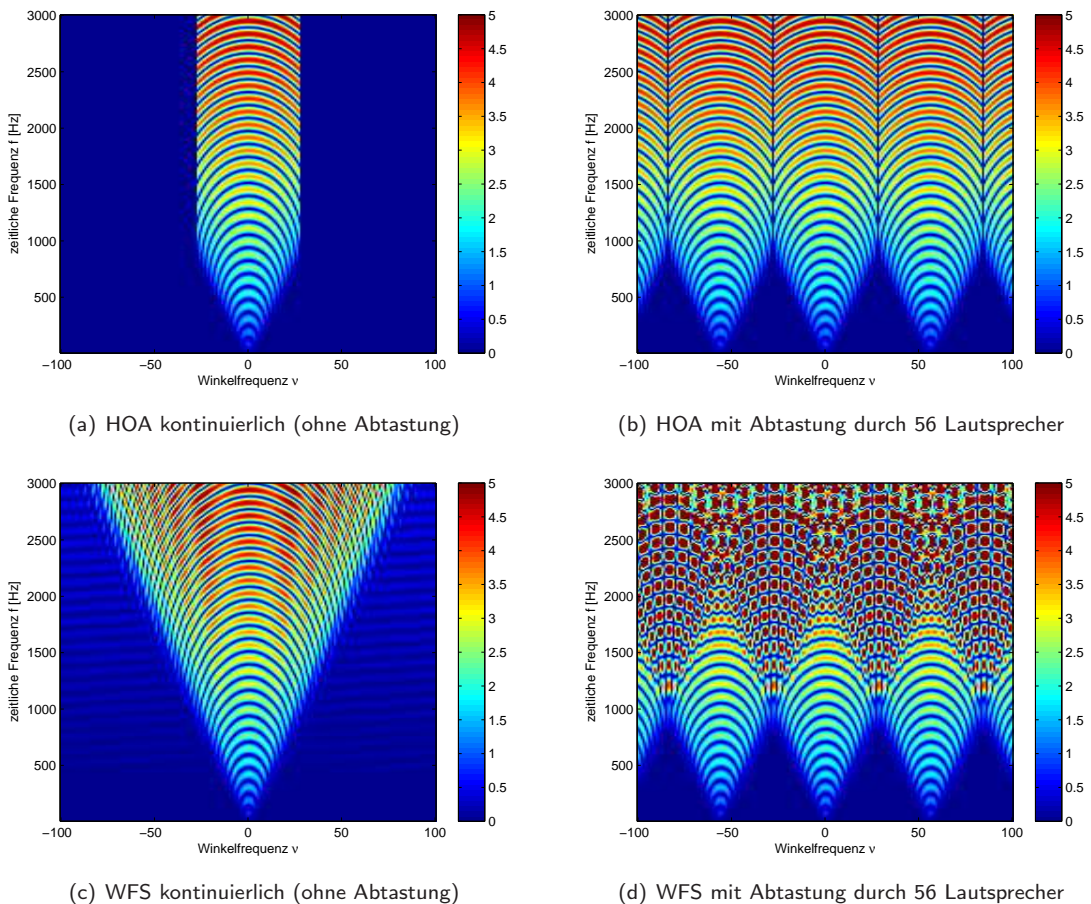


Abbildung 4: Absolutbetrag des Realteils der Winkelspektren der Ansteuerungsfunktionen $\hat{D}_{PW}(\nu, \omega)$ für die Verwendung von 56 Lautsprechern (für den Fall einer virtuellen ebenen Welle aus 0° Richtung)

Bei WFS gibt es aufgrund der überlappenden Spektren (siehe Abbildung 4) keinen Sweetspot, dafür ist das Aliasing aber räumlich gleichmäßiger verteilt. Die zeitliche Frequenz, ab der räumliches Aliasing auftritt, wird in diesem Fall hauptsächlich durch den Abstand zwischen den Lautsprechern der Wiedergabeordnung bestimmt. Da sich dieser Abstand bei Verwendung von weniger Lautsprechern vergrößert (siehe Abbildung 5), sinkt auch die Aliasingfrequenz.

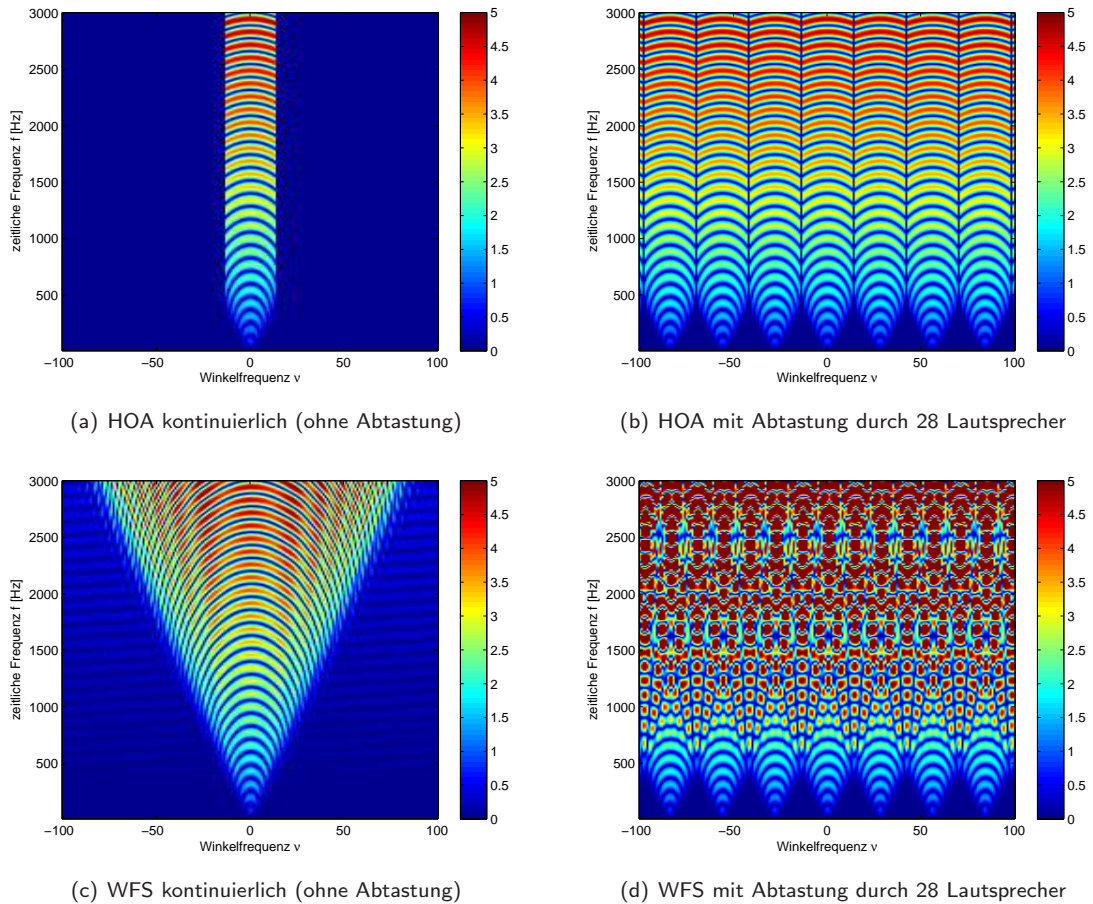


Abbildung 5: Absolutbetrag des Realteils der Winkelspektren der Ansteuerungsfunktionen $\hat{D}_{PW}(\nu, \omega)$ für die Verwendung von 28 Lautsprechern (für den Fall einer virtuellen ebenen Welle aus 0° Richtung)

2.6 Analytische Simulation der Schallfeldreproduktionsverfahren

Die im vorherigen Teil hergeleiteten Ansteuerungsfunktionen für WFS und HOA wurden in MATLAB implementiert. Diese Implementierung dient der Berechnung der Stimuli für den Hörversuch und der Simulation. Für den im Versuch verwendeten Fall einer ebenen Welle aus 0° Richtung werden nun Simulationen der reproduzierten monochromatischen Schallfelder und der Betragsfrequenzgänge für verschiedene Aufpunkte durchgeführt. Diese Positionen sind auch in der Darstellung der Schallfelder als Kreise eingezeichnet und entsprechen den verwendeten Hörpositionen (Center, Front, Side) im Versuch.

2.6.1 Reproduzierte Schallfelder HOA [SA08]

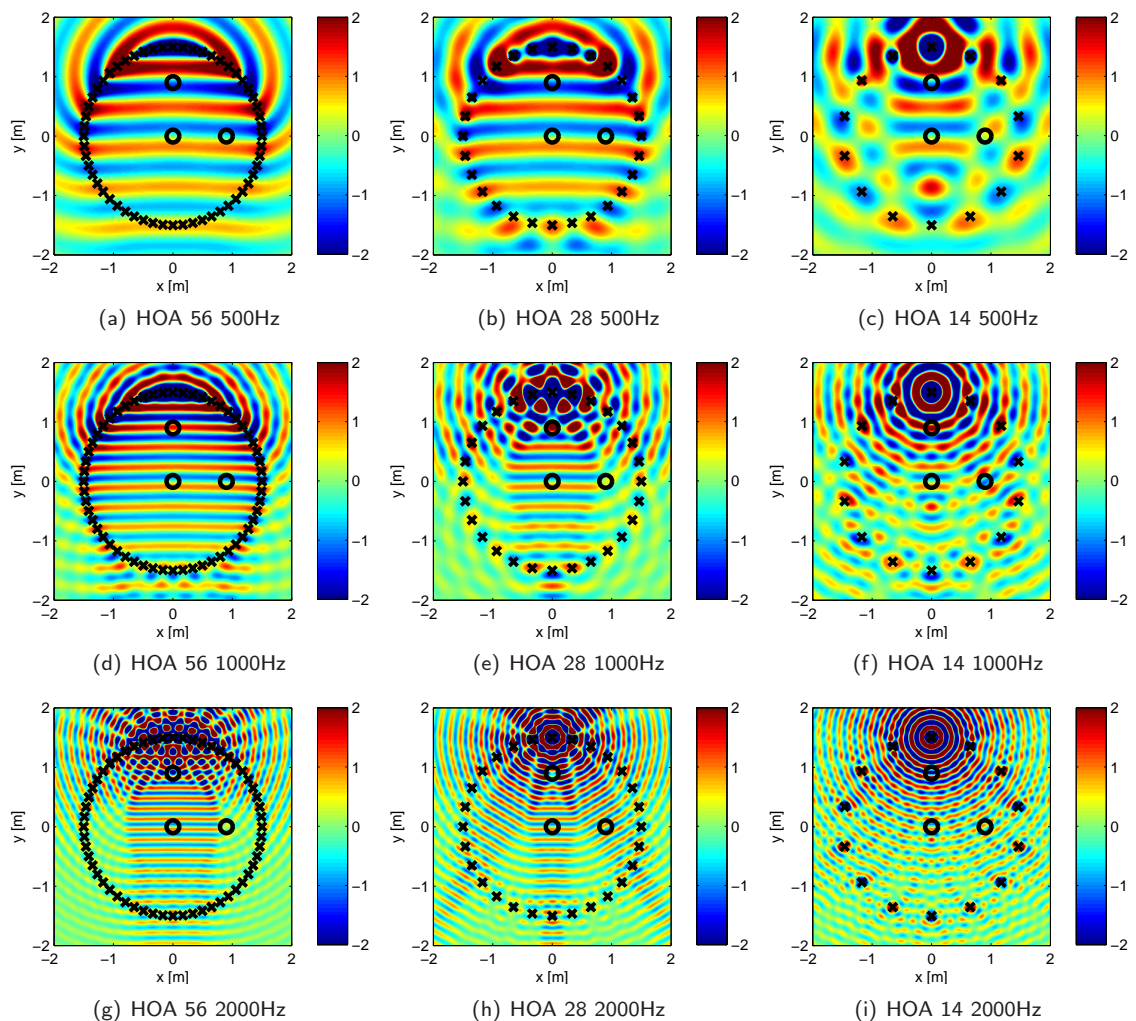


Abbildung 6: HOA: Schalldruck des reproduziertes Schallfeldes für verschiedenen Lautsprecheranzahlen und Frequenzen (für den Fall einer virtuellen ebenen Welle aus 0° Richtung)

Man erkennt deutlich die Bildung eines Sweetspots im Zentrum der Wiedergabeanordnung. Dieser Bereich verkleinert sich mit hohen Frequenzen und geringer Lautsprecheranzahl. Im Bereich der vorderen Hörposition ist die Bildung besonders starker Artefakte erkennbar.

2.6.2 Reproduzierte Schallfelder WFS [SA08]

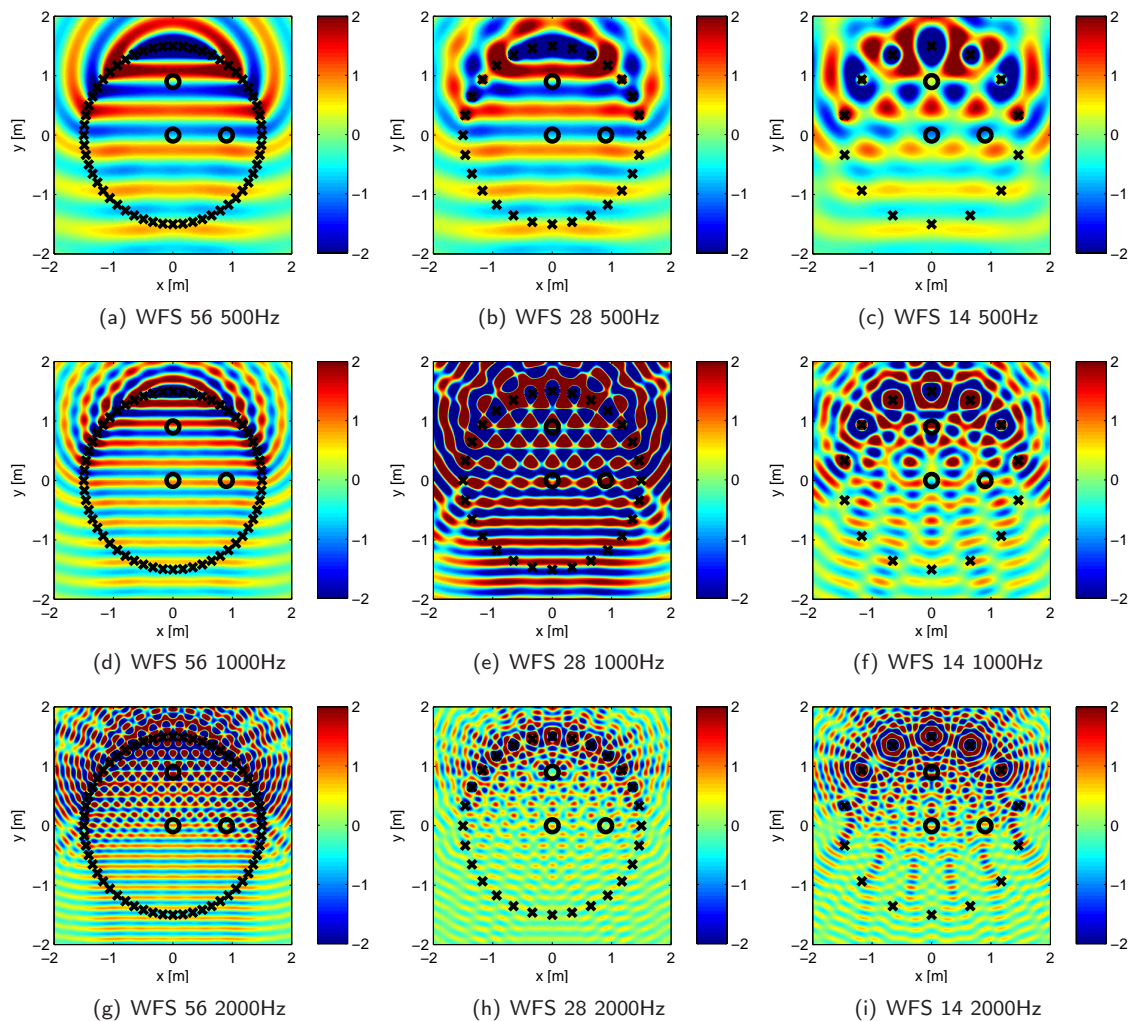


Abbildung 7: WFS: Schalldruck des reproduziertes Schallfeldes für verschiedenen Lautsprecheranzahlen und Frequenzen (für den Fall einer virtuellen ebenen Welle aus 0° Richtung)

Bei WFS ist auch die Zunahme der Artefakte bei hohen Frequenzen und geringer Lautsprecheranzahl und die besonders starke Ausprägung an der vorderen Hörposition erkennbar. Ein Sweetspot existiert hier nicht, allerdings ist die Reproduktion besser, je größer der Abstand zu den wiedergebenden Lautsprechern ist (Durch die Sekundärquellenauswahl liegt in diesem Fall nur an den Lautsprechern des vorderen Halbkreises ein Signal an).

2.6.3 Zeitliche Spektren HOA

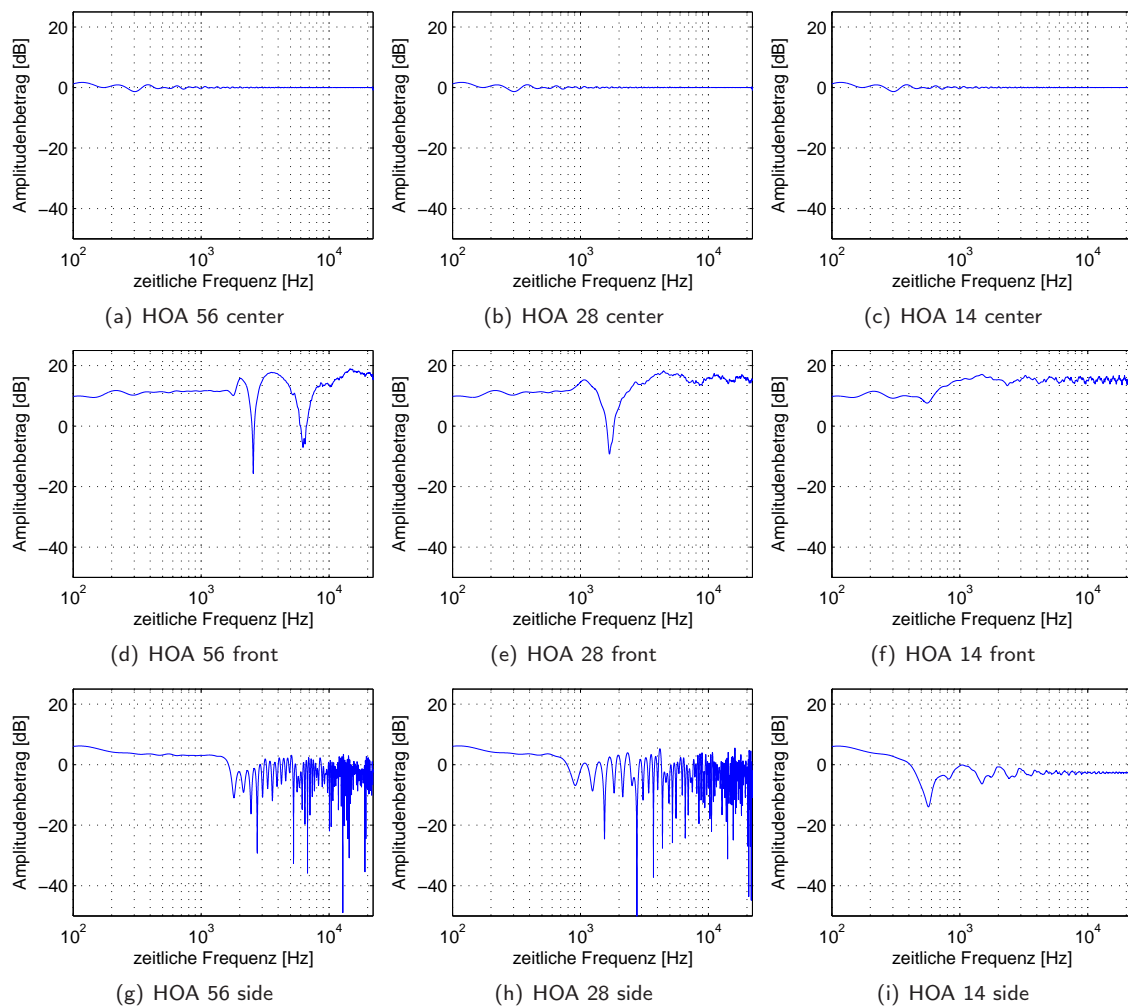


Abbildung 8: HOA: zeitliche Spektren an verschiedenen Orten für verschiedene Lautsprecheranzahlen (für den Fall einer virtuellen ebenen Welle aus 0° Richtung)

Bei den zeitlichen Spektren ist ein nahezu konstanter Frequenzgang für die zentrale Hörposition erkennbar. Für die vordere Position gibt es einzelne Einbrüche, die mit sinkender Lautsprecheranzahl verschwinden. An der Seite sind starke, sehr dichte Kammfilterstrukturen im oberen Frequenzbereich und eine Anhebung der tieferen Frequenzen zu erkennen. Bei der Verwendung von nur 14 Lautsprechern an dieser Position scheint der Frequenzgang auch oberhalb von ca. 3KHz nahezu konstant zu sein. Das ist aber, wie in Abbildung 10 zu erkennen ist, mit der günstigen Lage an der seitlichen Hörposition zu erklären.

Insgesamt zeigen die Frequenzgänge bei HOA eine hohe Abhängigkeit von der Hörposition.

Hier wird das vorher angesprochene Phänomen des scheinbar nahezu konstanten Frequenzgangs bei HOA mit 14 Lautsprechern an der seitlichen Hörposition untersucht. In Abbildung 9 ist die Konstellation der Lautsprecher im Vergleich zur Hörposition dargestellt. Der nahezu konstante Frequenzgang wird nur erreicht, da bei dieser Hörposition kein Lautsprecher direkt neben dem rechten Ohr ist und die beiden benachbarten Lautsprecher exakt den gleichen Abstand zur Hörposition haben.

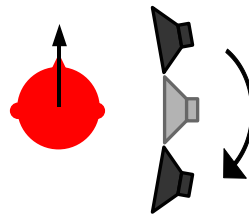
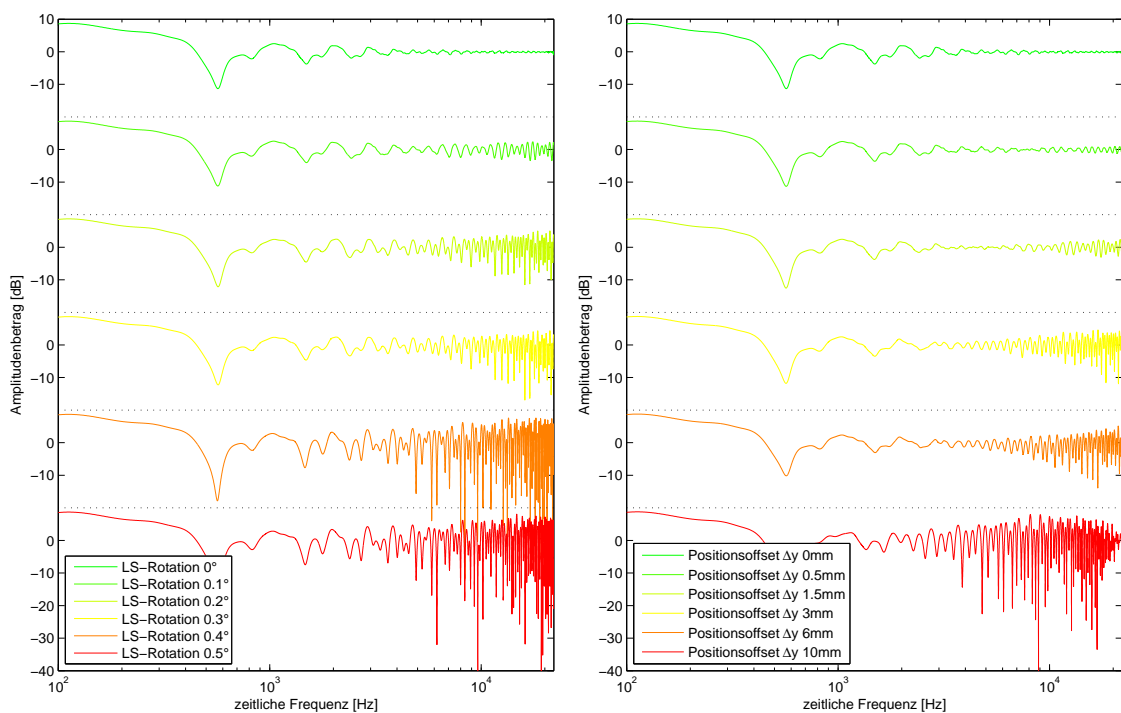


Abbildung 9: Konstellation der Lautsprecher und deren Änderung bei HOA mit 14 Lautsprechern an der seitlichen Hörposition

Variiert man diese Konstellation nur sehr leicht, z.B. durch minimale Drehung der Lautsprecheranordnung oder Veränderung der Hörposition, so sind sofort Kammfilterstrukturen erkennbar (siehe Abbildung 10). Da sich die ausschlaggebenden Abweichungen im Millimeter- und Zehntelgradbereich bewegen, ist anzunehmen, dass beim Hörversuch an dieser Position kein konstanter Frequenzgang zu erwarten ist.



(a) Abhängigkeit von der Rotation der Lautsprecheranordnung (LS-Rotation)

(b) Abhängigkeit von minimalen Veränderungen der Hörposition (Positionsoffset Δy)

Abbildung 10: Abhängigkeit des Frequenzgangs von der Rotation der Lautsprecheranordnung und von Veränderungen der Hörposition bei HOA mit 14 Lautsprechern an der seitlichen Hörposition

2.6.4 Zeitliche Spektren WFS

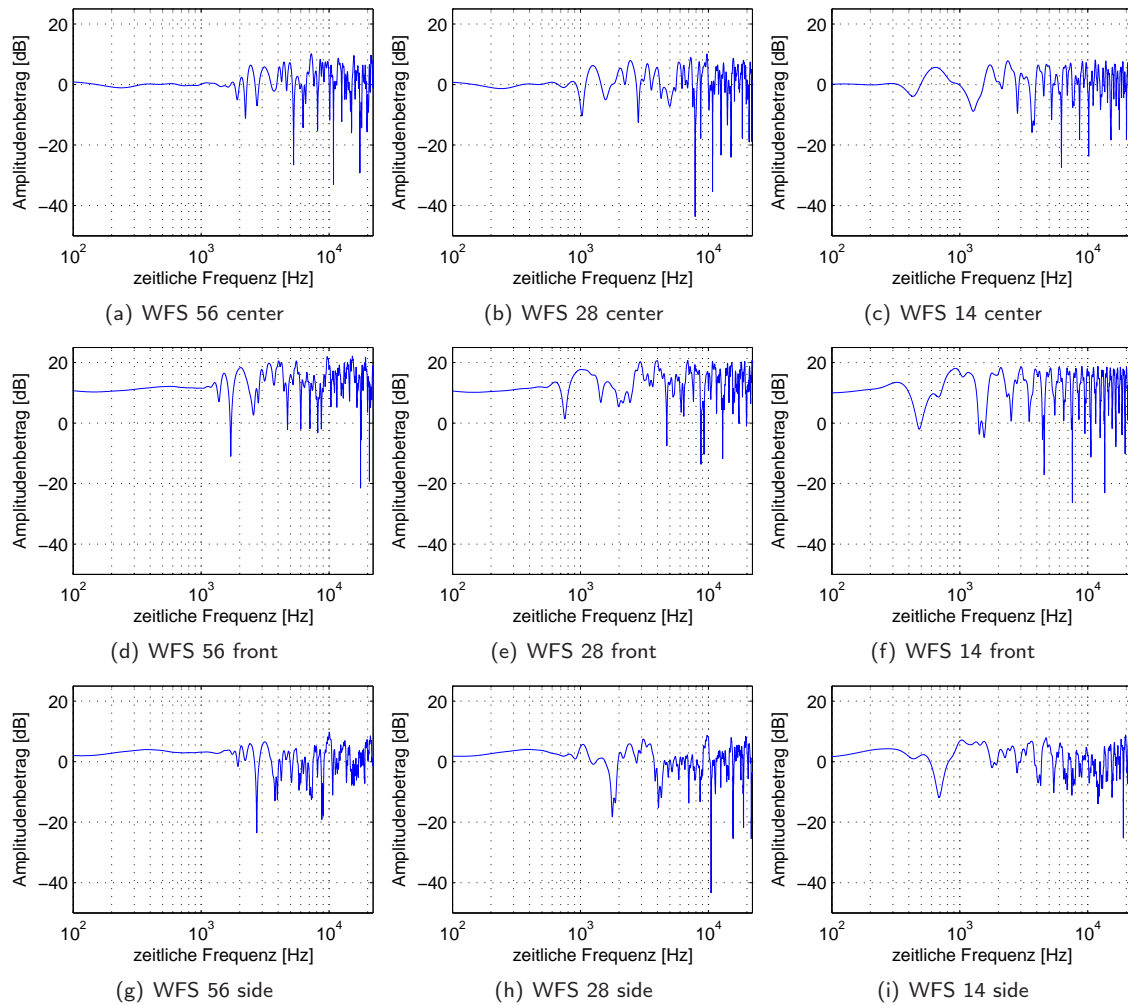


Abbildung 11: WFS: zeitliche Spektren an verschiedenen Orten für verschiedenen Lautsprecheranzahlen (für den Fall einer virtuellen ebenen Welle aus 0° Richtung)

Bei WFS sind die Betragsfrequenzgänge für alle Hörpositionen auf den ersten Blick sehr ähnlich. Die Abweichung vom konstanten Frequenzgang beginnt umso früher, je weniger Lautsprecher verwendet werden.

2.6.5 Lokale Variation der zeitlichen Spektren [AS09]

Ein wichtiger Punkt für die Wahrnehmung von Klangfarbenänderungen ist die Variation des Spektrum bei kleinen Bewegungen, wie z.B. Kopfdrehungen. In Abbildung 12 wurde die Variation der zeitlichen Spektren für die drei Hörpositionen simuliert.

