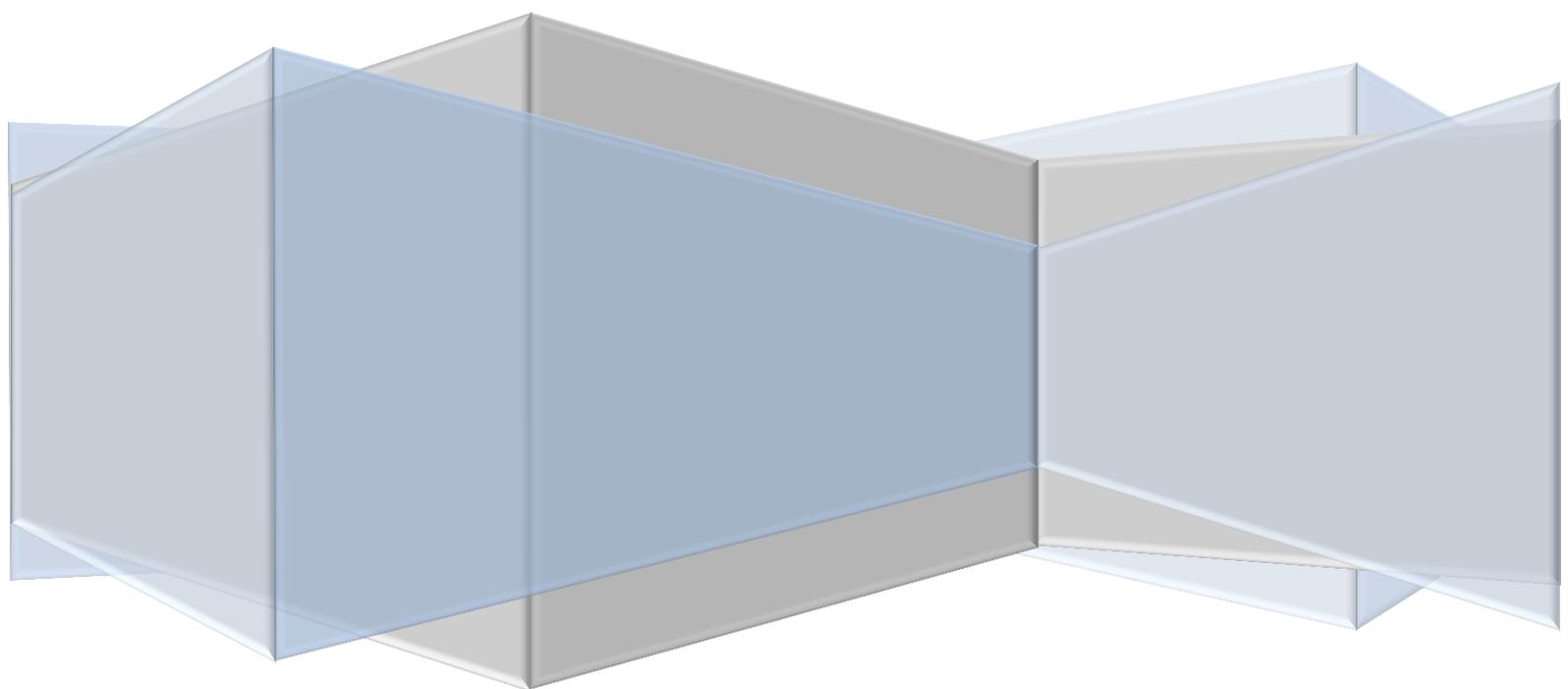


Autor: Tobias Klein

Bachelorarbeit

Softwaregestützte Hörversuche

Nach den Empfehlungen der
ITU-R BS.1116-1 und ITU-R BS.1534-1



Bachelorarbeit

Softwaregestützte Hörversuche

Nach den Empfehlungen der
ITU-R BS.1116-1 und ITU-R BS.1534-1

Vorgelegt von:

Tobias Klein

Baumgartenweg 4

71149 Bondorf

Matrikelnummer: 22170

SS 2013

Audiovisuelle Medien

an der Hochschule der Medien in Stuttgart

Eingereicht am:

29.08.2013

Erstprüfer: Herr Prof. Jens-Helge Hergesell

Zweitprüfer: Herr Prof. Oliver Curdt

Kurzfassung

Grundlage für diese Arbeit bilden die Dokumente ITU-R BS.1116-1 und ITU-R BS.1534-1, welche eine Empfehlung für die Durchführung eines Hörtests aussprechen und dadurch ermöglichen sollen, Hörtests vergleichbar zu machen. Außerdem wird ein Überblick und eine Beschreibung und Bewertung über verschiedene Softwares gegeben, mit welchen man Hörversuche durchführen kann.

Der praktische Teil dieser Arbeit umfasst die Durchführung und Dokumentation eines MUSHRA-Hörtests nach den Vorstellungen der ITU, welcher verschiedene MP3-Qualitätsstufen vergleicht.

Schlagwörter: Hörtest, ITU-R BS.1116, ITU-R BS.1534, MUSHRA, ABX, ABC/HR-Test

Abstract

The documents ITU-R BS.1116-1 and ITU-R BS.1534-1 which give a recommendation for performing a listening test and which at the same time allow to compare different listening tests are the basis for the present work.

In addition, they give a summary and a description and evaluation of various softwares allowing to perform listening tests.

The practical part of this work comprises the execution and documentation of a MUSHRA listening test which compares different quality levels of mp3 recordings and which was performed according the recommendations of the ITU.

Tags: listening test, ITU-R BS.1116, ITU-R BS.1534, MUSHRA, ABX, ABC/HR-test

Inhaltsverzeichnis

| | |
|-------------------------------------------------------------------------|--------|
| Vorwort..... | - 1 - |
| 1. Das Gehör | - 3 - |
| 1.1 Aufbau des menschlichen Ohrs..... | - 3 - |
| 1.1.1 Das äußere Ohr..... | - 3 - |
| 1.1.2 Das Mittelohr..... | - 4 - |
| 1.1.3 Das Innenohr | - 4 - |
| 1.2 Der Hörvorgang | - 5 - |
| 1.3 Lautstärkeempfinden | - 6 - |
| 2. Psychoakustik – der Verdeckungseffekt | - 7 - |
| 2.1 Verdeckung durch Tonlängen (Temporale Maskierung)..... | - 7 - |
| 2.2 Verdeckung durch Pegel und Tonhöhe (Gleichzeitige Maskierung)..... | - 8 - |
| 3. Standardisierte Testmethoden | - 10 - |
| 3.1 Triple stimulus-hidden reference-double blind test..... | - 12 - |
| 3.1.1 Auswahl und Selektion der Testpersonen..... | - 14 - |
| 3.1.2 Trainings- und Testphase der Testhörer für einen ABC/HR-Test..... | - 14 - |
| 3.2 Multi Stimulus test with Hidden Reference and Anchor..... | - 16 - |
| 3.2.1 Auswahl der Testhörer..... | - 17 - |
| 3.2.2 Trainingsphase der Testhörer für einen MUSHRA-Test | - 18 - |
| 3.2.3 Testmaterial..... | - 18 - |
| 3.3 Grundsätzlicher Testaufbau..... | - 19 - |
| 3.4 Anzahl der Testhörer | - 19 - |
| 3.5 Der Abhörraum | - 20 - |
| 3.6 Die Abhörmethoden | - 20 - |
| 3.7 Audiomaterial für Hörtests – SQAM-CD | - 21 - |
| 3.8 Dokumentation und Präsentation der Testergebnisse..... | - 23 - |
| 4. Weitere Testmethoden | - 25 - |

| | | |
|-------|---------------------------------------------------|--------|
| 4.1 | Absolutbewertung | - 25 - |
| 4.2 | ABX-Blindhörtest (Paar-Vergleichstest) | - 26 - |
| 5. | Statistik..... | - 27 - |
| 5.1 | Arithmetischer Mittelwert..... | - 27 - |
| 5.2 | Standardabweichung | - 28 - |
| 5.3 | Das Konfidenzintervall..... | - 28 - |
| 5.4 | Die Normalverteilung | - 29 - |
| 5.5 | Größe der Stichproben..... | - 29 - |
| 5.6 | Der Hypothesentest..... | - 30 - |
| 5.7 | Der p-Wert | - 30 - |
| 5.8 | t-Verteilung und t-Test..... | - 30 - |
| 5.9 | Binomialverteilung..... | - 31 - |
| 6. | Softwarebeschreibung: foobar2000 (ABX-Test) | - 34 - |
| 6.1 | Das ABX-Plug-In | - 34 - |
| 7. | Softwarebeschreibung: <i>STEP</i> | - 36 - |
| 7.1 | Systemvoraussetzungen und Installation..... | - 37 - |
| 7.2 | Benutzeroberfläche und Bedienung | - 37 - |
| 7.3 | Testerstellung..... | - 39 - |
| 7.3.1 | AB-Test in <i>STEP</i> | - 40 - |
| 7.3.2 | ABX-Test in <i>STEP</i> | - 40 - |
| 7.3.3 | Der ABC/HR-Test (BS.1116-1) in <i>STEP</i> | - 41 - |
| 7.3.4 | Der MUSHRA-Test in <i>STEP</i> | - 42 - |
| 7.4 | Testauswertung | - 43 - |
| 7.5 | Standard-konforme Hörtests mit <i>STEP</i> | - 43 - |
| 7.6 | Bewertung der Software <i>STEP</i> | - 43 - |
| 8. | Softwarebeschreibung: <i>MUHSRA</i> Test..... | - 44 - |
| 8.1 | Einstellungen und Erstellung von Hörtests | - 44 - |

| | | |
|--------|----------------------------------------------------------------------|--------|
| 8.2 | Das Training in <i>MUSHRATest</i> | - 46 - |
| 8.3 | Der Hörtest in <i>MUSHRATest</i> | - 47 - |
| 8.4 | Die Datendokumentation in <i>MUSHRATest</i> | - 48 - |
| 8.5 | Standard-konforme Hörtests mit <i>MUSHRATest</i> | - 49 - |
| 8.6 | Bewertung der Software <i>MUSHRATest</i> | - 50 - |
| 9. | Weitere Softwares zur Durchführung von Hörtests..... | - 51 - |
| 9.1 | Das WinABX-Tool | - 51 - |
| 9.2 | Das <i>ABC Hidden Reference-Audio Comparison Tool</i> (ABC/HR) | - 52 - |
| 9.3 | Die Software <i>scale</i> – Kurzbeschreibung zum MUSHRA-Test..... | - 53 - |
| 9.3.1 | Die Testerstellung in <i>scale</i> | - 54 - |
| 9.3.2 | Die Testdurchführung in <i>scale</i> | - 55 - |
| 9.3.3 | Die Testanalyse in <i>scale</i> | - 55 - |
| 9.3.4 | Bewertung der Software <i>scale</i> | - 56 - |
| 10. | Durchführung und Protokollierung eines MUSHRA-Hörversuches..... | - 57 - |
| 10.1 | Die Grundlagen der MP3-Datei..... | - 57 - |
| 10.2 | Auswahl der Audiobeispiele | - 58 - |
| 10.3 | Verwendete Softwares für den Hörtest..... | - 60 - |
| 10.3.1 | Cubase..... | - 60 - |
| 10.3.2 | <i>MUSHRATest</i> | - 60 - |
| 10.4 | Testpersonen | - 61 - |
| 10.5 | Die Abhörsituation | - 61 - |
| 10.6 | Einführungs- und Trainingsphase..... | - 62 - |
| 10.7 | Die Durchführung | - 64 - |
| 11. | Auswertung des durchgeführten MUSHRA-Tests..... | - 64 - |
| 11.1 | Postscreening der Testhörer | - 65 - |
| 11.2 | Erklärung der Box-Plot-Übersicht..... | - 65 - |
| 11.3 | Auswertung der Qualitätsstufen über alle Hörbeispiele | - 66 - |

| | | |
|---------|--------------------------------------------------------------------------|--------|
| 11.4 | Auswertung der Hörbeispiele mittels eines t-Tests | - 67 - |
| 11.5 | Auswertung der Hörbeispiele über die verschiedenen Qualitätsstufen | - 69 - |
| 11.6 | Abschlussbericht des MUSHRA-Hörtests | - 71 - |
| 12. | Fazit..... | - 72 - |
| 13. | Anhang..... | I |
| 13.1 | Audio-CD 1: Psychoakustische Effekte | I |
| 13.2 | CD und Trackinformationen zur SQAM-CD | II |
| 13.3 | Softwaredokumentation von foobar2000..... | VII |
| 13.3.1 | Systemvoraussetzungen und Installation..... | VII |
| 13.3.2 | Oberflächenaufbau | VIII |
| 13.3.3 | Die Playlists | IX |
| 13.3.4 | Menüpunkt File | X |
| 13.3.5 | Menüpunkt Edit | XI |
| 13.3.6 | Menüpunkt View | XII |
| 13.3.7 | Menüpunkt Playback | XV |
| 13.3.8 | Menüpunkt Library | XVI |
| 13.3.9 | Menüpunkt Help..... | XVI |
| 13.3.10 | Preferences..... | XVI |
| 13.3.11 | Metadaten | XVII |
| 13.3.12 | Dateiformatkonvertierung | XVIII |
| 13.3.13 | „ReplayGain“-Funktion | XX |
| 13.3.14 | Zusätzliche Plug-Ins in foobar2000..... | XXI |
| 13.4 | Audioausschnitte im MUSHRA-Hörversuch | XXII |
| 14. | Abkürzungsverzeichnis..... | XXIII |
| 15. | Abbildungsverzeichnis | XXIV |
| 16. | Literaturverzeichnis und Webhinweise | XXVI |
| 17. | Erklärung..... | XXX |

Vorwort

In einer Welt, in der es immer wichtiger ist, Daten auszutauschen und die Größe der Dateien eine maßgeblich Rolle für die Austauschgeschwindigkeit spielt, stellt sich in Bezug auf Audiodateien die Frage, ab wann eine Kompression der Daten für den Hörer als schädigend empfunden wird bzw. mit hörbaren qualitativen Einbußen verbunden ist. Dieser Problematik soll in der vorliegenden Arbeit unter anderem nachgegangen werden.

Hierfür geht es zunächst um die Arbeit und Funktion des menschlichen Gehörs, zudem werden verschiedene psychoakustische Effekte vorgestellt, welche sich auch Datenkompressionen (z.B. MPEG-Layer 3-Codec) zunutze machen.

Ferner beschäftigt sich die Arbeit mit verschiedenen MP3-Qualitätsstufen: Wie vergleicht man Audiofiles und nach welchen Kriterien beurteilt man deren Qualität? Mit diesen und ähnlichen Fragen beschäftigte sich die ITU und erarbeitete zwei Empfehlungen, durch welche sich zwei Standards etablierten, die es ermöglichen die Ergebnisse verschiedener Hörtests vergleichbar zu machen. Ein Schwerpunkt der vorliegenden Arbeit sind daher die Empfehlungen und Richtlinien der ITU, welche sich mit den Voraussetzungen für raum- und zeitunabhängige Hörtests beschäftigt.

Hörtests können analog mit einem Blatt, einem Stift und verschiedenen Hörbeispielen durchgeführt werden, allerdings gibt es im Zeitalter von Computern und digitalem Audiomaterial die Möglichkeit, Hörtests auf digitaler Ebene durchzuführen. Hierzu werden auf dem Markt verschiedene Softwares verwendet. Im Mittelpunkt der Arbeit steht die Erläuterung und Bewertung verschiedener Softwares, die Hörtests zur subjektiven Beurteilung von Audioqualität ermöglichen.

Im Einzelnen wird auf die Einhaltung der ITU-Richtlinien, den Aufbau und die Bedienung dieser Programme eingegangen.

Darüber hinaus werden die Prozesse und Arbeitsweisen der Programme verglichen. Anhand der hier ermittelten Unterschiede bzw. Vor- und Nachteile verschiedener Programme erfolgte die Auswahl der Software für die Durchführung eines Hörtests.

Dieser praktische Teil beschäftigt sich mit der Durchführung eines Hörtests nach den Richtlinien der ITU, welcher die subjektive Bewertung verschiedener MP3-Qualitätsstufen ermittelt. Um die Ergebnisse heranziehen zu können, ist es nötig die

Daten statisch auszuwerten. Daher werden zunächst die Grundlagen der Statistik erörtert und anschließend die Daten ausgewertet und präsentiert.

