

# Untersuchung zur Differenzierbarkeit digitaler Aufzeichnungsverfahren

## Hörvergleich *Direct Stream Digital (DSD)* und *High-Resolution-PCM (24bit/176,4kHz)*

Dominik Blech und Min-Chi Yang

Erich-Thienhaus-Institut (Tonmeisterinstitut), Hochschule für Musik Detmold

<http://www.hfm-detmold.de/hochschule/eti.html>

---

### 1. Einleitung

Die in technischer und klanglicher Diskussion nicht immer einvernehmlich, sondern vielmehr sehr kontrovers verlaufende Koexistenz von derzeit zwei digitalen Aufzeichnungsverfahren, „Puls Code Modulation“ (PCM) und „Direct Stream Digital“ (DSD) brachte auch zwei Formate hervor: Die auf PCM basierende **DVD-Audio** und die auf einem 1-Bit-Signal mit 64-facher CD-Abtastrate (=2,8224 MHz) basierende **Super Audio Compact Disc (SACD)**. Sowohl von Hersteller- als auch von Anwender- und Konsumentenseite werden hierbei immer wieder angeblich klangliche und dadurch deutlich wahrnehmbare Unterschiede herausgestellt. Welches der miteinander konkurrierenden Systeme sich langfristig durchsetzen wird, ist noch immer unklar.

Im Rahmen der vorliegenden Untersuchung soll mit Hilfe eines Hörtests eruiert werden, inwieweit Probanden bei einem ABX-Test einen Unterschied zwischen DSD und hochauflösendem PCM (176,4kHz/24Bit) wahrnehmen können. Schon an dieser Stelle sei darauf hingewiesen, dass sich im Laufe des Hörvergleichs immer deutlicher herauskristallisiert hat, wie wichtig ein Testverfahren in Form eines ABX-Doppel-Blindtests ist, da nur so – frei von Suggestion und unterbewusster Voreingenommenheit – gezeigt werden kann, was noch reproduzierbar wahrnehmbar ist – bekanntermaßen ist der Übergangsbereich zwischen auditiver Perzeption und Imagination sehr schmal.

Mit der Durchführung dieser Arbeit soll das bestehende Missverhältnis zwischen der umfangreichen theoretischen Datenlage und dem praktischen Nachweis anhand eines schon lange ausstehenden (offiziellen und den wissenschaftlichen Forderungen nach Reliabilität, Validität und Objektivität genügenden) Hörvergleichs verringert werden und hierzu ein objektiver und sachlicher Beitrag geleistet bzw. ein Anstoß für weitere (klangliche) Untersuchungen und Überlegungen gegeben werden.

### 2. Vorüberlegungen

Da der Hörvergleich auf das hinter dem reinen Tonträgerformat oder dessen Medium stehende digitale Kodierungsverfahren zurückgeführt werden soll, ist es notwendig, Aufnahme- und Wiedergabepfade zu realisieren, die – soweit möglich – auf gemeinsamen Audio-Komponenten basieren, um nicht unterschiedliches Equipment, sondern das Aufzeichnungsverfahren als solches klanglich zu vergleichen. Der unvermeidbare Schwachpunkt befindet sich bei den A/D- und D/A-Konvertern – eine Minimierung kann daher lediglich durch die Verwendung eines Wandlers desselben Herstellers und desselben Modells geschehen, welches beide Kodierungsverfahren unterstützt.

Weiterhin muss ein praktikabler Testablauf gewährleistet sein, der Latenzen, die z.B. durch die Synchronisation zweier unabhängiger Audio-Workstation-Systeme entstehen würden, beim Umschalten zwischen den beiden Ursprungsquellen (DSD und PCM) vermeidet. Das Differenzieren zwischen den Quellen auf rein zeitlicher Ebene wäre ansonsten denkbar und möglich. Als Lösung bietet sich das sogenannte Data-Bitmapping-Verfahren an, das bei der A/D-Konvertierung die Daten in einem Non-Audio-Format vollkommen verlustfrei – im vorliegenden Fall – in 24Bit/44,1kHz-Daten „verpackt“ und D/A-seitig diese nach demselben Algorithmus wieder „entpackt“.

Um einerseits der durchschnittlich größeren Hörerfahrung des potentiellen Testpublikums mit Stereowiedergabe gerecht zu werden, andererseits aber auch eventuell wahrnehmbare Auswirkungen auf räumliche Komponenten mit einzubeziehen, sollte der Hörvergleich sowohl in Stereo als auch in Surround stattfinden. Essentiell hierbei ist, dass die eigens dafür aufgenommenen Musik- und Klangbeispiele absolut identisch sind, folglich in keiner Weise bearbeitet wurden (durch Pegeländerung, Schnitt, Mix, etc.). Aus diesem Grund muss ebenfalls ein klangbeeinflussendes Mischpult umgangen, resultierend daraus das Audiosignal der entweder zwei oder fünf Mikrofone unverfälscht aufgezeichnet und anschließend entsprechend „geroutet“ über die Lautsprecher wiedergegeben werden.

Da jeder musikausübende und/oder -hörende Mensch seine ganz eigenen Hörerfahrungen, -erwartungen und -fokussierungen mit individuellen, instrumentenbezogenen Vorlieben besitzt, sollte nicht nur der Testpersonenkreis, sondern auch das Angebot an zur Verfügung stehenden Musik- und Klangbeispielen möglichst breit-

bandig gewählt werden. Aus diesem Individualitätsgedanken heraus ist es unumgänglich, jeden Probanden einzeln zu testen.

Es sollte durch Schaffung einer möglichst angenehm neutralen räumlichen, wie auch persönlichen Atmosphäre (mit optional einzulegenden Pausen) – ohne vorherige Beeinflussung des Kandidaten – versucht werden, dem Leistungsdruck, dem sich der Proband bei einem subjektiven Test automatisch aussetzt, weitestgehend entgegenzuwirken.