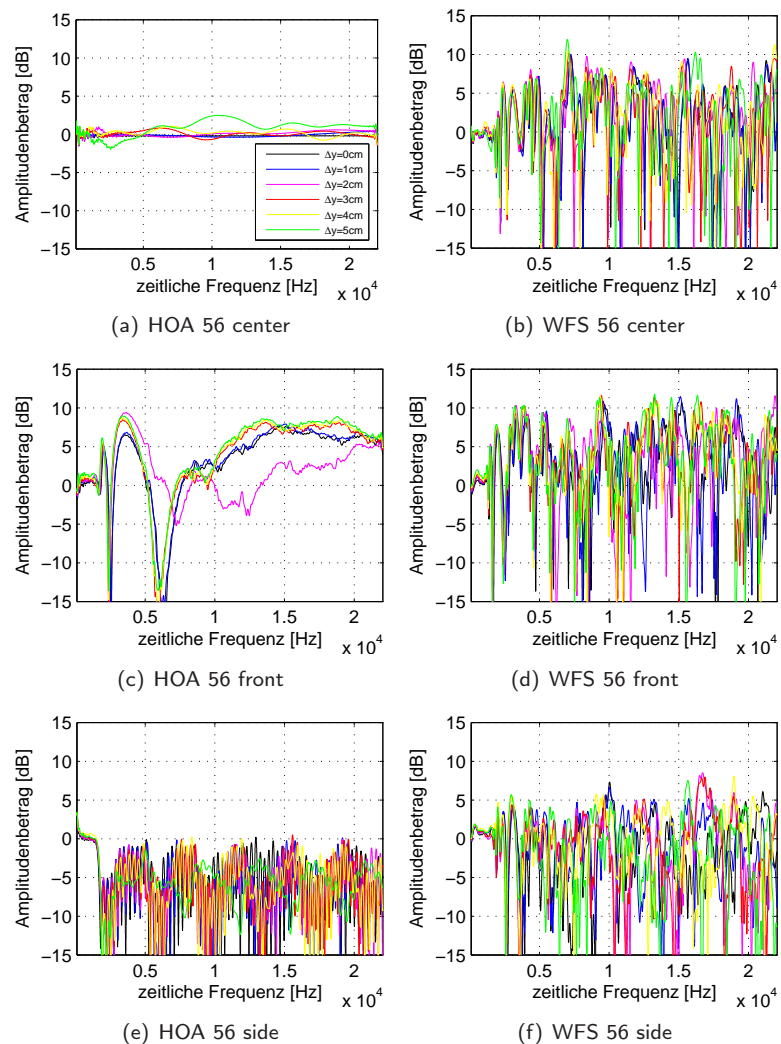


Abbildung 12: Variation der zeitliche Spektren (Verschiebung Δy von 0cm bis 5cm in 1cm Schritten) an den 3 Hörpositionen (für den Fall einer virtuellen ebenen Welle aus 0° Richtung)

Bei HOA erkennt man eine geringere Variation der Spektren als bei WFS, obwohl sich die Grundform der Spektren an den drei Positionen bei HOA deutlich unterscheidet. Bei WFS ist das genaue Gegenteil zu beobachten: Die einzelnen Hörpositionen unterscheiden sich auf den ersten Blick kaum, aber bereits bei kleinsten Variationen der Position kommt es zu Änderungen in der kammfilterartigen Struktur. Aufgrund dieser Simulationen sind bei HOA deutliche klangliche Unterschiede zwischen den Hörpositionen zu erwarten, aber keine gravierenden Änderungen der Klangfarbe bei Drehung des Kopfes. Dagegen wird bei WFS ein ähnlicher Klang an allen Hörpositionen erwartet, bei dem es aber bei Kopfdrehungen zu Änderungen in der Klangfarbe kommt.

3 Statistische Auswertungsverfahren für den Hörversuch

In diesem Kapitel werden die bei der Auswertung verwendeten statistischen Verfahren kurz beschrieben und anhand von einfachen Beispielen erklärt.

3.1 MATLAB: ANOVA1

Zum Vergleich verschiedener Messreihen wird die MATLAB-Funktion ANOVA1 verwendet. Sie liefert zwei Teilergebnisse: eine Varianzanalyse (engl. Analysis of Variance ANOVA) und eine graphische Darstellung von Median und IQR mittels Boxplots. Im Folgenden werden die beiden Berechnungsarten theoretisch und anschließend anhand eines Beispiels erklärt.

3.1.1 Varianzanalyse [Son07, Mat09]

Mit der Varianzanalyse oder ANOVA wird untersucht, wie sich die Varianzen zwischen verschiedenen Gruppen von Messwerten im Vergleich zu den Varianzen innerhalb der einzelnen Gruppen verhalten. Dadurch kann festgestellt werden, ob sich die Mittelwerte der Gruppen signifikant voneinander unterscheiden und die gewählte Gruppeneinteilung sinnvoll ist. Voraussetzung für die ANOVA ist eine Normalverteilung der Messwerte. Als Hypothese H_0 wird behauptet, dass keine Unterschiede zwischen den Gruppen festzustellen sind, also alle Messwerte einer einzigen Gruppe zugeordnet werden können. Am Ende der Berechnung erhält man den Wert $F_{observed}$, der eine Aussage liefert, mit welcher Signifikanz diese Hypothese akzeptiert oder abgelehnt wird. Die Berechnung erfolgt in mehreren Schritten:

Schritt 1:

Zuerst werden die quadratischen Abweichungen (Squared Sums) der Mittelwerte \bar{Y}_t der einzelnen Gruppen vom Mittelwert \bar{Y} über alle Gruppen berechnet und mit der Anzahl n_t an Elementen in der jeweiligen Gruppe multipliziert. Die erhaltenen Werte werden über alle Gruppen aufaddiert.

$$SS_G = \sum_{t=1}^k n_t \cdot (\bar{Y}_t - \bar{Y})^2 \quad (40)$$

Schritt 2:

Dann werden innerhalb der einzelnen Gruppen die quadratischen Abweichungen jedes Messwertes $Y_{k,i}$ zum Gruppenmittelwert \bar{Y}_k berechnet und aufsummiert. Wieder werden die Werte für alle Gruppen aufaddiert.

$$SS_{R,k} = \sum_{i=1}^{n_k} (Y_{k,i} - \bar{Y}_k)^2 \quad (41)$$

$$SS_R = \sum_{k=1}^k SS_{R,k} \quad (42)$$

Schritt 3:

Nun werden die beiden Abweichungen SS_G und SS_R durch die Anzahl an Freiheitsgraden dividiert. Die Anzahl an Freiheitsgraden für die Abweichungen zwischen den Gruppen SS_G berechnet sich durch die

Gruppenanzahl k als $df1 = k - 1$. Für N Messwerte insgesamt (in allen Gruppen zusammen) ergibt sich die Anzahl der Freiheitsgrade für die Abweichungen innerhalb der Gruppen zu $df2 = N - k$.

$$MS_G = \frac{SS_G}{df1} \quad (43)$$

$$MS_R = \frac{SS_R}{df2} \quad (44)$$

Schritt 4:

Teilt man MS_G durch MS_R erhält man den Wert $F_{observed}$. In Abhängigkeit von den Freiheitsgraden $df1$ und $df2$ lässt sich nun ein Wert F_α für ein gewünschtes Signifikanzniveau α aus einer F-Verteilungstabelle ablesen.

Ist der berechnete Wert $F_{observed} \geq F_\alpha(df1, df2)$, wird die Hypothese abgelehnt. Ist $F_{observed} < F_\alpha(df1, df2)$, so wird sie akzeptiert.

MatLab liefert in der Funktion ANOVA1 direkt das Signifikanzniveau für die Akzeptanz von H_0 und erspart so das Ablesen aus Tabellen.

3.1.2 Boxplot [Mat09]

Ein Boxplot ist eine graphische Darstellungsmöglichkeit für die Verteilung von Messwerten. Dabei wird nicht der Mittelwert, sondern der Medianwert als zentraler Zahlenwert benutzt. Den Median einer Folge der Länge N erhält man, indem man alle Elemente der Größe nach aufsteigend ordnet und das Element in der Mitte auswählt. Bei einer geraden Anzahl N an Zahlen, berechnet sich der Medianwert als arithmetischer Mittelwert der beiden mittleren Zahlen. Wie man in Abschnitt 3.1.3 erkennen kann, ist die Berechnung des Medianwertes sehr tolerant gegenüber Ausreißern. In MATLAB ist der Medianwert als waagerechte rote Linie im Boxplot gekennzeichnet.

Ein Maß für die Abweichung der Messwerte vom Medianwert, stellt der Interquartilsabstand (interquartile range, IQR) dar. Diesen Bereich erhält man, wenn man in der aufsteigend geordneten Zahlenfolge $N/4$ vom Medianwert nach unten (unteres Quartil) und oben (oberes Quartil) geht. Zwischen diesen beiden Werten liegen die Hälfte aller Messwerte. Der IQR ist im Boxplot durch waagerechte blaue Linien unterhalb und oberhalb des Medianwertes gekennzeichnet.

Die Einkerbungen (Notches) innerhalb des IQR zeigen denjenigen Bereich an, in dem der wahre Medianwert mit einer Wahrscheinlichkeit von 95% liegt. Vergleicht man die Boxplots zweier Gruppen von Messwerten, ist die Überlappung dieser Bereiche von großer Bedeutung: Überlappen sie nicht, so kann mit einer Wahrscheinlichkeit von mindestens 95% behauptet werden, dass auch die wahren Medianwerte der beiden Gruppen nicht gleich sind.

Die schwarzen Ausleger (Whiskers) außerhalb des IQR kennzeichnen einen Bereich, in dem noch Messwerte liegen. Die maximale Größe dieses Bereichs ist standardmäßig auf $1.5 \cdot \text{IQR}$ festgelegt. Werte außerhalb werden als Ausreißer gedeutet und durch rote Kreuze gekennzeichnet.

3.1.3 Beispiel

Zur Verdeutlichung der beiden obengenannten Berechnungsarten, wird hier ein kleines Beispiel durchgerechnet.

Vergleich: Unfallrisiko Sommer/Winter

Es wurde für 12 Jahre jeweils die Anzahl an schweren Verkehrsunfällen für eine Gemeinde im Juli (stellvertretend für den Sommer) und Dezember (stellvertretend für den Winter) desselben Jahres festgehalten. Folgende Werte ergaben sich:

Sommer: $x_S = [3 \ 4 \ 2 \ 3 \ 1 \ 3 \ 4 \ 2 \ 5 \ 3 \ 10 \ 2]$

Winter: $x_W = [6 \ 8 \ 5 \ 7 \ 0 \ 6 \ 4 \ 6 \ 8 \ 5 \ 7 \ 6]$

Die Hypothese H_0 lautet, dass die Anzahl an Verkehrsunfällen unabhängig von der Jahreszeit ist.

Varianzanalyse

$$SS_G = \sum_{t=1}^k n_t \cdot (\bar{Y}_t - \bar{Y})^2 = 12 \cdot (3.5 - 4.5833)^2 + 12 \cdot (5.6667 - 4.5833)^2 = 28.1675$$

$$SS_R = \sum_{i=1}^k SS_{R,i} = \sum_{i=1}^k \sum_{j=1}^{n_k} (Y_{k,i} - \bar{Y}_k)^2 = 59 + 50.6667 = 109.6667$$

$$MS_G = \frac{SS_G}{k-1} = \frac{SS_G}{2-1} = 28.1675$$

$$MS_R = \frac{SS_R}{N-k} = \frac{SS_R}{24-2} = 4.9849$$

$$F_{observed} = \frac{MS_G}{MS_R} = 5.6506$$

Für ein α von 5% (Verteilungstabellen gibt es meistens nur für Signifikanzniveaus von 10%, 5%, 2.5% und 1%) liegt $F_{0.05}(1, 22)$ bei 4.30. Da $F_{observed} = 5.6506 \geq 4.30$, wird die Hypothese mit einem Signifikanzniveau von $1 - \alpha = 95\%$ abgelehnt. Die Berechnung von MatLab zeigt die gleichen Zwischenergebnisse und liefert sogar ein α von 2.66% (siehe Abbildung 13). Dieses Ergebnis bedeutet, dass die Anzahl der Verkehrsunfälle mit einem Signifikanzniveau von 97.34% von der Jahreszeit abhängen.

ANOVA Table					
Source	SS	df	MS	F	Prob>F
Columns	28.167	1	28.1667	5.65	0.0266
Error	109.667	22	4.9848		
Total	137.833	23			

Abbildung 13: Ergebnisse der Varianzanalyse mit MatLab

Boxplot

Zur Bestimmung des Medianwertes und des IQR werden die Messwerte innerhalb der zwei Gruppen der Größe nach aufsteigend geordnet. Es ergeben sich folgende Werte:

Sommer: $x_S = [1\ 2\ 2\ 2\ 3\ 3\ 3\ 3\ 4\ 4\ 5\ 10]$ \Rightarrow Median: 3, IQR: $4 - 2 = 2$

Der Wert 10 ist außerhalb der Whiskers ($1.5 \cdot \text{IQR}$ oberhalb des oberen Quartils oder unterhalb des unteren Quartils $\rightarrow -1 \leq x \leq 7$) und wird als Ausreißer interpretiert.

Winter: $x_W = [0\ 4\ 5\ 5\ 6\ 6\ 6\ 6\ 7\ 7\ 8\ 8]$ \Rightarrow Median: 6, IQR: $7 - 5 = 2$

Hier ist der Wert 0 außerhalb der Whiskers ($\rightarrow 2 \leq x \leq 10$) und wird auch als Ausreißer interpretiert.

Da sich die Quartilsbereiche um die beiden Medianwerte für Sommer und Winter schon nicht überlappen, muss dies auch für die eingekerbten Bereiche (Notches) gelten. Deshalb kann man mit einer Wahrscheinlichkeit von mindestens 95% sagen, dass wahren Medianwerte nicht gleich sind, d.h. es gibt einen signifikanten Unterschied zwischen der Anzahl an schweren Verkehrsunfällen im Sommer und im Winter.

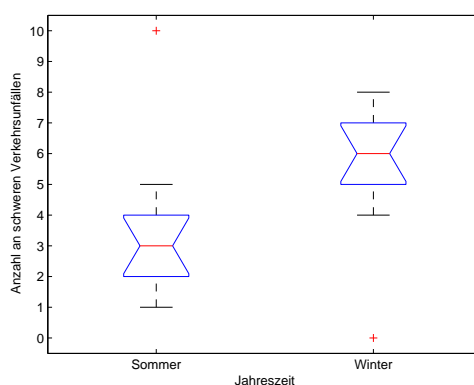


Abbildung 14: Ergebnisse des Boxplots

Obwohl Varianzanalyse und Boxplots verschiedene Berechnungen durchführen, kommt man mit beiden Methoden zu dem Ergebnis, dass Unfallrisiken jahreszeitabhängig sind.

3.1.4 Signifikanzmatrix

Die Signifikanzmatrix wurde als Erweiterung der Darstellung von Signifikanzniveaus für den Vergleich von mehreren Gruppen in Form eines vollständigen Paarvergleichs entwickelt. Dabei werden nur die interessanten Wertebereich farblich dargestellt: rot für Signifikanzniveaus kleiner als 5% (signifikante Unterschiedlichkeit) und grün für für Signifikanzniveaus größer als 95%. Alle anderen Signifikanzniveaus bleiben weiß. Die sich ergebende obere Dreiecksmatrix wird auf die ganze Matrix erweitert. Vergleiche, die nicht durchgeführt wurden, sind schwarz gekennzeichnet. Um die Kontraste zwischen signifikanten Unterschieden und Gleichheiten zu verstärken, wurden vor der Berechnung der Signifikanzniveaus die Ausreißer entfernt.

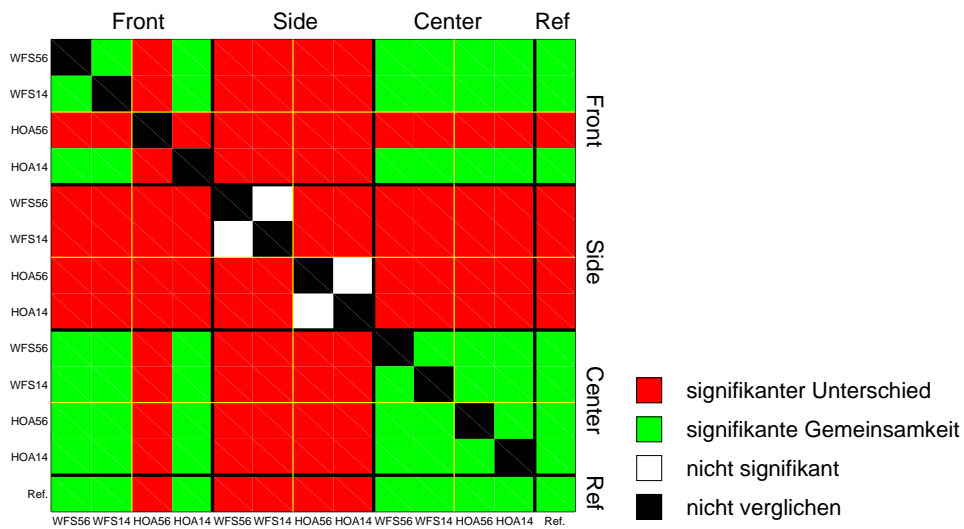


Abbildung 15: Beispiel für eine Signifikanzmatrix

Die einzelnen Zeilen setzen sich aus den Signifikanzniveaus aller Vergleiche mit dem links beschrifteten Verfahren an der rechts beschrifteten Hörposition zusammen. Analog dazu setzen sich die Spalten aus den Signifikanzniveaus aller Vergleiche mit dem untenstehenden Verfahren an der obenstehenden Hörposition zusammen.

3.2 Principle Component Analysis

3.2.1 Einführung in die PCA [Mat09, SPS]

Die PCA wird verwendet, um eine Dimensionsreduktion der Messdaten zu erreichen. Vergleicht man, wie hier im Hörversuch, Stimuli anhand mehrerer Parameter, ist es schwer, eine übersichtliche Darstellungsart all dieser Parameter zu finden. Die PCA versucht die Zusammenhänge zwischen den Parametern zu finden und die Parameter durch sogenannte Hauptkomponenten zusammenzufassen.

Im ersten Schritt müssen die Daten vorbereitet werden: Die Bewertungen jedes Parameters werden zunächst um den Mittelwert zentriert, d.h. es wird der Mittelwert berechnet und von den ursprünglichen Daten subtrahiert. Die neuen Daten mit einem Mittelwert von 0 werden nun durch die Standardabweichung dividiert. Nun wird die Korrelationsmatrix der normierten Parameter erstellt und ihre Eigenwerte und dazugehörigen Eigenvektoren berechnet. Der Eigenvektor des größten Eigenwerts ist die sog. erste Hauptkomponente (Principle Component). Der Betrag des Eigenwerts gibt dabei an, wieviele der Parameter (Dimensionen) durch diese einzelne Hauptkomponente dargestellt werden können. Teilt man diesen Wert durch die Gesamtanzahl aller Parameter, erhält man ein Maß dafür, wieviel Prozent der Varianz der Daten durch diese Komponente bereits erklärt werden. Verwendet man alle Hauptkomponenten, so sind 100% der ursprünglichen Varianz erklärbar. Allerdings macht das keinen Sinn, weil man dann genausoviele Parameter wie zu Beginn erhält. Um zu entscheiden, wie viele Hauptkomponenten verwendet werden, gibt es mehrere Möglichkeiten. Man erhöht die Anzahl bis:

- eine festgesetzte Dimension erreicht ist (aus Darstellungsgründen werden meist zwei Dimensionen verwendet) oder
- ein bestimmter Anteil der Varianz dargestellt werden kann (66-80%) oder
- der nächste Eigenwert kleiner als 1 wird und die dazugehörige Hauptkomponente damit weniger als einen der alten Parameter darstellen kann.

Durch eine Rotation kann der Varianzanteil der ausgewählten Hauptkomponenten angeglichen werden. Die Anwendung der PCA erfolgt in der Software PASW Statistics von SPSS, die anschließende Darstellung der Ergebnisse in MATLAB.

3.2.2 Beispiel zur PCA

Anhand eines einfachen Beispiels wird nun die Vorgangsweise bei der PCA erläutert. Dazu wurden 16 Fahrzeuge anhand von drei Parametern untereinander verglichen: Treibstoffverbrauch, Anzahl der Sitzplätze und Neupreis. Die Fahrzeuge stammen aus vier verschiedenen Fahrzeugklassen: Mittelklasse, Van, Cabrio und SUV. Je Klasse wurden vier Beispielfahrzeuge ausgewertet.

Die Normalisierung wird von der Software automatisch vorgenommen. Es ergibt sich folgende Korrelationsmatrix (siehe 1) für die drei Parameter.

Man kann bereits den Zusammenhang zwischen Preis und Verbrauch erkennen. Das lässt vermuten, dass die beiden Parameter gut zusammengefasst werden können.

In Tabelle 2 erkennt man, dass durch die erste Hauptkomponente bereits mehr als zwei der drei Parameter dargestellt werden können. Bei den beiden verbleibenden Hauptkomponenten sieht man den

	Verbrauch	Sitzplätze	Preis
Verbrauch	1.000	-0.461	0.810
Sitzplätze	-0.461	1.000	-0.325
Preis	0.810	-0.325	1.000

Tabelle 1: Korrelationsmatrix vor der Zusammenfassung

Komponente	vor der Rotation			nach der Rotation		
	Eigenwert	% der Varianz	% Gesamt	Eigenwert	% der Varianz	% Gesamt
1	2.094	69.794	69.794	1.764	58.786	58.786
2	0.731	24.377	94.171	1.062	35.384	94.171
3	0.175	5.829	100.000			

Tabelle 2: Erklärung der Varianz nach der Zusammenfassung

Rückgang der Eigenwerte und der geringere Anteil an der Varianz. Es wurde sich für eine PCA mit insgesamt zwei Hauptkomponenten entschieden. Damit sind 94.171% der ursprünglichen Varianz darstellbar. Durch Rotation (rechte Hälfte von Tabelle 2) werden die Varianzanteile gleichmäßiger auf die beiden Hauptkomponenten verteilt.

	PC 1	PC 2
Verbrauch	0.902	-0.305
Sitzplätze	-0.202	0.977
Preis	0.954	-0.117

Tabelle 3: Komponentenmatrix

Tabelle 3 zeigt die Darstellung der drei Parameter durch die beiden Hauptkomponenten. Auf diese Weise können nun die Parameter der einzelnen untersuchten Fahrzeuge in die Darstellung durch die Hauptkomponenten umgerechnet werden.

Die beiden Hauptkomponenten wurden anhand der Parameter folgendermaßen benannt:

- PC 1: hoher Preis, hoher Verbrauch
- PC 2: viele Sitzplätze.

In Abbildung 16 ist das Ergebnis der PCA dargestellt. Die blauen Vektoren zeigen die drei ursprünglichen Parameter im neuen Vektorraum der Hauptkomponenten. Pro Fahrzeugklasse wurde die beiden Mittelwerte (pro Hauptkomponente) und die dazugehörigen 95% Konfidenzintervalle. Die Darstellung erfolgt durch eine Ellipse um den Mittelwert mit den Konfidenzintervallen als große bzw. kleine Halbachsen. Die Farben sind den einzelnen Fahrzeugklassen zu geordnet:

- Mittelklasse
- Van
- Cabrio
- SUV.

Man erkennt deutlich, dass ein hoher Verbrauch mit einem hohen Preis einhergeht und die Anzahl an Sitzplätzen unabhängig davon ist. In der Darstellung sieht es so aus, als ob die SUVs mehr Sitzplätze als die Mittelklasse bieten würden, dabei wurde bei beiden von der gleichen Anzahl ausgegangen. Das liegt daran, dass die zweite Hauptkomponente nicht genau parallel zum Parameter Sitzplätze verläuft. Dennoch ist das Ziel der PCA, die übersichtliche Darstellung und Reduktion von mehrdimensionalen Vergleichen, erreicht. Die vier Fahrzeugklassen sind auf einen Blick deutlich erkennbar und unterscheidbar.

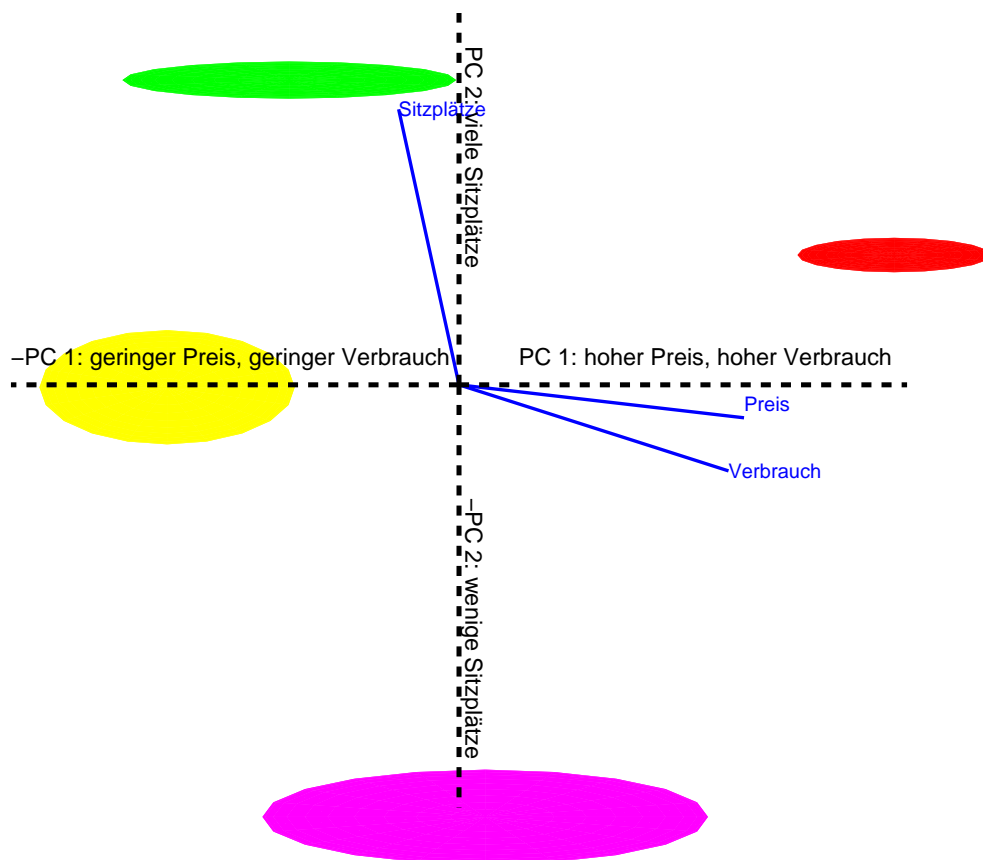


Abbildung 16: zweidimensionale PCA

4 Entwurf und Durchführung des Hörversuchs

Ziel des Versuchs ist der Vergleich von WFS und HOA mit verschiedenen Lautsprecheranzahlen und an verschiedenen Hörpositionen. Als Wiedergabesystem stand ein kreisförmiges Array mit 56 Lautsprechern (ELAC 301) und einem Radius von 1,5m zur Verfügung. Dieses System befindet sich im Raum "Pinta" der Deutschen Telekom Laboratories in Berlin. Da es fast unmöglich ist, alle Probanden an exakt die selbe Hörposition zu setzen und diese auch im Verlauf des Versuchs zu wechseln, wird eine dynamisch nachgeführte binaurale Simulation des Wiedergabesystems verwendet [GAS09].

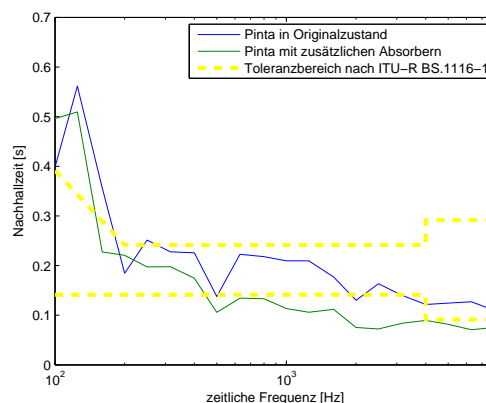


Abbildung 17: Nachhallzeit in Pinta im Vergleich zur ITU-Empfehlung

Der Raum mit seinen Dimensionen von ca. 4m x 4m x 2,8m ist laut ITU-R BS.1116-1 [ITU97] für Mehrkanalwiedergabe zu klein (dort sind mindestens 30 m² empfohlen). Bezogen auf die Nachhallzeit werden die Empfehlungen aber beinahe eingehalten (siehe Abbildung 17). Da der Einfluss des Raumes soweit wie möglich reduziert werden sollte, wurden zusätzliche Absorber angebracht um die Nachhallzeit zu verkürzen.

4.1 Generierung der Stimuli

Im ersten Schritt werden in MATLAB die Ansteuerungsfunktionen für verschiedene Szenarien (WFS/HOA, verschiedene Lautsprecheranzahlen) berechnet. Wenn man Audiomaterial mit den entsprechenden Ansteuerungsfunktionen (driving functions) faltet, erhält man die Signale für die einzelnen Lautsprecher. Um die Übertragungsfunktion zwischen den einzelnen Lautsprechern im Raum zu den Ohren der Versuchspersonen zu berücksichtigen, wird für jede Kopfposition und -orientierung die binaurale Raumpulsantwort (BRIR) des Wiedergabesystems benötigt.

4.1.1 Messung der BRIRs des Lautsprecherarrays

Der Autor konnte auf bereits bestehende Messdaten zurückgreifen. Der Messvorgang wird hier zum Verständnis trotzdem kurz erklärt. Zur Messung der BRIRs wurde das Messsystem "FABIAN" (Fast and Automatic Binaural Impulse response Acquisition) der TU-Berlin verwendet [LW06, LW07]. Der Messaufbau mit Lautsprecherarray und Messroboter ist in Abbildung 18 zu sehen.

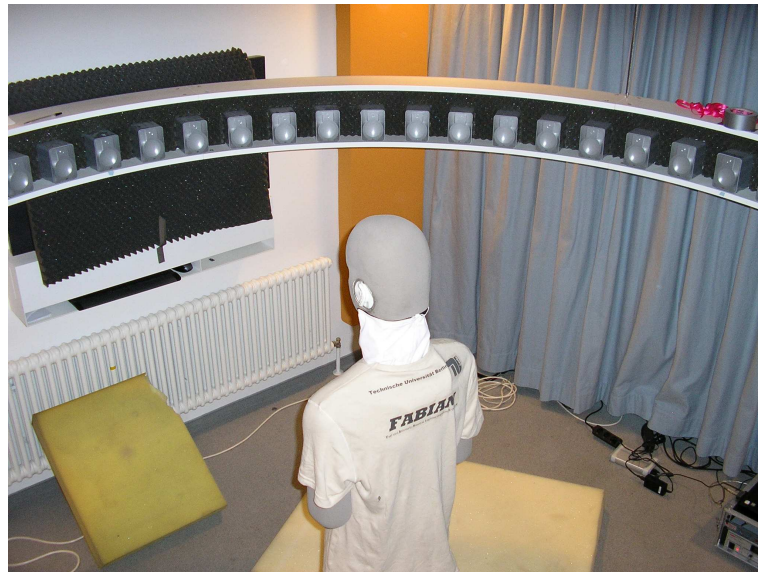


Abbildung 18: FABIAN Messsystem der TU Berlin

Für jede Hörposition wurden die binauralen Impulsantworten aller Lautsprecher nacheinander mit exponentiellen Sweeps gemessen. Entscheidend war dabei, dass diese Messungen für alle Kopfdrehungen von $\pm 80^\circ$ in 1° Schritten durchgeführt wurden, um später im Versuch die Abhängigkeit des Kopfhörersignals von der Kopfdrehung der ProbandInnen zu realisieren.

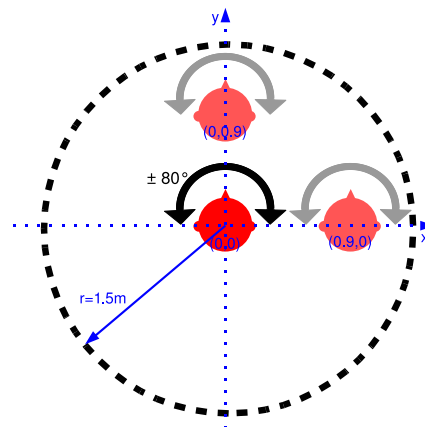


Abbildung 19: Messpositionen

Es wurden insgesamt drei Hörpositionen vermessen (siehe Abbildung 19): die zentrale Hörposition (Kopfzentrum im Ursprung), an der Seite rechts vom Ursprung (Kopfzentrum bei $(0,9\ 0)$) und vor der zentralen Hörposition (Kopfzentrum bei $(0\ 0,9)$). Die Hörpositionen werden im Folgenden "Center", "Side" bzw. "Front" genannt (vergleiche mit Abschnitt 2.6).

4.1.2 Binaural Room Scanning [GAS09]

Mit Hilfe der gemessenen BRIRs kann nun das Lautsprechersystem binaural für die 3 Hörpositionen und beliebige Kopfdrehungen (zwischen $\pm 80^\circ$) synthetisiert und somit auch WFS und HOA mit verschiedenen Lautsprecheranzahlen über Kopfhörer wiedergegeben werden. In Abbildung 20 ist das Schema der Berechnung dieser Binauralsignale dargestellt.