Im Anhang befinden sich verschiedene Daten und Beschreibungen zum durchgeführten Hörtest sowie eine Bedienungsanleitung des Musikmediaplayers *foobar2000* als Beispiel für eine Software, mit welcher sich ein ABX-Hörtest durchführen lässt.

1. Das Gehör

Mit dem Gehör lassen sich die Schallereignisse in der Natur, bestehend aus Frequenzgang und Schalldruck, im Gehirn zu Tonhöhe und Lautstärke verarbeiten. Das Ohr hat die Aufgabe die Druckschwankungen des Schallfeldes in Nervenreize umzuwandeln.¹ Ein gesundes Ohr kann Frequenzen zwischen 16 Hz und 20.000 Hz wahrnehmen, allerdings ist das Ohr für bestimmte Frequenzen unterschiedlich empfindlich. Das Ohr umfasst sowohl das Hörorgan, als auch das Gleichgewichtsorgan. Es besteht aus dem äußeren Ohr, dem Mittelohr und dem Innenohr.

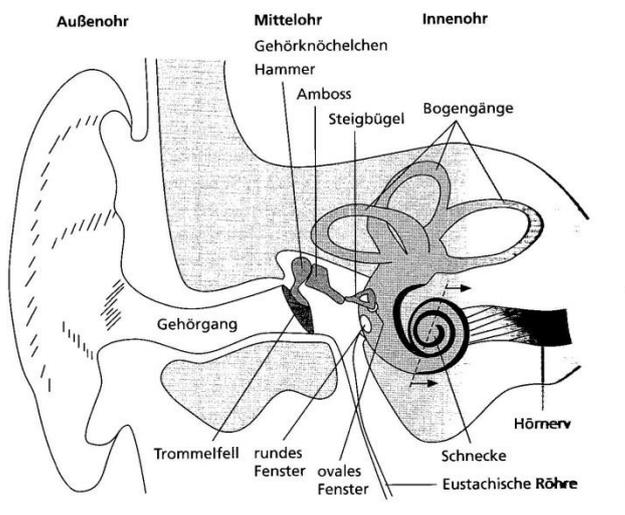


Abbildung A: Das menschliche Hörorgan²

1.1 Aufbau des menschlichen Ohrs

1.1.1 Das äußere Ohr

Das äußere Ohr besteht aus der Ohrmuschel und dem Gehörgang. Anders als bei Tieren, die mit ihren Stellmuskeln ihre Ohren dem Schallereignis zuwenden können, ist die Ohrmuschel beim Menschen im Grunde ohne Funktion. Sie besteht aus elastischen Knorpeln, die mit Haut überzogen sind. Der Gehörgang ist ein Rohr mit ca. 3 - 4 cm und einer Weite von 5 - 10 mm. Er verbindet die Ohrmuschel mit dem

¹ Vgl. (Webers, 2007 S. 93ff)

² (Kuttruff, 2004)

Trommelfell. Das äußere Ohr enthält die Ohrschmalzdrüsen, die ein hellgelbes Sekret liefern, welches zum Schutz vor Schmutzkörpern vorhanden ist.³

1.1.2 Das Mittelohr

Das Mittelohr (auch Paukenhöhle genannt) ist ein etwa 5 mm breiter Spalt, der durch die Ohrtrompete mit dem Nasen-Rachen-Raum verbunden ist. Im Mittelohr befindet sich das Trommelfell, das einen Durchmesser von ca. 10 mm besitzt.⁴ Es schließt den äußeren Gehörgang gegen die Paukenhöhle ab. Das Trommelfell ist eine etwa 0,1 mm dicke, ovale bis kreisförmige, bindegewebeartige Membran. Der Hammer, ein Teil der Gehörknöchelchen, ist mit dem sogenannten Hammergriff am oberen Teil des Trommelfells angewachsen. Die innere Wand der Paukenhöhle trennt das Innenohr vom Mittelohr ab und enthält zwei „Fenster“, welche Ovale und Rundes Fenster genannt werden. Es handelt sich hier einmal um den Kanal, in dem die Gesichtsnerven laufen und um einen Kanal, der in das Innenohr führt.⁵

1.1.3 Das Innenohr

„Das Innenohr, wegen seiner verzweigten Form auch als Labyrinth bezeichnet, lässt sich grob in die Schnecke (lat. Cochlea) und das Gleichgewichtsorgan (Vestibularorgan) unterteilen.“⁶

Das Innenohr, das wie ein Schneckengehäuse geformt ist, ist anders als das Außen- und Mittelohr nicht mit Luft, sondern mit den Flüssigkeiten Peri- und Endolymphe gefüllt, deren Dichte noch höher ist, als die von Wassers.

Der Steigbügel ist über das Ovale Fenster mit der mit Flüssigkeit gefüllten Ohrschnecke des Innenohres verbunden. Der innere Gehörgang ist etwa 1 cm lang und enthält die Hör- und Gleichgewichtsnerven, die Gesichtsnerven, die Innenohrschlagader und einen Teil der Innenohrvenen.⁷

Der Schneckengang enthält das eigentliche Hörorgan. Seine Grundmembran ist mit Hörsinneszellen bedeckt. Die Härchen der Sinneszellen ragen in die darüber liegende Deckmembran. Schon feine Erschütterungen des Schneckenganges reizen einen Teil

³ vgl. (Gontard)

⁴ vgl. (Euerle)

⁵ vgl. (Gontard)

⁶ (Heß)

⁷ vgl. (Gontard)

der Sinneshärchen. Die zugehörigen Sinneszellen melden den Reiz an den Hörnerv weiter.⁸

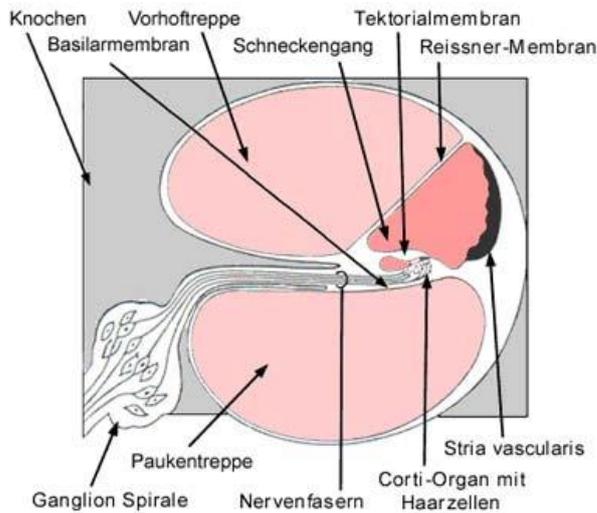


Abbildung B: Das Innenohr⁹

1.2 Der Hörvorgang

Eine Schallwelle erreicht das Hörorgan über die Ohrmuschel und den äußeren Gehörgang. Der Schalldruck versetzt das Trommelfell in Schwingungen, die sich über die Gehörknöchelchen des Mittelohrs übertragen. Die Gehörknöchelchen (Hammer, Amboss und Steigbügel) haben die Aufgabe, den ankommenden Schall zu verstärken und über das Ovale Fenster an den Eingang des Innenohrs weiterzuleiten. Diese Verstärkung entsteht durch die Übertragung der Schwingungen des Trommelfells auf eine kleinere Fläche (Ovales Fenster, Übergang in das Innenohr). Zudem ist die Verstärkung frequenzabhängig: Im Bereich der Resonanzfrequenz des Trommelfells ist die Verstärkung am größten, diese liegt bei ca. ein bis zwei kHz, also im Bereich der menschlichen Stimme.¹⁰

Ferner trägt die Hebelwirkung der Gehörknöchelchen dazu bei, dass die Schwingungen großer Amplitude und kleiner Kraft in Schwingungen kleiner Amplitude und großer Kraft umgewandelt werden.¹¹

⁸ vgl. (Euerle)

⁹ (Gontard)

¹⁰ vgl. (Gontard)

¹¹ vgl. (Webers, 2007 S. 93)

Die Schwingungen des Trommelfells werden über die Gehörknöchelchen und über das Ovale Fenster auf die Flüssigkeiten (Peri- und Endolymphe) übertragen, welche wiederum die Haarsinneszellen in Bewegung versetzt. Dies löst den Sinnesreiz aus, der dann über den Hörnerv zum Gehirn geleitet wird.

Das Gehirn verarbeitet die elektrischen Nervensignale zu Hörempfindungen und lernt von Geburt an, diese Signale zu erkennen und zu verstehen, also den wahrgenommenen Schall als Musik, Sprache oder Geräusch zu interpretieren. Die Hörfähigkeit eines gesunden Gehörs ist für tiefe, mittlere und hohe Töne unterschiedlich. Die Anzahl der Impulse im Gehirn ist maßgeblich für die empfundene Lautstärke.¹²

1.3 Lautstärkeempfinden

Das Gehör kann einen Schalldruckbereich von 20 μ Pa bis 20 Pa verarbeiten. Damit dieser Bereich sinnvoll dargestellt werden kann, wird dieser nach dem „Weber-Fechnerschen Gesetz“ mit der logarithmischen Einheit Bel bzw. Dezibel dargestellt. Der Bezugsschalldruck wurde auf 20 μ Pa bei 1 kHz mit einem Pegel von 0 dB festgelegt. Die Schmerzgrenze liegt bei ca. 120 dB (\cong 20 Pa). Das Schmerzgefühl kommt zustande, da das Gelenk zwischen Amboss und Steigbügel bei einem hohen Schalldruck seitlich ausknickt. Das Ausknicken stellt eine Art Überlastungsschutz für das empfindliche Innenohr dar.¹³

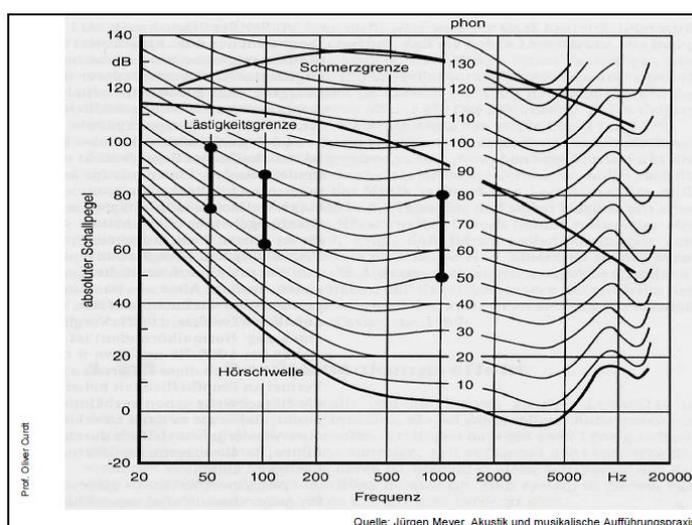


Abbildung C: Kurven gleicher Lautstärke¹⁴

¹² vgl. (Geers)

¹³ vgl. (Webers, 2007 S. 103f)

¹⁴ (Curdt, SS12)

In „Abbildung C: Kurven gleicher Lautstärke ist ersichtlich, dass der Mensch bei verschiedenen Frequenzen unterschiedlich empfindlich reagiert. Der Lautstärkeindruck des Menschen wird geprägt von der gesamten akustischen Energie (Leistung * Dauer). Bei mehreren Schallquellen unterschiedlicher Lautstärke wird die Gesamtlautstärke im Wesentlichen vom lautesten Ereignis geprägt.¹⁵

2. Psychoakustik – der Verdeckungseffekt

Gehört wird im Kopf: Ein Klangereignis wird beim Hören zahlreichen Datenkompressionsverfahren unterzogen. Dabei fallen einige scheinbar unwichtige Details bei der Wahrnehmung unter den Tisch. Im folgenden Abschnitt soll es darum gehen, auf welche Art und Weise das Gehör bzw. das Gehirn des Menschen arbeitet. Verschiedene Experimente der Psychoakustik zeigen, dass der Mensch beim Hörvorgang Schallereignisse teilweise anders wahrnimmt, als sie in der Natur tatsächlich auftreten. Dieses Phänomen macht sich die Musikindustrie zunutze, indem sie beispielsweise Datenkompressionen erstellt, die erstmals kaum einen subjektiven Qualitätsverlust aufweisen (MPEG-Layer-3-Format).

2.1 Verdeckung durch Tonlängen (Temporale Maskierung)

Die Lautstärke eines Signals ist nicht ausschließlich von der realen Lautstärke abhängig, sondern hängt ebenfalls mit der Länge des Signals zusammen.

Ein kurzes Knacken mit einer bestimmten Amplitude wird als leiser empfunden als ein anhaltender Ton gleicher Lautstärke.

Die Autoren Hugo Fastl und Eberhardt Zwicker haben sich ebenfalls mit dieser Thematik beschäftigt und anhand eines Versuches festgestellt, dass kurze Impulse leiser wahrgenommen werden, als längere (vgl. Abbildung D: Lautstärkeempfinden verschiedener Impulslängen S. - 8 -).

¹⁵ vgl. (Curdt, SS12)

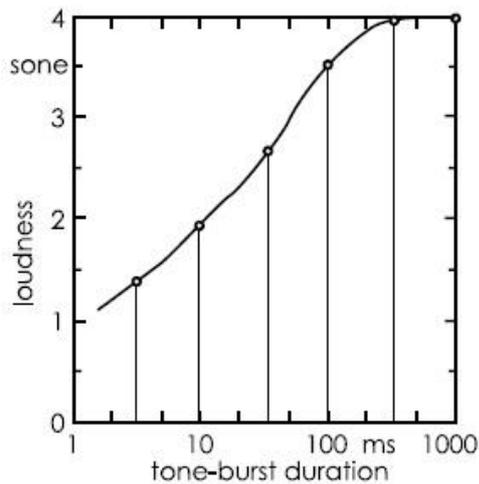


Abbildung D: Lautstärkeempfinden verschiedener Impulslängen ¹⁶

Das Hörbeispiel (vgl. 13.1 Audio-CD 1: Psychoakustische Effekte , S. I, Track 1) zu der Grafik ist wie folgt aufgebaut: Auf dem Hörbeispiel ist ein 3 KHz-Ton mit 60 dB zu hören, der in verschiedenen Längen gespielt wird. Der erste Ton ist immer 1000 ms lang, die weiteren Töne haben jeweils eine Länge von 1000, 300, 100, 30, 10 und 3 ms. Hier zeigt sich, dass unser Gehör kurze Impulse leiser empfindet als längere, obwohl alle Testtöne mit 60 dB abgespielt werden. Zusammenfassend lässt sich sagen: je kürzer der Ton, desto leiser wirkt er.¹⁷

Daraus lässt sich schließen, dass bei gleichzeitigem Abspielen eines längeren und kürzeren Tones, der längere dominiert bzw. je nach Verhältnis der Tonlängen der kürzere „verdeckt“ werden kann, da er vom menschlichen Gehirn nicht wahrgenommen wird. Man spricht dann von temporaler Maskierung.

2.2 Verdeckung durch Pegel und Tonhöhe (Gleichzeitige Maskierung)

Hier wird ein leises Geräusch von einem lauten Geräusch überlagert. Das leise Geräusch wird nicht mehr wahrgenommen. Ein Beispiel hierfür ist das Grundrauschen bei CD-Aufnahmen. Rauschen vor und nach einem gespielten Ton ist wahrnehmbar, nicht aber das Rauschen während des Tons.

Ein gleichmäßiger Dauerton oder Rauschen wird mit zunehmender Zeitdauer immer leiser erscheinen, das Gehör ermüdet und ordnet diesen Dauerschall als

¹⁶ (Zwicker, 1990, 1999, 2007 S. 457)

¹⁷ vgl. (Zwicker, 1990, 1999, 2007 S. 457)

unwichtigeres Hintergrundgeräusch ein.¹⁸ Ein anderes Beispiel hierzu wäre ein Presslufthammer in unmittelbarer Umgebung: Die Sprache in Unterhaltungslautstärke kann nicht mehr verstanden werden, sie wird vom Geräusch des Presslufthammers verdeckt und nicht mehr wahrgenommen.¹⁹

Innerhalb der Maskierung durch Pegel spielt auch die Tonhöhe eine Rolle:

„Ein hoher Ton verdeckt einen tieferen sehr leicht, wenn beide Töne nicht weit voneinander entfernt sind. Ein tiefer Ton verdeckt einen viel höheren nur, wenn er in sehr großer Lautstärke vorhanden ist. Aus diesen Tatsachen erklärt sich, dass man mit zunehmender Pegelstärke die Wiedergabe tiefer Töne immer besser wahrnimmt, da der Verdeckungseffekt, der durch höhere Töne verursacht wird, nachlässt.“²⁰

Zusammenfassend kann gesagt werden: Leise Töne werden durch naheliegende lautere verdeckt. Ein zuerst klingendes lautes Ereignis kann ein danach folgendes Ereignis verdecken. Ein nach einem leisen Ereignis erklingendes lauteres kann ebenso das erstere verdecken.