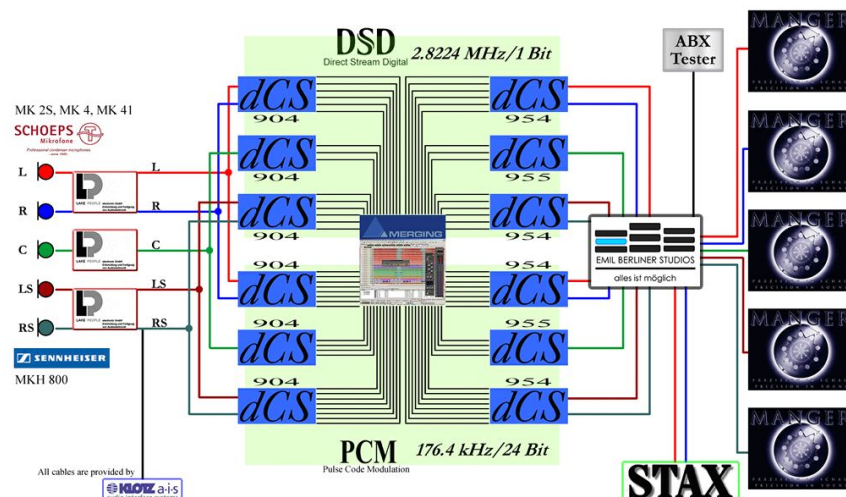
### 3. Versuchsaufbau

Wie schon im vorausgegangenen Kapitel erwähnt, ist eine grundsätzliche Bedingung für einen objektiven und technisch einwandfreien Hörvergleich, dass das zu vergleichende Musikmaterial vollkommen identisch und „unbehandelt“ ist. Das heißt, es darf weder im Pegel geändert, noch darf es editiert, verhallt oder auf eine andere Weise bearbeitet werden (in diesem Fall würde DSD rechnerintern in ein PCM-Signal gewandelt werden). Da solches Musikmaterial – wenn überhaupt vorhanden – nicht zugänglich war, wurde dieses zusammen mit Instrumentalisten der Hochschule für Musik Detmold als Stereo- und Surroundbeispiele in der dem Erich-Thienhaus-Institut angegliederten „Neuen Aula“ bei deaktivierter Klimaanlage unter optimierten Bedingungen eigens vor Beginn der Hörtest von den beiden Autoren parallel in DSD und in PCM (176,4kHz/24Bit) produziert.

Wie bereits erwähnt, wurden, um ein klangbeeinflussendes Mischpult umgehen zu können, die Stereo-Musikbeispiele mit je zwei, Surroundbeispiele mit je fünf Mikrofonen (mit erweitertem Frequenzgang bis 40 bzw. 50 kHz) der Firmen Schoeps (MK 2S, MK 4 und MK 41 mit den Mikrofonverstärkern CMC 6-Ug xt) und Sennheiser (MKH 800) aufgenommen – folglich war also jedem Lautsprecher ein Mikrofon zugeordnet. Die mit Lake People-Mikrofonvorverstärkern des Typs MIC AMP F/35II auf Line-Pegel angehobenen Signale wurden über 50 Meter lange, niederkapazitive Mikrofonskabel der Firma Klotz (Typ der Serie M1) in den Regieraum geführt, dort mittels eines Y-Adapters aufgesplittet und über dCS-Konverter (insgesamt 2 mal 3 zweikanalige dCS 904) einerseits in DSD, andererseits in PCM (24bit, 176,4kHz) A/D-gewandelt und nach dem Datenbitmapping-Verfahren der Konverter als 24bit/44,1kHz-Daten mit einem Pyramix „Virtual-Studio-System“ (Merging Technologies) gespeichert.

Anschließend wurde das Audiosignal ebenfalls über dCS-D/A-Wandler (insgesamt 2 mal 2 zweikanalige dCS 954 für L, R, LS, RS und 2 mal 1 zweikanaligen dCS 955 für C) in ein analoges Signal rückkonvertiert und auf eine von den Emil-Berliner-Studios entwickelte, hochwertige Stereo- und Surround-Monitorunit vom Typ MU 2000 gegeben. Auf diese konnte der Proband mittels einer ebenfalls bei den Emil-Berliner-Studios entwickelten ABX-Software zugreifen und jeweils zwischen den Signalen DSD und PCM auswählen. Mit Hilfe eines softwareseitig im ABX-Test implementierten, auf den Umschaltmoment zwischen zwei Signalen bezogenen Delays, wurde ein gleichzeitiges Erklingen beider Signale und somit eine den Probanden irreführende Pegel-Kaskadierung unterbunden. Als Wiedergabesystem wurden Schallsysteme der Firma Manger eingesetzt, die sich durch eine hervorragende Impulsgenauigkeit und einen Frequenzgang bis 35kHz auszeichnen. Hatte der Proband sich für ein Stereobeispiel entschieden, konnte er darüber hinaus über einen Kopfhörer der Firma Stax hören. Die gesamte Verkabelung wurde ausnahmslos mit neuen, hochwertigen Analog- und Digitalkabeln der Firma Klotz ausgeführt.

Mikrofon - LakePeople Mikrofonvorverstärker - dCS A/D Konverter 904 - Merging Pyramix Virtual Studio - dCS D/A Konverter 955/954 - EBS Monitor Unit & ABX Tester - Manger Schallsysteme



Um den unterschiedlichen Hörerfahrungen, -erwartungen und -fokussierungen auf musikalischer Seite gerecht werden zu können, sollte dem möglichst breitgefächerten Testpublikum während der Durchführung der Hörversuche selbstverständlich ein entsprechend vielfältiges Angebot an Musik- und Klangbeispielen zur Verfügung stehen. Die folgende Tabelle zeigt eine Übersicht der aufgenommenen Musikbeispiele:

STEREO	SURROUND	STEREO	SURROUND
<b>Harpsichord</b> F.Couperin – <i>Rondeau c-Moll</i> 3:21	<b>Harpsichord</b> Fr.Couperin – <i>Rondeau</i> 3:25	<b>Percussion Solo</b> <i>Maracas</i> 1:30 <i>Guiro</i> 1:30 <i>Windchime</i> 1:30 <i>Kastagnetten</i> 1:30	<b>Percussion Trio</b> N.J.Zivkovic – <i>aus Danza Barbara</i> 2:33 <i>Tutti-Bsp.Nr.1</i> 2:33 <i>Tutti-Bsp.Nr.2</i> 1:33
<b>Vocal</b> W.A.Mozart – <i>Le nozze di Figaro, Arie der Susanna „Deh vieni, non tardar“</i> 3:11	<b>Vocal</b> J.Strauss – <i>DieFledermaus, Lied der Adele, „Mein Herr Marquis“</i> 1:32	<b>Piano</b> D.Scarlatti – <i>Sonata, K.188 a-Moll</i> 2:38	<b>Piano</b> Fr.Chopin – <i>Études, op.25, Nr.11 a-Moll</i> 4:04
<b>Guitar</b> E.Clapton – <i>Signe</i> 2:06	<b>Guitar</b> Unbekannt – <i>Romance</i> 2:20	<b>Speech (Russian)</b> A.Puschkin – <i>aus Eugen Onegin</i> 2:08	<b>Speech (Russian)</b> A.Puschkin – <i>aus Eugen Onegin</i> 2:09
	<b>Jazz Trio</b> M.Manieri – <i>Sarah’s Touch</i> 4:32		<b>String Orchestra</b> E.Rautavaara – <i>Pelimannit „Fiddlers“</i> 1:08 <i>2.Satz (Presto)</i> 1:08 <i>5.Satz (Presto)</i> 1:16
<b>Oboe</b> G.Ph.Teleman – <i>Phantasie Nr.8, 2. Satz (Spirituoso)</i> 1:11	<b>Oboe</b> G.Ph.Teleman – <i>Phantasie Nr.8, 1. Satz (Largo)</i> 2:22 <i>2.Satz (Spirituoso)</i> 1:14	<b>Trumpet</b> <i>Blues Improvisation</i> 2:31	<b>Trumpet</b> <i>Blues Improvisation</i> 2:36
	<b>Organ</b> M.Reger – <i>Fuge d-Moll, op.135b</i> 5:10		<b>Violin</b> J.S.Bach – <i>Sonate Nr.1, BWV 1001, Adagio</i> 4:19

Die Musik- und Klangbeispiele wurden den Probanden in voller Spiellänge zur Verfügung gestellt. Die gesamte Bedienung sowie der Zugriff auf die ABX-Testsoftware etc., erfolgte vom Probanden mittels einer Steuerungseinheit, so dass dieser selber über den Ablauf und die zeitliche Einteilung des Hörvergleichs bestimmen konnte. Diese Möglichkeit war ein wichtiger Faktor, um den schon erwähnten Leistungsdruck des Probanden möglichst zu minimieren.