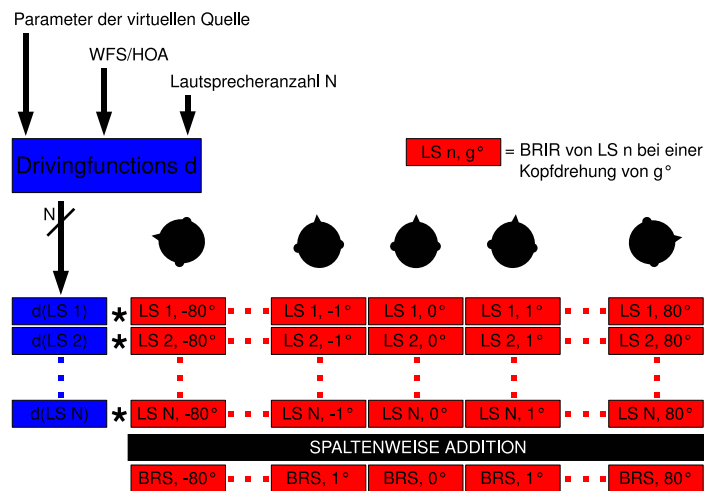


Abbildung 20: Berechnung der Signale für den BRS-Renderer für eine Hörposition

Die berechneten Ansteuerungsfunktionen (driving functions) für die einzelnen Lautsprecher werden mit der jeweiligen BRIR für diesen Lautsprecher für jede Kopfdrehung gefaltet. So erhält man für jede Kopfdrehung das Signal, was von diesem Lautsprecher an den Ohren ankommt. Diese Berechnung wird für alle verwendeten Lautsprecher durchgeführt und für jede Kopfdrehung zu je einer Gesamt-BRIR superponiert.

Zur Wiedergabe wird der SoundScapeRenderer [GAS08b] der Deutsche Telekom Laboratories im BRS (Binaural Room Scanning) Modus verwendet. Er wählt in Abhängigkeit von der Kopfdrehung der Probanden (über ein Head-Tracking-System bestimmt) die BRIR für diese Orientierung aus und faltet sie mit einem gewünschten Audiosignal. Der BRIR-Satz für jede Hörposition wird als 720-spurige WAV-Datei geladen: pro Kopfdrehung von 0° bis 359° zwei Spuren (linkes und rechtes Binauralsignal). Da die vermessenen BRIRs nur für Kopfdrehungen von $\pm 80^\circ$ vorhanden sind, wird für Winkel zwischen 80° und 180° bzw. -180° und -80° jeweils die selbe BRIR beibehalten. Dies ist aber in der Praxis kein Problem, da reine Kopfdrehungen ohne Bewegung des Torso auch keinen größeren Winkelbereich abdecken.

4.2 Versuchsaufbau

Der Versuchsaufbau ist auf zwei Räume aufgeteilt. Im Raum "Pinta", in dem sich auch das vermessene Lautsprechersystem (welches beim Versuch aber kein Audiosignal wiedergibt) befindet, sitzt die Versuchsperson. Diese sitzt innerhalb des Lautsprecherrings (siehe Abbildung 21), um einen passenden optischen Eindruck zum Gehörten zu erhalten. Über eine Maus wird die grafische Oberfläche des Ver-

suchsprogramms bedient. Als Kopfhörer wurde ein Lambda Pro New von Stax, sowie der zugehörige Verstärker SRM monitor verwendet. Der Frequenzgang des Kopfhörers wurde entzerrt. Auf dem Kopfhörer wurde der Empfänger des Head-Tracking-Systems (Polhemus Fastrak) angebracht. Der Sender wurde nahe über dem Kopf mit einem Mikrofonstativ fixiert. Um das elektro-magnetische Feld des Tracking-Systems nicht zu stören, wurde der Galgen des Stativs mit einem Rohr aus Pappe verlängert.



Abbildung 21: Versuchsaufbau im Raum "Pinta"

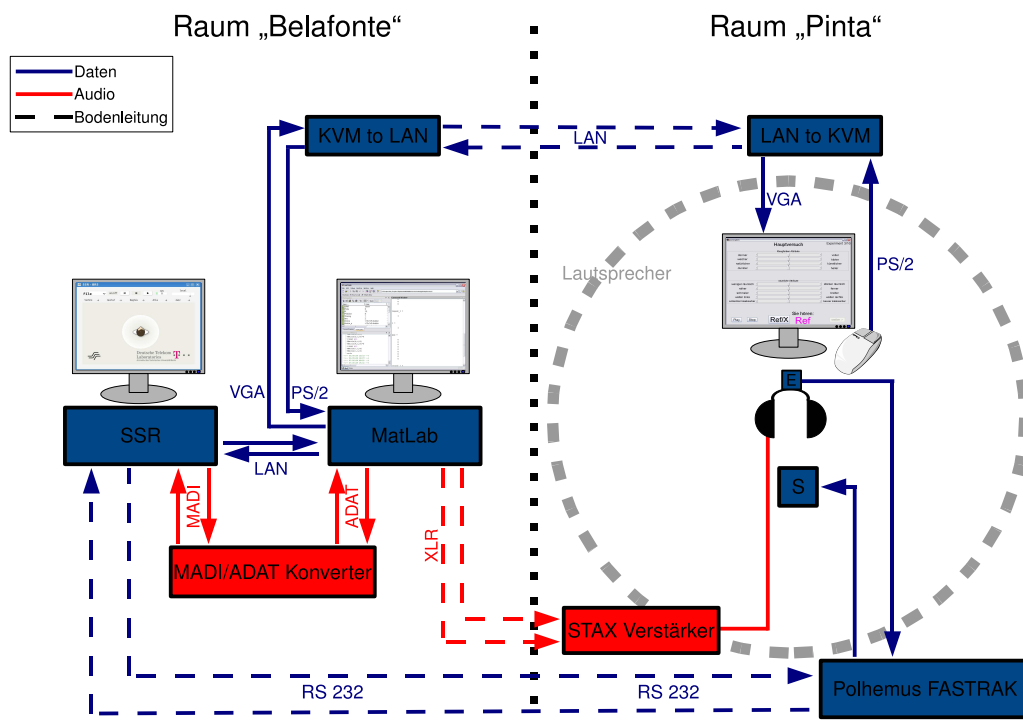


Abbildung 22: schematischer Versuchsaufbau

Zur Vermeidung von Störgeräuschen wurden alle Computer im Nebenraum "Belafonte" betrieben (siehe Abbildung 22). Die Verbindung zwischen beiden Räumen wurde über Leitungen im Boden realisiert (gestrichelte Linien in Abbildung 22). Auf dem Versuchsrechner lief sowohl die gesamte Versuchssteuerung für den Versuchsleiter als auch die grafische Oberfläche des Versuchsprogramms für die ProbandInnen unter MATLAB. Die Video- und Maussignale wurden zwischen beiden Räumen über Netzwerk (DAX-TEN Voyager Micro) übertragen. Die Klangbeispiele wurden aus diesem Rechner über ADAT/MADI (RME Multiface 2, RME ADI-648, RME HDSP MADI) an den Faltungsrechner übertragen. Auf diesem lief der SoundScapeRenderer [GAS08b] und faltete das Klangmaterial mit den entsprechenden BRIRs. Jedes Szenario (Wiedergabeverfahren, Lautsprecheranzahl, Hörposition) wurde dabei in der Software durch eine Quelle dargestellt [GAS08a]. Über Netzwerk bekam der Faltungsrechner vom Versuchsrechner die Befehle zum Umschalten zwischen den Quellen. Die Kopforientierung wurde über eine SR 232 Leitung vom Trackingsystem zum Faltungsrechner übertragen. Das dort entstandene Binauralsignal wurde wieder zurück über MADI/ADAT an das Multiface des Versuchsrechners übertragen, wo es in ein symmetrisches Analogsignal gewandelt und über zwei XLR-Leitungen zum Kopfhörerverstärker weitergeleitet wurde.

4.3 Auswahl der Stimuli/Klangbeispiele

Um die optimale Auswahl an Stimuli zu finden, wurden verschiedene Wiedergabeszenarien angehört und verglichen. Folgende Kriterien wurden für die Auswahl festgelegt:

- Beide Verfahren (WFS und HOA) sollen verglichen werden und zwar
- an allen drei vermessenen Hörpositionen und
- mit unterschiedlichen Lautsprecheranzahlen.

Im Minimalfall kommt man damit also auf 12 verschiedene Varianten. Für die virtuelle Quelle wurde eine ebene Welle aus 0° Richtung festgelegt. Die ebene Welle wurde gewählt, weil die Erforschung der Auswirkungen des räumlichen Aliasing bei WFS für Punktquellen in der Theorie gerade erst begonnen wurde [SA09]. Bei einem Einfall aus der 0° Richtung sind für die vordere Hörposition (Front) besonders ausgeprägte klangliche Artefakte zu erwarten (siehe Abbildungen 8 und 11). Da der Unterschied bei Verwendung von 56 (alle) und 28 Lautsprechern sehr gering wahrgenommen wurde, fiel die Wahl auf 56 und 14 Lautsprecher. Zusätzlich zu diesen 12 Varianten wurde noch eine Referenz erstellt: Bei den Vorversuchen handelte es sich dabei um einen einzelnen Lautsprecher in 0° Richtung für die zentrale Hörposition. Diese Referenz ist also unbeeinflusst von Wiedergabeverfahren und ist nur vom Klang des Lautsprechers selbst und dem Raum abhängig. Da alle anderen Varianten ebene Wellen wiedergeben, ist aber auch eine ebene Welle als Referenz wünschenswert. Deshalb wurde für den finalen Versuchsablauf eine ebene Welle mit HOA (83. Ordnung) und $56 \times 3 = 168$ Lautsprechern für die zentrale Hörposition erstellt. Für diese Position wird das Schallfeld im Bereich des Kopfes bis über 20KHz perfekt rekonstruiert. Da nur 56 und keine 168 Lautsprecher vermessen wurden, mussten die BRIRs für die fehlenden Lautsprecherpositionen aus den vorhandenen interpoliert werden.

Während des Probehörens viel auch auf, dass der vermessene Raum selbst Klangfarbenveränderungen hervorruft (zu sehen an den starken frühen Reflexionen des blauen Graphs in Abbildung 23). Deshalb wurde zusätzlich zu der gesamten BRIR-Länge auch eine gekürzte Version benutzt, die nur den Direkt-schall enthält.

Um bei allen Stimuli eine ähnliche Lautstärke zu erreichen, wurden die BRIR-Sätze alle auf den selben RMS-Pegel für die 0° Kopfausrichtung normalisiert.

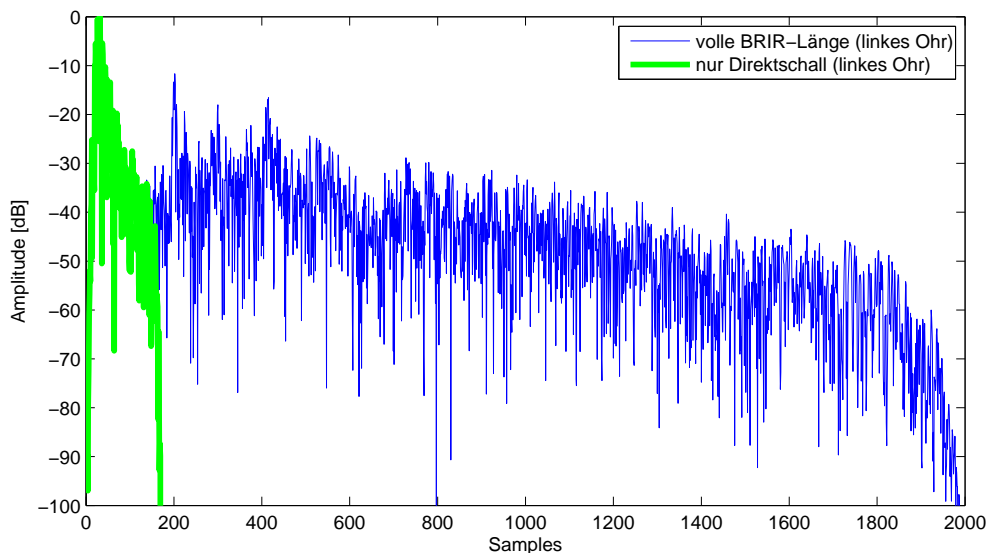


Abbildung 23: verwendete BRIR-Längen im Versuch (Lautsprecher 1, Kopfdrehung 0°, Center)

Damit möglichst belastbare Aussagen getroffen werden können, wurden die zu prozessierenden Klangbeispiele sorgfältig ausgewählt. Beim Probehören wurde zunächst gepulstes rosa Rauschen verwendet, da bei diesem Signal die Unterschiedlichkeit der Wiedergabeverfahren sehr deutlich zu hören war. Es gibt allerdings das Problem, dass bei diesem Signal die Angabe der Stärke der Artefakte und vor allem ihre Benennung mit Attributen sehr schwer fiel, weil das Signal nicht im Höralltag vorkommt. Es wurden deshalb Klangbeispiele gesucht, die Hörerfahrung mit dem analytischen Charakter des Rauschens kombinierten. Die Auswahl fiel auf folgende drei Beispiele:

- Sprache (deutsch, weibliche Stimme, trockene Aufnahme der Telekom)
- Cello ("Violoncello, melodious phrase" von der EBU Tech 3253 SQAM CD [EBU])
- Kastagnetten ("Castanets, single tone, rhythm" von der SQAM CD [EBU], trocken nachprogrammiert mit der ViennaSymphonicLibrary [Vie])

Sprache hat zwar nur einen schmalbandigen Grundton- und Formantbereich, zeichnet sich aber durch Explosivlaute und vor allem durch die größtmögliche Hörerfahrung aus. Beim Cello sind Klangfarbenveränderungen, auch im hochfrequenten Bereich, besonders stark wahrnehmbar. Um das zeitliche Verhalten der Wiedergabeverfahren zu testen, eignen sich besonders die sehr kurzen impulshaltigen Kastagnetten. Alle drei Klangbeispiele wurden auf eine ähnliche Länge von 8-10 Sekunden gekürzt. Der Schallpegel (lineare slow Bewertung) wurde für die Sprache und das Cello auf ca. 73dB eingemessen. Bei die Kastagnetten wurde perceptiv auf eine ähnliche Lautstärke eingepgelt, da aufgrund der transienten Struktur keine brauchbare Messung möglich ist.

Um alle zu untersuchenden Variablen abzudecken und gleichzeitig die Anzahl an Stimuli möglichst gering zu halten, wurde für den Vorversuch und finalen Versuchsablauf folgende Stimuli ausgewählt:

virtuelle Quelle	Hörposition	Verfahren	Ordnung	Lautsprecheranzahl
ebene Welle aus 0°	Front	WFS	∞	56
ebene Welle aus 0°	Front	WFS	∞	14
ebene Welle aus 0°	Front	HOA	27	56
ebene Welle aus 0°	Front	HOA	6	14
ebene Welle aus 0°	Side	WFS	∞	56
ebene Welle aus 0°	Side	WFS	∞	14
ebene Welle aus 0°	Side	HOA	27	56
ebene Welle aus 0°	Side	HOA	6	14
ebene Welle aus 0°	Center	WFS	∞	56
ebene Welle aus 0°	Center	WFS	∞	14
ebene Welle aus 0°	Center	HOA	27	56
ebene Welle aus 0°	Center	HOA	6	14
Referenz				

Tabelle 4: Stimuli pro Klangbeispiel und BRIR-Länge bei allen Versuchen [WFS bzw. HOA mit 14 Lautsprechern werden bei der Darstellung der Ergebnisse mit WFS14 bzw. HOA14 gekennzeichnet, mit 56 Lautsprechern WFS56 bzw. HOA56]

Pro Versuchsperson waren also diese 13 Stimuli mit jeweils zwei verschiedenen BRIR-Längen und je drei verschiedenen Klangbeispielen zu bewerten.

4.4 Vorversuche

Ähnliche Versuche sind in der Literatur nur wenige zu finden [Wit07, Weg05], meistens wurde jedoch das MUSHRA-Verfahren (**M**U**l**tiple **S**timuli with **H**idden **R**eference and **A**nchor) verwendet. Bei diesem Verfahren werden ca. acht Stimuli gleichzeitig mit einer Referenz hinsichtlich ihrer Qualität verglichen. Innerhalb dieser acht Stimuli ist wiederum die Referenz, sowie mehrere Ankersignale versteckt. Die Ankersignale werden so generiert, dass sie verschiedene fixe Abstufungen des Qualitätsverlustes im Vergleich zur Referenz darstellen und dadurch für die Versuchspersonen als Orientierung für die Bewertung dienen. Die Arbeit mit Ankersignalen ist allerdings nur möglich, wenn die Unterschiede bei allen Stimuli auf nur eine Art von Artefakten zurückzuführen sind, die sich nur im Grad der Ausprägung unterscheiden. Beim Probehören fiel bereits deutlich auf, dass die Art der Artefakte bei WFS und HOA unterschiedlich sind. Deshalb ist das MUSHRA-Verfahren beim qualitativen Vergleich von WFS und HOA nicht anwendbar.

4.4.1 Erster Vorversuch

Um die unterschiedliche Art der Artefakte zu erfassen, wurde der erste Vorversuch als vollständiger Paarvergleich mit anschließendem Multidimensional Scaling (MDS) [Hal01] konzipiert. Bei diesem Verfahren werden alle Stimuli paarweise miteinander verglichen und in Form einer Unterschiedlichkeitsmatrix angeschrieben. Diese Punkt-zu-Punkt-Abstände werden dann in einen n-dimensionalen Vektorraum transformiert. Die Anzahl n an Dimensionen bestimmt, wie genau die Abstände in dem Modell des Vektorraums repräsentiert werden können. Als Ergebnis erhält man also eine Art Karte, auf der man erkennen kann, in welchen Dimensionen sich welche Stimuli wieviel voneinander unterscheiden. Zum

Ende dieses Verfahrens müssen den gefundenen Dimensionen verschiedenen Attributen zur Beschreibung zugeordnet werden. Dieser Prozess stellt die eigentliche Herausforderung dieses Verfahrens in der Praxis dar. Die MDS ist ein beliebtes Verfahren beim Vergleich von Sprachkodierungen.

In jedem Paarvergleich werden nur die 13 Stimuli innerhalb einer der sechs Kombinationen aus BRIR-Länge und Klangbeispiel untereinander verglichen. Für den Vergleich von 13 Stimuli ergibt sich somit eine Gesamtzahl von 78 Einzelvergleichen.

Vor dem eigentlichen Versuch fanden mehrere Phasen des Trainings statt. Die erste Phase dient dem Kennenlernen des binauralen Wiedergabesystems (siehe Abbildung 24). Zu diesem Zweck wurde eine Demo der ViennaSymphonicLibrary (Prelude der Holberg Suite von Edvard Grieg, 2:32 [Viel]) mit dem BRIR-Satz der jeweiligen Referenz vorgespielt.



Abbildung 24: Grafische Oberfläche im Vorversuch: Kennenlernen des Wiedergabesystems

Im nächsten Schritt wurde das im Hauptversuch verwendete Klangbeispiel als Referenz und in drei verschiedenen Varianten vorgestellt (siehe Abbildung 25). Die drei Varianten stellten einen groben Überblick über die Art der Artefakte dar: WFS mit 14 Lautsprechern für die vordere Hörposition, WFS mit 56 Lautsprechern für die seitliche Hörposition und HOA mit 14 Lautsprechern für die seitliche Hörposition. Während der Wiedergabe des Klangbeispiels in einer einminütigen Schleife konnte beliebig zwischen den verschiedenen Varianten umgeschaltet werden.



Abbildung 25: Grafische Oberfläche im Vorversuch: Kennenlernen des verwendeten Klangbeispiels

Im Hauptversuch wurden jeweils die einzelnen Stimulipaare dargeboten. Auch hier konnte während der Wiedergabe umgeschaltet werden. Mit Hilfe eines Schiebereglers konnte die Versuchsperson den Unterschied stufenlos von "nicht unterscheidbar" (Wert 0) bis "sehr unterschiedlich" (Wert 100) skalieren (siehe Abbildung 26).

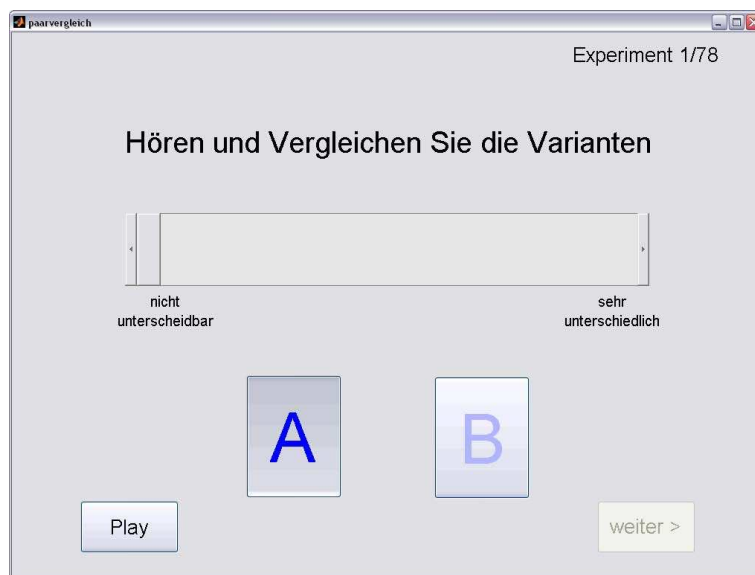


Abbildung 26: Grafische Oberfläche im Vorversuch: Vergleich

Am ersten Versuch nahmen fünf ProbandInnen teil. Alle waren Mitarbeiter der Audioabteilung der T-Labs. Von jeder Versuchsperson wurden alle drei Klangbeispiele mit der kurzen BRIR-Länge und je ein Beispiel mit der vollen BRIR-Länge verglichen.

Das MDS-Verfahren bestimmte sechs Dimensionen, um die einzelnen Unterschiedlichkeit mit geringem Fehler zu rekonstruieren. In Folge war es nicht möglich die dazu passenden Attribute eindeutig zuzuordnen. Dennoch ließ die Auswertung der letzten Zeile bzw. Spalte der Unterschiedlichkeitsmatrix, d.h. die Unterschiede jedes Stimulus zur Referenz schon interpretierbare Tendenzen (hauptsächlich Abhängigkeiten von der Lautsprecheranzahl und der Hörposition) erkennen (siehe Abbildung 27).

Gründe für das nicht zufriedenstellende Ergebnis lassen sich in der individuellen Berechnung der Gesamtunterschiedlichkeit und in dem zu kurzen Training vermuten. Folgende Schlüsse für die weitere Versuchsplanung wurden aus den Ergebnissen des ersten Vorversuchs gezogen:

- Das Training muss durch einen aktiven Teil, in dem die ProbandInnen das Bewerten üben, erweitert werden.
- Es wird kein vollständiger Paarvergleich durchgeführt, sondern nur Vergleiche mit der Referenz.
- Um mehrdimensionale Ergebnisse zu erhalten, wird die Gesamtunterschiedlichkeit in Einzelattribute aufgeteilt.

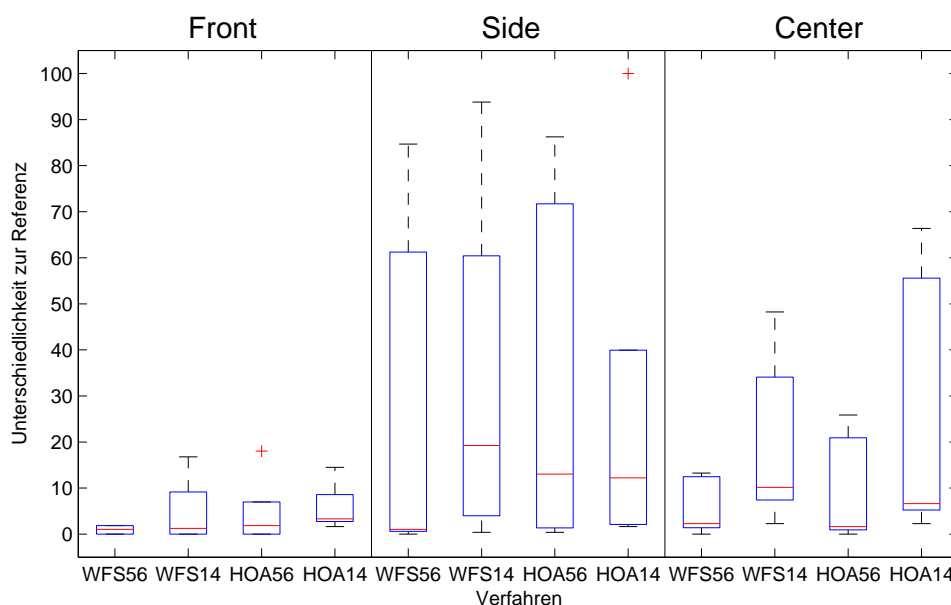


Abbildung 27: erster Vorversuch: Unterschiedlichkeit zur Referenz für Sprache mit kurzen BRIRs (Boxplot)

4.4.2 Weitere Vorversuche

Um das Versuchsdesign zu optimieren, wurden weitere Vorversuche durchgeführt, in denen die Erfahrungen aus dem ersten Vorversuch umgesetzt werden. Als erster Schritt wurden die Attribute der Unterschiedlichkeit festgelegt. Dazu wurde einerseits auf die Literatur [Wit07, BR99, Ber06, ITU03, CW07, CHC⁺07, GK04, Baa07] und andererseits auf die im ersten Vorversuch erlangte Hörerfahrung der ProbandInnen im Sinne der Repertory Grid Technique [BR99, Ber06] zurückgegriffen: Alle ProbandInnen hatten die Möglichkeit sich noch einmal die verschiedenen Klangbeispiele und Stimuli anzuhören. Daraus wurde jeweils eine Liste mit den subjektiv wichtigsten Attribute zur Beschreibung der Unterschiedlichkeit aufgestellt. Die Listen wurden anschließend diskutiert und mit den Attributen aus der Literatur verglichen. Dabei wurden neun Attributpaare festgelegt, vier für die Beschreibung der Unterschiedlichkeit in Bezug auf die Klangfarbe und fünf bezogen auf räumliche Unterschiede:

- dünner - voller
- weicher - härter
- natürlicher - künstlicher
- dunkler - voller
- weniger räumlich - stärker räumlich
- näher - ferner
- schmaler - breiter
- weiter links - weiter rechts
- besser lokalisierbar - schlechter lokalisierbar

Die weiteren Vorversuche wurden nicht zur Gewinnung von Messdaten genutzt, sondern zum Testen des Versuchsablaufs und zur Bestimmung der Zeit pro individuellem Paarvergleich: Da der Vergleich nun anhand von neun Attributen erfolgte (siehe Abbildung 28), erhöhte sich auch die Dauer pro Vergleichspaar auf ca. 1 Minute. Somit könnte innerhalb von 1.5 Stunden für alle sechs Kombinationen aus Klangbeispiel und BRIR-Länge jeweils 15 Paarvergleiche durchgeführt werden.

Das Training wurde durch eine zusätzliche Phase erweitert, in der der Bewertungsvorgang genau wie später im Hauptversuch geübt werden konnte.

4.5 Finaler Versuch



Abbildung 28: Grafische Oberfläche im finalen Versuch: Vergleich

Aus den Ergebnissen der Vorversuche wurde der finale Versuchsablauf festgelegt: Es gab pro ProbandIn sechs Teilversuche, bei denen die Paarvergleiche für die sechs Kombinationen aus Klangbeispiel und BRIR-Länge durchgeführt wurden. Die Teilversuche wurden auf zwei Termine aufgeteilt, um die Beeinflussung der Messdaten durch Ermüdungserscheinungen des Gehörs zu vermeiden. Vor Beginn des ersten Teilversuchs fand zusätzlich eine audiometrische Messung nach DIN EN 60645-1 mit der Software ear 2.0 digital PC-audiometer unter Verwendung einer ECHO IndigoI/O Soundkarte und eines Sennheiser HDA 200 Kopfhörers statt.

Jeder einzelne Teilversuch war folgendermaßen aufgebaut:

- Training zum Kennenlernen des Wiedergabesystems
- Training zum Kennenlernen des Klangbeispiels und der verwendeten BRIR-Länge anhand von 4 Varianten
- Training zum Kennenlernen des Bewertungsvorgangs anhand von fünf Paarvergleichen
- Hauptversuch, bestehend aus 16 Paarvergleichen

Die vier Varianten zum Kennenlernen des Klangbeispiels waren Referenz, WFS mit 14 Lautsprechern an der vorderen Hörposition, WFS mit 56 Lautsprechern an der seitlichen Hörposition und HOA mit 16 Lautsprechern an der seitlichen Hörposition. Für das Bewertungstraining wurden WFS und HOA mit 56 Lautsprechern an der vorderen Hörposition, WFS mit 14 Lautsprechern und HOA mit 56 Lautsprechern an der seitlichen Hörposition, sowie WFS mit 56 Lautsprechern an der zentralen Hörposition. Somit hatte jede Versuchsperson nach Ende des Trainings die meisten Varianten schon einmal gehört und bewerten müssen.

Die 16 Vergleiche im Hauptversuch setzten sich aus den 13 Varianten (Referenz, WFS/HOA mit 14/56 Lautsprechern an den drei Hörpositionen zusammen, siehe auch 4) zusammen. Der Vergleich zwischen Referenz und Referenz dient zur Überprüfung der ProbandInnen. Die restlichen drei Stimuli wurden zufällig aus den 13 ausgewählt. Durch die Wiederholung der Beispiele kann auch die Zuverlässigkeit der ProbandInnen überprüft werden. Alle Reihenfolgen, sowohl innerhalb der sechs Teilversuche, als auch die Reihenfolge der Klangbeispiele und BRIR-Längen wurden zufällig ausgewählt.

Die ProbandInnen erhielten vor Beginn der Versuche ein Handout, auf dem der Ablauf erklärt wurde (siehe Abbildungen 74 und 75 im Anhang). Aufgrund der großen Herausforderung an die ProbandInnen, im Vergleich neun Attribute zu skalieren, wurde der Hauptversuch als Expertenhörtest durchgeführt.

Die Versuche fanden vom 14. bis 29. April im Raum Pinta des Quality & Usability Lab der Deutschen Telekom Laboratories in Berlin statt. Insgesamt nahmen 15 (4 weibliche, 11 männliche) ProbandInnen im Alter zwischen 21 und 33 Jahren (medianes Alter 25 Jahre) teil. 14 der TeilnehmerInnen waren entweder StudentInnen oder AbsolventInnen des Studiengangs "Diplom Tonmeister" an der UdK Berlin. Als weiterer Proband nahm ein PhD Student am Fachgebiet Audiokommunikation der TU Berlin teil.

5 Auswertung und Interpretation des Hörversuchs

5.1 Eingangsvoraussetzungen

Vor Beginn der eigentlichen Auswertung muss die Erfüllung bestimmter Voraussetzungen überprüft werden. Dabei handelt es sich um Voraussetzungen an die ProbandInnen und an die erhaltenen Datensätze.

5.1.1 ProbandInnen

Wie bereits erwähnt wurde das Gehör jeder Versuchsperson audiometrisch vermessen. Als Toleranzschwelle für den Hörverlust wurde ein Wert von 15dB festgelegt. Wenn eine Versuchsperson diese Schwelle mehr als zwei Mal deutlich überschreitet, wird sie von der Teilnahme am Hörversuch ausgeschlossen. Aufgrund dieses Hörtests konnten alle ProbandInnen zum Hörversuch zugelassen werden. Die einzelnen Audiogramme sind im Anhang zu finden. Bei zwei ProbandInnen zeigte sich deutliche Unterschiede zwischen den beiden Ohren. Es konnte aber kein Zusammenhang zur Bewertung dieser ProbandInnen hergestellt werden. Auffällig ist auch der bei tiefen Frequenzen im Mittel etwas geringere Hörverlust des rechten Ohres im Vergleich zum linken. Das gilt vor allem für die tiefste Frequenz. Dieses Phänomen lässt sich vermutlich durch den Ablauf der Audiometrie erklären. Da die Messung auf dem linken Ohr bei der tiefsten Frequenz begonnen wurde, könnte das Ergebnis mit der anfänglich schlechteren Reaktionszeit zusammenhängen.