Das folgende Experiment soll den Verdeckungseffekt durch Pegel in der Praxis verdeutlichen:

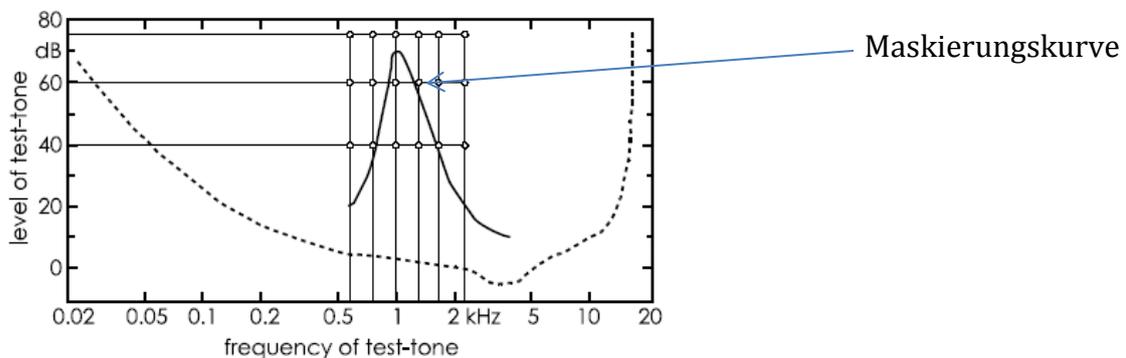


Abbildung E: Testtöne im schmalbandigen Rauschen ²¹

In Abbildung E ist eine Maskierungskurve dargestellt. Alle unterhalb dieser Linie liegenden Frequenzen sind unhörbar. Dem Schmalbandrauschen bei 1 kHz wird ein Pegel von 70 dB zugeordnet. Danach werden in drei Versuchsreihen die Testfrequenzen 600, 800, 1000, 1300, 1700 und 2300 kHz in verschiedenen Lautstärken (75 dB, 55 dB und 35 dB) abgespielt. Im Versuch 1 (75 dB) werden alle

¹⁸ vgl. (Dickreiter, 1997 S. 113)

¹⁹ vgl. (Kuttruff, 2004)

²⁰ (Rudolph)

²¹ (Zwicker, 1990, 1999, 2007)

Teiltöne wahrgenommen, in Versuch 2 (55 dB) verschwindet der Testton mit 1 kHz. In Versuch 3 (35 dB) verschwinden die Töne mit 800, 1000, 1300 und 1700 Hz. Die WAV-Datei zum Versuch des Verdeckungseffektes ist auf der CD im Anhang (vgl. 13.1 Audio-CD 1: Psychoakustische Effekte , S. I, Track 2) zu finden.²²

Ein alltägliches Beispiel für Verdeckung durch Pegel ist das Ticken eines Weckers: Bei Stille ist es gut hörbar, klingelt der Wecker, kann man es nicht mehr wahrnehmen, obwohl der Wecker weitertickt. Das Ticken wird also durch das Klingeln maskiert.

3. Standardisierte Testmethoden

Neue Arten der Kommunikation und neue Techniken der Digitaltechnik fordern neue Vorschriften und Richtlinien. Im Folgenden sollen nun verschiedene, durch die ITU standardisierte, Testmethoden zur Erfassung von Audioqualität erläutert werden. Die ITU ist die Internationale Fernmeldeunion mit dem Sitz in Genf, welche sich weltweit mit den Aspekten der Telekommunikation beschäftigt. Sie gliedert sich in zwei Bereiche:

Die **ITU-T** ist für die Radiokommunikation und für die Festlegung der Frequenzbereiche und Funkdienste verantwortlich.

Hierzu zählen für die subjektive Beurteilung von Audioqualität unter anderem die Empfehlungen der ITU-T P.800, P.810 oder P.830, welche sich mit der Qualität von Sprache am Telefon beschäftigen.

Die **ITU-R** erstellt Standards und Empfehlungen in den unterschiedlichsten Bereichen. Diese einheitlichen Verfahren ermöglichen zu jedem Zeitpunkt einen globalen Vergleich von Testergebnissen.²³

²² vgl. (Zwicker, 1990, 1999, 2007 S. 449)

²³ vgl. (ITU)

Es gibt fünf Dokumente und Richtlinien²⁴ der ITU-R, die sich mit der subjektiven Messung von Audioqualität beschäftigen:

| <u>Richtlinie</u> | <u>Inhalt</u> |
|-------------------|-------------------------------------------------------------------------------------------------|
| ITU-R BS.1284 | Allgemeine Hinweise für die Messung von subjektiver Audioqualität |
| ITU-R BS.1116 | Methoden für die subjektive Bewertung von kleinen, detailreichen Abweichungen von Audiomaterial |
| ITU-R BS.1285 | Methoden zur Vorauswahl von Teststücken für BS.1116-Tests |
| ITU-R BS.1286 | Methode für die subjektive Bewertung von Audiomaterial in Kombination mit visuellen Einflüssen |
| ITU-R BS.1534 | Methoden für die subjektive Bewertung von groben Abweichungen von Audiomaterial (BS.1534). |

Zur Messung subjektiver Audioqualität hat die ITU-R drei Testmethoden standardisiert (ITU-R BS.1116-1, ITU-R BS.1534-1 und ITU-R BS.1286). Diese Verfahren erlauben die Erfassung einer Störung durch gewisse Abweichungen und Verschlechterungen eines Signals. Somit lässt sich beispielsweise bestimmen welcher Codec die besten Ergebnisse im Hinblick auf eine geringe Datenmenge liefert bzw. ob Unterschiede von einer WAV- zu einer MP3-Datei hörbar sind.

Die festgelegten Bewertungsskalen, sowohl für Musik und Geräusche, als auch für TV-Programme erlauben eine statistische Prozessierungsmethode. Allerdings muss bei diesen Hörtests immer bedacht werden, dass die Art und der Ort der Ausführung (Raum, geometrische und akustische Eigenschaften) Einfluss auf das Testergebnis nehmen kann. Der Aufwand für Bewertungen und Auswertungen steigt mit dem Einfluss mehrerer Parameter, daher beschränkt man sich auf die sogenannte „Basic audio quality“, welche alle Veränderungen bzw. Beeinträchtigungen vereint. Zusätzlich können sogenannte Subparameter zur Analyse spezieller Aspekte hinzugezogen werden (z.B. „Stereophonic image quality“ (zweikanal-Beurteilung) oder mehrkanalige Tonbeispiele).²⁵

²⁴vgl. (ITU-R BS.1283-1, 1997-2003)

²⁵vgl. (ITU-R BS.1284, 1997-2003)

Neben den Empfehlungen der ITU, welche sich auf sogenannte Referenz-Vergleichstests (die Erkennung kleiner Qualitätsbeeinträchtigungen von Audiomaterial bzw. Audiosystemen durch einen direkten Vergleich mit einer Referenz²⁶) beziehen, gibt es weitere Teststrategien²⁷, mit denen es möglich ist, die subjektive Qualität von Audiomaterial einzuordnen:

Absolutbewertung: Wenn kein Vergleichsmaterial vorhanden ist, wird das Audiomaterial nach der Hörerfahrung der Testhörer beurteilt. Hier können verschiedene Faktoren, Wahrnehmbarkeiten und Effekte beurteilt werden. (siehe auch 4.1 Absolutbewertung auf S. - 25 -).

Paar-Vergleichstest: Hier werden zwei oder mehr Testobjekte miteinander verglichen – entweder miteinander oder jeweils zu einer Referenz – und anhand einer Testskala eingestuft. Normalerweise werden kurze Programmausschnitte (ca. 15 s) verwendet, welche mit der Referenz bzw. dem Testobjekt verglichen werden. Hier können absolute oder relative Bewertungen zustande kommen. (siehe auch 4.2 ABX-Blindhörtest (Paar-Vergleichstest) auf S. - 26 -).

3.1 Triple stimulus-hidden reference-double blind test

Bei der Testmethode nach ITU-R BS.1116-1, auch ABC/HR bzw. „Triple stimulus-hidden reference-double blind test“ genannt, werden subjektiv wahrgenommene Schädigungen des bearbeiteten Signals im Vergleich zum Originalsignal gemessen. Die Empfehlung muss aufgrund neuer Entwicklungen und Kodierverfahren ständig verbessert werden. Die Standardisierung dieser Bewertung für Audiosignale führt zur Möglichkeit des Austausches und der Kompatibilität von Testdaten. Mithilfe der Richtlinien der ITU-R BS.1116-1 können Tests durchgeführt werden, die sehr kleine hörbare Beeinträchtigungen messen.

Die ITU-R BS.1116-1 bildet die Grundlage für die „Triple stimulus-hidden reference-double blind Test“-Methode. Diese Methode hat sich als sehr sensibel und stabil erwiesen, um genaue Ergebnisse von kleinen Abweichungen zum Originalsignal zu erfassen. Eine Testperson hat die Auswahl zwischen drei Stimuli. Die bekannte

²⁶ vgl. (Dickreiter, 2008 S. 1166f)

²⁷vgl. (Dickreiter, 2008 S. 1166f)

Referenz ist immer Stimulus A. Hinter B und C verbirgt sich eine versteckte Referenz oder das Testsignal welches ggf. einen anderen Codec verwendet bzw. eine andere Qualitätsstufe ausweist.

Eines der beiden Stimuli B und C sollte keine Abweichungen vorweisen, da es identisch mit der Referenz ist. Der Proband hat die Aufgabe die beiden Stimuli B und C in Bezug auf die Referenz A zu bewerten. Die Abweichung von B mit A und C mit A wird mittels einer standardisierten 5-stufigen Skala erfasst und bewertet. Die Bewertungsskala richtet sich nach folgender Tabelle ²⁸:

| Störung | Note |
|--------------------------------------------------------------------|-------------|
| Nicht wahrnehmbar (Imperceptible) | 5.0 |
| Wahrnehmbar, aber nicht störend (Perceptible, but not annoying) | 4.0 |
| Wenig störend (Slightly annoying) | 3.0 |
| Störend (Annoying) | 2.0 |
| Stark störend (Very annoying) | 1.0 |

Sobald die Testperson die Bewertung eines Versuches beendet hat, sollte es möglich sein, direkt zum nächsten Versuch überzugehen. Hier werden die Stimuli B und C neu gemischt und der Hörer muss eine neue Bewertung vornehmen. Das Testtempo kann nach eigenem Ermessen gewählt werden, der Testhörer bekommt so viel Zeit, wie er benötigt, um eine Beurteilung durchzuführen.

Die Beschriftung der Bewertungsskala kann zu einer gewissen Tendenz führen, daher gibt es auch die Möglichkeit, die Bewertung lediglich mit den Zahlen 1-5 durchzuführen. Dies muss in einem Protokoll dokumentiert werden. ²⁹

²⁸ vgl. (ITU-R BS.1116-1, 1994-1997 S. 4ff)

²⁹ (ITU-R BS.1116-1, 1994-1997 S. 4ff)

3.1.1 Auswahl und Selektion der Testpersonen

Beim Hörtest mit minimalen Abweichungen (nach ITU-R BS.1116-1), werden Experten auf diesem Gebiet benötigt. Unter Experten versteht man „Personen mit guter Urteilssicherheit und geschultem, analytischem Hörvermögen.“³⁰

Die Expertenhörer sind diejenigen, denen es möglich ist, minimale Störungen oder Abweichungen zu erfassen. Je höher die Qualität eines Tests ist bzw. je geringer die Abweichungen der geschädigten Datei zum Original sind, desto wichtiger ist es, mit Experten zu arbeiten.³¹ Die Testhörer müssen über den Versuchsaufbau und Ablauf informiert werden und in einem Training vorab auf den Test vorbereitet werden. Jeder sollte ein individuelles Urteil, unbeeinflusst von anderen Testhörern, abgeben können.

Durch audiometrische Überprüfungen können Testpersonen vorab vom Test ausgeschlossen werden. Hier werden vorherige Erfahrungen und Leistungen geprüft und ausgewertet. Ein bestimmtes Training kann hier zu einer schnellen Selektion der „Unfähigen“ führen. Allerdings kann dies auch während des Tests passieren. Sollte es zum Ausschluss einer Testperson kommen, muss dies mit einer Begründung dokumentiert werden, damit auch spätere Beurteiler sich ein Bild der Situation verschaffen können. Die Postauswahlmethode erfolgt nach dem Test: Hier kann es passieren, dass extreme Abweichungen eines einzelnen im Vergleich zu den anderen Tests einen Ausschluss fordern. Ebenfalls kann es zu einem Ausschluss kommen, wenn der Eindruck entsteht, dass die Testperson keine genaue Angaben in Bezug auf den Hörtest gemacht hat. Allerdings ist hier zu bedenken, dass die Postauswahlmethode zwar das Ergebnis eindeutiger machen kann, jedoch empfindet jeder Mensch gewisse Artefakte anders und dies sollte auch berücksichtigt werden.³²

3.1.2 Trainings- und Testphase der Testhörer für einen ABC/HR-Test

Für einen Test bei welchem minimale Abweichungen vom Testsignal zum Referenzsignal unterschieden werden sollen, ist es nötig mit Experten zu arbeiten, die ein gewisses Training durlaufen haben. Das Training soll eine Vorbereitung auf den eigentlichen Test sein und soll Vertrauen schaffen bzw. eine Eingewöhnung in die

³⁰ vgl. (Dickreiter, 2008)

³¹ vgl. (ITU-R BS.1116-1, 1994-1997)

³² vgl. (ITU-R BS.1116-1, 1994-1997 S. 3)

Materie ermöglichen. Es soll die Hörer darauf vorbereiten, sich mit potenziellen Artefakten vertraut zu machen.

Die ITU gibt eine Möglichkeit vor, wie so ein Training konkret aussehen kann:³³

- Dauer bis zu 3 Stunden
- Vier Probanden (Testhörer), welche eine schriftliche Anweisung erhalten
- Ablauf:
 - Kurze Einführung in die Ziele
 - Vorspielen der ausgewählten Testauszüge, um den Hörern das Audiomaterial und die Soundpräsentation vorzustellen
 - Vorlesen der Bewertungsskala
 - Demonstration der Referenz und der Abweichung
 - Erklärung der Eigenschaften, die bewertet werden sollen
 - Einweisung in die Software
 - Einarbeiten in die gegebene Hardware
 - Einführung in die Bewertung von Audiomaterial

Anders als beim Hörtest wird beim Training auf einem Monitor angezeigt, welches der drei Stimuli die Referenz- und welches das veränderte Audiosignal ist. Somit bekommt der Testhörer ein Gefühl dafür, auf was geachtet werden soll. Die Audiosequenzen sollten wie beim eigentlichen Test 10 – 25 s lang sein, man kann hierfür Lautsprecher, Kopfhörer oder beides benutzen. Bevor es dann zum eigentlichen Hörtest geht, muss dem Testhörer die Bewertungsskala erklärt werden.

Im Anschluss daran sollte der eigentliche Hörtest stattfinden. Wie auch im Training bekommt der Hörer drei Stimuli A, B und C, die er miteinander vergleichen und bewerten soll.