Der gesamte Testaufbau wurde unter Beachtung absoluter Pegelgleichheit zwischen beiden Signalzweigen durchgeführt. Damit sämtliche A/D- und D/A-Konverter unter identischen Bedingungen arbeiteten, wurden sie vor Beginn der Tests exakt eingemessen.

#### 4. Hörraum

Der Hörraum wurde gemäß der Richtlinien für Abhörräume der European Broadcasting Union (EBU)<sup>1</sup> und der International Telecommunication Union (ITU)<sup>2</sup> bezogen auf Nachhallzeit, Hintergrundgeräusche und Referenz-Abhörpegel akustisch angepasst und eingemessen.

Entsprechend den Empfehlungen der ITU für die Lautsprecheranordnung von L, C, R, LS und RS bei Multi-Channel-Soundsystemen wurden die Manger-Lautsprechersysteme auf einem Kreis angeordnet. Dabei betrug die Basisbreite B zwischen L und R – sowie daraus folgend ebenso der Kreisradius – 2,20m. Die Lautsprechersysteme LS und RS wurden in einem 110°-Winkel zur Mittelsenkrechten (Referenzabhörpunkt – Center) äquidistant positioniert.

#### 5. Versuchsablauf

Der ABX-Test ist ein Testverfahren, das ermöglicht, zwei unterschiedliche Audiosignale A und B „blind“ miteinander zu vergleichen. Ziel des Testverfahrens ist, das Signal X, welches von einer ABX-Software mit Hilfe eines Zufallsgenerators nach jeder Entscheidung neu auf A oder B zugewiesen wird, mit den Originalen A und B klanglich zu vergleichen und anschließend dementsprechend zuzuordnen (X=A oder X=B). Zwischen den

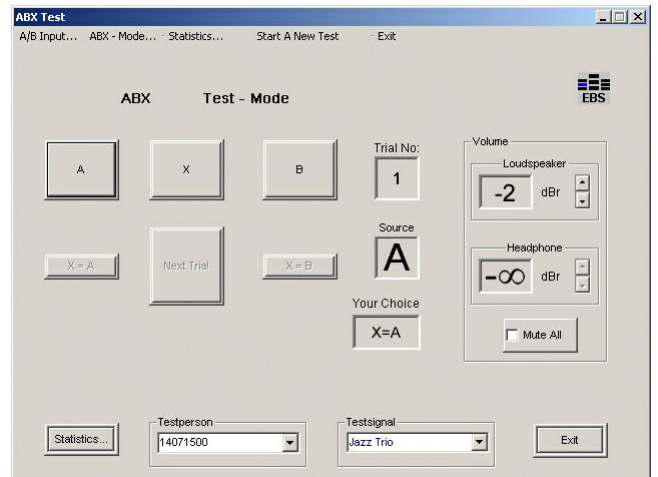
<sup>1</sup> EBU Tech.3276 – 2<sup>nd</sup> edition (5/1998).

<sup>2</sup> ITU-R BS.775 und ITU-R BS.1116-1.

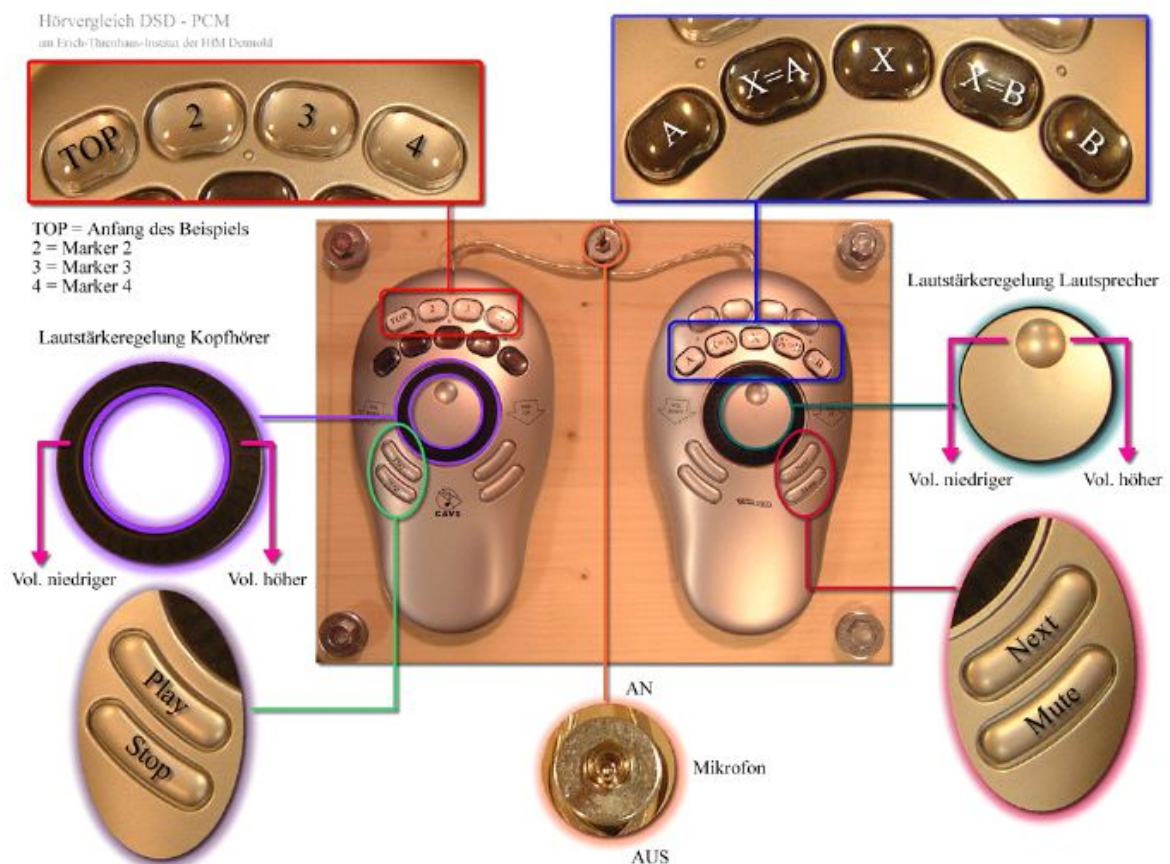
drei parallel ablaufenden Musikbeispielen A, B und X hat der Proband die Möglichkeit, frei und beliebig oft zu wechseln, so dass er keinem zeitlichen Entscheidungsdruck ausgesetzt ist. Um eine signifikante Aussage treffen und somit Zufallsentscheidungen in aller Regel ausschließen zu können, muss der Proband mindestens 16 dieser Entscheidungen durchlaufen. Anhand des statistischen Auswertungsverfahrens kann anschließend eine Aussage getroffen werden, ob ein Unterschied wahrgenommen wurde oder nicht. Das Testverfahren des Doppelblindtests wird von der ITU für die Durchführung solcher Hörvergleiche empfohlen.