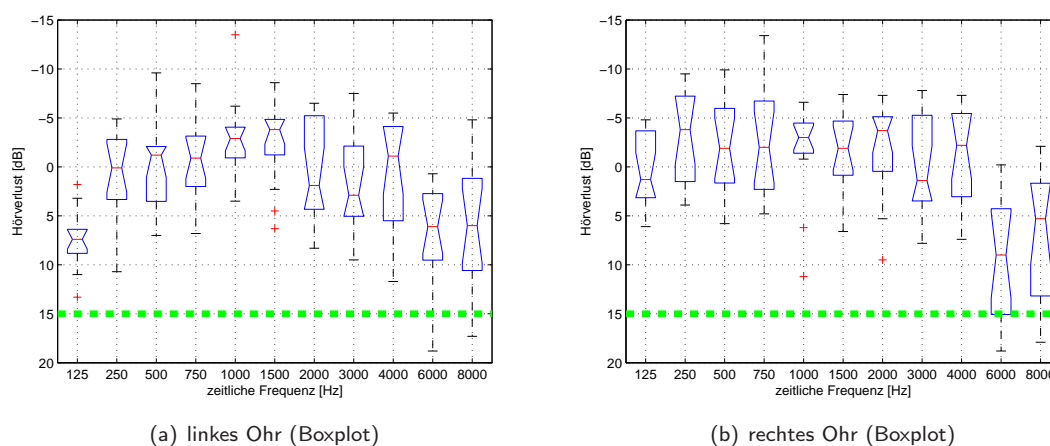


Abbildung 29: Ergebnisse der Audiometrie über alle ProbandInnen (grüne Linie = Toleranzschwelle)

Zusätzlich wurden die ProbandInnen nach ihrer Instrumentalerfahrung befragt: Alle ProbandInnen haben dabei mehrjährige (meistens deutlich mehr als 10 Jahre) Erfahrung auf meistens mehreren Instrumenten angegeben.

Eine Befragung zur Vertrautheit mit verschiedenen Wiedergabeverfahren (Stereo, Surround, WFS und HOA) hinsichtlich Hören, Aufnehmen und Mischen/Mastern wurde auch durchgeführt. Die Ergebnisse all dieser Befragungen sind ebenfalls im Anhang zu finden. Es zeigte sich, dass alle ProbandInnen ihre Erfahrungen im Bereich Stereo als recht gut einschätzen, während die Erfahrung sich bei Surround eher auf das Hören und Mischen/Mastern konzentrierte. Ein WFS-System hatte jede(r) schon einmal gehört. Erfahrungen im Bereich Aufnehmen und Mischen/Mastern gab es aber nicht. Dies gilt auch für HOA.

Dort ist aber der große Unterschied, dass alle Tonmeister(studentInnen) noch nie etwas von diesem System gehört hatten.

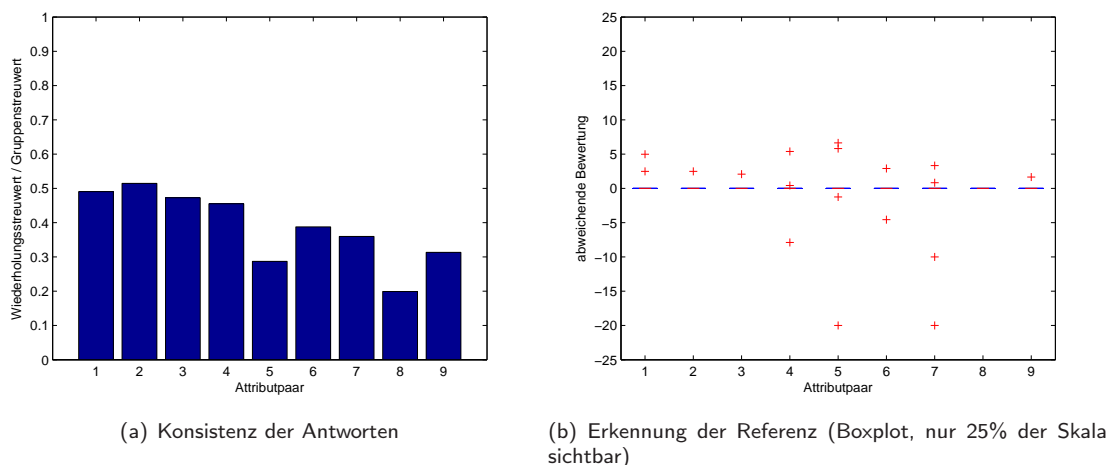


Abbildung 30: Zuverlässigkeit aller ProbandInnen

Aus den 13 Vergleichen in jedem Teilversuch wurden, wie oben beschrieben, drei zufällig ausgewählte wiederholt. Auf Basis dieser Daten kann ein Maß für die Konsistenz der Antworten errechnet werden: Zuerst wurden die Gruppenstreuwerte jeder Versuchsperson für die fünf Gruppen WFS56, WFS14, HOA56, HOA14 und Referenz (über alle Hörpositionen gemittelt) berechnet. Diese Werte zeigen, wie stark die Bewertung der jeweiligen Versuchsperson innerhalb jeder Gruppe streut. Nun kann die Streuung innerhalb der Wiederholungen mit diesem dazu passenden Gruppenstreuwert verglichen werden (Wiederholungsstreuwert/Gruppenstreuwert). Der Wert 1 würde bedeuten, dass die Bewertung des selben Stimulus in der Wiederholung genau soviel streut, wie alle Stimuli innerhalb der Gruppe zusammen. Werte deutlich größer 1 würden dagegen bedeuten, dass der Stimulus in der Wiederholung deutlich anders als vorher bewertet wurde und damit die Gruppenzugehörigkeit von der Versuchsperson nicht richtig erkannt wurde. Wie in Abbildung 30 (a) zu erkennen ist, liegt das Maß im Mittel fast immer unter 0.5. Die Maße für die einzelnen ProbandInnen sind im Anhang zu finden. Sie sind immer kleiner als 1.

Als weiteres Maß für die Zuverlässigkeit der ProbandInnen wurde die Erkennung der Referenz genutzt. Wie im Abbildung 30 (b) zu erkennen ist, wurde die Referenz, außer bei ein paar kleinen Ausreißern, als solche erkannt und bewertet. Die Bewertungen der einzelnen ProbandInnen für die Referenz sind ebenfalls im Anhang zu finden.

Aufgrund dieser Ergebnisse wurden die Messdaten aller ProbandInnen ausnahmslos in die Auswertung aufgenommen.

5.1.2 Normalverteilung

Da bei der Auswertung hauptsächlich mit Varianzanalysen gearbeitet wird, muss die Normalverteilung der Datensätze sichergestellt werden. Dazu wurde der Jarque-Bera-Test verwendet [Mat09]. Dieser Test vergleicht Asymmetrie und Kurtosis der Daten mit denen einer Normalverteilung. Man erhält als Ergebnis entweder einer Widerlegung der Hypothese auf 5% Signifikanzniveau, dass die Daten nicht normalverteilt sind oder man erhält keine Widerlegung dieser Hypothese.

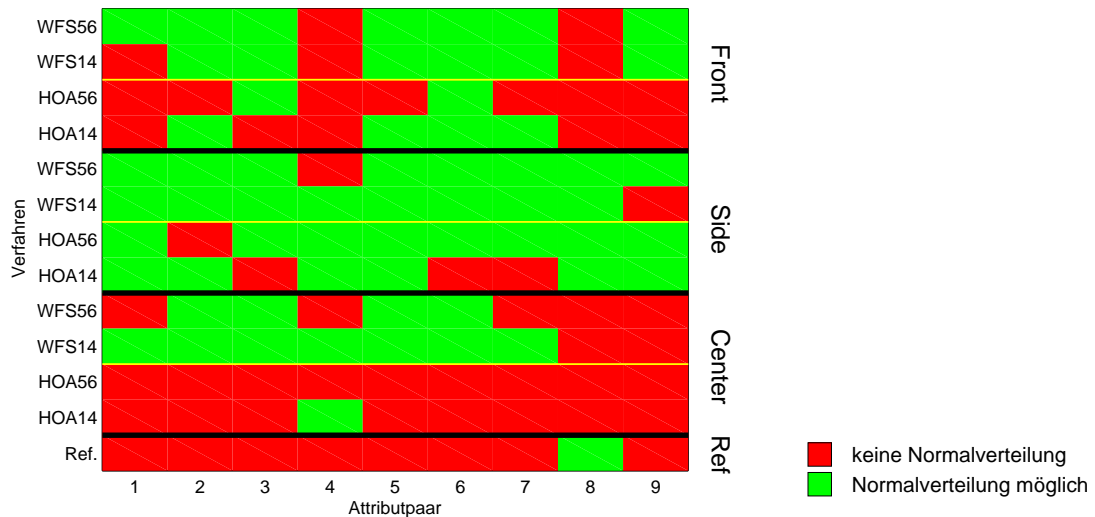


Abbildung 31: Jarque-Bera-Test ($\alpha=5\%$); rot: keine Normalverteilung, grün: Normalverteilung möglich

Der Test zeigt, dass in vielen Fällen eine Normalverteilung möglich ist. Ausnahme bilden vorallem die Bewertungen von HOA56, HOA14 für die zentrale Hörposition und der Referenz bei fast allen Attributpaaren. Das liegt daran, dass die Bewertungen in diesen Fällen nur aus sehr vielen Werten genau bei 0 und einzelnen Ausreißern bestehen. Dadurch können sehr große Werte für die Kurtosis entstehen. Das gilt auch für Attribut 8 (weiter links - weiter rechts), da dort an den Hörpositionen Front und Center ähnlich spitze Verteilungen entstehen.

Die gezeigten, nicht normalverteilten Fälle schränken die Anwendung der Varianzanalyse nicht ein, weil in diesen Fällen ohnehin nur eine sehr geringe Unterscheidbarkeit der Verfahren vorliegt.

5.2 Abhängigkeit von Klangbeispiel und BRIR-Länge

In diesem Abschnitt wird überprüft, ob die Bewertung von der Kombination aus Klangbeispiel und BRIR-Länge abhängt und alle Kombinationen bei der Auswertung getrennt betrachtet werden müssen. Zuerst wurde die Abhängigkeit mit Varianzanalysen überprüft. Dabei zeigten sich einige Gleichheiten (hauptsächlich beim gleichen Klangbeispiel mit unterschiedlicher BRIR-Länge) und auch Unterschiede. Aus diesen Ergebnissen lässt sich jedoch nicht schließen, ob die Stimuli komplett anders bewertet worden sind oder die Reihung der Stimuli nach ihrer Bewertung ähnlich bleibt und sich die verschiedenen Kombinationen nur durch eine Skalierung der Attribute bei der Bewertung unterscheiden. Um das zu überprüfen, wurden die Medianwerte für die einzelnen Stimuli für alle Kombinationen berechnet und quantisiert. Als Quantisierung wurde der Wertebereich von -100 bis 100 in elf Stufen der Breite ± 10 um die Werte ± 100 , ± 80 , ± 60 , ± 40 , ± 20 , 0 eingeteilt. Die Verfahren wurden nun nach der Größe ihrer quantisierten Medianwerte aufsteigend geordnet. Da in diesem Stadium der Auswertung die Zuordnung, welcher Medianwert zu welchem Verfahren gehört, nicht wichtig ist, wurden die Verfahren farblich kodiert. An der Übereinstimmung der Farbverläufe kann man so die Übereinstimmung der Reihungen der Verfahren für die verschiedenen Kombinationen aus Klangbeispiel und BRIR-Länge überprüfen.

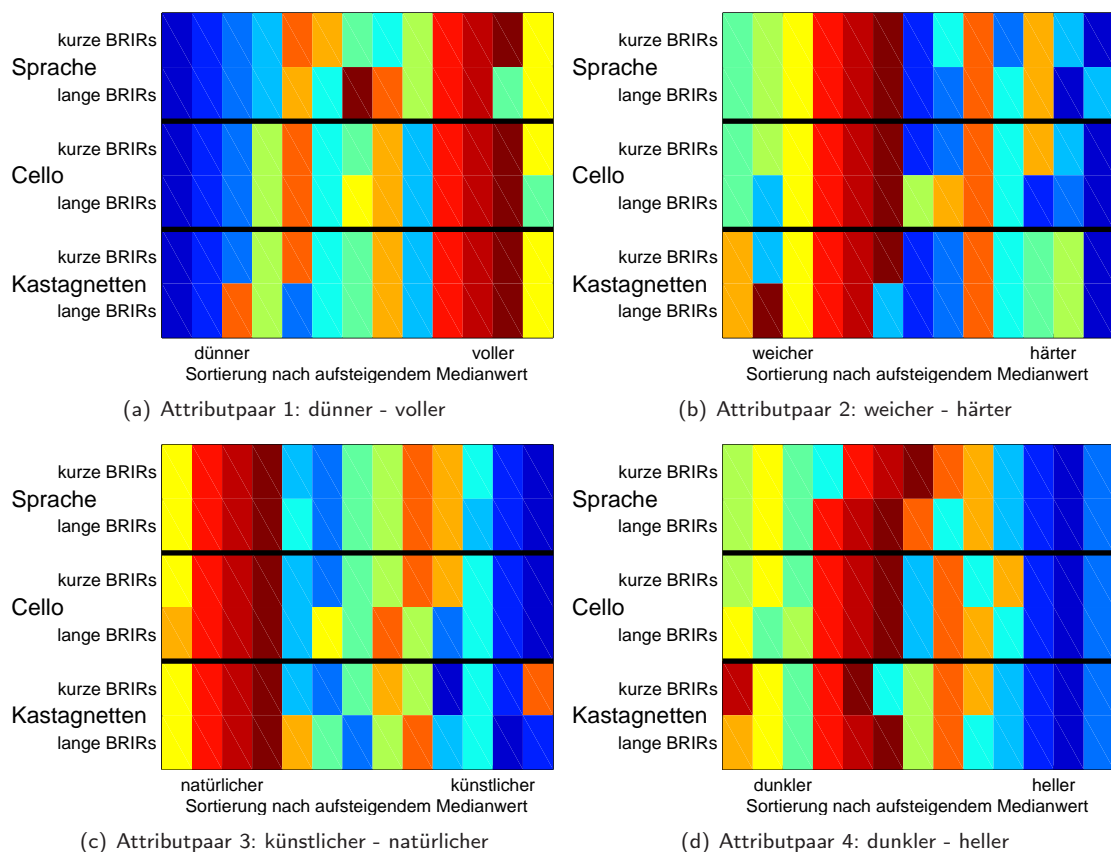


Abbildung 32: aufsteigende Sortierung der quantisierten Medianwerte für alle Klangbeispiele und BRIR-Längen: Klangfarbenattribute

Wie man erkennen kann, stimmen die Reihenfolgen bei den Klangfarbenattributen gut überein.

Auch bei den räumlichen Attributen gibt es eine gute Übereinstimmung. Eine Ausnahme macht dabei Attribut 6 (näher - ferner). Das lässt sich damit begründen, dass die Reihenfolgen innerhalb jeder einzelnen Kombination nicht signifikant sind (siehe dazu die einzelnen Signifikanzmatrizen im Anhang).

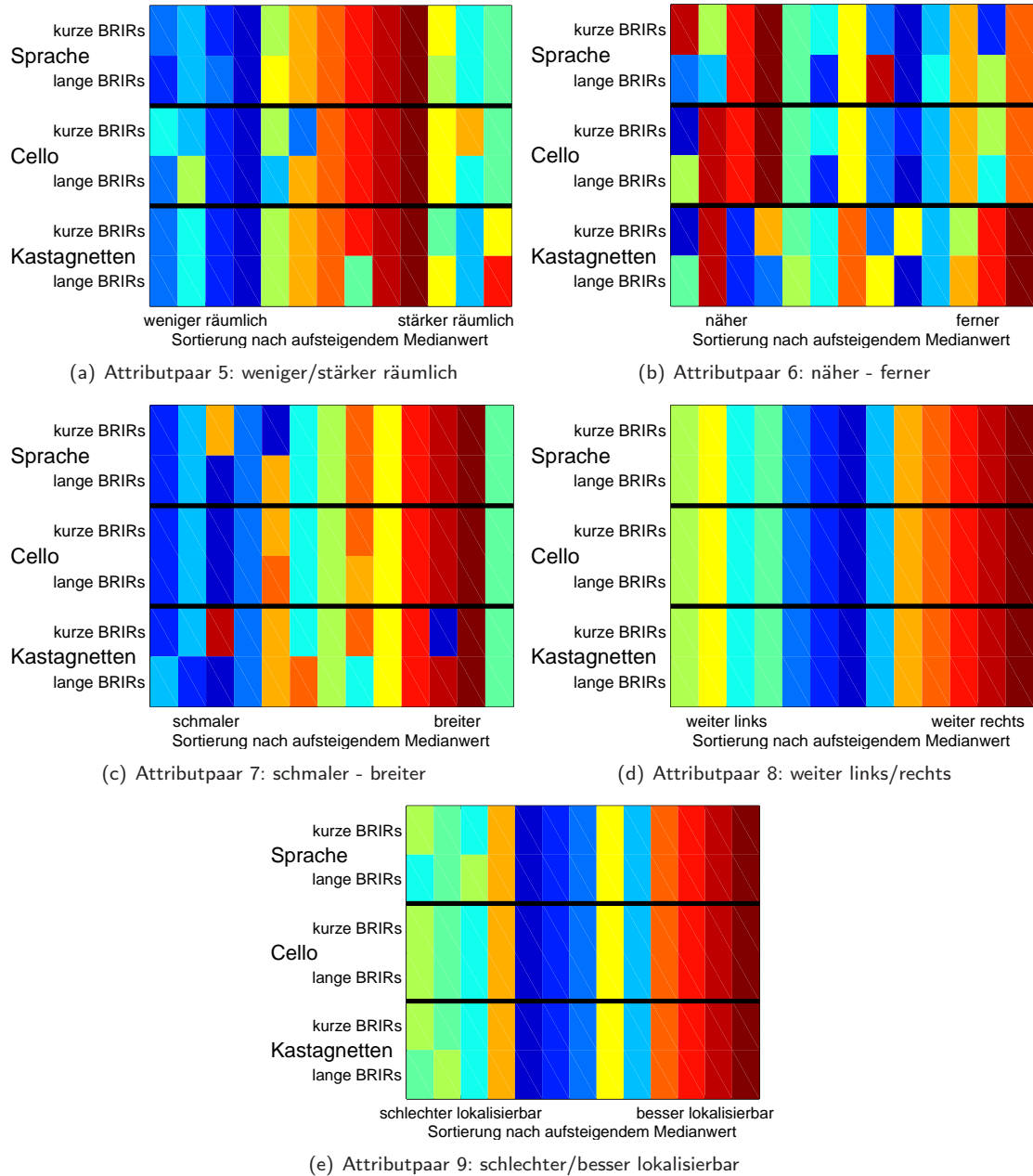


Abbildung 33: aufsteigende Sortierung der quantisierten Medianwerte für alle Klangbeispiele und BRIR-Längen: räumliche Attribute

Aufgrund dieser Übereinstimmung, werden die Kombinationen aus Klangbeispiel und BRIR-Länge nicht getrennt, sondern gemeinsam betrachtet.

5.3 Auswertung der einzelnen Attributpaare

Im Folgenden werden die Ergebnisse der einzelnen Attributpaare ausgewertet. Dies geschieht anhand von einem Boxplot der abgegebenen Bewertungen und der dazugehörigen Signifikanzmatrix.

5.3.1 Attributpaar 1: dünner - voller

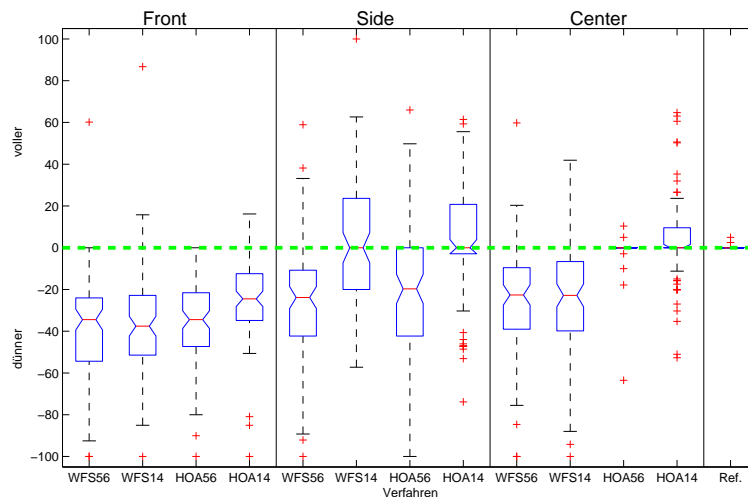


Abbildung 34: Bewertung für Attributpaar 1 über alle Klangbeispiele und ProbandInnen (Boxplot)

Insgesamt sind in der Signifikanzmatrix sehr viele Unterschiede und einige Gemeinsamkeiten erkennbar. Im Mittel wird die vordere Hörposition als wesentlich dünner im Vergleich zur Referenz bewertet. Dabei liefern WFS56, WFS14 und HOA56 ähnliche Ergebnisse, für HOA14 wurde der Klang weniger dünn bewertet. An der seitlichen Hörposition sind die Verfahren mit mehr Lautsprechern dünner als die mit weniger Lautsprechern. Im Zentrum wurden WFS56 und WFS14 deutlich dünner als HOA56 und HOA14 bewertet. Die beiden letztgenannten sind außer durch die Anzahl der Ausreißer und die Größe des IQR nicht von der Referenz zu unterscheiden.

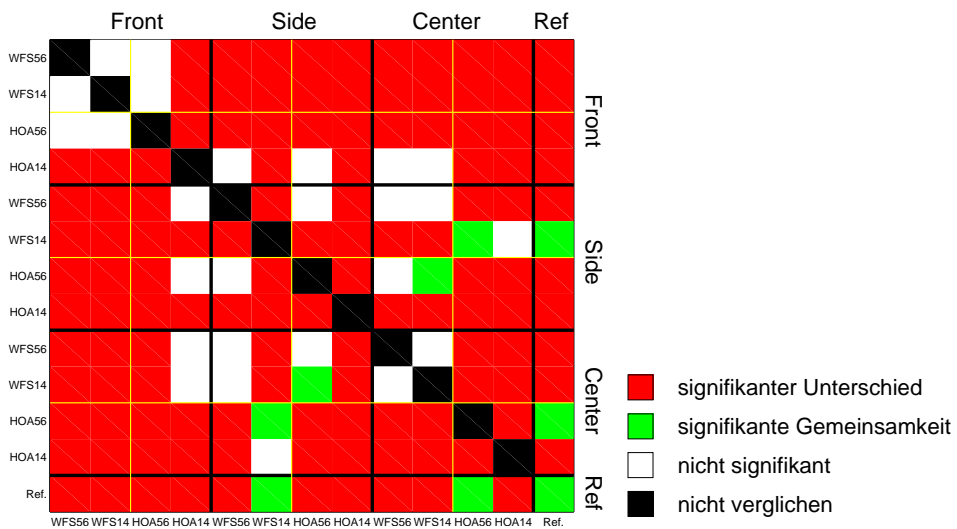


Abbildung 35: Signifikanzmatrix für Attributpaar 1 über alle Klangbeispiele und ProbandInnen

5.3.2 Attributpaar 2: weicher - härter

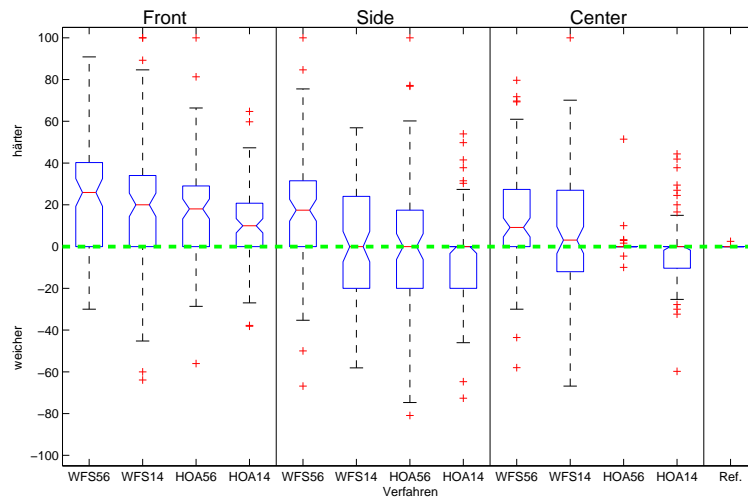


Abbildung 36: Bewertung für Attributpaar 2 über alle Klangbeispiele und ProbandInnen (Boxplot)

Auch bei diesem Attributpaar wurden die vorderen Hörpositionen als am stärksten von der Referenz abweichend bewertet. Die Stärke der Abweichung in Richtung "härter" nimmt von WFS56 über WFS14 und HOA56 zu HOA14 hin ab. An der seitlichen Hörposition klingt WFS an härtesten, WFS14 und HOA56 ähnlich hart und HOA14 etwas weicher. Für das Zentrum gibt es die selben Tendenzen wie beim ersten Attributpaar: WFS56 und WFS14 wurden härter empfunden als HOA56, HOA14 und die Referenz.

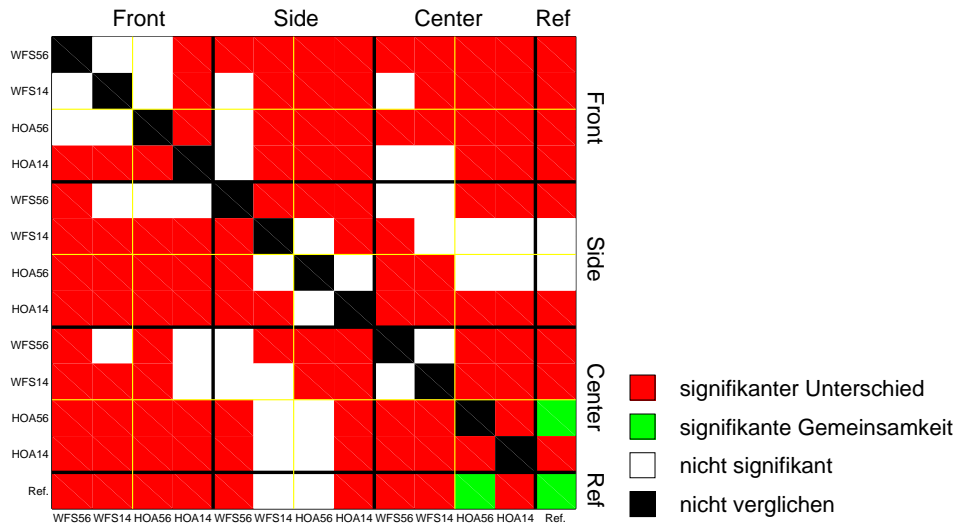


Abbildung 37: Signifikanzmatrix für Attributpaar 2 über alle Klangbeispiele und ProbandInnen

5.3.3 Attributpaar 3: natürlicher - künstlicher

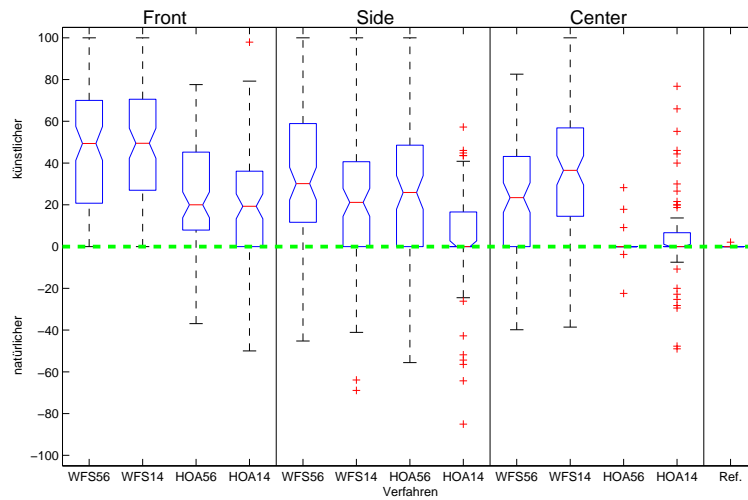


Abbildung 38: Bewertung für Attributpaar 3 über alle Klangbeispiele und ProbandInnen (Boxplot)

Wie schon für die beiden vorherigen Attributpaaren zeigt sich auch die größte abweichende Bewertung von der Referenz für die vordere Hörposition. Dabei sind aber diesmal die beiden WFS56 und WFS14 deutlich künstlicher als die beiden HOA56 und HOA14 bewertet worden. An der seitlichen Hörposition wiederholt sich auch die Tendenzen des zweiten Attributpaares, außer dass hier HOA14 in Richtung "künstlicher" tendiert. Auch im Zentrum ist die bekannte Konstellation wiederzufinden: WFS56 und WFS14 wurden deutlich künstlicher bewertet als HOA56, HOA14 und die Referenz, wobei WFS14 am künstlichsten eingestuft wurde.

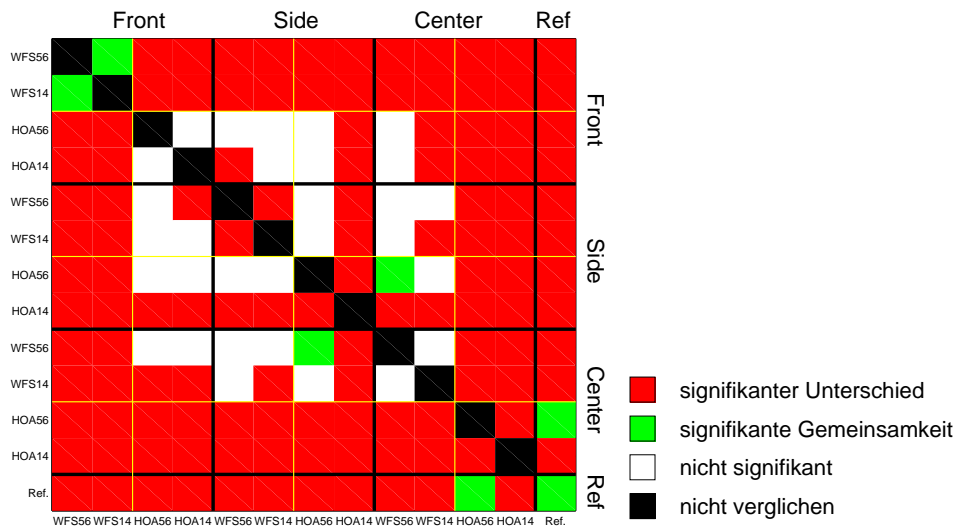


Abbildung 39: Signifikanzmatrix für Attributpaar 3 über alle Klangbeispiele und ProbandInnen

5.3.4 Attributpaar 4: dunkler - heller

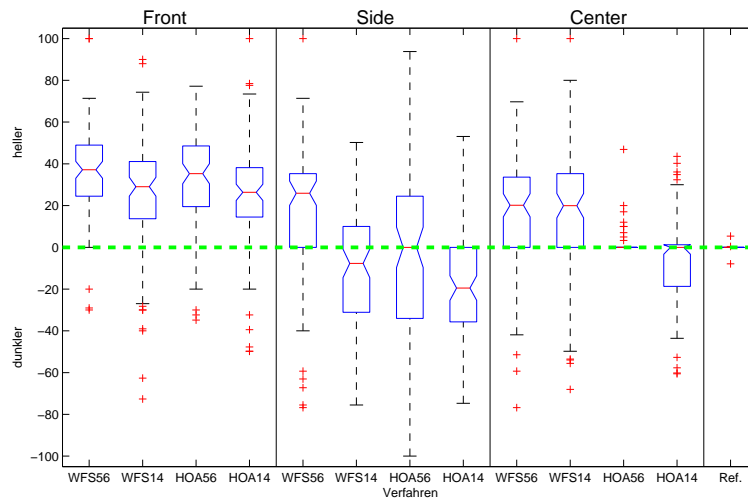


Abbildung 40: Bewertung für Attributpaar 4 über alle Klangbeispiele und ProbandInnen (Boxplot)

Wie bei allen Attributpaaren zur Beschreibung der Klangfarbe, sind auch hier wieder die größten Abweichungen von der Referenz für die vordere Hörposition zu verzeichnen. Dabei werden HOA14 und WFS14 als dunkler empfunden als HOA56 und WFS56. An der seitlichen Hörposition wird zwar WFS56 als heller, jedoch WFS14 als dunkler empfunden. HOA56 zeigt eine große Streuung um die Referenz, während HOA14 im Vergleich dazu als deutlich dunkler empfunden wird. Im Zentrum findet sich die bekannte Konstellation zwischen den Verfahren wieder.