Bei jedem Vergleich ist das Referenzsignal doppelt enthalten, somit sollte bei der Skalenbewertung bei einem Stimuli (B oder C) die Bewertung 5.0 auftreten. Wenn B und C besser bzw. gleich als die Referenz empfunden werden, dann ist dies ein Hinweis, dass die beiden Signale keine signifikanten Unterschiede zueinander aufweisen. Ziel eines Trainings soll sein, die Testperson zu einer möglichst

³³ vgl. (ITU-R BS.1116-1, 1994-1997 S. 45ff)

differenzierten Bewertung zu bringen. Während dem Test dürfen keine Absprachen mit und zwischen den Probanden über die Skalen abgehalten werden.³⁴

3.2 Multi Stimulus test with Hidden Reference and Anchor

Die Empfehlung ITU-R BS.1534-1 beschäftigt sich mit der Qualität verschiedener Kodierungstechniken für Audiomaterial, welches mit mittlerer Audioqualität, bevorzugt im Internet und Konsumentenbereich auftritt. Anders als in der ITU-R BS.1116-1 geht es hier nicht um die Abweichungen von minimalen Veränderungen von zwei Audiosignalen sondern um grobe Qualitätsveränderungen von Audiodateien wie sie vor allem in der Internetübertragung und im Rundfunk vorkommen. Bei derartigen Dateien mit schlechterer Audioqualität liefert die MUSHRA – Methode nach ITU-R BS.1534-1 zuverlässige Ergebnisse. Zur Beurteilung wird eine kontinuierliche Qualitätsskala von 0 - 100 (mangelhaft bis ausgezeichnet), bei welcher die grundlegende Audioqualität mit allen zusätzlichen Faktoren (Stereobreite, Auflösung usw.) beurteilt wird, benutzt. Es sollte mindestens eine Bewertung mit 100 Punkten geben, da hier ebenfalls ein verstecktes Referenzsignal im Vergleichsmaterial vorhanden ist. Die verwendeten Testsignale sollten eine Länge von ca. 20 s nicht überschreiten, da ansonsten die Testhörer recht schnell ermüden.

In jedem Test muss neben der Referenz eine versteckte Referenz und ein Vergleichspunkt (Anchor-point) mit einem Lowpass bei 3,5 kHz vorhanden sein. Es können auch noch weitere Anchor-points verwendet werden, beispielsweise mit einer Frequenzbandbeschränkung bei 7kHz.

Die Zielgebiete unterscheiden sich darin, dass die ITU-R BS.1116-1 sich mit der Evaluierung von hochqualitativen Audiocodecs bei unbeträchtlichen Störungen beschäftigt, während die MUSHRA-Methode für die Beurteilung von Signalen mit mittleren sowie starken Verzerrungen und somit bei Beispielen mit groben Qualitätsunterschiede einsetzbar ist. Die Testperson hat die Möglichkeit zwischen allen Signalen in beliebiger Reihenfolge zu vergleichen, allerdings muss in jedem Durchgang die Referenz als Vergleich zur Verfügung stehen. Bei der MUSHRA-Methode handelt es sich im Gegensatz zur ABC/HR-Methode um einen Vergleich, bei welchem bis zu 15 Testsignale (12 Testsignale + Referenz + versteckte Referenz + Anchor-point) gleichzeitig verglichen und beurteilt werden können.

³⁴Vgl. (ITU-R BS.1116-1, 1994-1997 S. 23f)

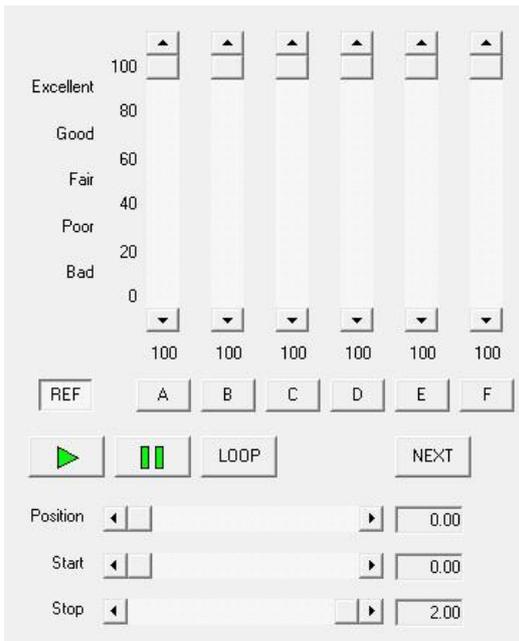


Abbildung F: MUSHRA Oberfläche³⁵

Der Hörer hat die Benutzeroberfläche (siehe Abbildung F: MUSHRA Oberfläche) vor sich. Sollte es keine Software geben, kann auch mit einem Papier und einer gedruckten Skala gearbeitet werden. Jeder Ablauf sollte dokumentiert werden, um mögliche, große Abweichungen zu erkennen. Eine einfache Möglichkeit hierfür ist das Mitschneiden eines Videos während des Tests. Sollte dann während oder nach dem Test erkennbar sein, dass eine große Abweichung oder Fehleinschätzung vorhanden ist, lässt sich leicht mit einer Prüfung des Videomaterials nachvollziehen, ob es sich hier um menschliches Versagen oder um einen Hardwaredefekt der Geräte handelte.

3.2.1 Auswahl der Testhörer

Anders als bei der Richtlinie ITU-R BS.1116-1 ist es bei der MUSHRA-Methode nicht nötig, ausschließlich mit Experten zu arbeiten, da hier keine minimalen Qualitätsverluste erkannt werden müssen. Die Testhörer sollten allerdings Erfahrungen im kritischen Hören von Audiomaterial mitbringen, da diese zu einem schnelleren und glaubwürdigeren Ergebnis kommen, als absolut unerfahrene Hörer. Wie in der ABC/HR-Methode können auch hier Testhörer bereits vor oder nach dem

³⁵ (Klein(Screenshot))

Test ausgeschlossen werden (siehe 3.2.1 Auswahl der Testhörer auf S. - 17 -). Ein gesundes und funktionsfähiges Gehör wird bei beiden Methoden vorausgesetzt.³⁶

3.2.2 Trainingsphase der Testhörer für einen MUSHRA-Test

Der Testhörer sollte vor dem eigentlichen Test eine Trainingsphase durchlaufen. Anders als im „Triple stimulus-hidden reference-double blind Test“ ist dieser jedoch nicht so aufwändig und zeitintensiv. Das Training gliedert sich in zwei Phasen:

- In der ersten Phase können alle Soundbeispiele angehört werden. Wie im späteren Test sind auch hier alle Qualitätsstufen inklusive der Referenz und den Anchor-points vorhanden. Während dieser Phase werden keine Bewertungsskalen angezeigt.
- In der zweiten Phase sollen sich die Probanden mit der Software und deren Umgang vertraut machen. Es soll deutlich gemacht werden, wie die vorhandenen Beispiele genutzt werden und wie die Qualität beurteilt werden kann. Das Ergebnis sollte nicht mit anderen Testteilnehmern diskutiert werden, jeder Testhörer sollte sich selbständig in die Materie einarbeiten. Das Training fließt nicht in die Bewertung des Tests ein.³⁷

3.2.3 Testmaterial

Das zu verwendende Testmaterial ist nicht fest definiert, allerdings würde es sich anbieten mit typischem „Broadcast-Material“ zu arbeiten, welches in der Übertragung genutzt wird. Die Suche nach passendem Material kann sehr zeitintensiv sein. Es ist sehr wichtig, dass das ausgesuchte Material genau definiert wurde und eine genaue Beschreibung für die zu bewertenden Kriterien vorliegt. Zudem sollten die verschiedenen Testobjekte in ihrer Lautstärke angepasst werden, um den psychoakustischen Effekt (Lauter = Besser) zu vermeiden.³⁸

³⁶ vgl. (ITU-R BS.1534, 2001-2003 S. 5f)

³⁷ vgl. (ITU-R BS.1116-1, 1994-1997 S. 3ff)

³⁸ vgl. (ITU-R BS.1116-1, 1994-1997 S. 3ff)

3.3 Grundsätzlicher Testaufbau

Die Dokumente ITU-R BS.1116-1 und ITU-R BS.1534-1 gliedern sich in verschiedene Bereiche. Im Versuchsaufbau der Dokumente wird beschrieben, dass die Tests streng nach den Empfehlungen und so formal als möglich durchgeführt werden sollen. Eine sorgfältige Durchführung der Tests ist von Nöten, damit keine unkontrollierten Faktoren in den Test miteinfließen und das Ergebnis beeinflussen. Der Test muss so aufgebaut werden, dass diese Störfaktoren sofort ersichtlich sind und ggf. berücksichtigt oder verändert werden können. Sollten dauerhafte Störfaktoren während dem Test vorhanden sein, dann sollten diese protokolliert und im Ergebnis berücksichtigt werden. Die Hörtests müssen so konstruiert werden, dass die Probanden nicht überfordert werden und ihr Gehör nicht überstrapaziert wird; das bedeutet, die Anzahl der durchzuführenden Hörbeispiele sollte gut überlegt werden. Es ist wichtig, dass keine visuellen Einflüsse das Ergebnis beeinflussen. Möchte man ein Ergebnis erzielen, dass mit visuellen Komponenten arbeitet, sollte man die Empfehlung ITU-R BS.1286 berücksichtigen.

Für den Testablauf ist es wichtig, dass unbeschädigtes Audiomaterial zur Verfügung steht, welches der Testperson unbekannt ist. Zudem sollte der Test so aufgebaut werden, dass es für den Probanden nicht ersichtlich ist, was als nächstes folgt. Durch die Unterschiede zwischen „unbeschädigtem“ und qualitativ schlechterem bzw. „geschädigtem“ Signal entstehen die fünf Stufen der Bewertungsskalen je nach angewandter Empfehlung.³⁹

3.4 Anzahl der Testhörer

Eine Anzahl von Testhörern kann vorab festgelegt werden, wenn eine statistische Grundlage vorliegt und eine Varianz eingeschätzt werden kann. Die Erfahrung der letzten Jahre hat gezeigt, dass eine Anzahl von 20 Testhörern bei beiden Methoden ausreicht, um ein nennenswertes Ergebnis zu erhalten.⁴⁰

³⁹ vgl. (ITU-R BS.1116-1, 1994-1997 S. 3ff)

⁴⁰ vgl. (ITU-R BS.1116-1, 1994-1997 S. 3)

3.5 Der Abhörraum

Die Auswahl geeigneter Abhörsituationen ist ein wesentlicher Bestandteil für die Kontrolle von Audiomaterial. Für Vergleichstests, wie in der ITU-R BS.1116-1 empfohlen, eignen sich bestimmte Standards, um ein repräsentatives Ergebnis erzielen und auswerten zu können. Die geometrischen und akustischen Eigenschaften des Abhörraumes sowie die Anordnung der Lautsprecher und die Auswahl der Hörplätze spielen eine große Rolle. Eine transparente und neutrale Abhörmöglichkeit sollte gegeben sein. Der Abhörraum sollte einen natürlichen Raumeindruck vorweisen, jedoch sollte der Anteil der Rauminformationen nur so stark wahrnehmbar sein, dass die zu beurteilenden Signale nicht überdeckt oder verfälscht werden.

Die Raumgröße für eine Mono- oder Stereowiedergabe sollte bei ca. 20-60 m² liegen. Für eine Multikanalabhörsituation werden ca. 30-70 m² benötigt. Besonders gut eignen sich rechteckige oder trapezförmige Raumabmessungen. Durch ein nichtganzzahliges Verhältnis von Breite zu Länge können stehende Wellen vermieden werden. Durch das Anbringen von Absorbern oder reflektierenden Flächen ist das Auftreten von Flatterechos oder Reflexionen zu vermeiden. Der Nachhall sollte möglichst diffus sein, die mittlere Nachhallzeit zwischen 200 Hz und 4 kHz sollte zwischen 0,2 und 0,4 s liegen.⁴¹

3.6 Die Abhörmethoden

Der Amplitudenfrequenzgang der Lautsprecher soll im Frequenzbereich 40 Hz bis 16 kHz nicht mehr als 3 dB überschreiten. Die Frontlautsprecher dürfen im Bereich zwischen 250 Hz und 2 kHz maximal eine Abweichung von 1 dB aufweisen. Die Präsenzwirkung der Lautsprecher, das Bündelungsmaß, soll dafür sorgen, dass ein wesentlicher Teil der Schallenergie als Direktschall den Hörer erreicht. Hier liegen die Grenzen zwischen 500 Hz und 10 kHz, bei einem Bündelungsmaß von 6 dB bis 12 dB. Der Klirrfaktor sollte bei 90 dB/SPL bei Frequenzen unter 250 Hz bei 3 % liegen, bei Frequenzen über 250 Hz bei 1 %. Die Abklingzeit richtet sich nach der Frequenz und sollte folgenden Richtwert nicht überschreiten:

⁴¹Vgl. (ITU-R BS.1116-1, 1994-1997 S. 12f)

$$t < 5/f$$

t = Abklingzeit

f = Frequenz

Zwischen den zwei Lautsprechern (bei Stereowiedergabe) herrscht ein Zeitversatz, dieser sollte 100 μ s nicht überschreiten.⁴²

Hörtests sollten im Normalfall über Lautsprecher erfolgen, sollten die räumlichen Bedingungen eine qualitativ hochwertige Wiedergabe ausschließen, können auch professionelle Kopfhörer zum Einsatz kommen, da hier die Raumbedingungen keinen Einfluss haben.⁴³

3.7 Audiomaterial für Hörtests – SQAM-CD

Die wesentliche Voraussetzung für die Qualitätsbeurteilung ist das Testmaterial. Bei den Testdateien sollte es sich um Programmausschnitte mit einer maximalen Länge von ca. 20 Sekunden handeln. Die Empfehlung ITU-R BS.1284 gibt verschiedene Ratschläge zur Auswahl der Hörbeispiele. Die Vergleichbarkeit von Hörtests beruht auf identischem Testmaterial. Die EBU hat hierzu eine CD mit verschiedenen Audiomaterialien zusammengestellt, welche verschiedene Testsignale für die subjektive Qualitätswahrnehmung in sämtlichen Bereichen beinhaltet. Die Vielzahl des gesammelten Materials soll es ermöglichen, Testsignale für fast jede Aufgabe zu finden. Die Auszüge auf der CD sind speziell für Testzwecke erstellt worden. Die CD beinhaltet 70 Tracks. Mit der CD wird auch ein Handbuch geliefert, in welchem eine Tabelle mit näheren Erläuterungen zu den einzelnen Tracks aufzeigt ist (s. Anhang 13.2 CD und Trackinformationen zur SQAM). Das Audiomaterial der CD lässt sich in vielen Fällen besser benutzen als selbst gesichtetes Material. Es gibt insgesamt neun Anwendungsgebiete für die sich die einzelnen Tracks gut oder schlecht verhalten, in einer Tabelle lässt sich ablesen welche Tracks für welche Testgebiete geeignet sind.⁴⁴

⁴² vgl. (ITU-R BS.1116-1, 1994-1997) und (Dickreiter, 2008 S. 1158ff)

⁴³vgl. (Dickreiter, 2008 S. 1164)

⁴⁴ (3253, September 2008)

- **AD/DA- Wandler**

Das Konvertieren von analogem zu digitalem Material oder umgekehrt kann Rauschen oder Verzerrungen verursachen. Beim Testen der Wandler kommt es vor allem in tiefen Frequenzen zu Artefakten.

- **Abtastfehler**

Bei hochfrequentem Audiomaterial (Frequenzen über der Hälfte der Samplingfrequenz) können leicht hörbare Artefakte beim Wandeln von analog zu digital entstehen.

- **Bitfehler**

Bitfehler im digitalen System können feine Hintergrundgeräusche oder sehr störende vordergründige Effekte sein. Jedes System reagiert anders auf zufällige Fehler und Häufungen von Fehlern, dies ist abhängig von seiner Fehlerkorrektur und Fehlerverdeckung.

- **Bitratenreduktion**

Es gibt verschiedene Techniken, um die Bitrate zu minimieren. Diese Techniken (Dynamikeinschränkungen, Teilcodierung oder eine Differential Pulse Code Modulation) variieren sehr stark. Das auf der CD zusammengestellte Material ist so gewählt, dass diese Schwächen ggf. aufgedeckt werden können.

- **Dynamikumfang**

Testmaterial mit einem großen Dynamikumfang ist nützlich um die Effekte von Dynamikveränderungen durch einen Limiter oder einen Kompressor deutlich zu machen.