Im Fall des vorliegenden Hörvergleichs war A entweder DSD und B komplementär dazu PCM (176,4kHz/24Bit) oder umgekehrt dazu A = PCM und B = DSD (die Zuordnung von A und B auf DSD oder PCM blieb selbstverständlich während eines Testdurchgangs pro Proband unverändert). An Stelle von 16 musste der Proband aufgrund der höheren Aussagekraft jedoch 20 dieser Entscheidungen durchlaufen.

Der Hörvergleich unterteilte sich in zwei Phasen: Zunächst gab es eine Lernphase, innerhalb derer sich der Proband – nach einer genauen Erklärung des Testablaufs und einer technischen Einweisung durch einen der beiden den Test durchführenden Autoren – an die relativ einfache Bedienung der Steuerungsmodule für die ABX-Software wie auch die Software als solche gewöhnen und nach eigenem Interesse in das zur Verfügung gestellte Musikmaterial (Stereo und/oder Surround) hineinhören konnte. Der Proband hatte die Möglichkeit, über definierte Markerpositionen gezielt musikalische Abschnitte wiederholt zu hören. Im Gegensatz zur späteren Testphase erhielt der Proband während der Lernphase nach einer getroffenen Entscheidung eine Bestätigung, ob er X richtig zugeordnet hatte oder nicht. Somit war – gesamt gesehen – eine Simulation der „realen“ Testsituation möglich.



Bildschirmfenster der ABX-Software, die der Proband mit Hilfe eines Steuerungsmoduls (s.u.) bedienen konnte.



Steuerungsmodul für die Bedienung der ABX-Software: Verwendet wurden zwei auf einem Tablett montierte und modifizierte „Shuttle Pro“-Bedienungen von Contour.

Da es aus statistischen Gründen notwendig ist, dass die 20 Entscheidungen eines ABX-Tests mit jeweils nur einem Musikbeispiel bei gleicher Wiedergabeart durchgeführt werden, musste sich der Proband während der Lernphase für ein Musikbeispiel (Stereo oder Surround) und Wiedergabe über Lautsprecher oder Kopfhörer (Stereo) entscheiden. Um Ermüdungserscheinungen des Gehörs schon vor dem eigentlichen Testdurchlauf zu vermeiden, sollte diese Phase allerdings nicht länger als ca. 20-25 Minuten (einschließlich Erklärungen) dauern.

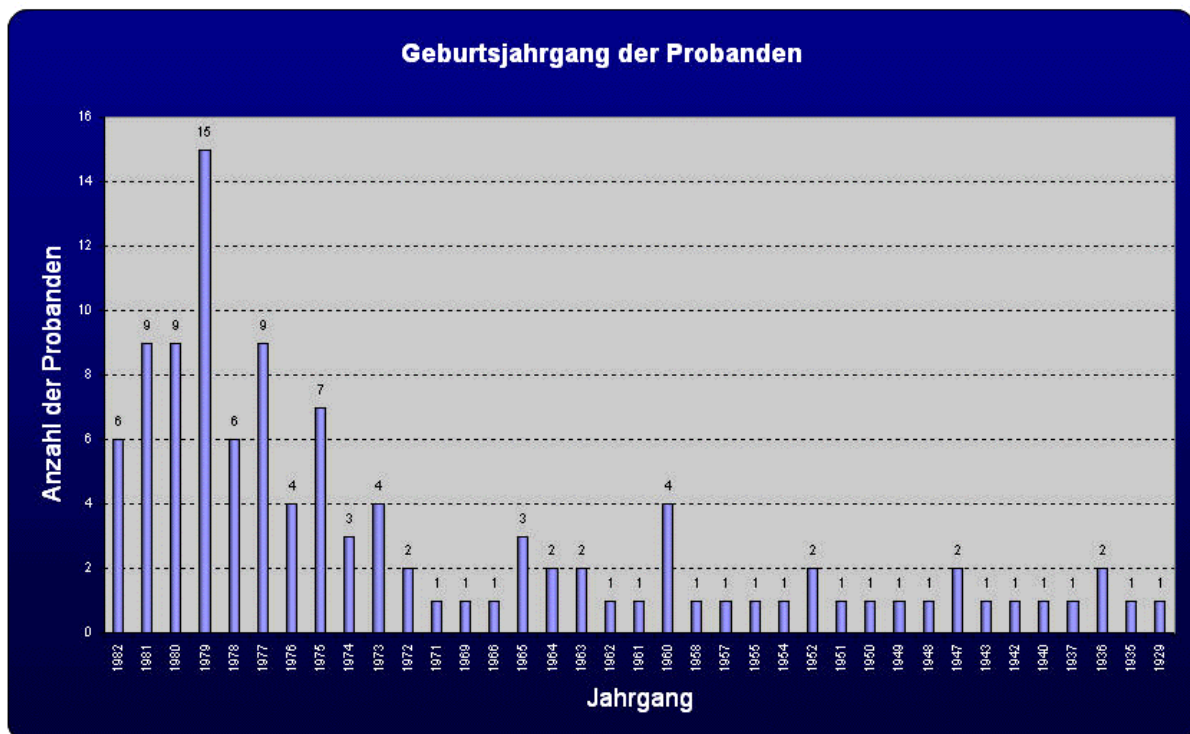
Nach einer optionalen kleinen Pause wurde die zweite Phase, der eigentliche Hörvergleich, mit dem zuvor gewählten Beispiel durchgeführt. Das Ergebnis erfuhr der Proband erst nach Beendigung des Hörvergleichs und dem Ausfüllen zweier Fragebögen mit Fragen zu seiner Person, seinen Musik-Hörgewohnheiten und dem Testablauf.