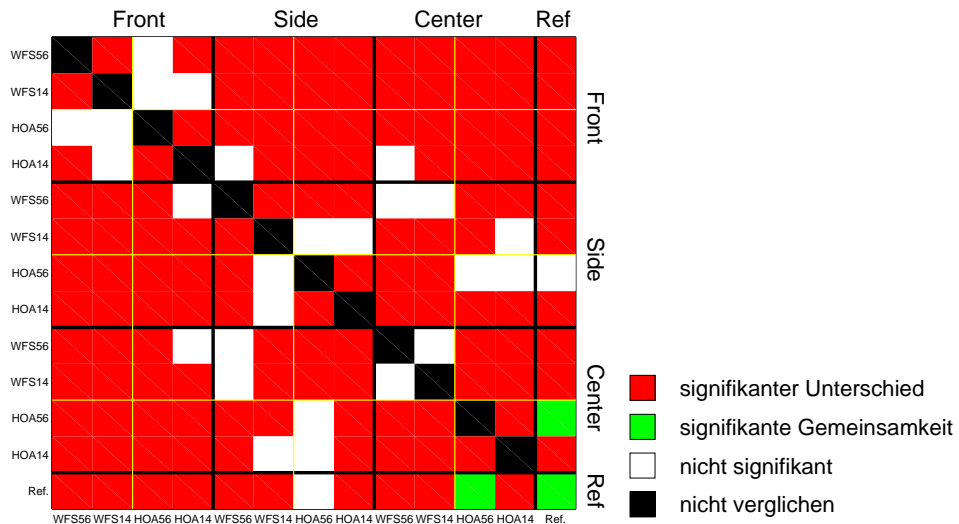


Abbildung 41: Signifikanzmatrix für Attributpaar 4 über alle Klangbeispiele und ProbandInnen

Insgesamt lässt sich sagen, dass sich die vordere Hörposition in Bezug auf Klangfarbe am stärksten von der Referenz unterscheidet, dieser Unterschied bei WFS56 und WFS14 meist größer als bei HOA56 und HOA14 ist. An der seitlichen Hörposition zeigt sich dagegen eine bemerkenswerte Abhängigkeit von der Lautsprecheranzahl: Je weniger Lautsprecher desto geringer ist der Unterschied zur Referenz. Im

Zentrum erkennt man die Ausprägung eines Sweetspots bei HOA56 und HOA14. Beide sind garnicht oder kaum von der Referenz zu unterscheiden. Für WFS56 und WFS14 zeigen sich im Vergleich dazu deutliche Unterschiede.

5.3.5 Attributpaar 5: weniger/stärker räumlich

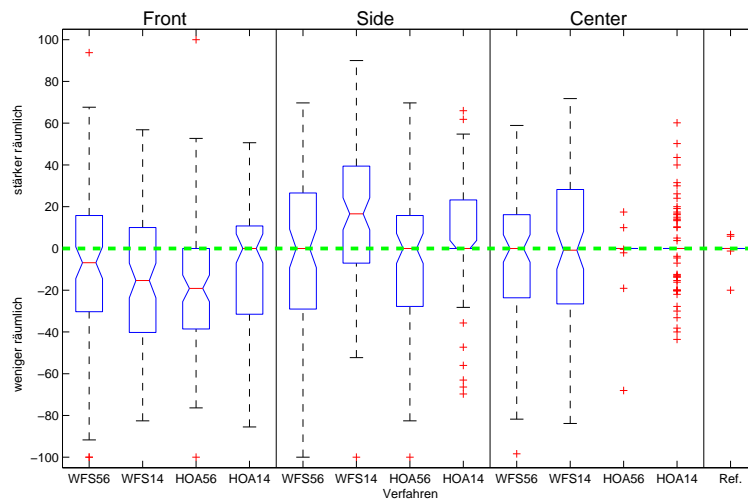


Abbildung 42: Bewertung für Attributpaar 5 über alle Klangbeispiele und ProbandInnen (Boxplot)

Bei diesem Attributpaar zeigt sich eine wesentlich geringere Anzahl an signifikanten Daten als bei den bisher betrachteten. Für die vordere Hörposition ist die geringste Räumlichkeit zu beobachten. An der seitlichen Hörposition werden die Verfahren mit weniger Lautsprechern eher als räumlich empfunden. Für das Zentrum zeigt sich ein großer IQR um die Referenz bei WFS56 und WFS14, HOA56 und HOA14 zeigen einen IQR von 0, wobei bei HOA14 sehr viele Ausreißer zu erkennen sind.

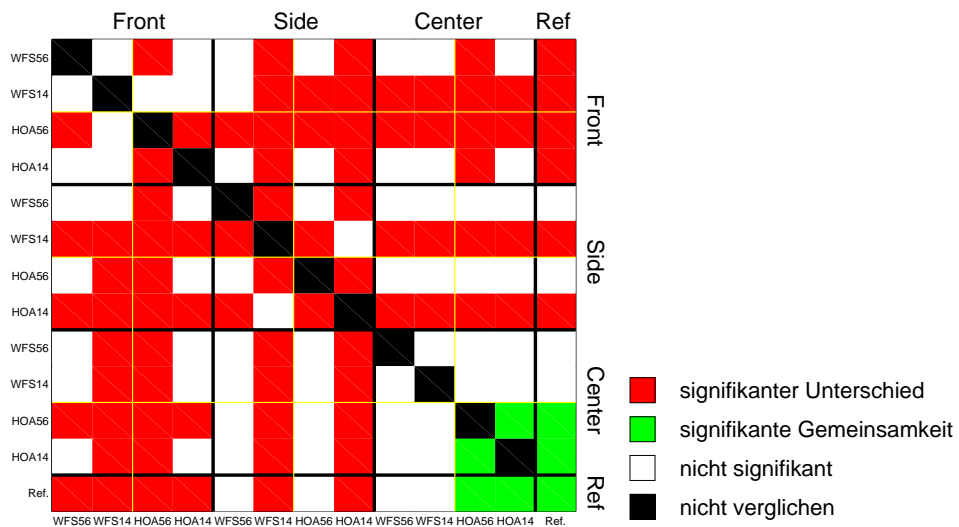


Abbildung 43: Signifikanzmatrix für Attributpaar 5 über alle Klangbeispiele und ProbandInnen

5.3.6 Attributpaar 6: näher - ferner

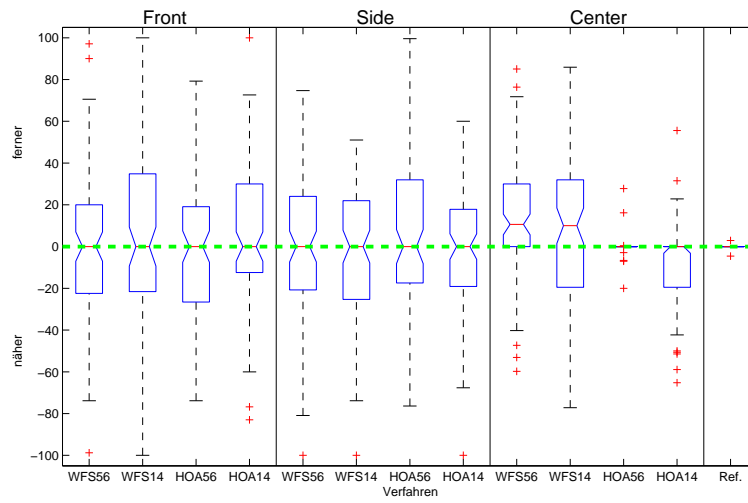


Abbildung 44: Bewertung für Attributpaar 6 über alle Klangbeispiele und ProbandInnen (Boxplot)

Bei diesem Attributpaar sind noch weniger signifikante Werte zu beobachten. Für die vordere und seitliche Hörposition sind nur große Streuungen um den Referenzwert erkennbar. Lediglich im Zentrum ist ein Unterschied zwischen WFS56, WFS14 und HOA56, HOA14 zu erkennen.

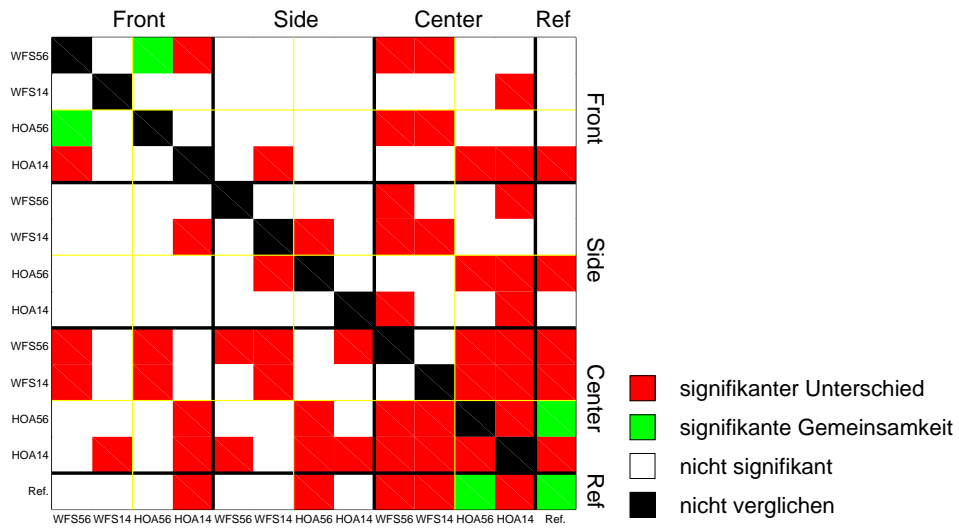


Abbildung 45: Signifikanzmatrix für Attributpaar 6 über alle Klangbeispiele und ProbandInnen

5.3.7 Attributpaar 7: schmaler - breiter

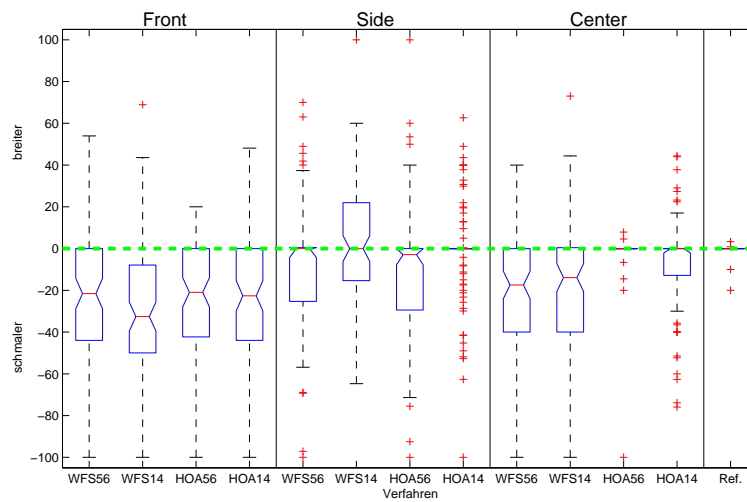


Abbildung 46: Bewertung für Attributpaar 7 über alle Klangbeispiele und ProbandInnen (Boxplot)

Bei Attributpaar 7 sind wieder deutlich mehr signifikante Daten vorhanden. Die vordere Hörposition wurde schmaler als die Referenz bewertet, wobei WFS14 am schmalsten ist. Für die seitliche Hörposition hauptsächlich eine Streuung um 0 mit der Tendenz, dass die Verfahren mit mehr Lautsprechern eher schmaler wahrgenommen wurden. Im Zentrum wurden WFS56 und WFS14 schmaler empfunden als die anderen Verfahren.

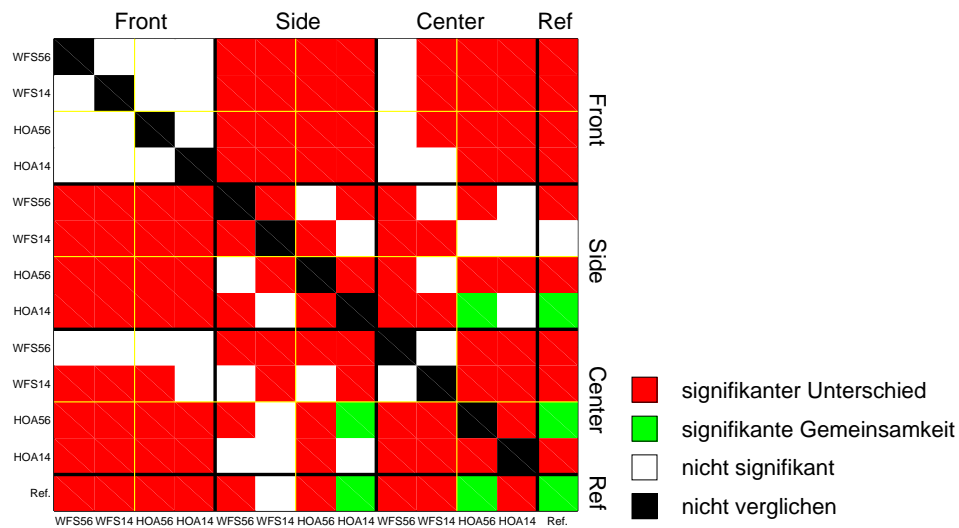


Abbildung 47: Signifikanzmatrix für Attributpaar 7 über alle Klangbeispiele und ProbandInnen

5.3.8 Attributpaar 8: weiter links - weiter rechts

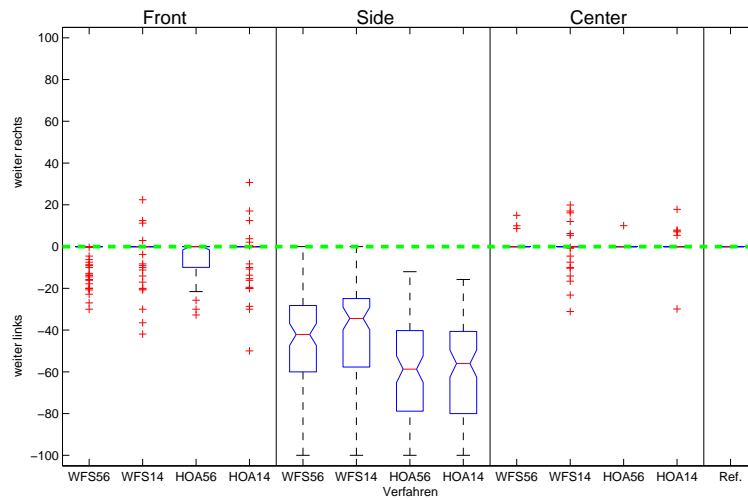


Abbildung 48: Bewertung für Attributpaar 8 über alle Klangbeispiele und ProbandInnen (Boxplot)

Hier reicht schon fast ein Blick auf die Signifikanzmatrix aus. Für die vordere und zentrale Hörposition gibt es fast nur Ausreißer und immer einen Median bei 0. Die seitlichen Hörpositionen sind davon deutlich unterscheidbar, während HOA56 und HOA14 deutlich mehr zur Seite ausgelenkt sind als WFS56 und WFS14.

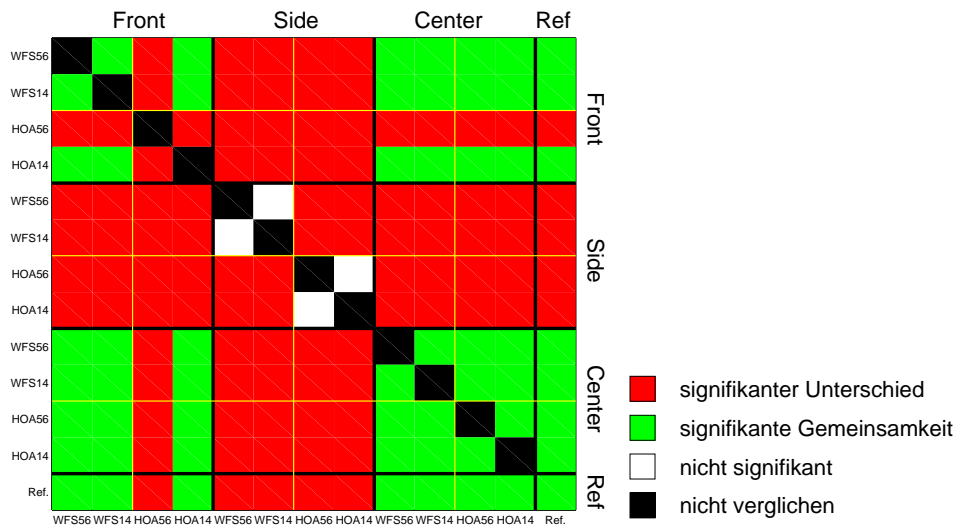


Abbildung 49: Signifikanzmatrix für Attributpaar 8 über alle Klangbeispiele und ProbandInnen

5.3.9 Attributpaar 9: schlechter/besser lokalisierbar

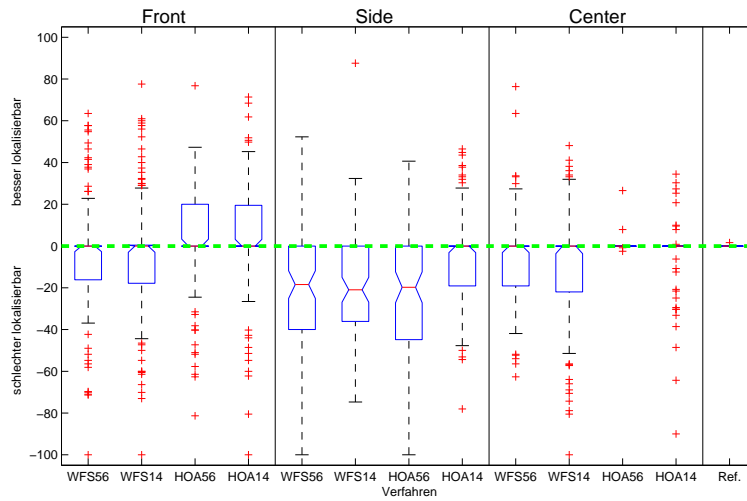


Abbildung 50: Bewertung für Attributpaar 9 über alle Klangbeispiele und ProbandInnen (Boxplot)

Für die vordere Hörposition sind nur schwache Tendenzen um die 0 erkennbar, wobei WFS56 und WFS14 eher schlechter und HOA56 und HOA14 eher besser lokalisierbar sind. An der seitlichen Hörposition herrscht eine schlechtere Lokalisierbarkeit, bei HOA14 aber nur tendenziell. Das Zentrum verhält sich hier ähnlich wie die vordere Position, wobei die Streuung und Anzahl der Ausreißer im Zentrum geringer ist.

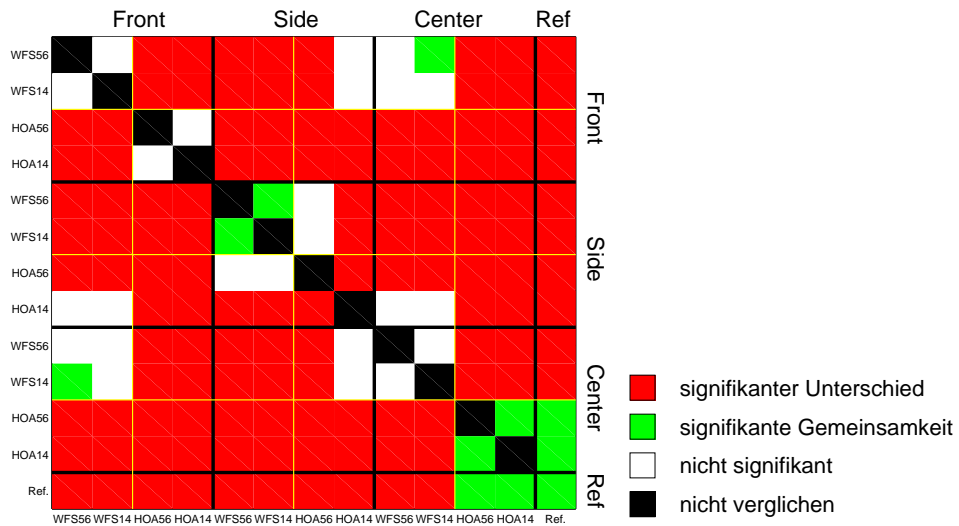
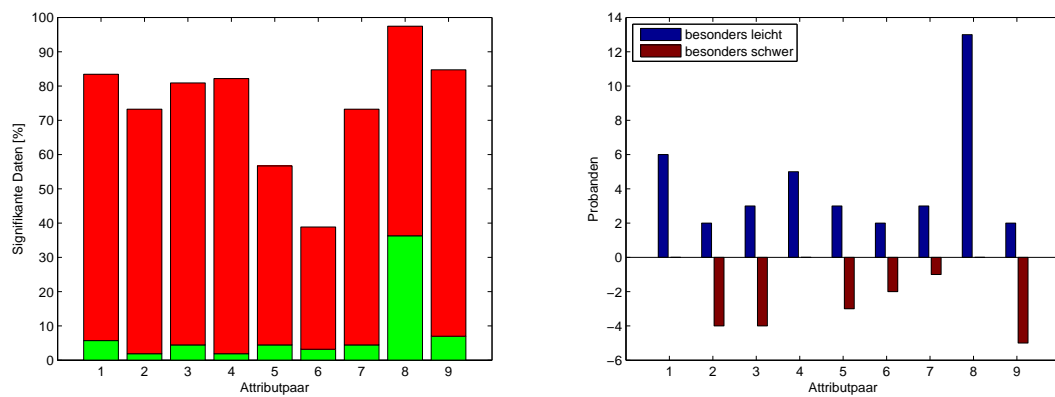


Abbildung 51: Signifikanzmatrix für Attributpaar 9 über alle Klangbeispiele und ProbandInnen

Bei den räumlichen Attributen sind die Unterschiede zwischen den Verfahren an der selben Hörposition nicht sehr groß. Es ist eher eine Abhängigkeit von der Hörposition zu erkennen. Zwei der Attributpaare liefern kaum signifikante Daten.

5.3.10 Attributsignifikanz

Wie sich im vorherigen Abschnitt herausstellte, unterscheiden sich die verschiedenen Attributpaare durch die Anzahl an signifikanten Messwerten. In Abbildung 52(a) sind die Signifikanzmatrizen der einzelnen Attributpaare zusammengefasst. Dabei wird der Anteil an signifikanten Unterschieden (rot dargestellt) und Gemeinsamkeiten (grün dargestellt) in den Datensätzen aufaddiert. Der sich daraus ergebende Wert kann als Maß für die Aussagekraft des jeweiligen Attributpaares beim Vergleich der Verfahren herangezogen werden.



(a) Anzahl der signifikanten Unterschiede (rot) und Gemeinsamkeiten (grün) pro Attributpaar

(b) Bewertung der Attributpaare

Abbildung 52: Vergleich der Attributpaare bezüglich Signifikanz und Bewertung

Man kann erkennen, dass Attributpaar 5 und vor allem Attributpaar 6 deutlich weniger signifikante Daten liefern. Die geringe Aussagekraft der Ergebnisse dieses Attributpaares war bereits in den vorherigen Stadien der Auswertung erkennbar.

Zusätzlich wurden die ProbandInnen gefragt, bei welchen Attributpaaren ihnen die Auswertung besonders leicht bzw. besonders schwer gefallen ist (siehe Abbildung 52(b)). Es sind zwischen der Einschätzung der ProbandInnen und den Ergebnissen Zusammenhänge erkennbar: Die Paare 1, 4 und 8 wurden nie als besonders schwer eingestuft und lieferten auch besonders viele signifikante Daten. Dieser Zusammenhang gilt auch für Attributpaar 7, dass eher leicht als schwer zu bewerten war. Bei den restlichen Paaren (2,3,5,6,9) wurden mindestens so oft schwer wie leicht eingestuft. Unter diesen befinden sich auch die beiden Paare, die die wenigsten signifikanten Daten liefern.

5.4 Zusammenfassung der Attributpaare durch die PCA

Als Abschluss der Auswertung wird nun der Zusammenhang zwischen den Attributpaaren untersucht, um diese anschließend zusammenzufassen. Ziel ist es, die Klangfarben, sowie die räumlichen Attribute mittels einer PCA in einem jeweils zweidimensionalen Vektorraum darzustellen. Dadurch wird eine übersichtliche Darstellung der Konstellation der Stimuli erreicht.

	voller	härter	künstlicher	heller
voller	1.000	-0.470	-0.563	-0.508
härter	-0.470	1.000	0.396	0.459
künstlicher	-0.563	0.396	1.000	0.337
heller	-0.508	0.459	0.337	1.000

Tabelle 5: Klangfarbenattribute: Korrelationsmatrix vor der Zusammenfassung

In Tabelle 5 ist die Korrelationsmatrix der Klangfarbenattribute dargestellt. Da alle vier Attributpaare viele signifikante Daten liefern, werden auch alle bei der PCA berücksichtigt.

Komponente	vor der Rotation			nach der Rotation		
	Eigenwert	% der Varianz	% Gesamt	Eigenwert	% der Varianz	% Gesamt
1	2.372	59.293	59.293	1.591	39.766	39.766
2	0.689	17.231	76.524	1.470	36.758	76.524
3	0.550	13.754	90.278			
4	0.389	9.722	100.000			

Tabelle 6: Klangfarbenattribute: Erklärung der Varianz nach der Zusammenfassung

Bei einer Dimensionsreduktion auf zwei Dimensionen können bereits über 75% der Varianz mit den beiden Komponenten erklärt werden. Durch die Rotation wird der ungleichmäßige Anteil der beiden Komponenten ausgeglichen.

	PC 1	PC 2
voller	-0.496	-0.697
härter	0.749	0.297
künstlicher	0.146	0.934
heller	0.873	0.154

Tabelle 7: Klangfarbenattribute: Komponentenmatrix

In Tabelle 7 sind die Klangfarbenattribute in den Koordinaten der neuen Basis aus den beiden Komponenten dargestellt.

Die sich ergebenden Komponenten wurden anhand der Attribute benannt:

- PC 1: härter, heller, dünner
- PC 2: künstlicher, dünner

Für die Anwendung der PCA auf die räumlichen Attribute wurden die beiden Paare 5 (weniger/stärker räumlich) und 6 (näher - ferner) entfernt, da diese kaum signifikante Werte liefern. Dadurch wird die Dimensionsreduktion vereinfacht.

	breiter	weiter rechts	besser lokalisierbar
breiter	1.000	-0.123	-0.279
weiter rechts	-0.123	1.000	0.212
besser lokalisierbar	-0.279	0.212	1.000

Tabelle 8: räumliche Attribute: Korrelationsmatrix vor der Zusammenfassung

Komponente	vor der Rotation			nach der Rotation		
	Eigenwert	% der Varianz	% Gesamt	Eigenwert	% der Varianz	% Gesamt
1	1.416	47.189	47.189	1.258	41.948	41.948
2	0.883	29.436	76.625	1.040	34.678	76.625
3	0.701	23.375	100.000			

Tabelle 9: räumliche Attribute: Erklärung der Varianz nach der Zusammenfassung

Da die Abweichungen von der Referenz beim Attributpaar 8 (weiter links - weiter rechts) zum überwiegenden Teil mit negativen Werte angegeben wurden, ist es sinnvoll, den Attributvektor um 180° zu drehen und seinen Namen von "weiter rechts" auf "weiter links" zu ändern.

	PC 1	PC 2
breiter	0.869	0.072
weiter rechts	-0.082	0.966
weiter links	0.082	-0.966
besser lokalisierbar	-0.705	0.320

Tabelle 10: räumliche Attribute: Komponentenmatrix

Die beiden räumlichen Komponenten wurden folgendermaßen benannt:

- PC 1: breiter, schlechter lokalisierbar
- PC 2: weiter links

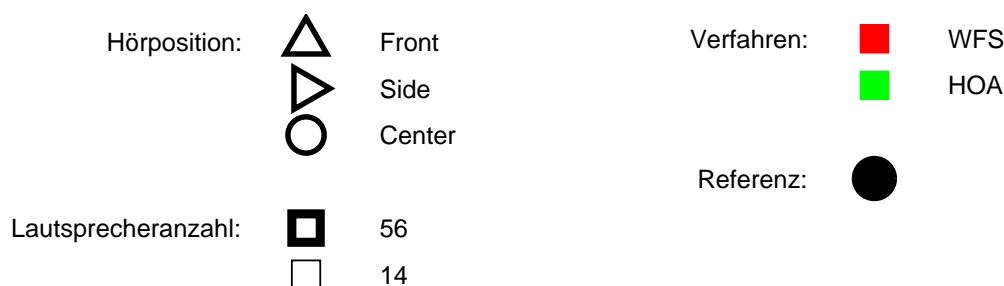


Abbildung 53: Legende zur Darstellung der PCA-Ergebnisse

Im Folgenden werden die Ergebnisse der beiden PCAs anhand der Abbildungen 54 und 55 diskutiert. Dabei werden zuerst alle vier Verfahren innerhalb einer Hörposition miteinander verglichen.

1. Klangfarbe

An der Hörposition **Center** ist HOA56 praktisch nicht von der Referenz unterscheidbar. Auch im Vergleich zu HOA14 sind die Unterschiede gering. WFS wirkt dagegen künstlicher und dünner. Auch hier hat die Lautsprecheranzahl keinen sehr großen Einfluss. Trotzdem kann man sagen, dass für diese Hörposition die Verfahren mit mehr Lautsprechern näher an die Referenz herankommen als das selbe Verfahren mit weniger Lautsprechern.

Für die Hörposition **Front** wurden WFS und HOA gleichermaßen deutlich härter und heller als die Referenz bewertet, wobei HOA aber als deutlich natürlicher und voller im Vergleich zu WFS empfunden wurde. Auch hier kann man eine Abhängigkeit von der Lautsprecheranzahl feststellen: Je mehr Lautsprecher verwendet werden, desto größer wird der Abstand zur Referenz.

Dieses Phänomen ist an der Hörposition **Side** besonders stark. HOA wird dort generell als etwas weicher und dunkler gehört, wobei die Version mit 56 Lautsprechern deutlich künstlicher und dünner wirkt. Bei WFS ist Abhängigkeit von der Lautsprecheranzahl bezüglich des Klangcharakters auch deutlich: WFS14 ist eher weicher und dunkler, WFS56 eher härter und heller. Beide sind dabei künstlicher als die Referenz.

Als Ergebnis der PCA zur Klangfarbe kann man festhalten, dass WFS im Vergleich zur Referenz für alle Hörpositionen künstlicher, dünner, heller und härter klingt. Der Klangcharakter ist dabei auch relativ unabhängig von der Lautsprecheranzahl (mit Ausnahme von WFS14 Side). Dagegen klingt HOA im Mittel eher weicher und dunkler, wobei sich die vorderen Hörpositionen klanglich deutlich in Richtung WFS bewegen. Dadruch zeigt sich eine größere Abhängigkeit von der Hörposition bei HOA im Vergleich zu WFS.

2. Räumliche Attribute

Für die beiden Hörpositionen **Center** und **Front** sind die Unterschiede zwischen WFS und HOA eher gering. Dabei wurden bei Front alle Verfahren eher schmaler und besser lokalisierbar wahrgenommen und bei Center im Vergleich dazu eher breiter und schlechter lokalisierbar (WFS56 Center tendiert dabei eher in Richtung Front).

Davon deutlich getrennt ist die Hörposition **Side**: Die Stimuli wurden dort viel weiter links wahrgenommen. Man kann auch erkennen, dass diese seitliche Auslenkung bei HOA größer ist als bei WFS.