- **Frequenzantwort**

Die Testdaten mit einem großen Frequenzspektrum sind wichtig, um Systeme zu testen, die eine geringe Bandbreite haben. So können eventuelle Klangfärbungen erkannt werden.

- **Übersteuerung nach der Bearbeitung**

Signale, die „gut angesteuert“ aufgenommen wurden, können nach dem Bearbeiten durch einen EQ o.ä. einer Verzerrung oder einer Klangfärbung unterliegen. Mit den auf der CD enthaltene Testdaten lassen sich solche Verfärbungen erkennen und einschätzen.

- **Eigenrauschen**

Alle dynamischen Anpassungen und Systeme in der Bearbeitung beinhalten ein gewisses Eigenrauschen. Das Testmaterial wurde so ausgesucht, dass das Eigenrauschen durch die wirklichen Signale nicht oder nur minimal verdeckt wird. Verschiedene Audiodateien verhalten sich für bestimmte Geräte besser, da nicht jedes Gerät ein identisches Grundrauschen besitzt.

- **Stereobreite**

Dieses Testmaterial eignet sich, um die Ortung von Geräuschen und Instrumenten zu identifizieren. Die vorhandene Stereobreite kann durch bestimmte Systeme mittels Amplituden oder Phasenveränderungen verfälscht werden.⁴⁵

3.8 Dokumentation und Präsentation der Testergebnisse

Für die Darstellung der Ergebnisse ist es wichtig, dass alle Informationen schnell ablesbar sind und auch für Nichtexperten verstanden werden können. Der Testbericht sollte neben dem Grund des Tests und den Methoden auch eine klare Struktur enthalten. Der Test sollte alle Ergebnisse beinhalten, sodass diese jederzeit nachvollzogen werden können. Folgende Punkte sollten enthalten sein:

- **Grafische Übersicht der Ergebnisse**
- **Die Beschreibung und Auswahl der Testhörer**

Die Auswahl der Testhörer kann das Ergebnis in eine bestimmte Bahn lenken, daher ist es wichtig die Auswahl der Testpersonen zu dokumentieren.

⁴⁵ (3253, September 2008)

- **Die Beschreibung und Auswahl des Audiomaterials**
- **Informationen über die Software.**
- **Informationen zum Testaufbau**
- **Dokumentation der physikalischen Bedingungen (Raumgröße, Akustik, Schallübertragung, Hardwaregeräte)**
Hier macht es einen Unterschied, wie die Beschaffenheit der Räume ist und ob Lautsprecher oder Kopfhörer verwendet wurden, daher sollte dies sehr genau dokumentiert werden.
- **Der Versuchsaufbau: Training, Einführung, Testablauf, Dokumentation**
- **Datenverarbeitung: Datenanalyse und Statistiken**
- **Abschlussbericht**

In der Dokumentation sollten des Weiteren folgende Punkte auftauchen:

- **Die Anzahl der Auswerter und Beobachter**
- **Die Übersicht über alle Testelemente**
- **Der Durchschnitt und das 95%-Konfidenzintervall (Statistik)**

4. Weitere Testmethoden

Neben den Empfehlungen der ITU (ITU-R BS.1116-1 und ITU-R BS.1534-1) gibt es noch weitere Testmethoden, welche keine Standardisierung durch die ITU aufweisen.

4.1 Absolutbewertung

Bei der Absolutbewertung stehen keine Vergleichsobjekte zur Verfügung. Eine Testperson bewertet die Audioqualität lediglich auf Grundlage der Hörerfahrung und Hörgewohnheit. Es kann nach bestimmten Effekten gefragt werden:

| Hauptparameter | Teilparameter | Skale | Bestimmende Faktoren (Beispiele) |
|------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------|-------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Raumeindruck (Spatial impression) | - Homogenität des Raumschalls - Nachhalldauer, Hallbalance - empfundene Raumgröße - Tiefenstaffelung | II | Nachhallzeit T Direktschallmaß Frequenzgang der Nachhallzeit Richtungsdiffusität |
| Stereofoner Eindruck (Stereo/Surround impression) | - Richtungsgleichgewicht - Richtungsstabilität - Abbildungsbreite - Lokalisierungsschärfe | II | Pegeldifferenzen Phasen-/Laufzeitdifferenzen Übersprechabstand |
| Durchsichtigkeit (Transparency) | - Registerdurchsichtigkeit - Zeitdurchsichtigkeit - Textverständlichkeit - Durchmischung | II | Deutlichkeit Silbenverständlichkeit raumakustische Diffusität |
| Akustisches Gleichgewicht (Sound balance) | - Lautstärkegleichgewicht - Dynamikbereich - Klangfarbengleichgewicht | II | Mischungsverhältnis Dynamik, Kompression |
| Klangfarbe (Timbre) | - Klangfarbe Direktschall - Klangfarbe Raumschall - Klangeinsatz | III | Amplitudenfrequenzgang nichtlineare Verzerrungen Einschwingvorgänge |
| Störgeräusche (Noise) | - akustische Störwirkungen - elektrische Störwirkungen - Codierstörungen - Aufnahmetechnik - raumakustische Irregularitäten | I | Spiel-, Publikums-, Klima- geräusche Rauschen, Knacke, Brummen Bitfehler, Vor-/Nachechos, Regelvorgänge, Schnittfehler, Echos, Resonanzen, Einzelre- flexionen |
| Akustischer Gesamteindruck (Basic audio quality / Main impression) | (summarische Bewertung aller Haupt- und Teilparameter) | III | |

Abbildung G: Bewertungsparameter für die Absolutbewertung ⁴⁶

⁴⁶ (Dickreiter, 2008 S. 1166)

Die Skalenbewertung teilt sich in drei Teile auf ⁴⁷:

Typ I: Wahrnehmbarkeit von Beeinträchtigung

Typ II: Vergleich von Varianten

Typ III: absolute Bewertung der Qualität

4.2 ABX-Blindhörtest (Paar-Vergleichstest)

Der ABX-Blindhörtest ist eine vereinfachte Form des „Triple stimulus-hidden reference-double blind-Tests“. Anders als beim ABC/HR-Test müssen hier keine Bewertungen abgegeben werden. Es werden lediglich eine Referenz und zwei weitere Buttons angegeben, unter welchen sich jeweils eine versteckte Referenz und das veränderte Audioformat befinden. Der Testhörer hat die Aufgabe herauszufinden, wo sich die versteckte Referenz befindet.

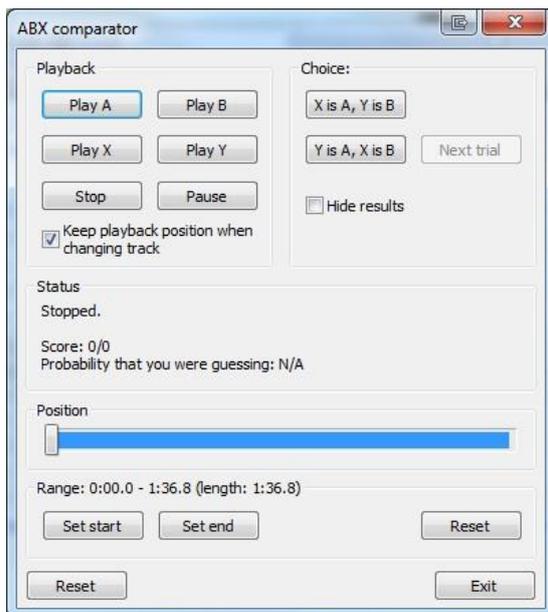


Abbildung H: ABX-Test mit foobar2000 ⁴⁸

In der Software *foobar2000* werden die zwei verschiedenen Testdateien den Buchstaben A und B sowie X und Y zugeordnet. Der Testhörer hat dann die Aufgabe zu entscheiden ob „X = A“ und „Y = B“ oder „X = B“ und „Y = A“ ist. Dieser Test wird mehrfach hintereinander durchgeführt. Je öfter der Testhörer das Ergebnis richtig

⁴⁷ (Dickreiter, 2008 S. 1116)

⁴⁸ (Klein(Screenshot))

erkannt hat, desto wahrscheinlicher ist es, dass ein Unterschied zwischen den zwei Dateien vorhanden ist.

5. Statistik

Ergebnisse von subjektiven Qualitätsparametern zeigen oft sehr unterschiedliche individuelle Ergebnisse auf. Deshalb müssen für ein aussagekräftiges Ergebnis mindestens 20 Personen befragt werden, um eine statistische Sicherheit zu erlangen. Für das Ergebnis werden im Allgemeinen die Stichproben mit Hilfe des arithmetischen Mittels und des sogenannten Konfidenzintervalls des Mittelwerts beschrieben, jeweils bezogen auf eine statistische Sicherheit von 95% (MUSHRA-Test). Bei einem Hörtest nach der ITU-R BS.1116-1 kommen komplexere und detailliertere Analysen zum Zuge (z.B. ANOVA).⁴⁹

5.1 Arithmetischer Mittelwert

Als Mittelwert (Durchschnittswert) wird meistens das arithmetische Mittel benutzt. „Gegeben seien Daten bezüglich eines quantitativen Merkmals, also eine Stichprobe aus miteinander vergleichbaren Zahlenwerten. Das arithmetische Mittel ist dann der Quotient aus der Summe aller Werte geteilt durch die Anzahl der Werte.“⁵⁰

$$\bar{u}_{jk} = \frac{1}{N} \sum_{i=1}^N u_{ijk}$$

u_i = Bewertung der Testperson i bei verschiedenen Testbedingungen und einem Audiobeispiel k

N = Anzahl aller Testpersonen⁵¹

u_j und u_k werden für jeden Testdurchlauf und jede Testsequenz einzeln ermittelt.

⁴⁹vgl. (Dickreiter, 2008 S. 1174)

⁵⁰ (Bosch, 2002 S. 26)

⁵¹ Vgl. (ITU-R BS.1534, 2001-2003 S. 12)

5.2 Standardabweichung

Im Idealfall ergeben sich bei der Auswertung gleiche Werte aller Testpersonen, was jedoch in der Praxis nicht vorkommt. Die Standardabweichung beschreibt das Maß der Streuung der Werte um das arithmetische Mittel. Sie gibt somit auch an, wie stark die Stichproben miteinander übereinstimmen.⁵² Die Standardabweichung ist die durchschnittliche Entfernung bzw. der durchschnittliche Abstand zum arithmetischen Mittel.⁵³

Die Standardabweichung wird folgendermaßen berechnet:

$$S_{jk} = \sqrt{\sum_{i=1}^N \frac{(\bar{u}_{jk} - u_{ijk})^2}{(N-1)}} \quad 54$$

Für eine bessere Übersicht empfiehlt es sich, die einzelnen Testsequenzen darzustellen, anstatt nur die Mittelwerte und Konfidenzintervalle aufzuzeigen.⁵⁵

5.3 Das Konfidenzintervall

Aus einer Datenerfassung tauchen verschiedene Resultate häufiger auf als andere Ergebnisse, die genau oder weitgehend mit der Grundgesamtheit übereinstimmen. Mit dem Konfidenzintervall wird die Tatsache berücksichtigt, dass die Stichprobenergebnisse voneinander abweichen und somit Hinweis darauf geben, wie hoch die Abweichung erwartungsgemäß ausfallen wird. Konfidenzintervalle können verschieden breit sein.⁵⁶ Die ITU empfiehlt ein Signifikanzniveau von 95%, das bedeutet, dass der Mittelwert zu 95% in dem vom Konfidenzintervall beschriebenen Bereich liegt.⁵⁷

Für eine Bewertung kann das Konfidenzintervall ebenfalls ein geeignetes Mittel sein. Überlappen sich die Konfidenzintervalle zweier Ergebnisse, dann kann man darauf schließen, dass sie sich nicht signifikant unterscheiden.

⁵² (Bosch, 2002)

⁵³ Vgl. (Rumsey, 2010 S. 63)

⁵⁴ Vgl. (ITU-R BS.1534, 2001-2003)

⁵⁵ Vgl. (ITU-R BS.1534, 2001-2003 S. 12f)

⁵⁶ Vgl. (Rumsey, 2010 S. 70)

⁵⁷ Vgl. (ITU-R BS.1534, 2001-2003 S. 12f)

Die Berechnung erfolgt mit folgender Formel:

$$\left[\bar{u}_{jk} - \delta_{jk}, \bar{u}_{jk} + \delta_{jk} \right]$$

bei

$$\delta_{jk} = t_{0.05} \frac{S_{jk}}{\sqrt{N}}$$

Der t-Wert beschreibt das Signifikanzniveau, $t_{0,05}$ steht für den Wert von 95%.

5.4 Die Normalverteilung

Jede Datenmenge hat einen Mittelwert. Sind die Daten normalverteilt, dann sammeln sich die Daten um diesen Mittelpunkt nach der glockenförmigen oder Gaußschen Wahrscheinlichkeitskurve.⁵⁸ Je weiter sich die Datenergebnisse vom Mittelpunkt entfernen, desto weniger Datenpunkte sind dort zu finden. Optisch lässt sich diese Verteilung als Kirchturmglöcke beschreiben. In der Statistik wird diese Kurve Normalverteilung genannt.⁵⁹

5.5 Größe der Stichproben

Durch einen größeren Stichprobenumfang, variieren die Ergebnisse weniger und sind stichfester. Bei kleinen Stichprobengrößen fällt der Standardfehler grob aus. Je mehr Daten erhoben werden, desto weniger Abweichungen erhält man von Stichprobe zu Stichprobe. Das bedeutet zusammenfassend, je größer die Stichprobe, desto kleiner ist die Fehlergrenze.⁶⁰

⁵⁸ (Fitch, 2012)

⁵⁹ Vgl. (Rumsey, 2010 S. 68)

⁶⁰ Vgl. (Rumsey, 2010)

5.6 Der Hypothesentest

In der Statistik dreht es sich häufig um Hypothesen, also Annahmen, die man bestätigen oder zurückweisen möchte. Ein Beispiel: Sind die Erbsen der einen Sorte süßer, als die der anderen

Hypothesentests werden eingesetzt, um eine Annahme über die Grundgesamtheit zu bestätigen bzw. zu verwerfen. Sobald ein Ergebnis „statistisch signifikant“ ist, handelt es sich um einen Hypothesentest. Dieser Test ist ein statistisches Verfahren, mit welchem Daten im Hinblick auf eine Behauptung über die Grundgesamtheit gesammelt werden. Wenn ein Hypothesentest korrekt durchgeführt wird, kann man damit der Wahrheit sehr nahe kommen, ohne sie zu kennen. Allerdings muss die Gesamtheit aller Daten betrachtet werden, denn sobald die Entscheidung auf einer Stichprobe basiert, kann der Hypothesentest falsche Schlussfolgerungen liefern.⁶¹

5.7 Der p-Wert

Die Annahme die geprüft werden soll, wird als Nullhypothese bezeichnet. Alle Hypothesentests benutzen einen p-Wert, um das Gewicht des Beweises zu bemessen, das heißt um zu ermitteln, was die Daten über die Grundgesamtheit aussagen.

Der p-Wert ist eine Zahl zwischen 0 und 1, welcher die Aussagekraft der Daten widerspiegelt. Ist der p-Wert klein, spricht diese gegen die Nullhypothese. Der p-Wert wird benutzt, um das Gewicht des Beweises zu bemessen. Ab einem p-Wert von unter 0,05 spricht man von einer Signifikanz zwischen den gemessenen Daten.⁶²

5.8 t-Verteilung und t-Test

Bei kleinen Stichproben (unter 30 Einheiten) benutzt man die sogenannte t-Verteilung anstelle der Z-Verteilung. Die t-Verteilung ist die etwas flachere Variante der Standardnormalverteilung (Z-Verteilung). Der p-Wert steigt ebenfalls bei sinkender Stichprobengröße.⁶³ Mittels des t-Tests lässt sich die Urteilszuverlässigkeit der einzelnen Personen abschätzen.