## 6 Deskriptive und interpretative Auswertung der Versuchsergebnisse

Ziel des vorliegenden Hörvergleichs ist, herauszufinden, ob es Probanden nachweisbar möglich ist, zwischen den zwei digitalen Kodierungsverfahren DSD und PCM (176,4kHz/24Bit) zu differenzieren. Die mathematische Auswertung basiert auf dem stochastischen Modell der Binomialverteilung. Sowohl von der ITU als auch von der ABX-Company wird dabei ein Signifikanz-Niveau von 5% empfohlen. Bei 20 während eines Testdurchlaufes zu treffenden Entscheidungen lässt sich aus dieser Vorgabe berechnen, dass der Proband mindestens 75% (bei 20 Entscheidungen sind nur 5%-Schritte möglich) korrekte Antworten gegeben haben muß. Die Wahrscheinlichkeit hierfür beträgt bei Betrachtung des Tests als Zufallsexperiment  $p = 0,021$  bzw. 2,1%, wobei die Wahrscheinlichkeit, 14 von 20 Entscheidungen korrekt zu treffen, fast 6% beträgt ( $p = 0,0591$ ). In Übereinstimmung mit internationalen Konventionen wurde die Grenze der kritischen Wahrscheinlichkeit bei diesem Hörvergleich auf 15 richtige Antworten festgelegt.

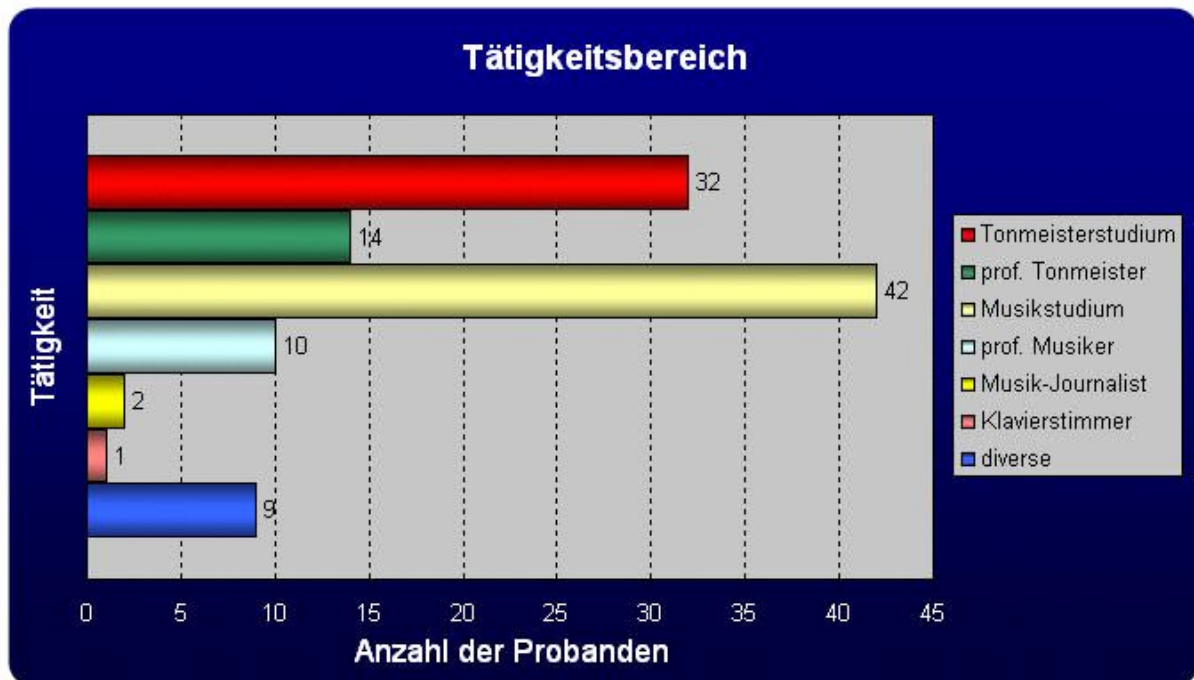
Die Hörtests fanden innerhalb von 28 Testtagen statt. In diesem Zeitraum konnten 145 Tests mit 110 Testpersonen durchgeführt werden (manche Probanden führten den Test mit zwei Musikbeispielen entweder direkt aneinander anschließend oder auch an unterschiedlichen Tagen durch). Dadurch wird auf jeden Fall den ITU-Richtlinien entsprochen, die besagen, dass bereits aufgrund der Ergebnisse von 20 Personen Schlussfolgerungen gezogen werden können.

Das Testpublikum bestand aus 43 weiblichen und 67 männlichen Testpersonen, deren Altersstruktur aus Grafik Nr.5 ersichtlich wird. Im arithmetischen Mittel ergibt sich daraus ein Durchschnittsalter der Probanden von 32,9 Jahren.

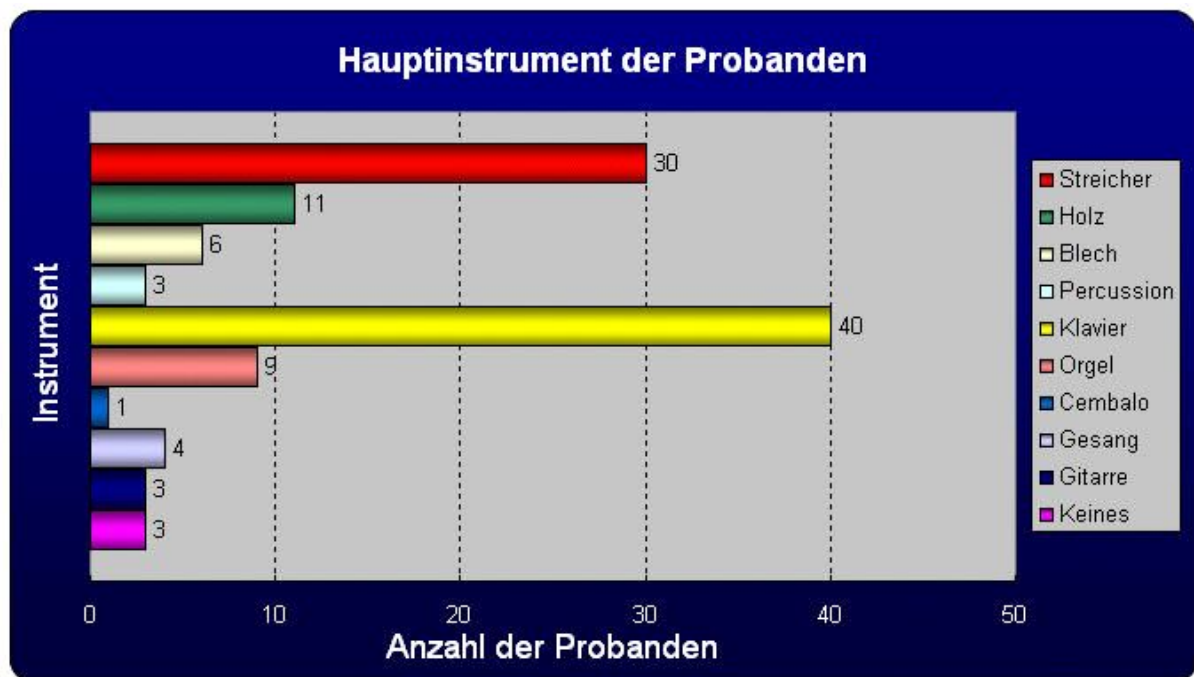


Altersstruktur der Testpersonen.