Zusammenfassend kann man sagen, dass die räumlichen Attribute eher von der Hörposition als vom Verfahren oder der Lautsprecheranzahl abhängen. Für die seitlich Hörposition wird allerdings auch die Abhängigkeit vom verwendeten Verfahren deutlich.

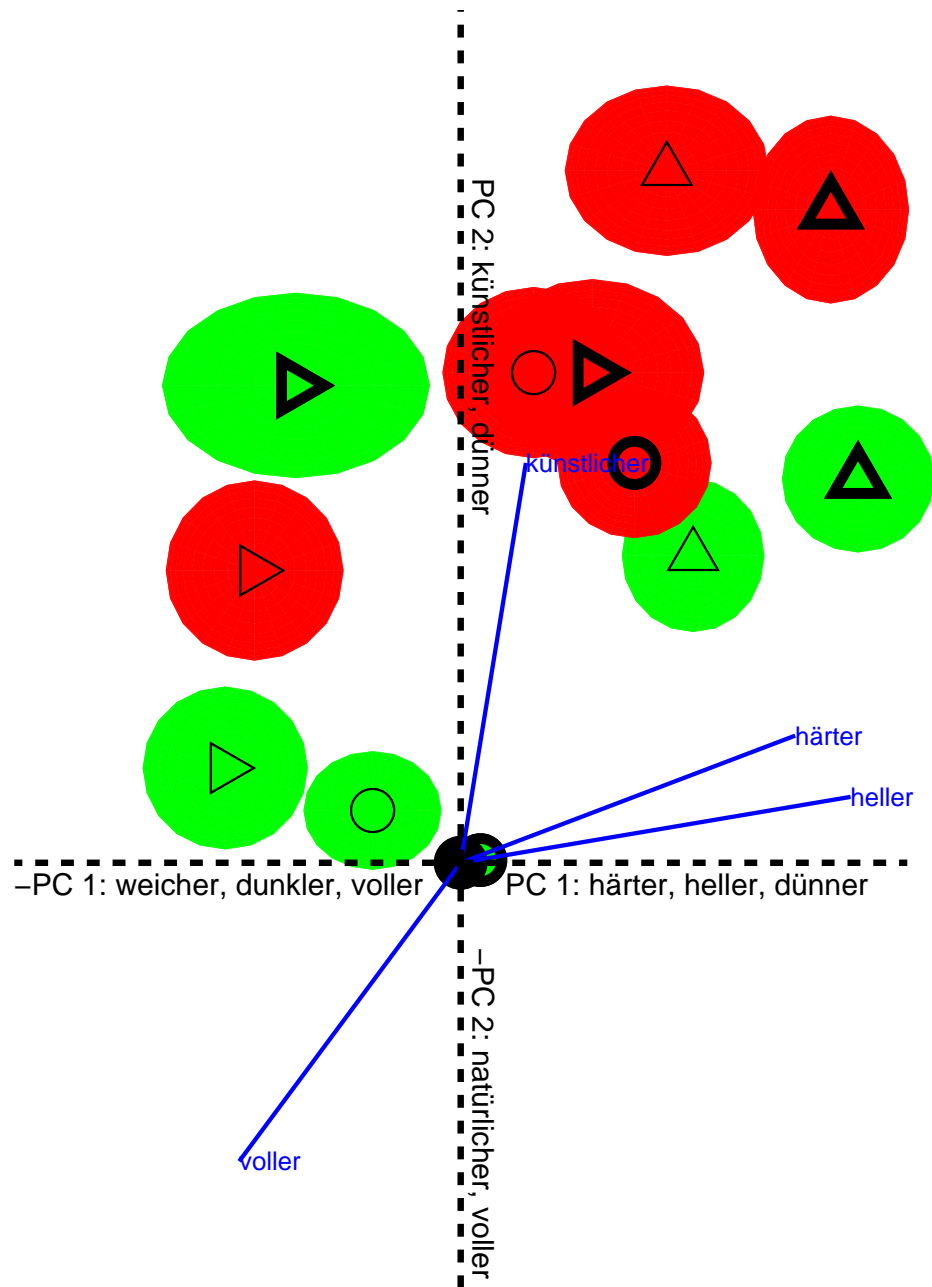


Abbildung 54: zweidimensionale PCA für Klangfarbe

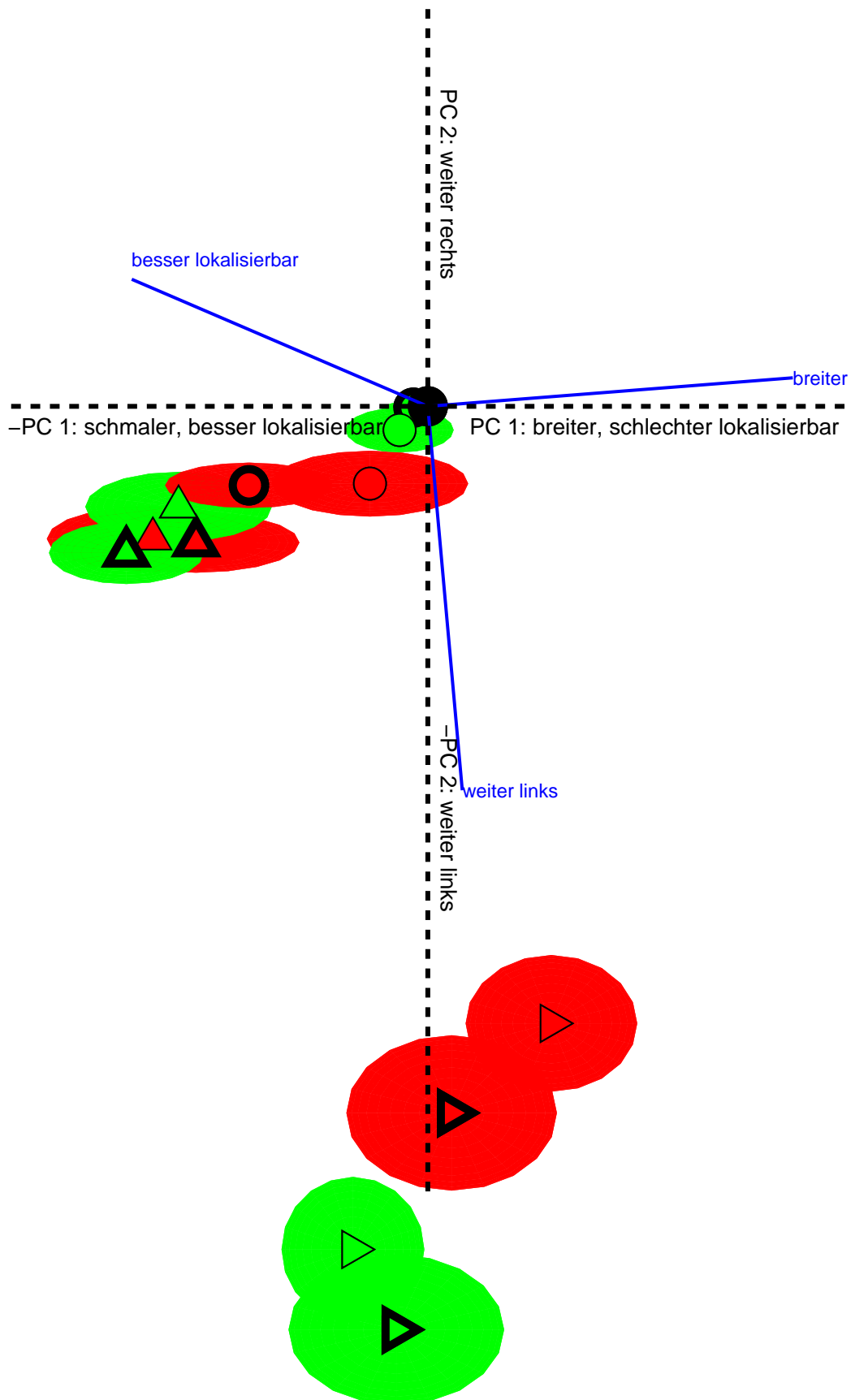


Abbildung 55: zweidimensionale PCA für räumliche Attribute

5.5 Interpretation der Ergebnisse und Vergleich mit der Theorie und Simulation

5.5.1 Klangfarbe

Der Klangcharakter der WFS wird im Vergleich zur Referenz als künstlicher, dünner, heller und härter empfunden. Der Charakter wird bei allen Hörpositionen beibehalten und ist, außer bei WFS14, auch relativ unabhängig von der Lautsprecheranzahl. Dieses Verhalten passt zu den Simulationen, wobei sich, aufgrund der gleichmäßigen Verteilung der Wiedergabebefehle, für alle Hörpositionen und Lautsprecheranzahlen ein sehr ähnlicher Frequenzgang ergibt (siehe Abbildung 11).

Bei HOA wird dagegen eine größere Abhängigkeit von der Hörposition festgestellt: Während im Zentrum kaum eine Unterscheidung von der Referenz erkennbar ist, klingt die seitliche Hörposition dunkler und weicher, aber die vordere eher schon wie WFS. Wie in den Simulationen ist also ein ausgeprägter Sweetspot erkennbar (siehe Abbildungen 8 und 6).

Sowohl bei WFS als auch HOA ist erkennbar, dass die seitliche Hörposition eher dunkler und weicher empfunden wird. Das lässt sich mit dem Frequenzgang an dieser Hörposition erklären, der in den Simulationen eine Überhöhung der tiefen Frequenzen erkennen lässt (siehe Abbildungen 8 und 11).

Für beide Verfahren ist eine Tendenz in der Abhängigkeit von der Lautsprecheranzahl erkennbar: Im Zentrum kommen die Verfahren mit mehr Lautsprechern am nächsten an die Referenz. Das lässt sich mit der höheren Aliasingfrequenz erklären. Für die beiden dezentralen Hörpositionen gilt allerdings das Gegenteil. Auf den ersten Blick steht dies im Widerspruch zur Theorie, kann aber durch die Nähe zu den Lautsprechern erklärt werden: Besonders bei der vorderen Hörposition wird bei geringer Lautsprecheranzahl nie das gewünschte Schallfeld erreicht. Das reproduzierte Schallfeld ähnelt eher dem einer Punktquelle, besonders bei HOA (siehe Abbildungen 6 und 7). Wenn die Ohren nun, wie bei der verwendeten vorderen Hörposition, nur 60cm von den vorderen Lautsprechern entfernt sind, verschlechtert sich die Klangqualität, je mehr Lautsprecher beteiligt sind. An der seitlichen Hörposition lässt sich das ähnlich erklären. Das heißt also: Wenn es für diese äußeren Hörpositionen nicht gelingt, die Charakteristik des gewünschten Schallfeldes zu reproduzieren, dann sollten für eine bessere Klangqualität eher weniger Lautsprecher verwendet werden.

5.5.2 Räumliche Attribute

In Bezug auf räumliche Attribute sind die Unterschiede zwischen WFS und HOA, sowie die Abhängigkeit von der Lautsprecheranzahl kaum erkennbar. Das stärkste Unterscheidungskriterium liefert die Hörposition: Die zentrale Hörposition ist am besten mit der Referenz vergleichbar. Die vorderen Hörpositionen sind durch die Nähe zu den Lautsprechern eher besser lokalisierbar und wirken deshalb schmaler. WFS56 Center macht dabei eine saubere Trennung zwischen Front und Center zu, weil die Werte eher zur vorderen Hörposition tendieren.

Die seitliche Hörposition ist jedoch sehr deutlich von den beiden anderen getrennt. Alle Stimuli wurden viel weiter links wahrgenommen. Bei HOA ist diese seitliche Auslenkung noch größer als bei WFS. Das lässt sich damit erklären, dass die Form der ebenen Welle bei WFS an der seitlichen Hörposition besser erhalten bleibt als bei HOA, wo sie bei hohen Frequenzen eher einer Punktquelle ähnelt (siehe Abbildungen 7 und 6).

6 Zusammenfassung

In dieser Arbeit werden Wellenfeldsynthese (WFS) und Nearfield Compensated Higher-Order Ambisonics (HOA) durch Hörversuche miteinander verglichen. Es wird mittels nachgeführter Binauralsynthese ein reales Lautsprechersystem nachgebildet. Im Versuch werden die beiden Schallfeldreproduktionsverfahren verglichen; dabei variieren die Hörpositionen und Lautsprecheranzahlen. Die Hörversuche liefern folgende Ergebnisse:

WFS und HOA zeigen deutliche Unterschiede. Der charakteristische Klang von WFS lässt sich mit den Attributen dünn, hell, hart und künstlich bezeichnen. Dieser wird unabhängig von Lautsprecheranzahl und Hörposition weitgehend beibehalten. Bei HOA gibt es dagegen einen ausgeprägten Sweetspot im Zentrum der Wiedergabeanordnung, wobei der empfundene Klang auch an den dezentralen Hörpositionen weniger von der Referenz abweicht als bei WFS. Der Klang wird dabei an der Seite eher dunkler und weicher, während er an der vorderen Position eher in Richtung des typischen WFS Klangs tendiert. Bei beiden Verfahren wird an den dezentralen Hörpositionen die Klangqualität bei geringerer Lautsprecheranzahl besser empfunden.

Bezüglich der räumlichen Qualität gibt es zwischen den beiden Verfahren nur einen Unterschied an der seitlichen Hörposition: Bei WFS gleicht die wahrgenommene Richtung besser der wiedergegebenen als bei HOA. Die Form der ebenen Welle wird bei WFS also besser beibehalten.

Insgesamt wird in den Versuchen WFS eine bessere räumliche Qualität und HOA eine bessere klangliche Qualität zugeschrieben. Dieses Ergebnis passt sehr gut mit den Erwartungen aus den analytischen Simulationen zusammen. Um allgemeingültigere Aussagen treffen zu können, sollten allerdings noch weitere Hörversuche durchgeführt werden, die auch die Wiedergabe in Räumen mit deutlich längerer Nachhallzeit, sowie die Lokalisationseigenschaften der beiden Verfahren untersuchen.

7 Danksagung

Das Gelingen dieser Arbeit war nur durch die Hilfe und Unterstützung vieler möglich.

Allen voran danke ich meinen Betreuern der Deutsche Telekom Laboratories in Berlin Jens Ahrens und Dr. Sascha Spors, sowie Franz Zotter und Dr. Alois Sontacchi vom IEM in Graz. Ich danke auch dem Quality & Usability Lab für die Möglichkeit, den praktischen Teil meiner Arbeit dort durchzuführen. Mein Dank gilt allen ProbandInnen des Hörversuchs und der Vorversuche. Für die Hilfe bei der Versuchsplanung danke ich besonders Matthias Geier, Dr. Alexander Raake und Marcel Wältermann. Für die entspannende Mittagspause bedanke ich mich bei den KollegInnen der Mittagsrunde und dem Restaurant "Moon Thai 2". Ich bedanke mich bei Boris Müller für die kurzfristige Vermittlung der Wohnung in Berlin. Mein Dank gilt auch allen KollegInnen und DozentInnen, die mich auf dem Weg durch mein Studium begleitet haben.

Ganz besonders möchte ich mich bei meinen Eltern, Doris und Peter, und meinem Bruder Andreas für die Unterstützung und Motivation während des gesamten Studiums bedanken.

Literatur

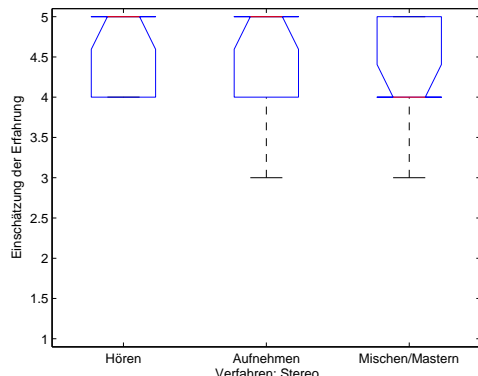
- [AS72] M. Abramowitz and I.A. Stegun. *Handbook of Mathematical Functions*. Dover Publications, 1972.
- [AS08a] J. Ahrens and S. Spors. An analytical approach to sound field reproduction using circular and spherical loudspeaker distributions. *Acta Acustica united with Acustica*, 94(6):988–999, November-December 2008.
- [AS08b] J. Ahrens and S. Spors. Analytical driving functions for higher-order ambisonics. In *IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, April 2008.
- [AS09] J. Ahrens and S. Spors. Alterations of the temporal spectrum in high-resolution sound field reproduction of different spatial bandwidths. In *126th AES Convention*, Munich, Germany, May 2009.
- [Aug04] Tobias Augustin. Zur Wahrnehmbarkeit von Klangfarbenänderungen bei Wellenfeldsynthese. Master's thesis, Institut für Rundfunktechnik GmbH, 2004.
- [Baa07] Marije A.J. Baalman. Reproduction of arbitrarily shaped sound sources with Wave Field Synthesis - physical and perceptual effects. In *122nd AES Convention*, Vienna, Austria, May 2007.
- [BDG⁺07] Stéphanie Bertet, Jérôme Daniel, Laëtitia Gros, Etienne Parizet, and Olivier Warusfel. Investigation of the Perceived Spatial Resolution of Higher Order Ambisonics Sound Fields: A Subjective Evaluation Involving Virtual and Real 3D Microphones. In *30th AES Int. Conference*, 2007.
- [BDPW08] Stéphanie Bertet, Jérôme Daniel, Etienne Parizet, and Olivier Warusfel. Investigation on the restitution system influence over perceived Higher Order Ambisonics sound field: a subjective evaluation involving from first to fourth order systems. In *Acoustics-08, Paris*, 2008.
- [BdVV93] A.J. Berkhout, D. de Vries, and P. Vogel. Acoustic control by wave field synthesis. *J. Acoust. Soc. Am.*, 93(5):2764–2778, May 1993.
- [Ber88] A.J. Berkhout. A holographic approach to acoustic control. *J. Audio Eng. Soc.*, 36:977–995, 1988.
- [Ber06] Jan Berg. Identification of quality attributes of spatial audio by repertory grid technique. *J. Audio Eng. Soc.*, 54(5):365–379, May 2006.
- [Bla83] J. Blauert. *Spatial Hearing - The Psychophysics of Human Sound Localization*. The MIT-Press, Cambridge, MA, 1983.
- [BR99] Jan Berg and Francis Rumsey. Identification of perceived spatial attributes of recordings by repertory grid technique and other methods. In *106th AES Convention*, Munich, Germany, 1999.

- [BV95] J. S. Bamford and J. Vanderkooy. Ambisonic sound for us (an analysis of imaging in ambisonics, stereo and dolby surround systems). In *99th AES Convention*, New York, October 6-9 1995.
- [CHC⁺07] Sylvain Choisel, Patrick Hegarty, Flemming Christensen, Benjamin Pedersen, Wolfgang Ellermeier, Jody Ghani, and Wookeun Song. A listening test system for automotive audio, part 4: Comparison of attribute ratings made by expert and non-expert listeners. In *123rd AES Convention*, New York, NY, October 2007.
- [CW07] Sylvain Choisel and Florian Wickelmaier. Evaluation of multichannel reproduced sound: Scaling auditory attributes underlying listener preference. *J. Acoust. Soc. Am.*, 121(1):388–400, January 2007.
- [Dan00] Jérôme Daniel. *Représentation de Champs Acoustiques, Application à la Transmission et à la Reproduction de Scènes Sonores Complexes dans un Contexte Multimédia*. PhD thesis, Université de Paris, 2000.
- [Dan03] J. Daniel. Spatial sound encoding including near field effect: Introducing distance coding filters and a viable, new ambisonic format. In *AES 23rd International Conference*, Copenhagen, Denmark, May 2003.
- [DNM03] Jérôme Daniel, Rozenn Nicol, and Sébastien Moreau. Further investigations of high order ambisonics and wavefield synthesis for holophonic sound imaging. In *114th AES Convention*, Amsterdam, The Netherlands, March 22-25 2003.
- [EBU] EBU. Tech 3253 - Sound Quality Assessment Material (SQAM). http://www.ebu.ch/en/technical/publications/tech3000_series/tech3253/index.php.
- [FZS08] Matthias Frank, Franz Zotter, and Alois Sontacchi. Localization experiments using different 2d ambisonics decoders. In *25. Tonmeistertagung*, Leipzig, 2008.
- [GAS08a] Matthias Geier, Jens Ahrens, and Sascha Spors. ASDF: Ein XML Format zur Beschreibung von virtuellen 3D Audioszenen. In *DAGA*, Dresden, Germany, March 2008.
- [GAS08b] Matthias Geier, Jens Ahrens, and Sascha Spors. The soundscape renderer: A unified spatial audio reproduction framework for arbitrary rendering methods. In *124th AES Convention*, Amsterdam, Netherlands, May 2008.
- [GAS09] Matthias Geier, Jens Ahrens, and Sascha Spors. Binaural monitoring of massive multichannel sound reproduction systems using model-based rendering. In *DAGA*, 2009.
- [Ger72] M. A. Gerzon. Periophony (with-height sound reproduction). In *2CE AES Convention*, Munich, Germany, March 14-16 1972.
- [GK04] Catherine Guastavino and Brian F. G. Katz. Perceptual evaluation of multi-dimensional spatial audio reproduction. *J. Acoust. Soc. Am.*, 116(2):1105–1115, August 2004.
- [Hal01] Joseph L. Hall. Application of multidimensional scaling to subjective evaluation of coded speech. *J. Acoust. Soc. Am.*, 110(4):2167–2182, October 2001.

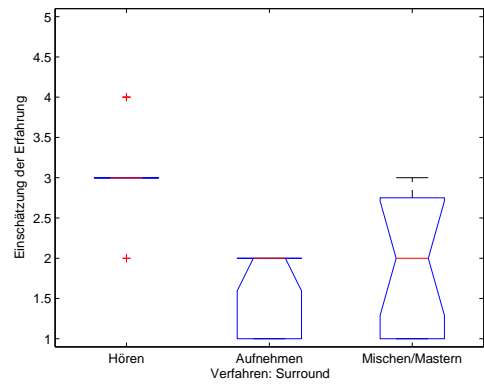
- [Hub02] Thomas Huber. Zur Lokalisation akustischer Objekte bei Wellenfeldsynthese. Master's thesis, Institut für Rundfunktechnik GmbH, München, 2002.
- [ITU97] ITU. ITU-R BS.1116-1 Methods for the subjective assessment of small impairments in audio systems including multichannel sound systems, 1994-1997.
- [ITU03] ITU. ITU-R BS.1284-1 General methods for the subjective assessment of sound quality, 1997-2003.
- [LW06] Alexander Lindau and Stefan Weinzierl. FABIAN - An instrument for software-based measurement of binaural room impulse responses in multiple degrees of freedom. In *24. Tonmeistertagung*, Leipzig, 2006.
- [LW07] Alexander Lindau and Stefan Weinzierl. FABIAN - Schnelle Erfassung binauraler Raumimpulsantworten in mehreren Freiheitsgraden. In *DAGA*, Stuttgart, 2007.
- [Mat09] The Mathworks. MATLAB Statistics Toolbox: Documentation. <http://www.mathworks.com/access/helpdesk/help/helpdesk.html>, 2009.
- [Pul97] V. Pulkki. Virtual sound source positioning using vector base amplitude panning. *J. Audio Eng. Soc.*, 45(6):456–466, June 1997.
- [SA08] Sascha Spors and Jens Ahrens. A comparison of wave field synthesis and higher-order ambisonics with respect to physical properties and spatial sampling. In *125th AES Convention*, San Francisco, CA, USA, October 2-5 2008.
- [SA09] Sascha Spors and Jens Ahrens. Spatial sampling artifacts of wave field synthesis for the reproduction of virtual point sources. In *126th AES Convention*, Munich, Germany, May 2009.
- [Son07] A. Sontacchi. Versuchsdesign in der Psychoakustik. Vorlesungsunterlagen, Institut für Elektronische Musik und Akustik, Graz, 2007.
- [SPS] SPSS. SPSS Statistics 17.0 Algorithms. <http://support.spss.com/Student/Studentdefault.asp>.
- [SR06] S. Spors and R. Rabenstein. Spatial aliasing artifacts produced by linear and circular loudspeaker arrays used for wave field synthesis. In *120th AES Convention*, Paris, France, May 2006.
- [SRA08] S. Spors, R. Rabenstein, and J. Ahrens. The theory of wave field synthesis revisited. In *124th AES Convention*, Amsterdam, The Netherlands, May 17-20 2008.
- [The90] G. Theile. On the performance of two-channel and multi-channel stereophony. In *88th AES Convention*, 1990.
- [Vie] ViennaSymphonicLibrary. <http://vsl.co.at/>.
- [WA01] D. Ward and T. Abhayapala. Reproduction of a plane-wave sound field using an array of loudspeakers. *IEEE Transactions on Speech and Audio Processing*, 9:697–707, 2001.

- [Weg05] Dominik Wegmann. Zu Unterschieden in der Hörereigniswahrnehmung bei Wellenfeldsynthese und Stereophonie im Vergleich zum natürlichen Hören. Master's thesis, Institut für Hörtechnik und Audiologie, Fachhochschule Oldenburg, 2005.
- [Weia] Eric W. Weisstein. Divergence Theorem. <http://mathworld.wolfram.com/DivergenceTheorem.html>.
- [Weib] Eric W. Weisstein. Green's Identities. <http://mathworld.wolfram.com/GreensIdentities.html>.
- [Wik] Wikipedia. Gaußscher Integralsatz. http://de.wikipedia.org/wiki/Gaußscher_Integralsatz.
- [Wil99] E.G. Williams. *Fourier Acoustics: Sound Radiation and Nearfield Acoustical Holography*. Academic Press, 1999.
- [Wit07] Helmut Wittek. *Perceptual differences between wavefield synthesis and stereophony*. PhD thesis, Department of Music and Sound Recording School of Arts, Communication and Humanities, University of Surrey, 2007.

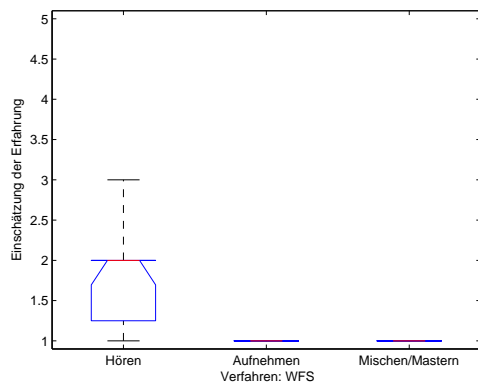
Anhang



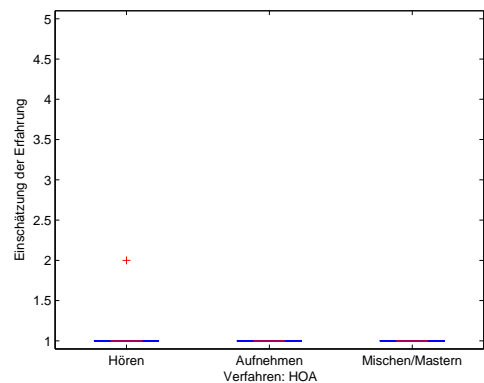
(a) Stereo



(b) Surround



(c) WFS



(d) HOA

Abbildung 56: Erfahrung mit verschiedenen Aufnahme-/Wiedergabeverfahren

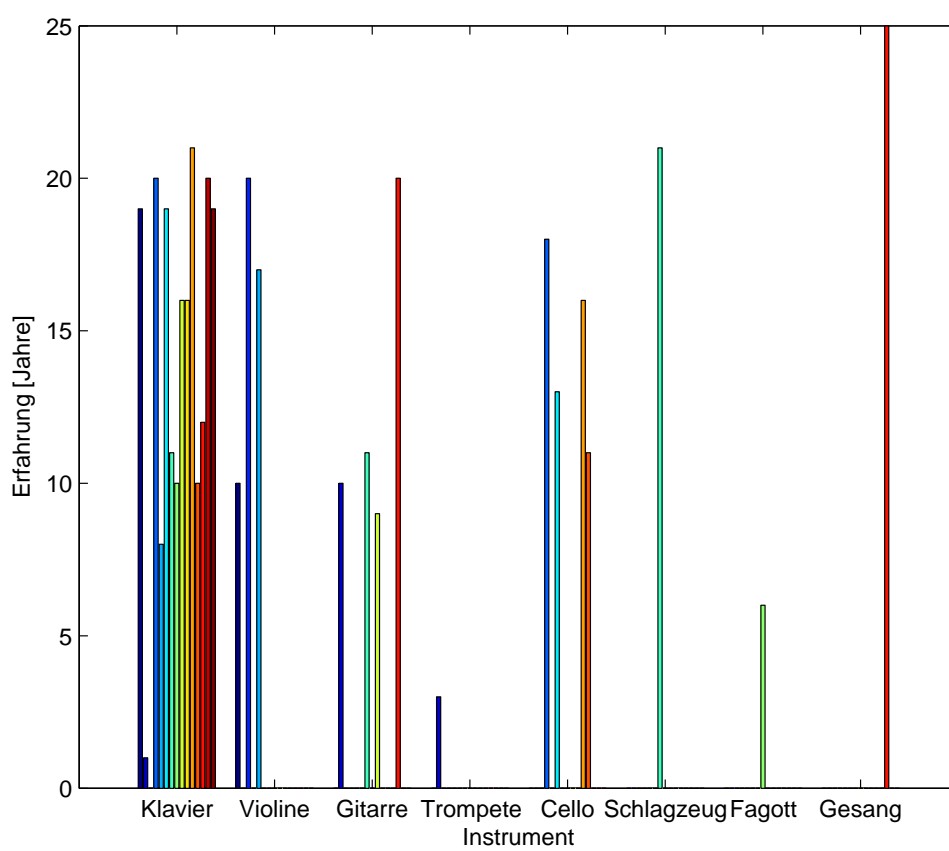
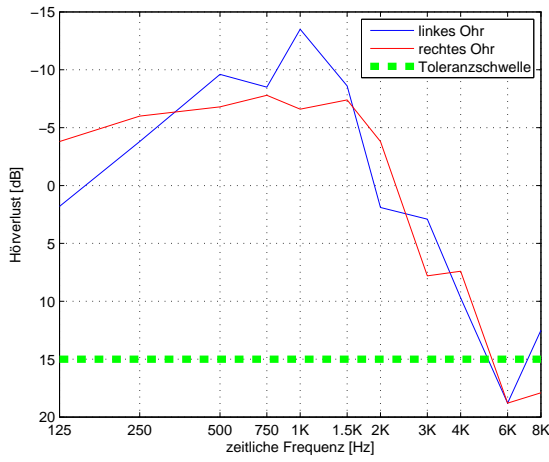
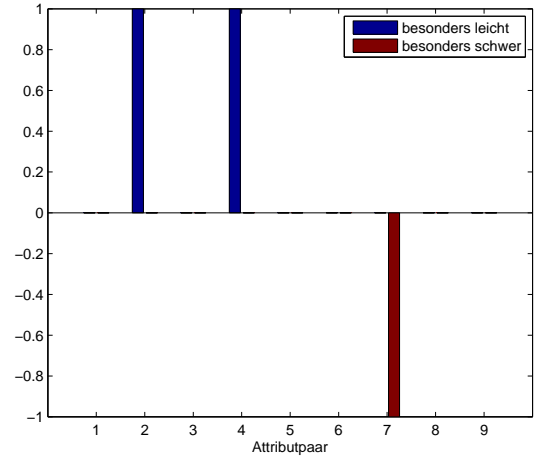


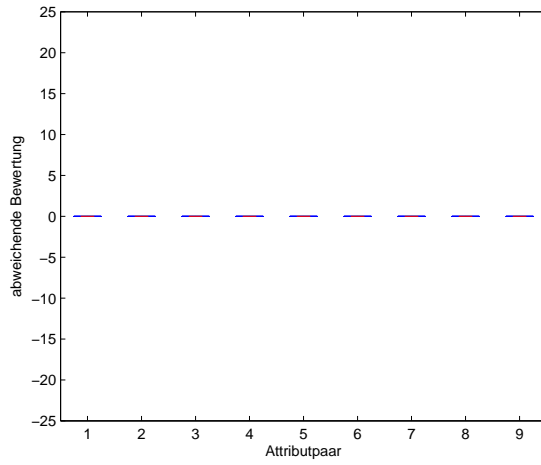
Abbildung 57: Instrumentalerfahrung der ProbandInnen