⁶¹ Vgl. (Rumsey, 2010 S. 72)

⁶² Vgl. (Rumsey, 2010 S. 72)

⁶³ Vgl. (Rumsey, 2010 S. 244)

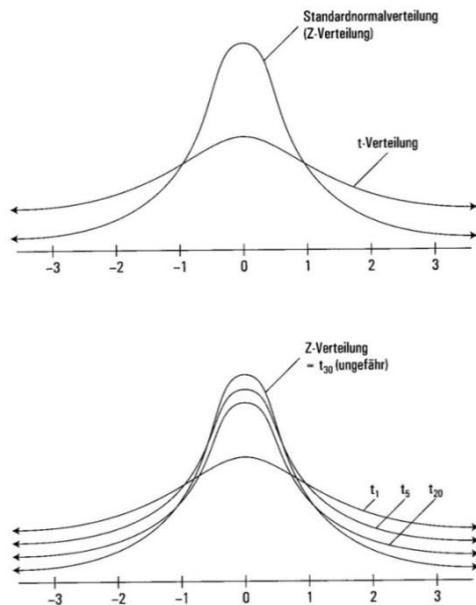


Abbildung I: Verteilung bei kleinen Stichproben ⁶⁴

5.9 Binomialverteilung

Sobald ein Experiment immer dieselben Antwortmöglichkeiten besitzt (A und nicht A) entsteht die Binomialverteilung; z.B. beim Werfen einer Münze ist bei jedem Wurf Kopf oder Zahl möglich. Mit der Binomialverteilung lässt sich berechnen, wie hoch die Wahrscheinlichkeit ist, 7-mal von 10 Würfeln Zahl zu werfen. Voraussetzungen für die Binomialverteilung sind:

- Das Ereignis A, in jedem Telexperiment, hat immer dieselbe Wahrscheinlichkeit p.
- Die einzelnen Experimente sind voneinander unabhängig, der zweite Versuch wird vom ersten nicht beeinflusst.

Betrachtet man den Fall von zwei aufeinanderfolgenden Fragen mit jeweils zwei Antwortmöglichkeiten, lässt sich die Wahrscheinlichkeit berechnen, wie hoch die Wahrscheinlichkeit ist, dass geraten wird. Dies lässt sich durch das Produkt der Einzelwahrscheinlichkeiten berechnen:

$$p=0,5 \quad p * p = p^2 \quad 0,5*0,5=0,25 \quad \Rightarrow 25\%$$

⁶⁴ (Rumsey, 2010 S. 241)

Mit diesem Wissen lässt sich auch recht schnell klären, wie hoch die Wahrscheinlichkeit ist, beispielsweise bei 2 von 8 Fragen richtig zu raten:

$$p^2 * (1-p)^8 = \binom{8}{2} * p^2 * (1-p)^6$$

$$\binom{8}{2} = 28 \text{ Möglichkeiten für 2 richtige Lösungen unter 8 Aufgaben}$$

$$\binom{8}{2} * 0,5^2 * (1 - 0,5)^6 = 0,11$$

Die allgemeine Gleichung für die Binomialverteilung lautet:

$$(x + a)^n = \sum_{k=0}^n \binom{n}{k} x^k a^{n-k}$$

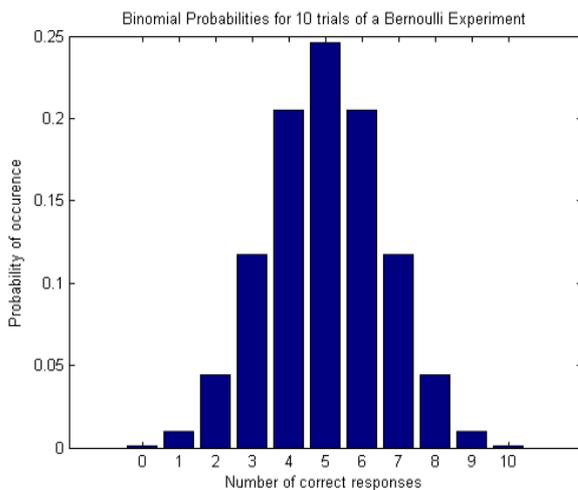


Abbildung J: Binomialverteilung ⁶⁵

Betrachtet man die Software *foobar2000*, mit der man die Möglichkeit besitzt, einen ABX-Test zu kreieren, welcher die Wahrscheinlichkeit anzeigt, dass die Testperson geraten hat, muss dafür der Rechengvorgang etwas geändert werden und die Anwendung der Summenverteilung kommt zum Zuge. ⁶⁶

⁶⁵ (Statistical Analysis of ABX Results Using Signal Detetion Theory, 2009)

⁶⁶ Vgl. (Oestreich, et al., 2012 S. 186ff)

Die Summenverteilung verwendet man, wenn verschiedene Ereignisse unterschiedliche Wahrscheinlichkeiten aufweisen. Dabei werden die Häufigkeiten beginnend mit den kleinsten Ausprägungen in aufsteigender Reihenfolge addiert.

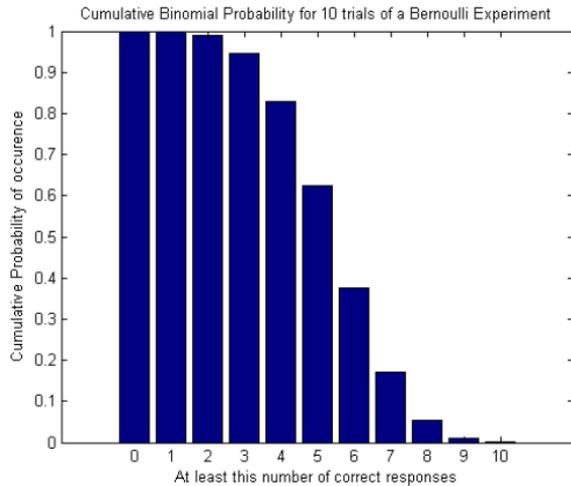


Abbildung K: Summenverteilung ⁶⁷

Beispiel: Bei 10 Hörversuchen und keiner richtigen Antwort, liegt die Wahrscheinlichkeit bei 100%, dass der Hörer geraten hat. Hingegen ergibt sich bei 10 richtigen Ergebnissen eine Wahrscheinlichkeit von 0,1%, dass der Proband geraten hat. ⁶⁸

⁶⁷ (Statistical Analysis of ABX Results Using Signal Detection Theory, 2009)

⁶⁸ vgl. (Statistical Analysis of ABX Results Using Signal Detection Theory, 2009)

6. Softwarebeschreibung: foobar2000 (ABX-Test)

foobar2000 ist ein Freeware-Audioplayer für Windows. Er wurde von Peter Pawlowski entwickelt, welcher als freiberuflicher Arbeiter bei der Firma „Nullsoft“ arbeitet. Der *foobar2000*-Media Player, der primär nichts mit einem Hörtest bzw. einem ABX-Test zu tun hatte, besitzt ideale Eigenschaften, die man von einem Media Player erwartet.

foobar2000 unterstützt sämtliche Audioformate, beispielsweise MP3, MP4, AAC, CD Audio, WMA, Vorbis, Opus, FLAC, WavPack, WAV, AIFF, Musepack, Speex, AU, SND. Neben diesen Standardformaten können auch Archive wie ZIP oder RAR abgespielt werden.

Neben mehreren Wiedergabelisten ist es in *foobar2000* möglich, sogenannte „Gapless-Playbacks“ abzuspielen, das bedeutet, dass eine Wiedergabe ohne Unterbrechung möglich ist. Dies ist beispielsweise bei einem klassischen Konzert sinnvoll, bei dem man pausenlos Musik genießen möchte. Zudem lässt sich die Oberfläche optisch individuell gestalten. Ein weiterer Vorteil der Software ist die „Replay-Gain-Funktion“. Mit dieser lässt sich eine individuelle Lautstärkeanpassung verschiedener Titel realisieren. Durch die „FreeDB“-Unterstützung lassen sich Metadaten von Onlinedatenbanken bequem verändern. Das Streamen von Audiodateien macht eine Beschallung im ganzen Netzwerk möglich.

6.1 Das ABX-Plug-In

Neben der Funktion als Audioplayer, lässt sich *foobar2000* mit verschiedenen Plug-Ins erweitern. In *der* Software lässt sich ein sogenanntes ABX-Plug-In installieren. Es handelt sich hier um den ABX-Blindtest, mit welchem ausgewertet werden kann, ob bestimmte Faktoren (z.B. Audiocodec) einen subjektiven Einfluss auf die Qualität eines Stückes haben.

Nachdem das Plug-In installiert wurde, werden die zwei Audioquellen, welche man vergleichen möchte, markiert und mit einem Rechtsklick auf „Utilities/ABX Two Tracks...“ ausgeführt.

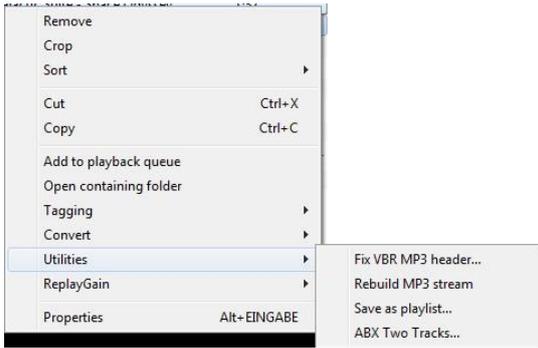


Abbildung L: ABX-Test in foobar2000 aufrufen ⁶⁹

Es öffnet sich ein Optionsfenster, in dem es die Möglichkeit gibt, die Lautstärke der beiden Audiobeispiele anzupassen. Es ist wichtig, dass die Testdateien aus einer Quelle stammen und somit bereits die gleiche Lautstärke haben, da ansonsten lautere Passagen subjektiv als besser empfunden werden. Wenn die Replayinformationen berücksichtigt werden sollen, kann hier ein Haken gesetzt werden. Zudem gibt es die Möglichkeit die DSP-Einstellungen, die der Benutzer für seine Playbacks bzw. seine Songs anwendet aus- bzw. anzuschalten. Auch hier empfiehlt es sich, sich auf das Originalsignal zu verlassen. Nach einer kurzen Berechnung steht der ABX-Komparator zur Verfügung.

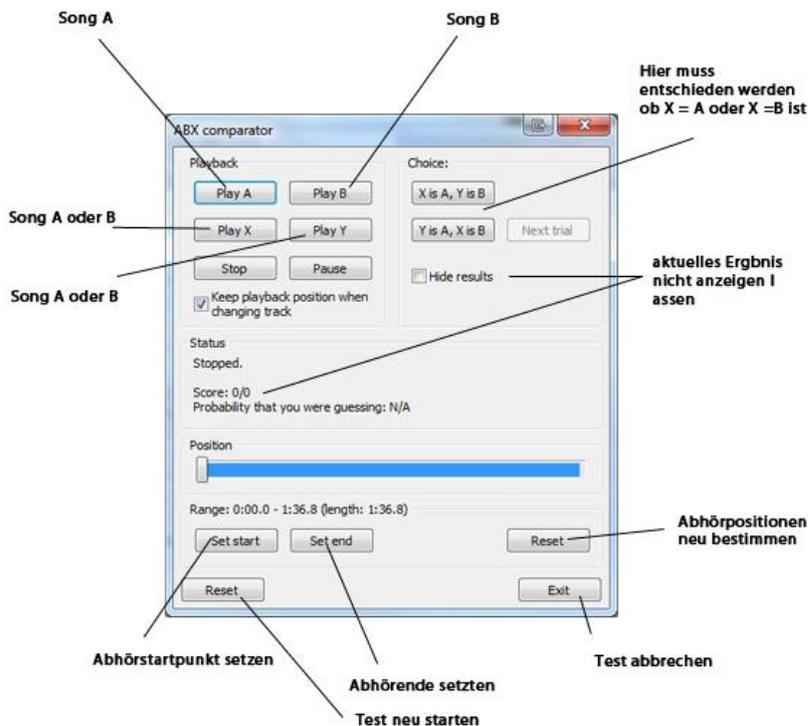


Abbildung M: ABX-Komparator in foobar2000 mit Erklärungen ⁷⁰

⁶⁹ (Klein(Screenshot))

Der ABX-Komparator hat lediglich die Aufgabe, die Dateien miteinander zu vermischen. Es steht Song A und Song B zur Auswahl. Diese Audiodateien werden ebenfalls auf X und Y kopiert. Die Aufgabe der Testperson ist es, festzustellen, ob A/X und B/Y oder A/Y und B/X identisch sind. Der Benutzer kann sich die Testdateien solange anhören, wie er möchte. Sobald er sich entschieden hat, kann er seine Entscheidung unter „Choice“ auswählen. Nach einem Klick auf „Next Trial“ geht der Ablauf von vorne los. Die Files werden neu gemischt und es müssen die identischen Songs zugeordnet werden. Im Statusfenster lassen sich die Anzahl der Durchgänge und die Erfolgsquote des Users anzeigen. Score 6/6 bedeutet, dass bei sechs Durchgängen sechs Mal die jeweiligen Kopien richtig zugeordnet wurden. Zudem berechnet *foobar2000*, wie hoch die Wahrscheinlichkeit ist, dass dieses Ergebnis durch Raten erzielt wurde. Je höher dieser Wert ist, desto häufiger hat die Testperson geraten. Somit kann davon ausgegangen werden, dass zwischen Song A und B kaum Unterschiede herrschen und somit kaum ein Qualitätsverlust/gewinn festzustellen ist. Eine detailliertere Erklärung der Ratewahrscheinlichkeit findet sich auf Seite - 31 -5.9 oben Binomialverteilung.

7. Softwarebeschreibung: *STEP*

Bei der Software *STEP (Subjective Training and Evaluation Program)* handelt es sich um eine Software, die speziell entwickelt wurde, um Hör- und Vergleichstests durchzuführen. Die Software wurde in der Firma Audio Research Labs mittlerweile auf Version 1.08 entwickelt. Allerdings wurde das Programm seit 2010 nicht weiterentwickelt. *STEP* ist im Gegensatz zu *foobar2000* nicht kostenlos, sondern muss für 299\$ erworben werden. Auf der Homepage gibt es eine Testversion zum Download, mit welcher man mittels Testfiles das Programm und seine Eigenschaften prüfen und testen kann. Die *STEP*-Software unterstützt Hörtests nach den Empfehlungen folgender Dokumente:

⁷⁰ (Klein(Screenshot))

- ITU-R BS.1116 für geringe Qualitätsabweichungen
- ITU-R BS.1534 für grobe Qualitätsabweichungen
- ITU-R BS.1294 für generelle Testmethoden

Die Software bietet verschiedene Bewertungsmöglichkeiten von 5- oder 7-stufigen Skalen an sowie einen einfachen ABX-Test. Zudem bietet *STEP* für alle Testmöglichkeiten verschiedene Trainingsprogramme an. Es können hier sowohl Stereo- als auch Multikanaltests durchgeführt werden. ⁷¹

7.1 Systemvoraussetzungen und Installation

Die Software *STEP* benötigt laut Hersteller ein Windows XP oder 2000 System. Einen Prozessor mit mindestens 1 GHz, 256 MB Arbeitsspeicher und eine 2-Kanal oder 6-Kanal Soundkarte. Die Testversion von *STEP* lässt sich nach einer Registrierung auf der Homepage downloaden und installieren:

<http://audioresearchlabs.com/step/download.php> ⁷²

Bei der Testversion können nur mitgelieferte Beispieltest durchgeführt werden.