Grafik Nr.6 verdeutlicht, aus welchem beruflichen Betätigungsfeld sich die Testhörerschaft zusammensetzte. Da fast alle Probanden ausübende Musiker waren, bietet Grafik Nr.7 ergänzend dazu eine Übersicht über deren jeweiliges Hauptinstrument. Aus beiden Grafiken lässt sich erkennen, dass das Testpublikum in der Mehrzahl auf jeden Fall (professionell-) musikalisches und kritisch-analytisches Hören gewohnt war.

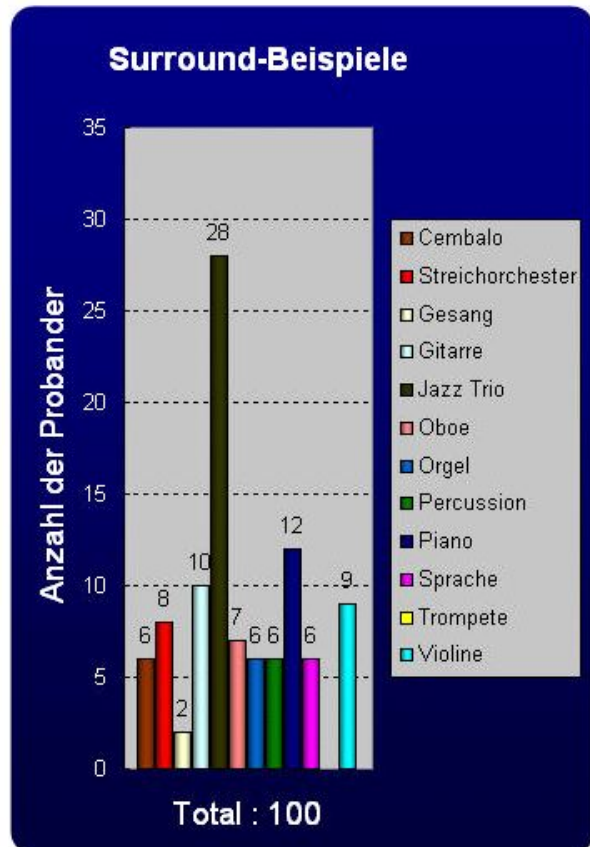
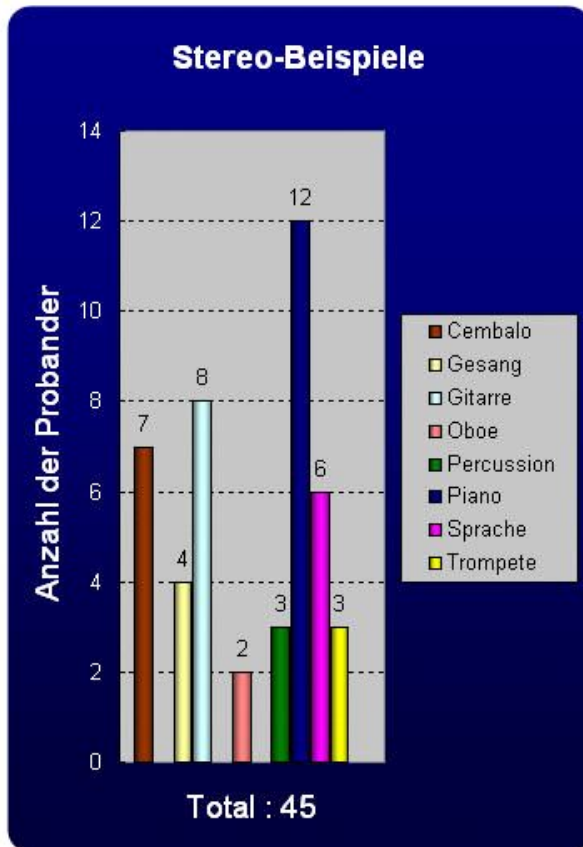


Beruflicher Tätigkeitsbereich der Probanden.



Hauptinstrument der Probanden.

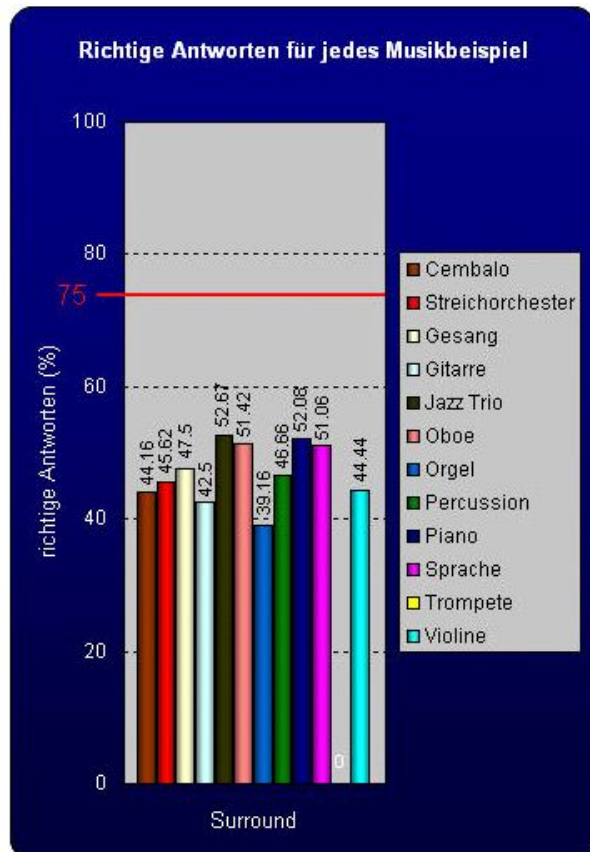
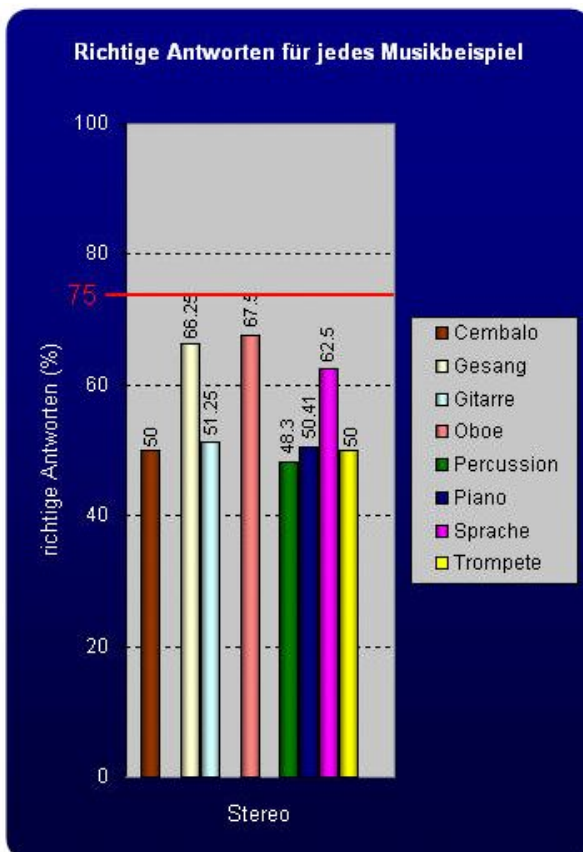
Die 145 durchgeführten Tests teilen sich in einem Verhältnis von 1:2,2 auf 45 Stereo-Beispiele (davon 30 mit Kopfhörerbildung) und 100 Surround-Beispiele auf. Aus den Grafiken Nr.8a und 8b lässt sich entnehmen, mit welchem Verteilungsverhältnis sich die Probanden für die 20 zur Auswahl stehenden Musikbeispiele in der eigentlichen Testphase entschieden haben.