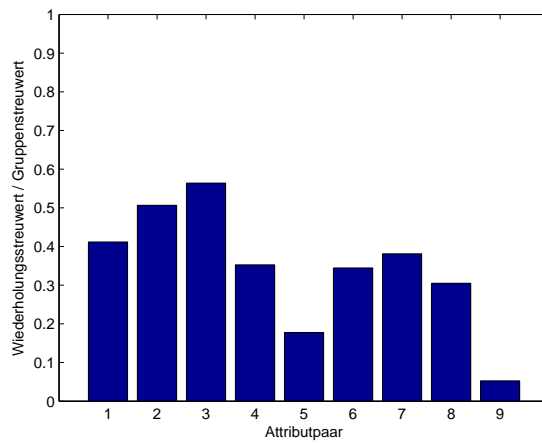
(a) Hörverlust



(b) Bewertung der Attributpaare

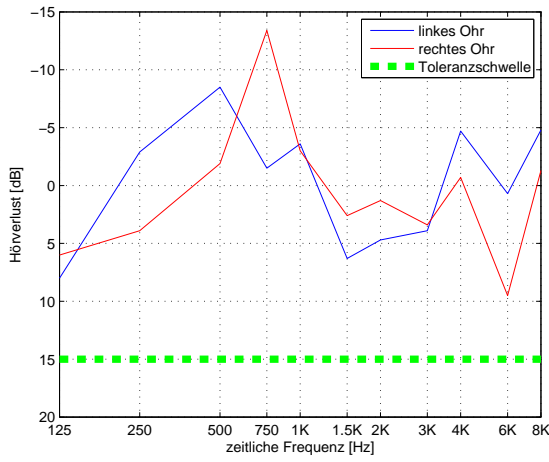


(c) Erkennung der Referenz (Boxplot, nur 25% der Skala sichtbar)

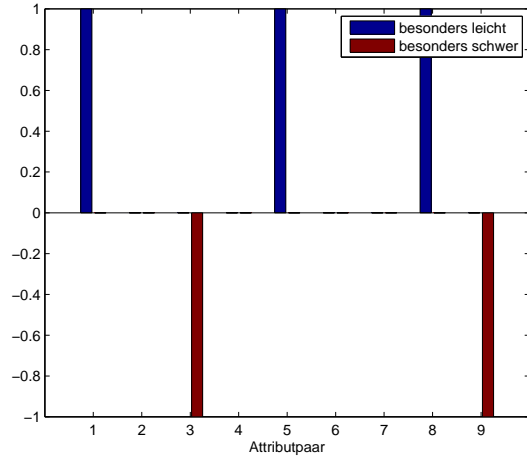


(d) Konsistenz der Antworten

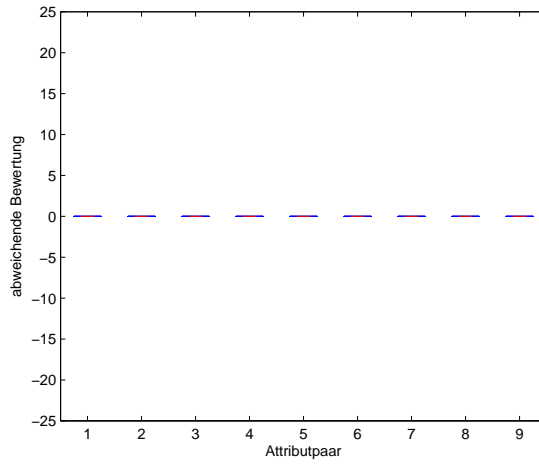
Abbildung 58: Werte für ProbandIn 2



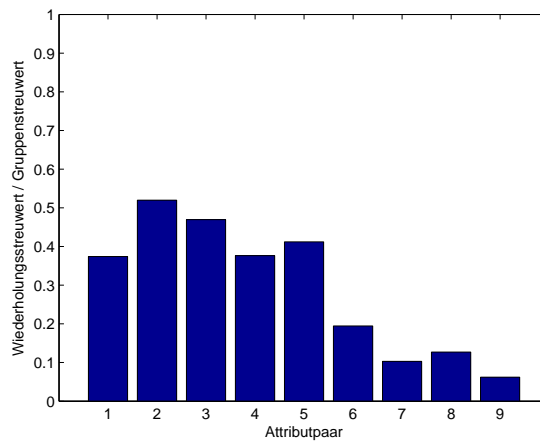
(a) Hörverlust



(b) Bewertung der Attributpaare

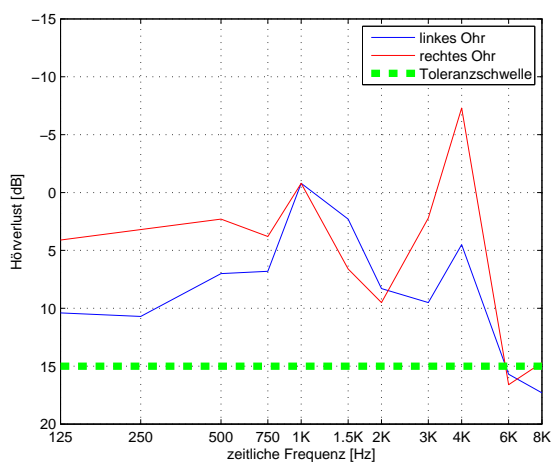


(c) Erkennung der Referenz (Boxplot, nur 25% der Skala sichtbar)

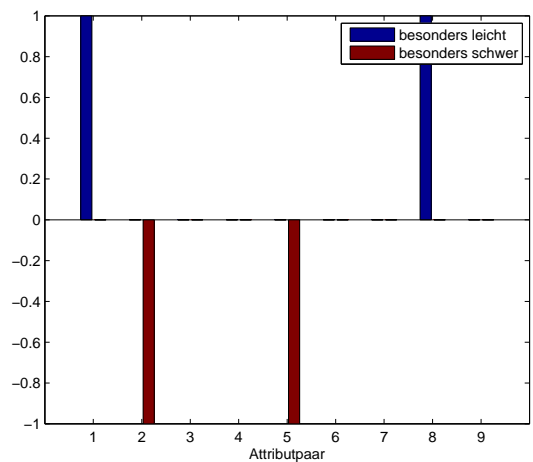


(d) Konsistenz der Antworten

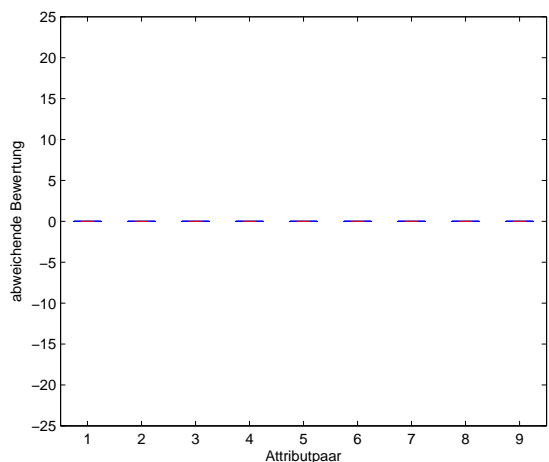
Abbildung 59: Werte für ProbandIn 3



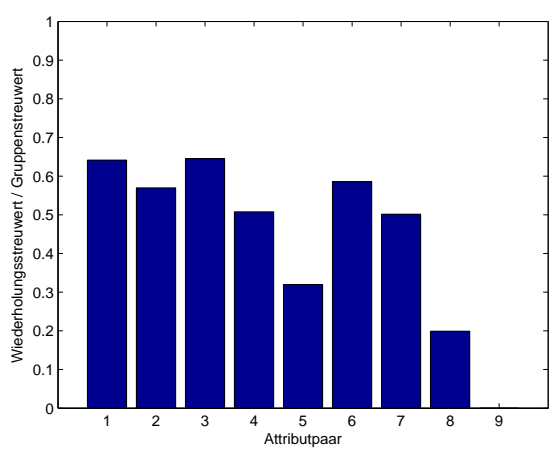
(a) Hörverlust



(b) Bewertung der Attributpaare

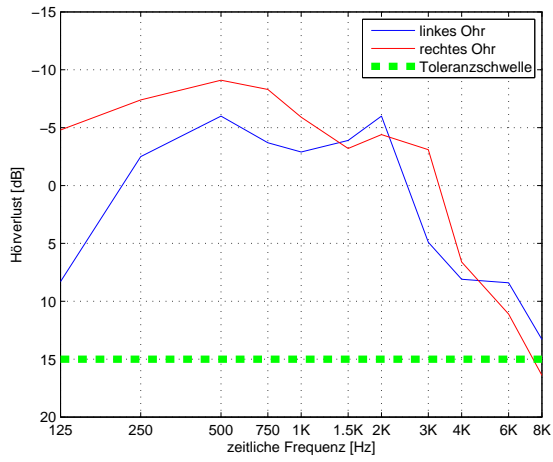


(c) Erkennung der Referenz (Boxplot, nur 25% der Skala sichtbar)

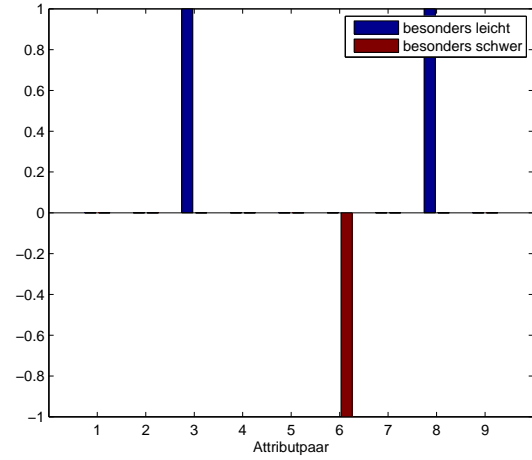


(d) Konsistenz der Antworten

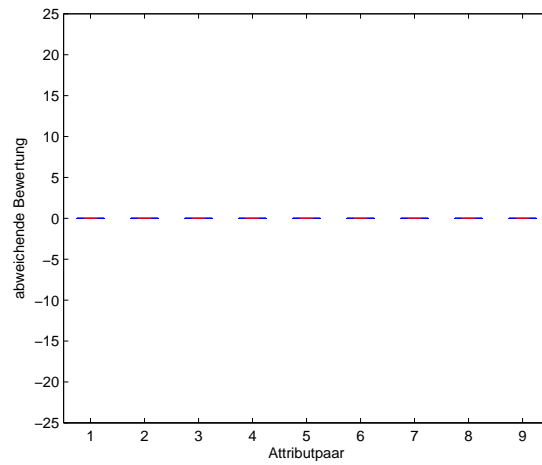
Abbildung 60: Werte für ProbandIn 4



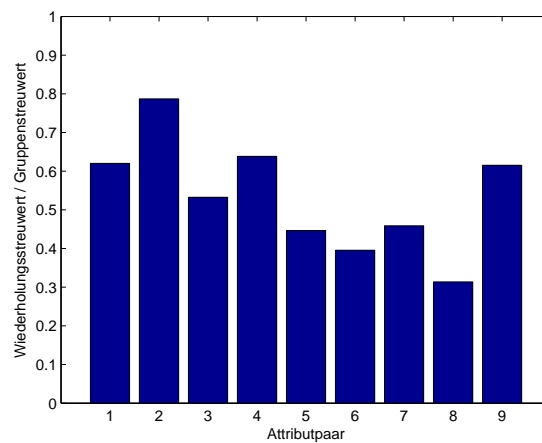
(a) Hörverlust



(b) Bewertung der Attributpaare

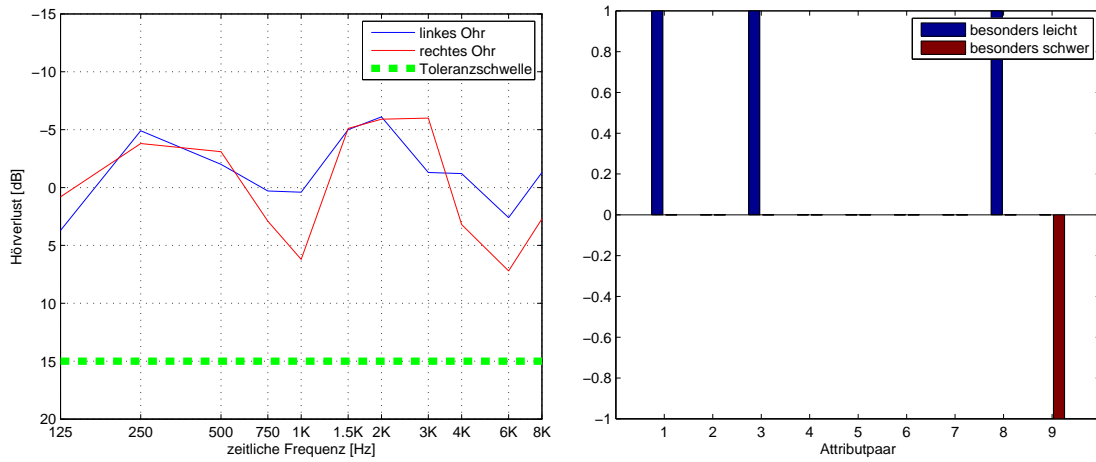


(c) Erkennung der Referenz (Boxplot, nur 25% der Skala sichtbar)



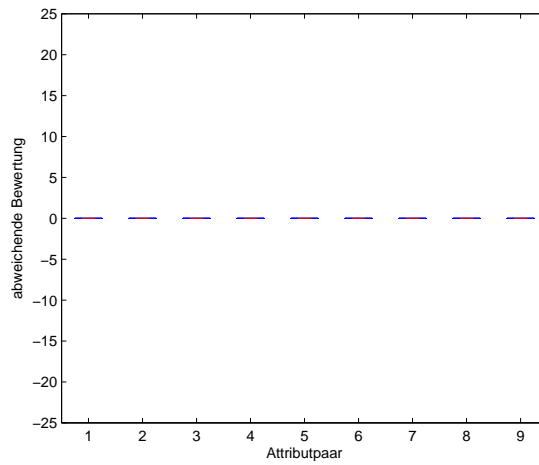
(d) Konsistenz der Antworten

Abbildung 61: Werte für ProbandIn 5

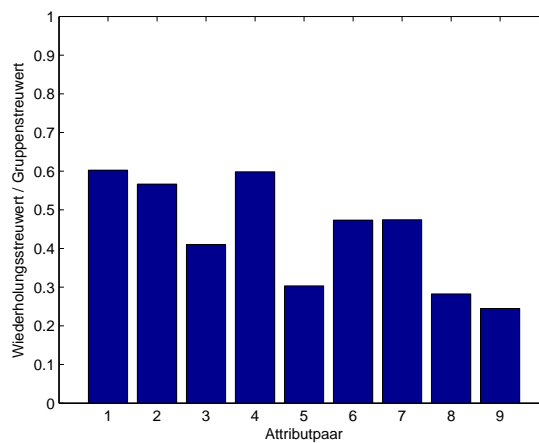


(a) Hörverlust

(b) Bewertung der Attributpaare

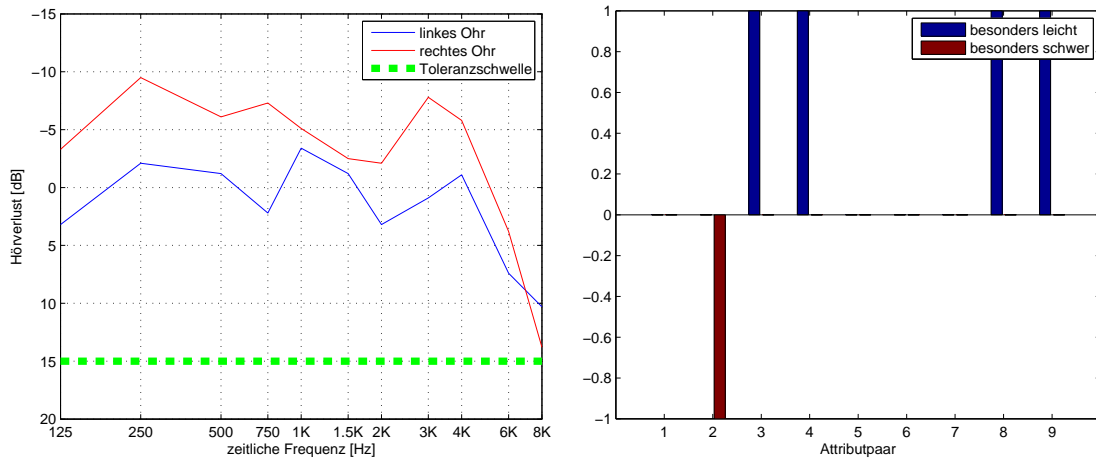


(c) Erkennung der Referenz (Boxplot, nur 25% der Skala sichtbar)



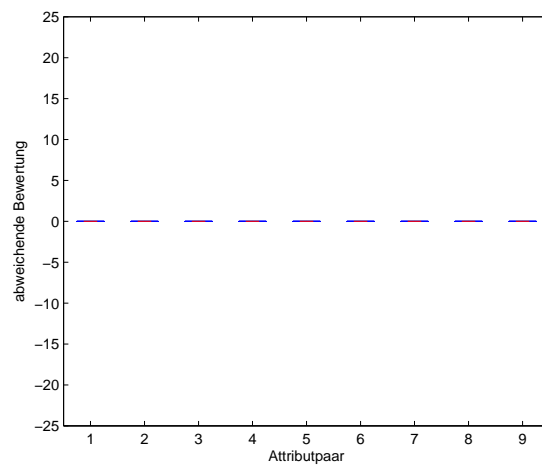
(d) Konsistenz der Antworten

Abbildung 62: Werte für ProbandIn 6

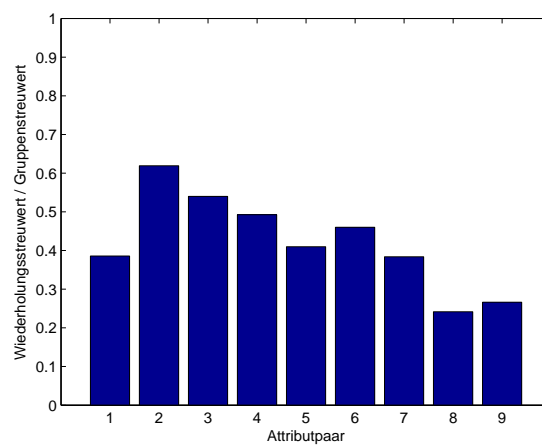


(a) Hörverlust

(b) Bewertung der Attributpaare

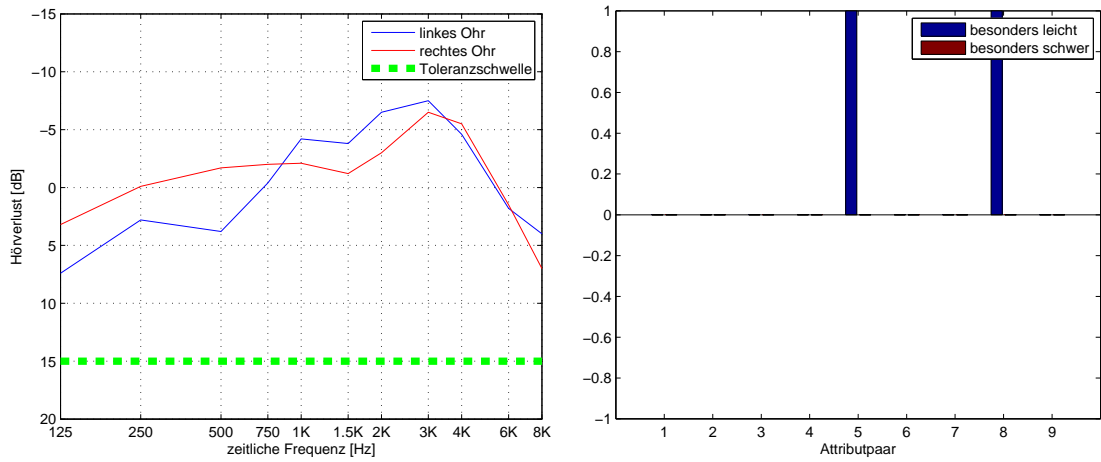


(c) Erkennung der Referenz (Boxplot, nur 25% der Skala sichtbar)



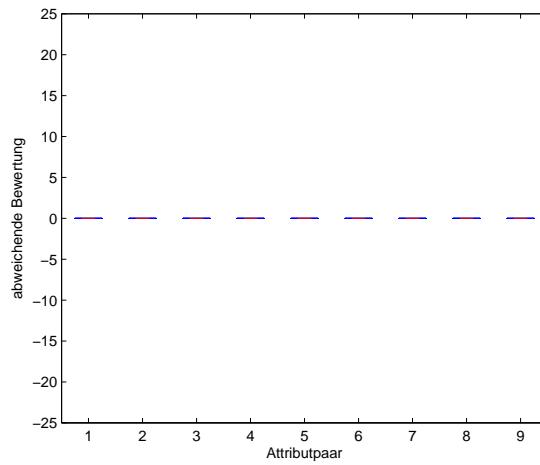
(d) Konsistenz der Antworten

Abbildung 63: Werte für ProbandIn 7

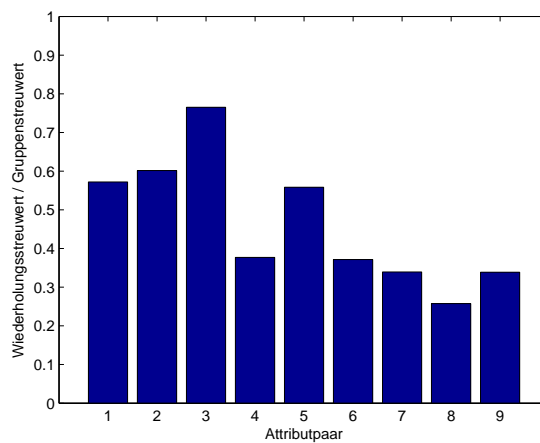


(a) Hörverlust

(b) Bewertung der Attributpaare

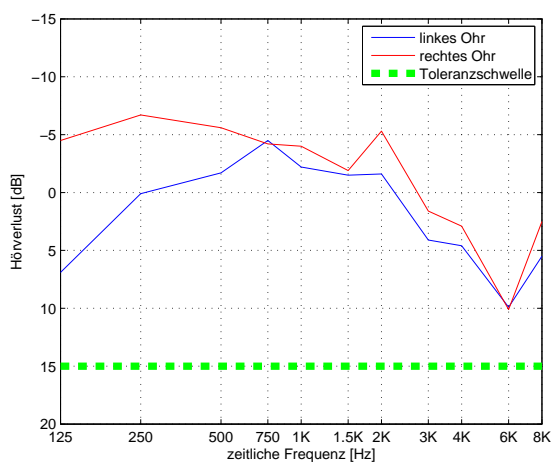


(c) Erkennung der Referenz (Boxplot, nur 25% der Skala sichtbar)

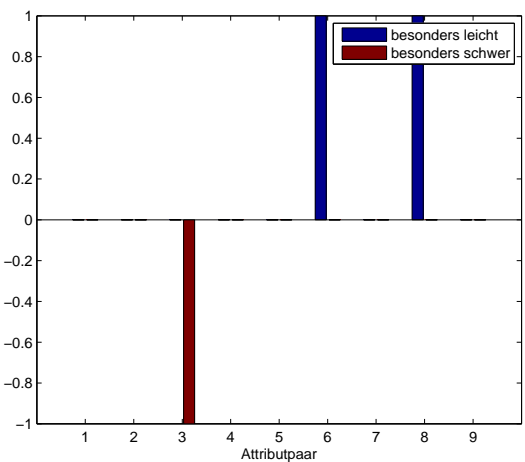


(d) Konsistenz der Antworten

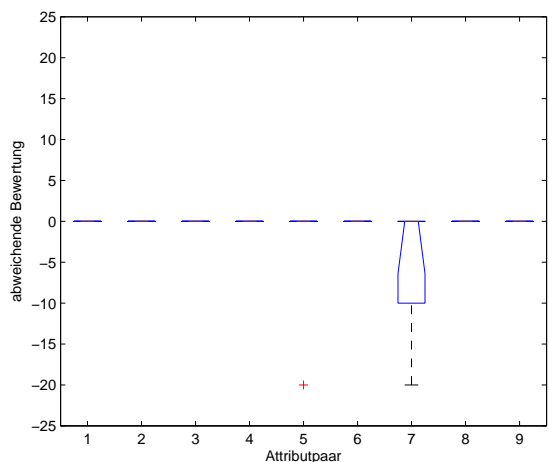
Abbildung 64: Werte für ProbandIn 8



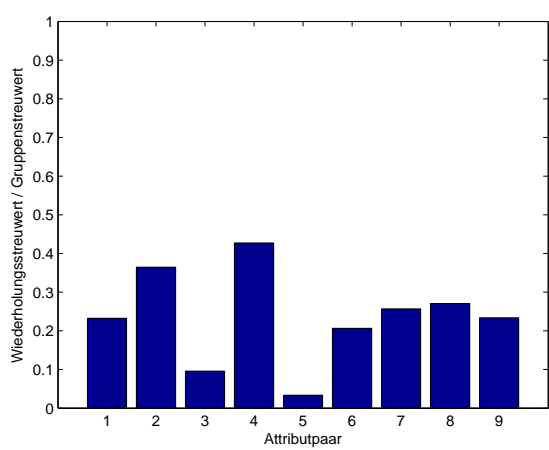
(a) Hörverlust



(b) Bewertung der Attributpaare

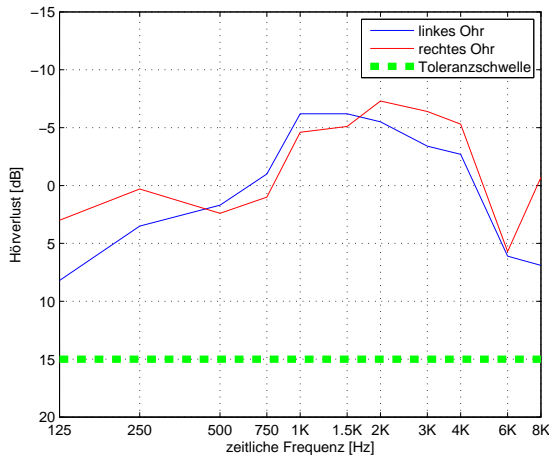


(c) Erkennung der Referenz (Boxplot, nur 25% der Skala sichtbar)

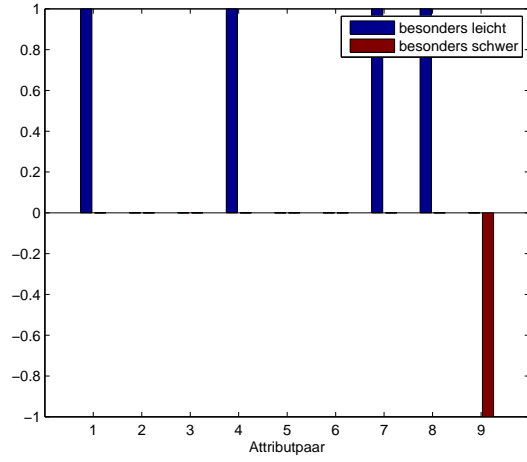


(d) Konsistenz der Antworten

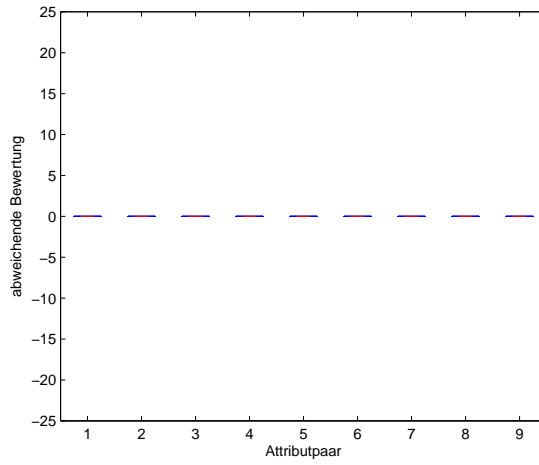
Abbildung 65: Werte für ProbandIn 9



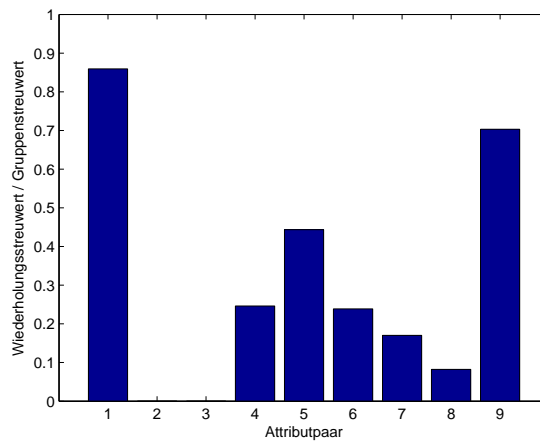
(a) Hörverlust



(b) Bewertung der Attributpaare

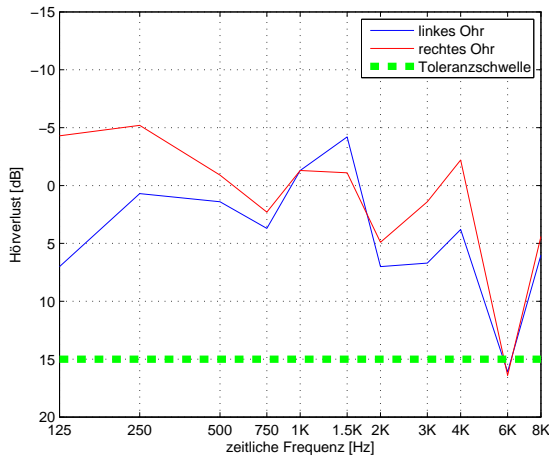


(c) Erkennung der Referenz (Boxplot, nur 25% der Skala sichtbar)

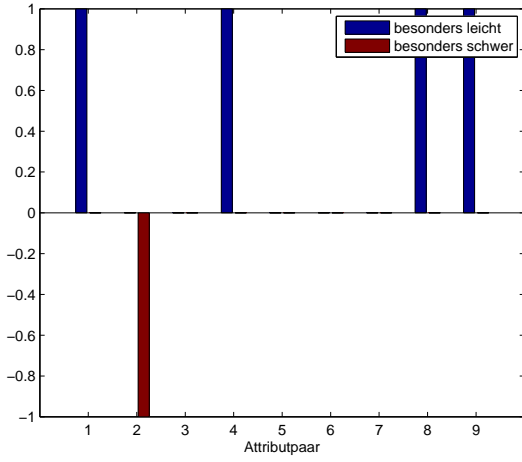


(d) Konsistenz der Antworten

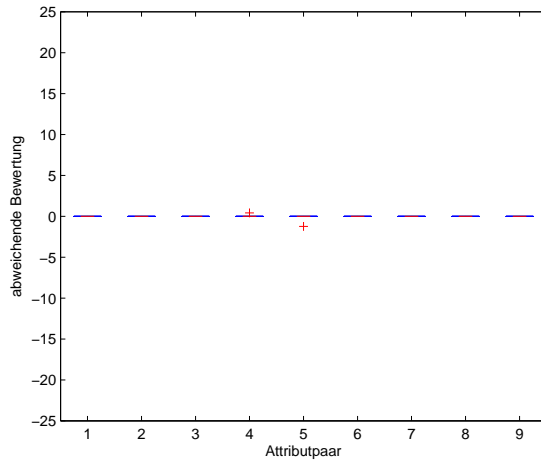
Abbildung 66: Werte für ProbandIn 10



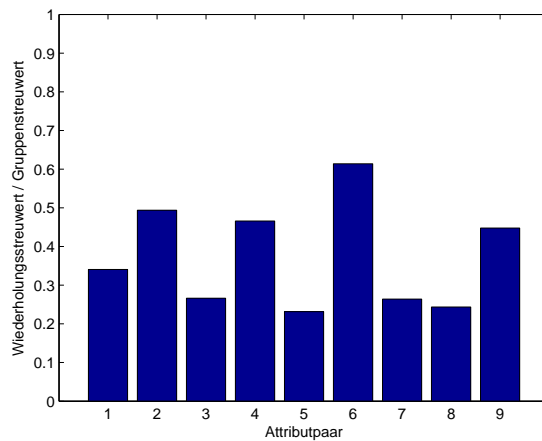
(a) Hörverlust



(b) Bewertung der Attributpaare

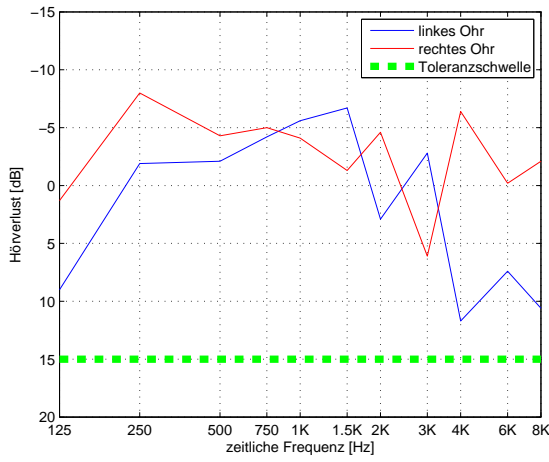


(c) Erkennung der Referenz (Boxplot, nur 25% der Skala sichtbar)