7.2 Benutzeroberfläche und Bedienung

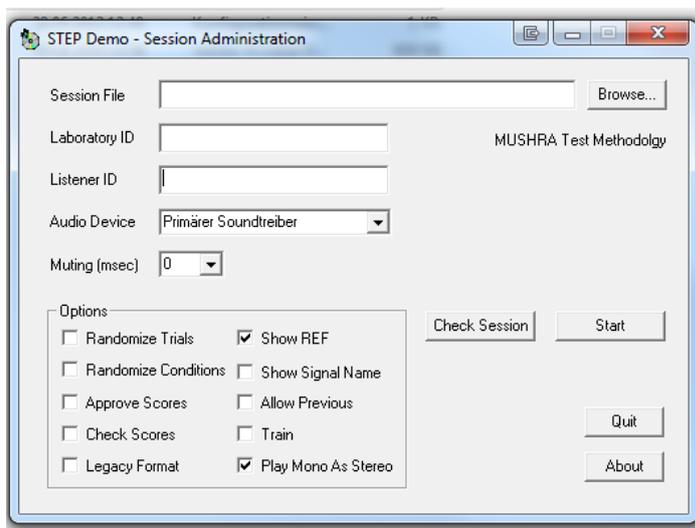


Abbildung N: *STEP*-Oberfläche ⁷³

⁷¹ vgl. (Audio Research Labs, September 2010)

⁷² Stand: 08.06.2013

⁷³ (Klein(Screenshot))

Nach dem Start öffnet sich das Bedienfeld: Hier müssen nun verschiedene Eintragungen erfolgen. Über Browse lässt sich die Testsession laden. Die ID-Eintragungen dienen dem späteren Verwalten und Überblicken der Probanden. Im Bereich Audio-Device muss die Soundkarte ausgewählt werden. Mit der Muting-Funktion lässt sich einstellen, wie lange *STEP* auf stumm schalten soll, wenn die Hörbeispiele gewechselt werden. Bei 0 ms wechselt *STEP* ohne Pause zum nächsten Abhörpunkt. Unter Options lassen sich folgende Einstellungen verändern⁷⁴:

| | |
|----------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Randomize Trials | Zufällige Anordnung der verschiedenen Beispiele von Hörtests. |
| Randomize Conditions | Zufällige Anordnung der verschiedenen Qualitätsstufen (Referenz, Anchorreferenz usw.) eines Tests. |
| Approve Scores | Hier kann beim Fortschreiten zum nächsten Programmausschnitt eine zusätzliche Anfrage angezeigt werden, bei der der Proband bestätigen muss, dass er wirklich zum nächsten Beispiel gelangen möchte. |
| Check Scores | Mit dieser Funktion überprüft <i>STEP</i> , ob bei einem ABC/HR-Test mindestens ein Programmausschnitt mit 5.0 bewertet wurde. |
| Legacy Format | Mit dieser Funktion können alte <i>STEP</i> -Sessions aus früheren Versionen verwendet werden. |
| Show REF | Hier kann der Referenzausschnitt angezeigt bzw. ausgeblendet werden. |
| Show Signal Name | Hier können die Dateinamen der einzelnen Textdateien angezeigt werden. Dies kann in einer Trainings-session sinnvoll sein. |
| Allow previous | Hier kann man zu einem bereits abgeschlossenen Szenario zurückgehen. Dies ist ebenfalls im Trainingsmodus sinnvoll. |
| Train | Hier kann die jeweilige Session im Trainingsmodus gestartet werden |
| Play Mono as Stereo | Hier werden Monodateien auf dem rechten und linken Kanal abgebildet. |

⁷⁴ vgl. (Audio Research Labs, September 2010 S. 9f)

Nicht alle Optionen sind für alle Sessions geeignet:

| Option | BS-1116 | MUSHRA | AB Comparison | ABX Comparison | Training |
|----------------------|---------|--------|------------------|-------------------|----------|
| Randomize Trials | X | X | X | X | |
| Randomize Conditions | X | X | X | X | |
| Approve Scores | X | X | X | X | |
| Check Scores | X | | | | |
| Show REF | X | X | | | X |
| Show Signal Name | | | | | |
| Allow previous | | | | | X |
| Train | | | | | |
| Play Mono as Stereo | X | X | X | X | X |

Abbildung O: Optionenauswahl in verschiedenen Sessions in STEP⁷⁵

7.3 Testerstellung

STEP benötigt sogenannte Sessionfiles, mit welchen die Tests durchgeführt werden. Diese Sessionfiles müssen extern über den Texteditor erstellt werden. Diese Datei gibt dem Programm bekannt, welches die Referenz und welches die Vergleichsmaterialien sind. Sessionfiles müssen in einer bestimmten Struktur geschrieben und später mit der Endung .asi abgespeichert werden. In der Startoberfläche von STEP lässt sich diese Datei dann auf Fehler prüfen und ausführen.⁷⁶

| Line in Session file | Description |
|------------------------------|------------------------|
| session=MUSHRA | Session type |
| # Signal_1 | Displayed signal name |
| originals/sig1.wav Ref | Open reference |
| originals/sig1.wav HR | Hidden reference |
| coder1/sig1.wav Sys1 | System under test Sys1 |
| coder2/sig1.wav Sys2 | System under test Sys2 |
| Coder3/sig1.wav Sys3 | System under test Sys3 |
| anchors/sig1_lp70.wav LP70 | 7.0 kHz lowpass anchor |
| anchors/sig1_lp35.wav LP35 | 3.5 kHz lowpass anchor |
| # Signal_2 | Displayed signal name |
| originals/sig2.wav Ref | Open reference |
| originals/sig2.wav HR | Hidden reference |
| coder1/sig2.wav Sys1 | System under test Sys1 |
| coder2/sig2.wav Sys2 | System under test Sys2 |
| Coder3/sig2.wav Sys3 | System under test Sys3 |
| anchors/sig2_lp70.wav LP70 | 7.0 kHz lowpass anchor |
| anchors/sig2_lp35.wav LP35 | 3.5 kHz lowpass anchor |

Abbildung P: MUSHRA-Sessionfile in STEP⁷⁷

⁷⁵ (Audio Research Labs, September 2010 S. 10)

⁷⁶ Vgl. (Audio Research Labs, September 2010)

⁷⁷ (Klein(Screenshot))

7.3.1 AB-Test in STEP

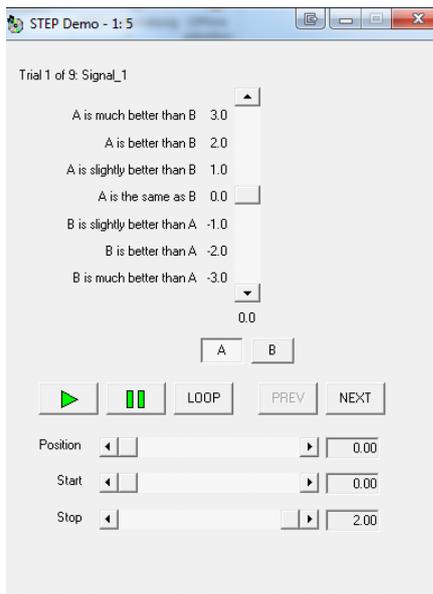


Abbildung Q: AB-Test in STEP ⁷⁸

In *STEP* kann ein AB-Test durchgeführt werden. Hier lassen sich zwei Programmausschnitte miteinander vergleichen und zueinander bewerten. Durch verschiedene Optionen können zusätzliche Buttons eingestellt werden (siehe hierzu 7.2 Benutzeroberfläche und Bedienung auf S. - 37 -). Da es von der ITU keine ausdrücklichen schriftlichen Aussagen zu einem AB-Test gibt, können die Daten lediglich statistisch ausgewertet werden.

7.3.2 ABX-Test in STEP

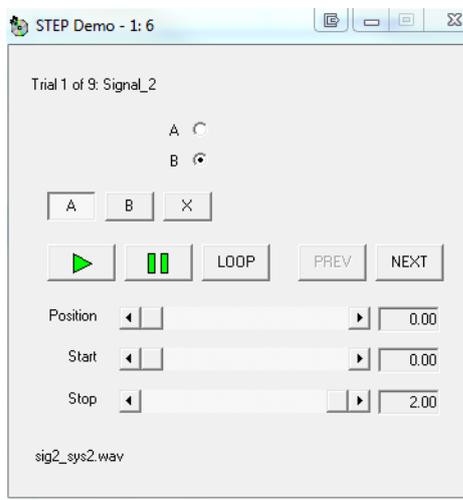


Abbildung R: ABX-Test in STEP ⁷⁹

⁷⁸ (Klein(Screenshot))

Der ABX-Test in *STEP* (siehe 4.2 ABX-Blindhörtest S. - 26 -) bietet alle notwendigen Eigenschaften, um einen ABX-Test durchzuführen. Hier können über die Optionen im Hauptfenster ebenfalls hilfreiche Zusatzfunktionen hinzugefügt werden. Da es von der ITU keine schriftliche Empfehlung zu einem ABX-Test gibt, können die Daten lediglich statistisch ausgewertet werden.

7.3.3 Der ABC/HR-Test (BS.1116-1) in *STEP*

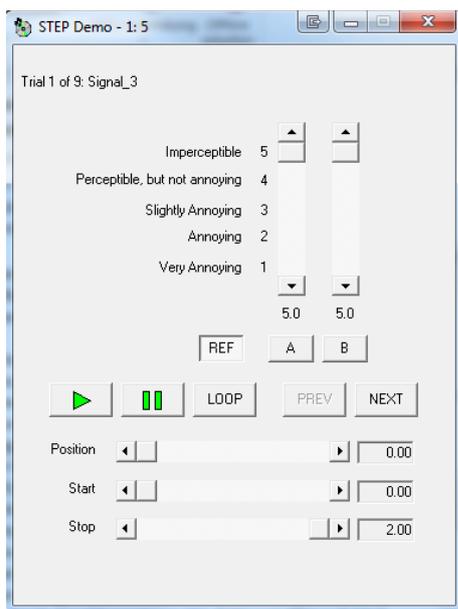


Abbildung S: ABC/HR (BS. 1116-1) in *STEP*⁷⁹

In diesem Programmmodus können die Testhörer einen Test nach den Empfehlungen der ITU-R BS.1116-1 durchführen. Der Proband erhält zwei Programmausschnitte, wobei einer mit der Referenz identisch ist. Diese Ausschnitte müssen in Bezug auf die Referenz bewertet werden. Einer dieser beiden Ausschnitte (A oder B) muss den Skalenwert 5.0 erreichen. Mit dem Klicken auf Next werden die zwei Audiofiles, (jeweils eine Referenz und ein qualitativ schlechteres Audiofile) neu gemischt und der Testhörer muss abermals entscheiden, welches die versteckte Referenz ist. *STEP* arbeitet exakt nach den Vorgaben der ITU. Wie empfohlen, werden hier drei Stimuli angeboten, von denen eine die Kopie der Referenz ist. Die ITU empfiehlt zudem, nach dem Festlegen der Stimuli recht schnell zur nächsten Aufgabe zu gelangen. Die Bewertungsskala entspricht ebenfalls den Vorstellungen der ITU-Skalen. Wählt man

⁷⁹ (Klein(Screenshot))

⁸⁰ (Klein(Screenshot))

in den *STEP*-Optionen die Funktion Train an, lässt sich ein Training, wie es die ITU empfiehlt, durchführen. Der Testhörer kann die Testfiles miteinander vergleichen, zudem kann er in den Optionen einstellen, dass die Beschriftung der Testfiles angezeigt werden soll. Somit lässt sich sehr gut trainieren, auf was im späteren Test geachtet werden muss. Die Bewertungsskala ist, wie von der ITU beschrieben, im Trainingsmodus ausgeblendet.

7.3.4 Der MUSHRA-Test in *STEP*

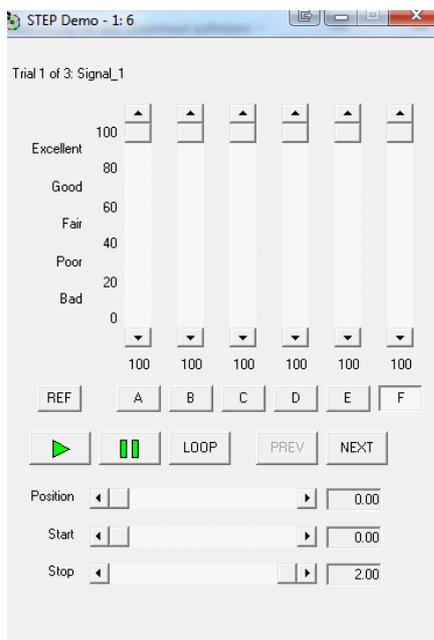


Abbildung T: MUSHRA in *STEP* ⁸¹

Der Hörer hat die Möglichkeit während dem Abspielen der Testdateien zwischen den einzelnen Qualitätsstufen hin- und herzuwechseln und sich so durch schnelles Ändern des abgespielten Audiomaterials ein Bild über die Qualität zu machen, jeweils in Bezug auf die Referenz. *STEP* bietet auch hier Eigenschaften, die nach der Vorstellung der ITU umgesetzt wurden. Es können im Trainingsmodus alle Programmausschnitte, nach Wunsch mit Beschriftung, abgespielt werden und mit der Referenz ohne eine Bewertungsskala verglichen werden. Im Test hat der Hörer die Möglichkeit, während dem Abspielen der Testsequenzen zwischen den einzelnen Stimuli zu wechseln. So kann er sich genau auf die Qualitätsunterschiede konzentrieren.

⁸¹ (Klein(Screenshot))

7.4 Testauswertung

STEP legt für jeden Hörer eine Text- und eine Excel-Datei an. In diesen Dokumenten befinden sich die Auswertungen bzw. die Ergebnisse des Probanden. In der Excel-Datei sind bereits vorgefertigte Formeln eingetragen, um statistische Werte zu berechnen. In einer zweiten Datei sind ebenfalls die Ergebnisse des Tests dokumentiert, somit bietet *STEP* in Verbindung mit Excel eine übersichtliche Methode, um die Ergebnisse zu überblicken und auszuwerten. Neben dem Durchschnitt zeigt Excel auch sofort die Standardabweichung und die Anzahl der Testhörer an. In der Textdatei „Scores“ findet man alle Ergebnisse des jeweiligen Testhörers.

7.5 Standard-konforme Hörtests mit *STEP*

STEP hält sich an alle Vorgaben der ITU, alle Tests können normgerecht durchgeführt und ausgewertet werden. *STEP* bietet neben verschiedenen Optionen (Trainingsphase, Referenzanzeige, usw.) eine komfortable Oberfläche, mit der sich die Tests individuell gestalten und durchführen lassen.

7.6 Bewertung der Software *STEP*

STEP besitzt eine einfache, komfortable und selbsterklärende Oberfläche, mit der sich verschieden Testverfahren (AB, ABX, MUSHRA, ABC/HR) durchführen lassen, welche alle nach den Empfehlungen der ITU arbeiten. In Tests mit der MSUHRA- Methode wäre es hilfreich gewesen die „Check Scores“-Funktion einzubauen, da dem Testhörer bereits vorab bekannt ist, dass sich in den Testfiles eine versteckte Referenz befindet. Somit wäre der Hörer dazu gezwungen, eins der MSUHRA-Files mit 100% zu bewerten, was zu präziseren Ergebnissen führen würde.

Darüber hinaus ist die Erstellung eines Tests mit dem Texteditor etwas umständlich und hätte komfortabler gelöst werden können, beispielsweise mit einer GUI, um die Files auszuwählen.

Die Testauswertung dagegen ist sehr komfortabel, da durch die Kooperation mit Excel sofort statistische Werte entnommen werden können und nicht mit einem weiteren Programm gearbeitet werden muss.

8. Softwarebeschreibung: *MUHSRATest*

Die Software *MUHSRATest* wurde von Roman Kosobrodov entwickelt und programmiert. Das Programm wurde in C++ geschrieben und läuft als selbständige Applikation auf dem PC bzw. MAC-Computer. Die Installationsdatei lässt sich kostenlos im Internet downloaden: <http://mushra.kosobrodov.net/download.html> ⁸². Mit *MUHSRATest* lassen sich ausschließlich MUSHRA-Tests nach Empfehlungen der ITU-R BS.1534 durchführen. Nach dem Download der Installationsdatei lässt sich das Programm auf dem Computer installieren. Es gibt zwei Dateien, die nach der Installation gestartet werden können.

8.1 Einstellungen und Erstellung von Hörtests

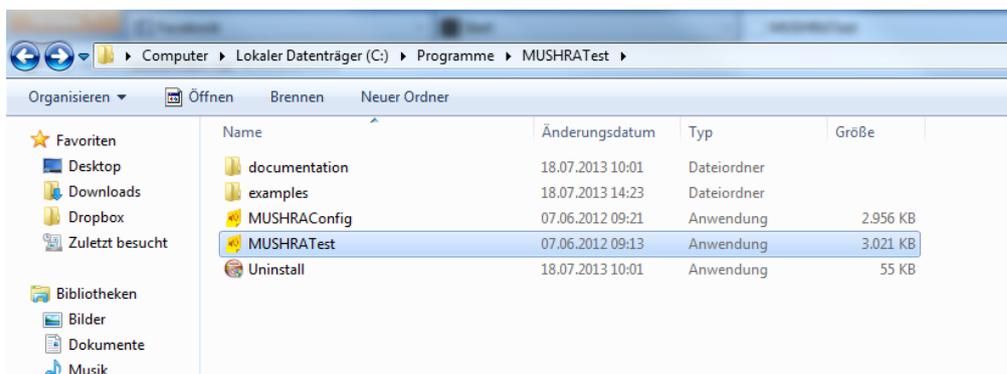


Abbildung U: *MUHSRATest* – Dateien ⁸³

Mit der Datei „MUSHRAConfig“ lässt sich die Software einrichten, einstellen und die Tests können vorbereitet werden. Neben den Einstellungen für die Soundkarte müssen hier die Tests erstellt werden. In der Rubrik „TestConfigurator“ muss für jeden Testhörer eine „Subject“-Einstellung mit dem jeweiligen Pfad, wo sich die Testdateien befinden, erzeugt werden.