Verteilung der durchgeführten Tests auf die zur Verfügung stehenden Musikbeispiele.

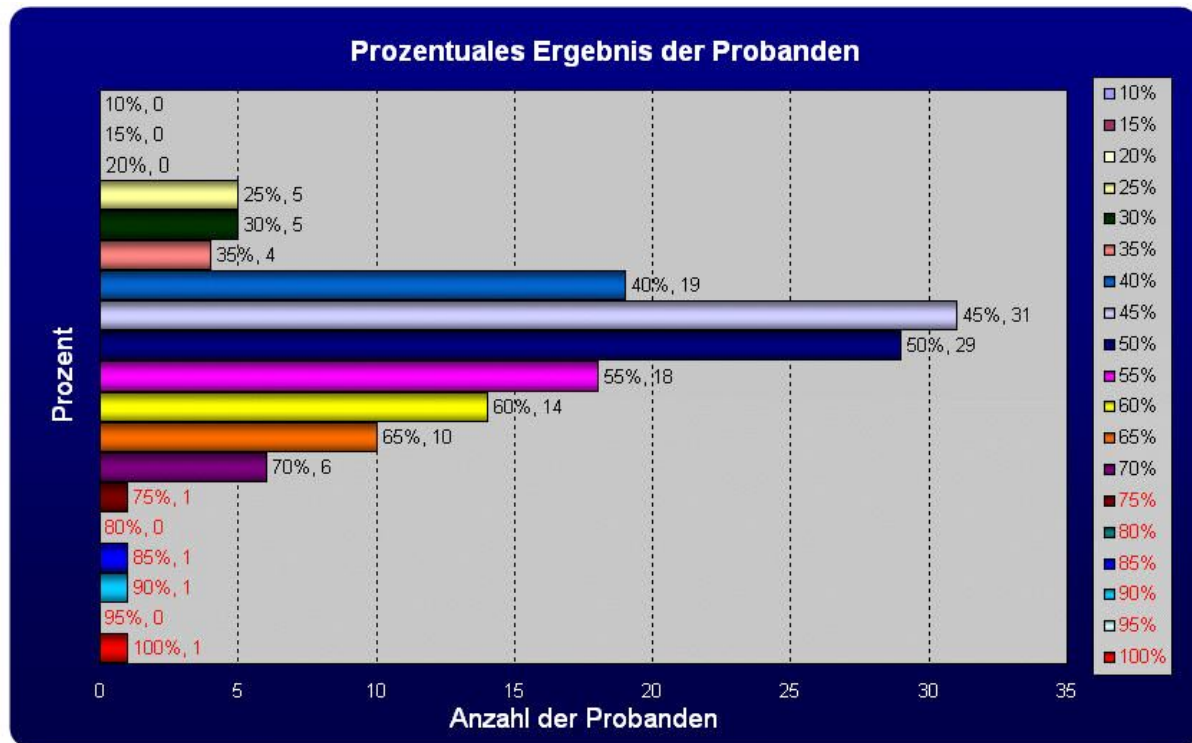
Auffällig ist, dass das Jazz Trio überdurchschnittlich oft als Testbeispiel gewählt wurde. Begründet wurde das von den Probanden damit, dass durch die deutliche Panorama-Verteilung der Instrumente (Klavier ↔ L, R; Bass ↔ C; Schlagzeug ↔ LS, RS) die Aufnahme sehr durchsichtig sei und gleichzeitig sowohl unterschiedliche klangliche als auch räumliche Aspekte beinhalte und somit gute Anhaltspunkte zum Hören biete.

Anzahl der richtigen Antworten im arithmetischen Mittel für jedes Musikbeispiel:



Wie sich die Anzahl der richtigen Antworten pro durchgeführtem Test im arithmetischen Mittel auf die einzelnen, verfügbaren Musikbeispiele in Stereo und Surround verteilt, ist aus den Grafiken Nr.9a und 9b ersichtlich. Die horizontal rot eingetragene Linie markiert bei 75% die untere Grenze der mindestens 15 richtigen Entscheidungen, bei denen die Eintrittswahrscheinlichkeit im Bereich der kritischen Wahrscheinlichkeit von 5% liegt.

Aus Grafik Nr.10 lässt sich die Verteilung der 145 prozentualen Ergebnisse der Probanden entnehmen, die im Verlaufe der Testdurchführung erzielt wurden.



Anzahl der Probanden in Abhängigkeit von deren erreichtem prozentualen Ergebnis (Anzahl der richtigen Entscheidungen).

Aus der Übersichtsgrafik Nr.10 wird deutlich, dass 4 Testpersonen im Bereich der kritischen Wahrscheinlichkeit liegen (rot unterlegt). Dies entspricht lediglich 2,76% der Gesamttests. Gemeinsam ist diesen von vier unterschiedlichen Probanden erzielten Ergebnissen, dass sie mit Stereo-Beispielen stattgefunden haben, die über Kopfhörer, also unter weitestgehender Ausblendung von ablenkenden Umwelt- und Raumeinflüssen, wiedergegeben wurden. Jeder dieser Probanden hatte dabei allerdings ein anderes Beispiel gewählt:

- **Oboe:** Stereo bei Kopfhörerwiedergabe **75%** richtige Antworten →  $p = 0,0207$
- **Sprache:** Stereo bei Kopfhörerwiedergabe **85%** richtige Antworten →  $p = 0,0013$
- **Gitarre:** Stereo bei Kopfhörerwiedergabe **90%** richtige Antworten →  $p = 0,0002$
- **Gesang:** Stereo bei Kopfhörerwiedergabe **100%** richtige Antworten →  $p < 0,000001$

Leider wurden alle 4 Tests innerhalb der letzten vier Testtage absolviert, so dass es aufgrund einer hundertprozentigen Frequentierung der Testtermine in dieser Zeit nicht mehr möglich war, mit diesen Personen jeweils einen nachfolgenden Verifikationstest durchzuführen.