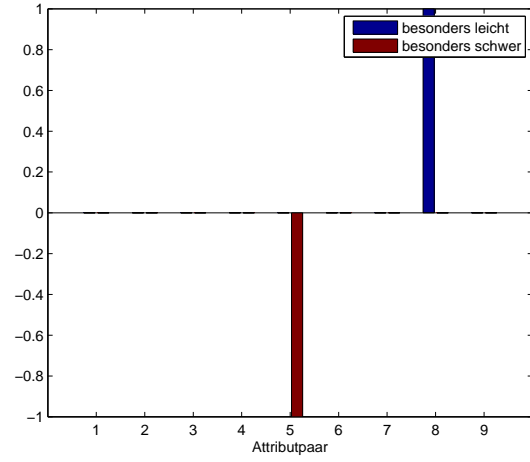


(d) Konsistenz der Antworten

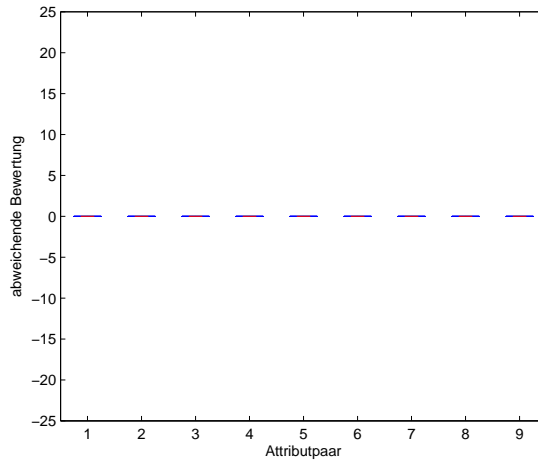
Abbildung 67: Werte für ProbandIn 11



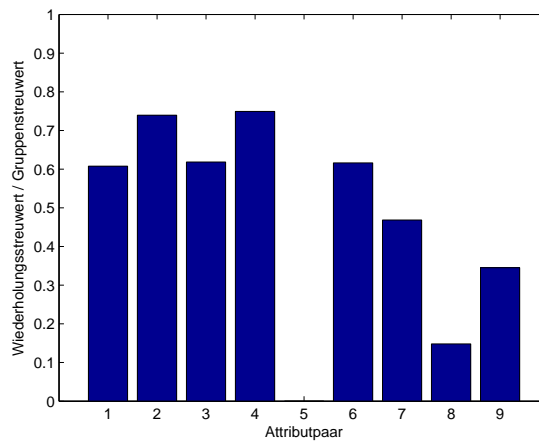
(a) Hörverlust



(b) Bewertung der Attributpaare

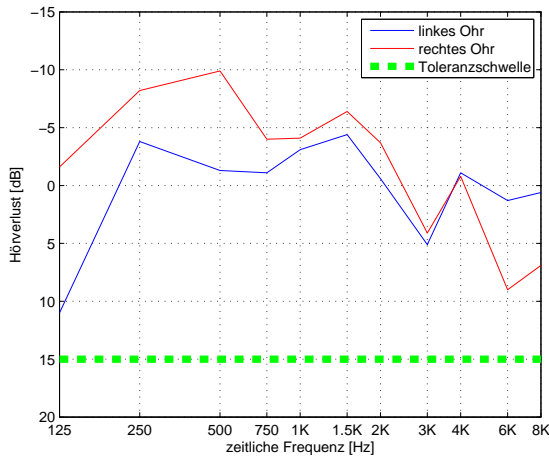


(c) Erkennung der Referenz (Boxplot, nur 25% der Skala sichtbar)

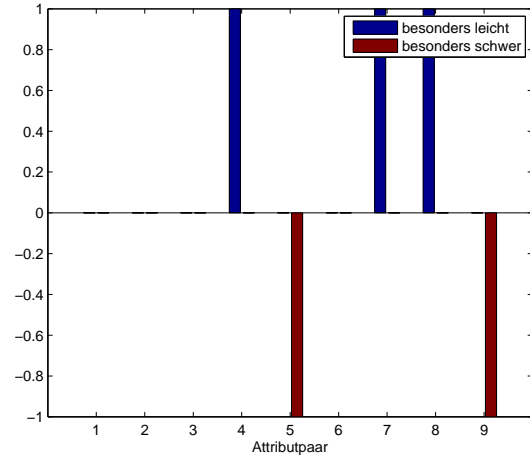


(d) Konsistenz der Antworten

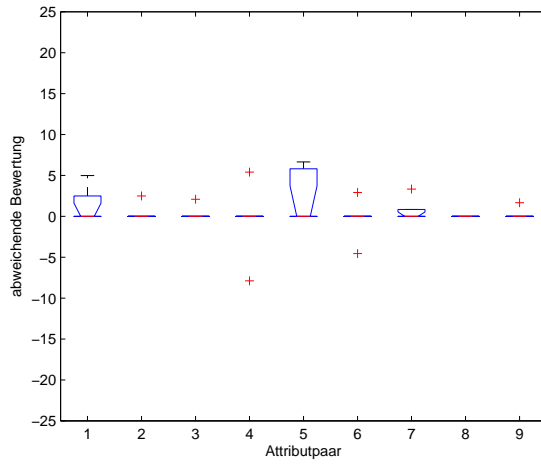
Abbildung 68: Werte für ProbandIn 13



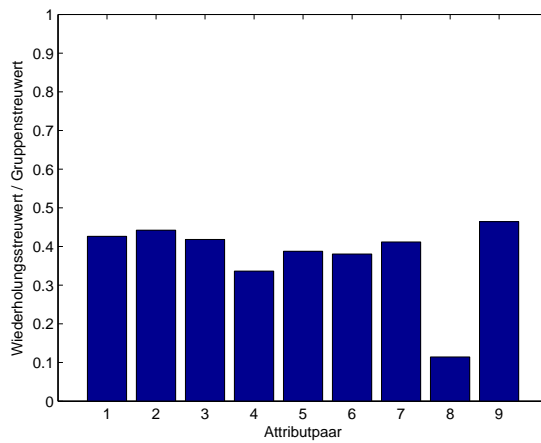
(a) Hörverlust



(b) Bewertung der Attributpaare

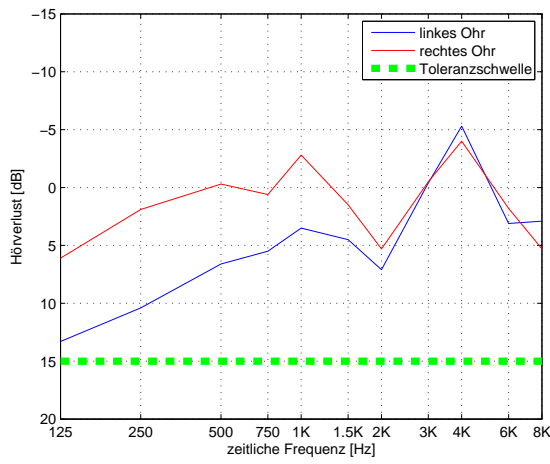


(c) Erkennung der Referenz (Boxplot, nur 25% der Skala sichtbar)

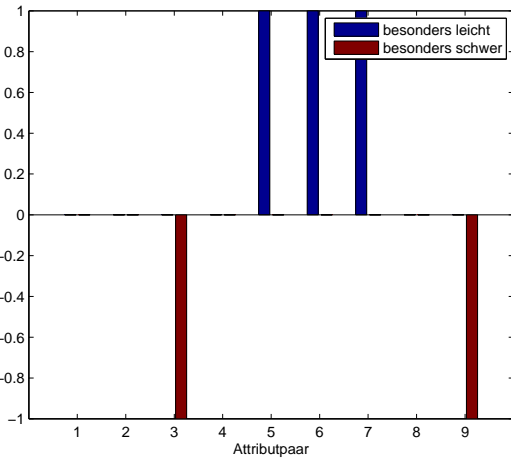


(d) Konsistenz der Antworten

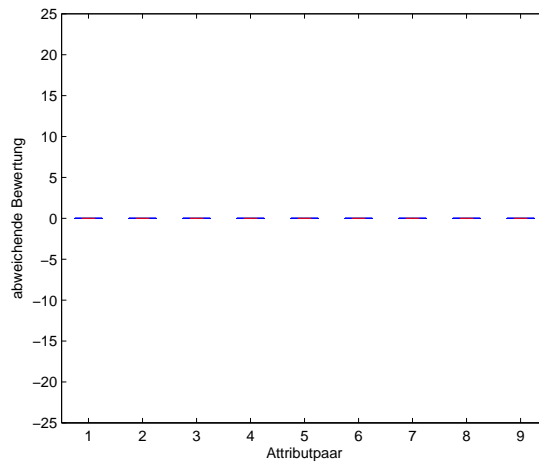
Abbildung 69: Werte für ProbandIn 14



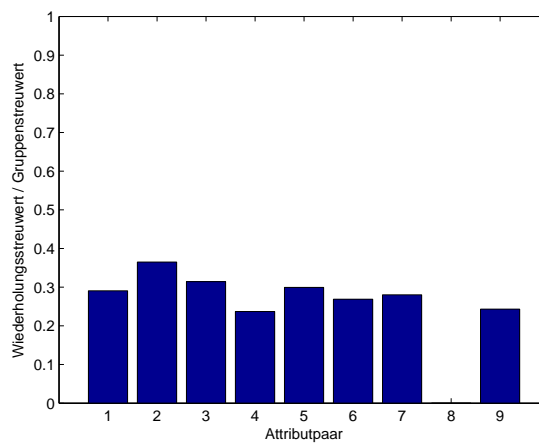
(a) Hörverlust



(b) Bewertung der Attributpaare

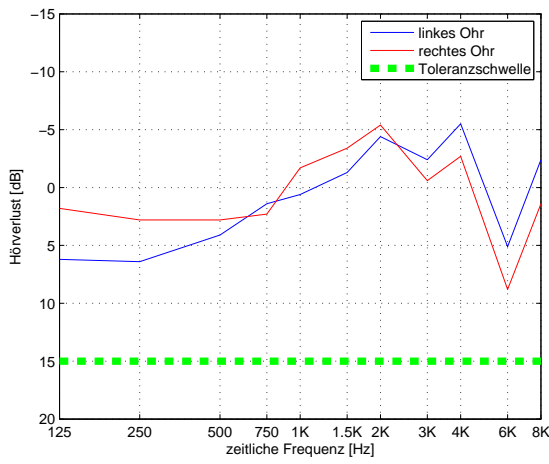


(c) Erkennung der Referenz (Boxplot, nur 25% der Skala sichtbar)

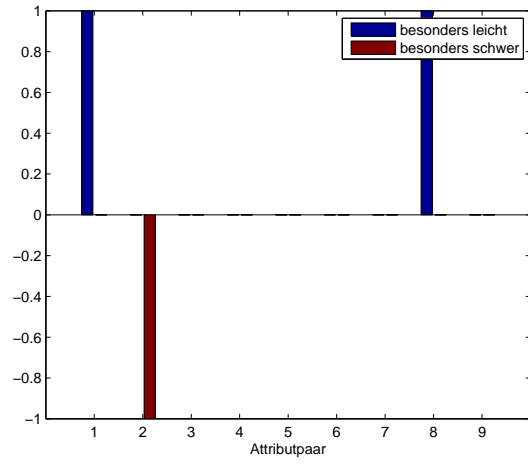


(d) Konsistenz der Antworten

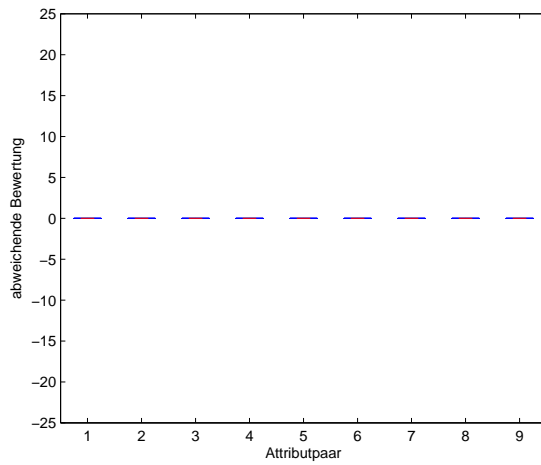
Abbildung 70: Werte für ProbandIn 15



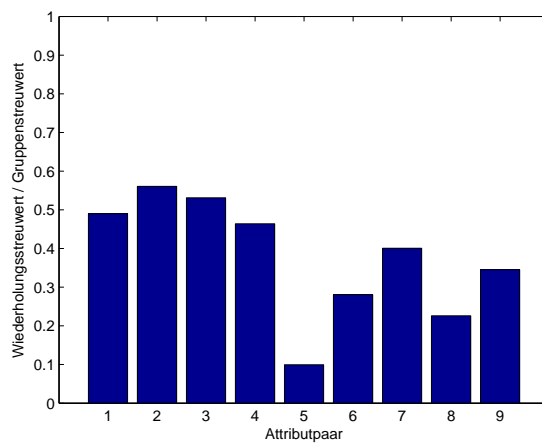
(a) Hörverlust



(b) Bewertung der Attributpaare

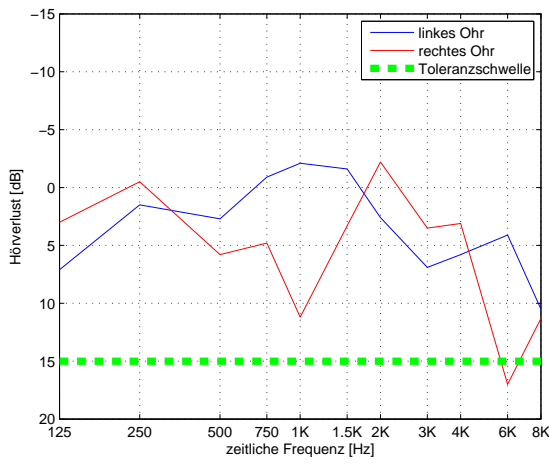


(c) Erkennung der Referenz (Boxplot, nur 25% der Skala sichtbar)

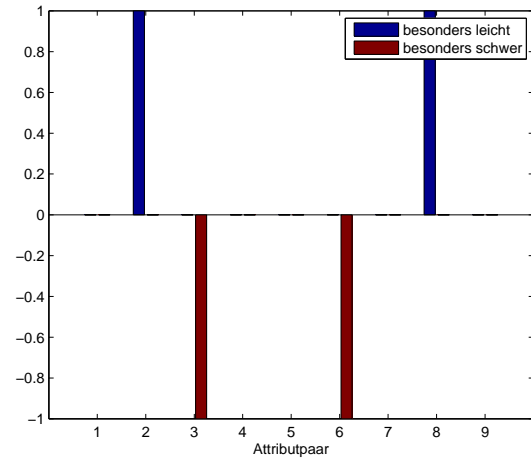


(d) Konsistenz der Antworten

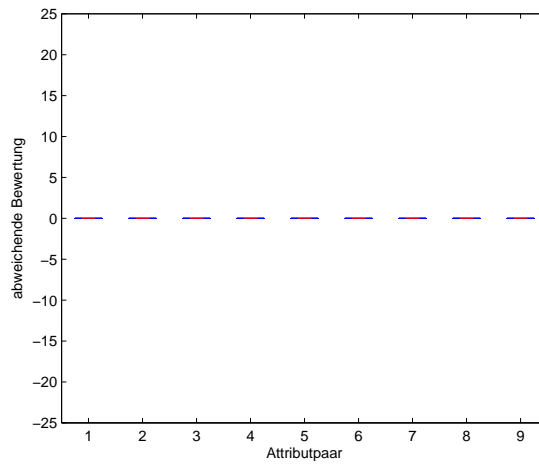
Abbildung 71: Werte für ProbandIn 16



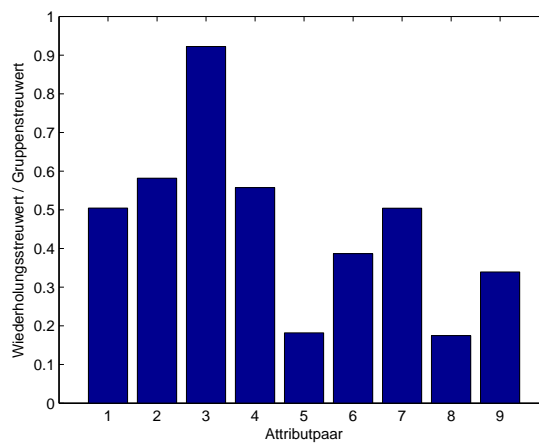
(a) Hörverlust



(b) Bewertung der Attributpaare

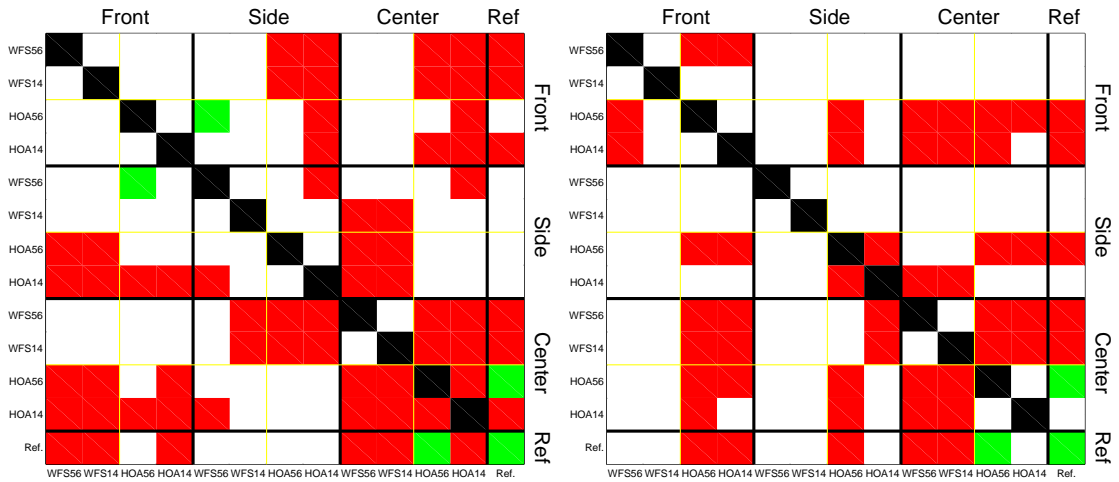


(c) Erkennung der Referenz (Boxplot, nur 25% der Skala sichtbar)



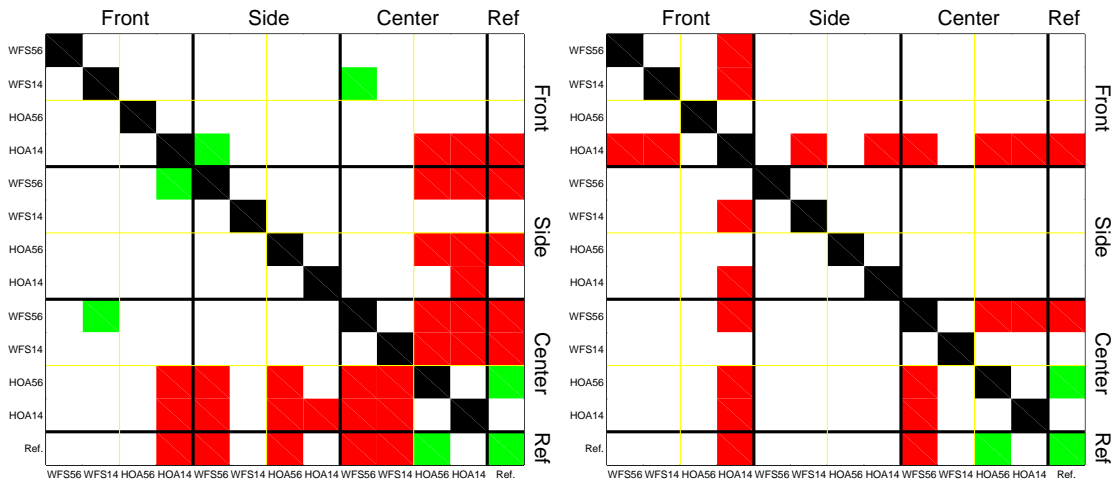
(d) Konsistenz der Antworten

Abbildung 72: Werte für ProbandIn 17



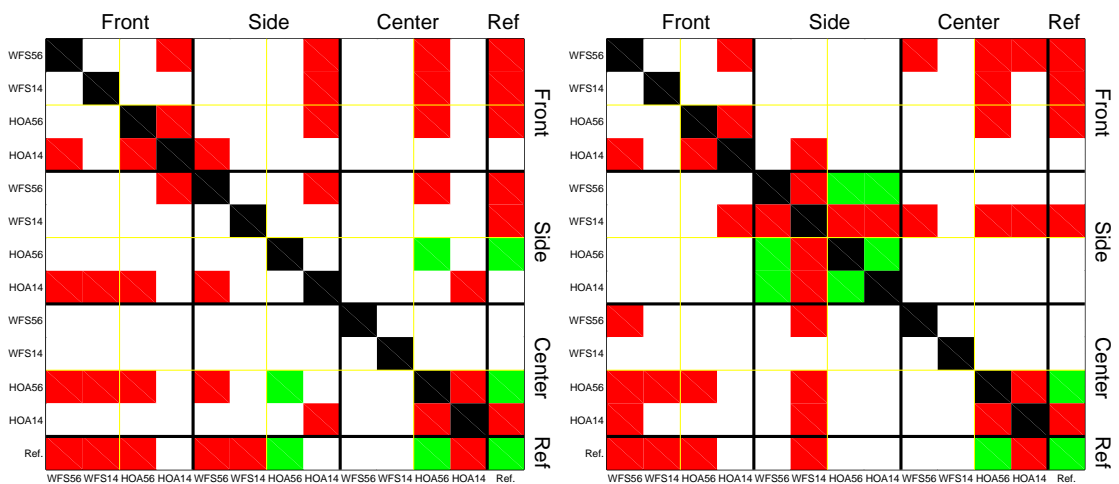
(a) Sprache, direkt

(b) Sprache, voll



(c) Cello, direkt

(d) Cello, voll



(e) Kastagnetten, direkt

(f) Kastagnetten, voll

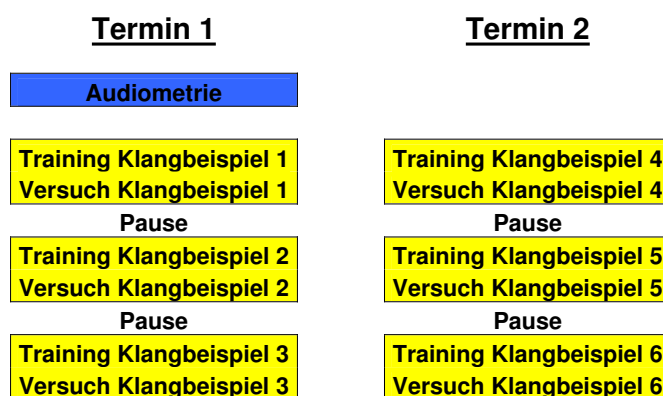
Abbildung 73: Signifikanzmatrizen für Attributpaar 6 (näher - ferner) über alle ProbandInnen für alle einzelnen Kombinationen aus Klangbeispiel und BRIR-Länge

Hörversuch: Vergleich von Schallfeldreproduktionsverfahren

Vielen Dank, dass Sie an diesem Experiment teilnehmen! Bitte nehmen Sie sich die Zeit, um die Anweisungen durchzulesen. Sollten Sie Fragen haben, wenden Sie sich bitte an den Versuchsleiter.

Ihre Aufgabe bei diesem Hörversuch ist es, die dargebotenen Stimuli mit einer gegebenen Referenz zu vergleichen. Dabei wird eine Kopfhörersimulation des Lautsprechersystems in diesem Raum benutzt. Die Unterschiedlichkeit ist in einzelne Attribute aufgeteilt (siehe zweite Seite). Als Klangbeispiele dienen 3 verschiedene Samples (Sprache, Cello, Kastagnetten) mit jeweils 2 verschiedenen Raumanteilen.

Der Gesamtversuch ist auf zwei Termine aufgeteilt, die einen fast identischen Ablauf haben. An jedem Termin werden Vergleiche für 3 Klangbeispiele durchgeführt. Zwischen diesen 3 Teilen sind kurze Pausen geplant.



Vor Beginn jedes Hauptversuchs finden kurze Trainingsphasen statt, um sich mit der Hörumgebung und der Skalierung der Attribute vertraut zu machen.

Die Grundelemente der Bedienoberfläche sind sowohl im Training als auch im Hauptversuch gleich:

Mit einem Klick auf <Play> startet die Wiedergabe des Klangbeispiels in einer ca. einminütigen Schleife. Ein erneuter Klick, auch während der laufenden Wiedergabe, lässt das Beispiel wieder am Anfang starten. Über die Taste <Stop> stoppt die Wiedergabe. Sie kann über <Play> wieder gestartet werden. Die Taste in der rechten unteren Ecke der jeweiligen Oberfläche bringt sie zum nächsten Beispiel bzw. zum nächsten Teil des Versuchs. Um zwischen der Referenz und den verschiedenen Varianten umzuschalten klicken Sie während der Wiedergabe einfach auf die entsprechende Taste. Im Bewertungstraining und im Hauptversuch geschieht das Umschalten über eine einzige Taste. Neben dieser Taste wird angezeigt, welche Variante gerade gehört wird. Dabei steht <Ref> für die Referenz und <X> für die zu vergleichende Variante.

Abbildung 74: Handout zum Versuch (Seite 1/2)

Folgende Attributpaare werden im Versuch zum Vergleich benutzt:

Klangfarben Attribute:

dünnere – vollere

weicher – härter

natürlicher – künstlicher

dunkler – heller

räumliche Attribute:

weniger räumlich – stärker räumlich

näher – ferner

schmäler – breiter

weiter links – weiter rechts

schlechter lokalisierbar – besser lokalisierbar

1. Wie alt sind Sie? _____ Jahre																				
2. Haben Sie schon einmal an einem Hörversuch teilgenommen? <input type="checkbox"/> ja <input type="checkbox"/> nein																				
2a. Falls ja: Welcher Art war dieser Versuch? <input type="checkbox"/> Gehörtest (z.B. beim HNO-Arzt) <input type="checkbox"/> Beurteilung von Klangqualität <input type="checkbox"/> Beurteilung von räumlicher Qualität <input type="checkbox"/> Beurteilung von eigenem Geschmack/Gefallen <input type="checkbox"/> sonstiges: _____																				
3. Welche Studiengänge haben Sie absolviert bzw. studieren Sie gerade? <table border="1" style="width: 100%; border-collapse: collapse;"> <tr> <td style="width: 40%; height: 20px;">_____</td> <td style="width: 15%;"><input type="checkbox"/> fertig</td> <td style="width: 45%;">aktuelles Semester: _____</td> </tr> <tr> <td style="height: 20px;">_____</td> <td><input type="checkbox"/> fertig</td> <td>aktuelles Semester: _____</td> </tr> <tr> <td style="height: 20px;">_____</td> <td><input type="checkbox"/> fertig</td> <td>aktuelles Semester: _____</td> </tr> </table>	_____	<input type="checkbox"/> fertig	aktuelles Semester: _____	_____	<input type="checkbox"/> fertig	aktuelles Semester: _____	_____	<input type="checkbox"/> fertig	aktuelles Semester: _____											
_____	<input type="checkbox"/> fertig	aktuelles Semester: _____																		
_____	<input type="checkbox"/> fertig	aktuelles Semester: _____																		
_____	<input type="checkbox"/> fertig	aktuelles Semester: _____																		
4. Welche Instrumente spielen Sie? <table border="1" style="width: 100%; border-collapse: collapse;"> <tr> <td style="width: 60%; height: 20px;">_____</td> <td style="width: 40%;">seit _____ Jahren</td> </tr> <tr> <td style="height: 20px;">_____</td> <td>seit _____ Jahren</td> </tr> <tr> <td style="height: 20px;">_____</td> <td>seit _____ Jahren</td> </tr> </table>	_____	seit _____ Jahren	_____	seit _____ Jahren	_____	seit _____ Jahren														
_____	seit _____ Jahren																			
_____	seit _____ Jahren																			
_____	seit _____ Jahren																			
5. Bewerten Sie Ihre Erfahrung mit den gegebenen Verfahren auf einer Skala von 1 (keine) bis 5 (sehr viel Erfahrung): <table border="1" style="width: 100%; border-collapse: collapse; text-align: center;"> <thead> <tr> <th style="width: 25%;"></th> <th style="width: 15%;">Stereo</th> <th style="width: 15%;">Surround</th> <th style="width: 15%;">Wellenfeldsynthese</th> <th style="width: 15%;">Ambisonics</th> </tr> </thead> <tbody> <tr> <td style="text-align: left;">Hören</td> <td></td> <td></td> <td></td> <td></td> </tr> <tr> <td style="text-align: left;">Aufnehmen</td> <td></td> <td></td> <td></td> <td></td> </tr> <tr> <td style="text-align: left;">Mischen/Mastern</td> <td></td> <td></td> <td></td> <td></td> </tr> </tbody> </table>		Stereo	Surround	Wellenfeldsynthese	Ambisonics	Hören					Aufnehmen					Mischen/Mastern				
	Stereo	Surround	Wellenfeldsynthese	Ambisonics																
Hören																				
Aufnehmen																				
Mischen/Mastern																				
6a. Bei welchen Attributpaaren fiel Ihnen die Bewertung besonders leicht? <hr style="border: 0; border-top: 1px solid black; margin-bottom: 5px;"/> <hr style="border: 0; border-top: 1px solid black; margin-bottom: 5px;"/> <hr style="border: 0; border-top: 1px solid black;"/>																				
6b. Bei welchen Attributpaaren fiel Ihnen die Bewertung besonders schwer? <hr style="border: 0; border-top: 1px solid black; margin-bottom: 5px;"/> <hr style="border: 0; border-top: 1px solid black; margin-bottom: 5px;"/> <hr style="border: 0; border-top: 1px solid black;"/>																				
7. Haben Sie noch weitere Kommentare zu diesem Hörversuch? <hr style="border: 0; border-top: 1px solid black; margin-bottom: 5px;"/> <hr style="border: 0; border-top: 1px solid black; margin-bottom: 5px;"/> <hr style="border: 0; border-top: 1px solid black;"/>																				

Abbildung 76: Fragebogen zum Versuch