In der umfangreichen schriftlichen Ausarbeitung der Arbeit wird der ausführliche Versuch einer möglichen Erklärung dieser vier – außerhalb des deutlich erkennbaren Trends liegenden – Testergebnisse unternommen (zumal bei 100 durchgeführten Surround-Tests kein einziges Ergebnis im Bereich des Signifikanzniveaus lag). Dieser soll hier in Kürze dargestellt werden:

Systembedingt produziert das Non-Audio-Format bei Funktionen, die direkt oder auch indirekt eine „Stop“- oder „Play“-Funktion beinhalten, beim DSD-Modus ein im Gegensatz zum PCM-Modus teilweise minimal unterschiedlich klingendes, sehr kurzes Knackgeräusch, das vermutlich nur durch einen Fade-In oder Fade-Out auf digitaler Ebene hätte vermieden werden können. Dies würde einerseits natürlich der Voraussetzung widersprechen, dass das Material in keiner Weise bearbeitet sein darf; der essentiellere Aspekt jedoch wäre andererseits, dass ein solcher nur schwer in einem Non-Audio-Format realisierbar ist. Trotz intensiver Beschäftigung mit diesem Problem und Beratung durch einige der den Hörvergleich unterstützenden Firmen, sowie einer

Umstrukturierung des ursprünglich geplanten Rechnerkonzepts, war zum damaligen Zeitpunkt eine Lösung durch eine zusätzliche Soft- oder Hardware nicht greifbar und wäre auch momentan nur schwer umzusetzen. Dass dieses Knacken nicht bewusst zur Kenntnis genommen wurde und der Test somit in keiner Weise seiner Aussagekraft entzogen wird, zeigt die Auswertung der Fragebögen, in denen ein Knackgeräusch ausnahmslos nicht erwähnt wurde.

Betrachtet man das Testergebnis im Zusammenhang gerade mit der der Wahrnehmungspsychologie entstammenden „Dämpfungstheorie der selektiven auditiven Aufmerksamkeit“ nach Treisman, die heute als weitestgehend anerkannt gilt, könnten auch klangliche Elemente wie ein Knacken als – tonmeisterspezifisch – semantische Inhalte gelten und eine Entscheidungsfindung unbewusst oder mitbewusst beeinflussen. Da alle vier im Bereich der kritischen Wahrscheinlichkeit liegenden Probanden Tonmeisterstudenten im Alter zwischen 25 und 28 Jahren waren und die jeweiligen Musikbeispiele bezeichnenderweise über Kopfhörer gehört und sich demzufolge noch einmal ausschließlich auf das Hören eingelassen haben (störende Raum- bzw. Umwelteinflüsse werden deutlicher ausgeschlossen), könnte diese Bedeutungsebene hier besonders wirksam werden.

Es sei jedoch ausdrücklich betont, dass die Gültigkeit der Ergebnisse dieser vier Testpersonen keinesfalls angezweifelt wird und somit auch der Versuch einer Erklärung dieser vier Fälle selbstverständlich nicht in die deskriptive Auswertung der Arbeit mit einfließen soll.

Die Schlussfolgerung aus den Ergebnissen der Auswertung muss also lauten, dass bei 4 Fällen im Bereich der kritischen Wahrscheinlichkeit aufgrund der zuvor getroffenen Entscheidungsregel die Hypothese H (es existieren keine wahrnehmbaren Unterschiede zwischen der Quelle A und der Quelle B) verworfen und die Gegenhypothese G (es existieren wahrnehmbare Unterschiede zwischen A und B) angenommen werden kann. In diesen Fällen kann davon ausgegangen werden, dass die Probanden einen Unterschied zwischen den Quellen A und B wahrgenommen haben.

Dahingegen kann bei 141 von 145 Testergebnissen (entspricht 97,24%) die Hypothese H nicht abgelehnt, sondern muss beibehalten werden. In diesen Fällen kann die Vermutung geäußert werden, dass ein Unterschied zwischen den Quellen A und B vom Probanden nicht wahrgenommen wurde.

Die Grafiken Nr. 9a und 9b zeigen sehr anschaulich, wie sich die prozentuale Verteilung bei den einzelnen Musikbeispielen in der Nähe des Zufalls-Niveaus von 50 Prozent bewegt (bei den Surround-Beispielen noch erheblich deutlicher als bei den Stereo-Beispielen). Bestärkt werden diese Beobachtungen, wenn alle falschen und richtigen Entscheidungen zusammen addiert werden: So ergibt sich bei insgesamt 145 Tests (multipliziert mit jeweils 20 Entscheidungen) eine Gesamt-Entscheidungszahl von 2900. Mit Hilfe der Grafik Nr.10 lässt sich eine Verteilung von 1454 richtigen zu 1446 falschen Entscheidungen berechnen, wobei bei Betrachtung des Tests als Zufallsexperiment eine bemerkenswerte Genauigkeit bezüglich des statistischen Erwartungswertes (arithmetischer Mittelwert einer Zufallsvariablen) von 1450 richtigen bzw. falschen Entscheidungen erreicht wird – bei nur 4 (nicht zu verwechseln mit den 4 signifikanten Testergebnissen) von diesem Erwartungswert abweichenden „Differenzentscheidungen“ entspricht das einer Abweichung von 0,28%.

Ebenso konnte eine Unterscheidung perkussiver Elemente (Percussion und Cembalo), die sehr schnelle Signalanstiegsflanken enthalten, nicht vorgenommen werden.

## FAZIT

Aufgrund der Resultate des vorliegenden Hörvergleichs hat sich gezeigt, dass selbst mit hochwertigstem Equipment unter optimalen Abhörbedingungen und unterschiedlichster Hörfokussierungen bzw. Hörerfahrungen der Probanden in der Regel keine signifikanten Unterschiede zwischen DSD und High-Resolution-PCM (24bit/176,4kHz) hörbar sind, sich demzufolge die These aufstellen ließe, dass keines der beiden getesteten Systeme unter Anwendung wissenschaftlicher Kriterien für sich in Anspruch nehmen kann, klanglich höherwertige hörbare Ergebnisse zu erzielen. Dieser Realität sollte sich die Argumentation über Klangqualität auch auf Seiten der PR-Abteilungen stellen.

Lediglich bei 4 der 145 durchgeführten Tests liegt das Ergebnis im Bereich des Signifikanzniveaus von 5%, übertragen auf die Gesamtauswertung also nur 2,76% der Testergebnisse. Während diese vier Tests sich ausschließlich auf Stereobeispiele bei Kopfhörerwiedergabe beziehen, lag bei 100 durchgeführten Surround-Tests, also unter Einbeziehung räumlicher Komponenten, kein Testergebnis im Bereich der kritischen Wahrscheinlichkeit.

Weniger nachweisbar in Form einer mathematischen Auswertung, jedoch umso einprägender war das hohe Maß an Frustration, das viele Probanden, die in der Mehrzahl professionelles und kritisch-analytisches Hören gewohnt waren, während der Durchführung der Tests empfanden und das sie auf für sie nicht annähernd zu erkennende klangliche Unterschiede zurückführten. Ein weiteres, häufiges Diskussionsthema war ebenso die Wahrnehmung von *vermeintlichen* Unterschieden: Hörillusion – ein Thema, das sicherlich speziell für Tonmeister eine bedeutende Rolle spielen und über dessen Ursache – soll effektiv und sowohl musikalisch als auch klanglich gewinnbringend gearbeitet werden – persönliche Klarheit bestehen sollte.