⁸² Stand 28.07.2013

⁸³ (Klein(Screenshot))

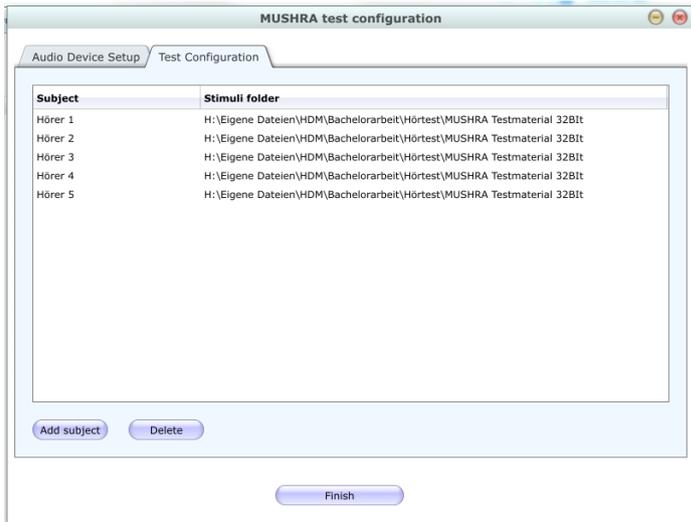


Abbildung V: Tests erstellen in *MUSHRATest* ⁸⁴

Alle Testdateien müssen im WAV-Format in 44,1 kHz und 32 Bit-float vorliegen. Zudem müssen sich alle Testdateien in einem Ordner befinden. Dieser Ordner kann für die verschiedenen Szenarien pro Test weitere Unterordner besitzen.

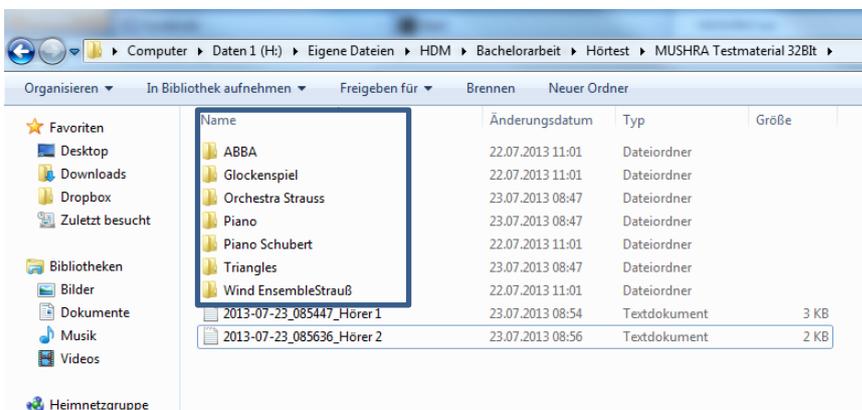


Abbildung W: Ordnerstruktur *MUSHRATest* ⁸⁵

Im aktuellen Fall befinden sich im angegebenen Ordner sieben Unterordner, welche die verschiedene Beispiele beinhalten. In jedem dieser Ordner müssen gleich viele Dateien vorhanden sein. Zudem muss eine Datei das Wort „reference“ enthalten, damit *MUSHRATest* erkennt, welches die Referenzdatei ist. Die Beschriftung aller anderen Stimuli ist egal.

⁸⁴ (Klein(Screenshot))

⁸⁵ (Klein(Screenshot))

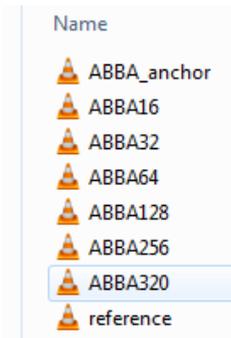


Abbildung X: Ordnerstruktur 2 in *MUSHRATest* ⁸⁶

Diese Struktur ist absolut wichtig und einzuhalten, da *MSUHRTest* ansonsten die Files nicht findet oder nicht abspielen kann.

8.2 Das Training in *MUSHRATest*

Nach dem alle Konfigurationen erstellt wurden, lässt sich die Datei „*MUSHRATest*“ starten und bedienen. Zuerst muss ausgewählt werden, welcher konfigurierte „Subject“ verwendet werden soll. Danach startet das sogenannte Training. Dieses Training gliedert sich in zwei Stufen. In Part A werden dem Hörer alle Stimuli aufgelistet:

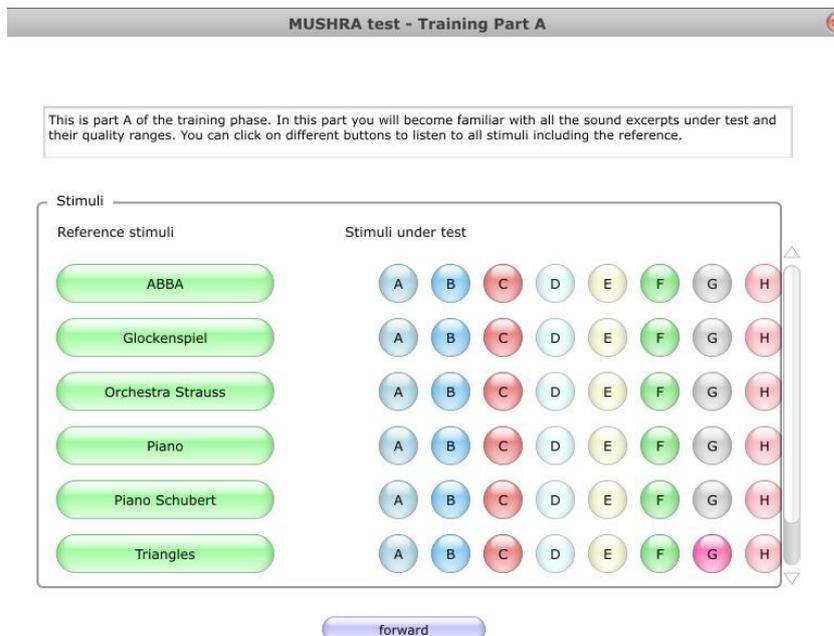


Abbildung Y: Training Part A in *MUSHRATest* ⁸⁷

⁸⁶ (Klein(Screenshot))

⁸⁷ (Klein(Screenshot))

Alle Beispiele sind sortiert und der Hörer kann sich mit dem Material vertraut machen. Bei gleicher Beschriftung in der Ordnerstruktur erkennt *MUSHRATest* beispielsweise, dass Stimulus A beispielsweise immer die MP3-16 Bit-Variante ist und G immer die „Anchor Hidden Referenz“ darstellt. Somit kann sich der Hörer mit dem Bewusstsein, was sich unter den Buttons versteckt, mit den Programmausschnitten vertraut machen. Der Hörer kann sich so viel Zeit nehmen, wie er möchte. Nach einem Klick auf „forward“ gelangt der Proband in den zweiten Teil der Trainingsphase.

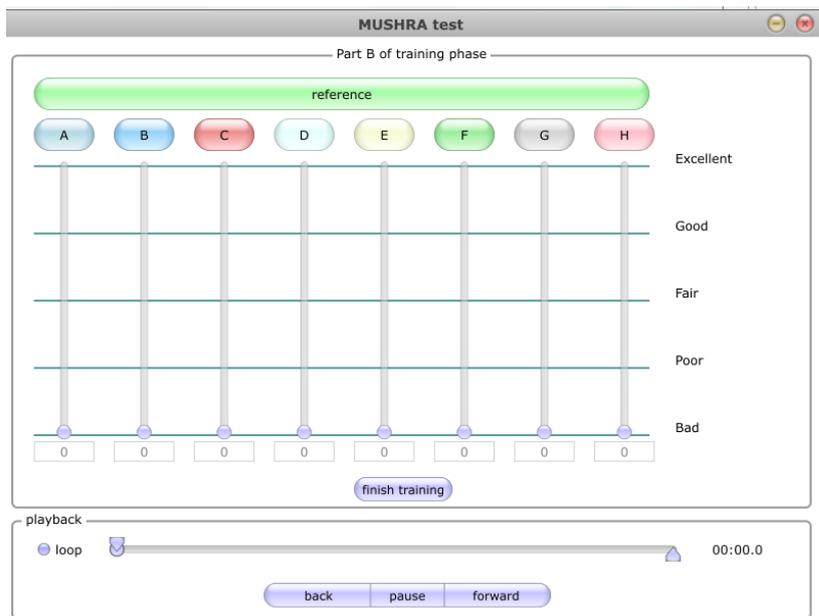


Abbildung Z: *MUSHRATest* Training Part B ⁸⁸

8.3 Der Hörtest in *MUSHRATest*

Im zweiten Teil der Trainingsphase, welcher identisch aufgebaut ist, wie der Test später, soll der Hörer die Software kennenlernen und vom Leiter des Tests auf die Funktionen hingewiesen werden. Der Testhörer hat die Möglichkeit, sich alle acht Stimuli plus Referenz so oft anzuhören, wie er möchte. Danach muss mit den Schiebereglern festgelegt werden, wie die verschiedenen Qualitätsstufen empfunden wurden. Mit „back“, „pause“ und „forward“, kann der jeweilige Programmschnitt gespult bzw. pausiert werden. Mit dem „loop-Button“ wird der ausgewählte Stimulus in einer Dauerschleife gespielt. Zudem kann der Anfang und das Ende jedes Ausschnittes individuell gekürzt werden. Da unter den verwendeten Stimuli eine Referenz versteckt ist, ist der Benutzer verpflichtet einen Stimulus mit „Excellent“ zu

⁸⁸ (Klein(Screenshot))

bewerten, ansonsten leitet MUSHRATest nicht zum nächsten Beispiel weiter und meldet einen „incorrect input“.

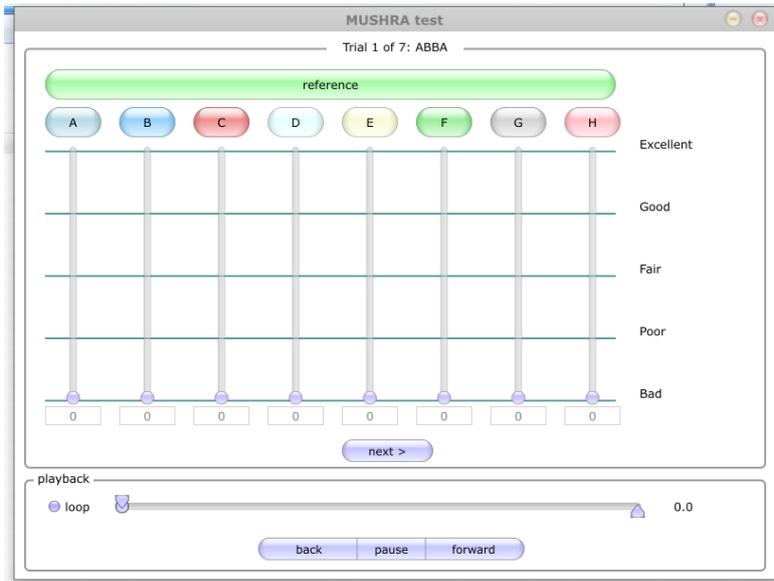


Abbildung AA: Hörtest mit *MUSHRATest* ⁸⁹

Es kann immer nur der Stimulus in seiner Qualitätsstufe bewertet werden, welcher auch aktuell hörbar ist, das bedeutet, wenn der Proband Stimulus A hört, hat er nicht die Möglichkeit Stimulus F zu bewerten. Nachdem der Hörer seine Bewertungen für alle acht Stimuli eingestellt hat, gelangt er mit einem Klick auf Next zum nächsten Hörbeispiel.

8.4 Die Datendokumentation in *MUSHRATest*

Alle Ergebnisse der durchgeführten Hörtests werden im Ordner, in welchem sich auch das Testmaterial befindet, gespeichert. Jeder eingetragene „Subject“ erhält eine eigene Textdatei mit der Dokumentation seiner eingeschätzten Werte.

⁸⁹ (Klein(Screenshot))

```

2013-07-23_085447_Hörer 1 - Editor
Datei Bearbeiten Format Ansicht ?
# date: 23 Jul 2013 8:54
# subjectID: Hörer 1
#
#-----
# Test: ABBA
# stimulus      response
ABBA_anchor.wav 10
ABBA32.wav      22
ABBA256.wav     39
ABBA64.wav      13
ABBA128.wav     88
ABBA320.wav    100
ABBA16.wav      12
reference.wav   73
#-----
# Test: Glockenspiel
# stimulus      response
Glockenspiel_anchor.wav 38
Glockenspiel32.wav      65
Glockenspiel_reference.wav 51
Glockenspiel256.wav    100
Glockenspiel128.wav    31
Glockenspiel16.wav     100
Glockenspiel320.wav    13
Glockenspiel64.wav      3
#-----
# Test: Orchestra Strauss
# stimulus      response
orchestra_strauss_anchor.wav 10
orchestra_strauss_256.wav    100
orchestra_strauss_reference.wav 29
orchestra_strauss_64.wav     100
orchestra_strauss_128.wav    24
orchestra_strauss_320.wav    54
orchestra_strauss_16.wav     56
orchestra_strauss_32.wav    100
#-----
# Test: Piano

```

Abbildung BB: Datendokumentation in *MUSHRATest* ⁹⁰

MUSHRATest speichert den Tag und die Uhrzeit des Hörtests. Zudem finden sich hier für jeden Programmausschnitt die eingestellten Werte des Probanden. Für eine Auswertung der Testergebnisse ist das Übertragen der Daten in Excel oder ein drittes Programm von Nöten.

8.5 Standard-konforme Hörtests mit *MUSHRATest*

Die Software arbeitet sehr exakt nach den Vorgaben der ITU-R BS.1534-1. Die MSUHRA-Methode wird angewendet, um größere Qualitätsunterschiede zu erkennen und einzuschätzen. Die Auswahl und Selektion der Probanden ist dem Testleiter überlassen. Die Anforderungen der ITU-R BS:1534-1 werden in der Software *MUSHRATest* exakt beachtet und verwirklicht. Die Software weist eine kontinuierliche Bewertungsskala auf, die der Empfehlung der ITU-R BS.1534-1 entspricht. Die ITU-R BS.1534-1 gibt vor, dass die einzelnen Programmausschnitte in beliebiger Reihenfolge verglichen werden können. Hier bietet *MUSHRATest* eine hervorragende Möglichkeit, wodurch man während dem Abspielen zwischen den einzelnen Ausschnitten wechseln kann. So ist ein direkter Vergleich möglich, ohne dass der Programmausschnitt jedes Mal von Beginn an startet. Die Trainingsphase kann direkt in der Software absolviert werden und gliedert sich wie von der ITU empfohlen

⁹⁰ (Klein(Screenshot))

ebenfalls in zwei Teile, in denen jeweils verschiedene Trainingsabschnitte erläutert werden (siehe 8.2 Das Training auf S. - 46 -).

8.6 Bewertung der Software *MUSHRATest*

Die Software ist schlicht und übersichtlich aufgebaut. Es bedarf einer kurzen Einarbeitung, um die Tests konfigurieren zu können. Die eigentliche Anwendung der Tests ist sehr simpel und selbsterklärend. Die Software macht einen optisch guten und übersichtlichen Eindruck, es werden nur die nötigsten Buttons angezeigt, keine komplizierten Vorgänge oder Anwendungsmöglichkeiten. Zudem besitzt die Software alle notwendigen Eigenschaften um einen MUSHRA-Test nach der ITU-R BS:1534 durchzuführen. Ein weiterer Vorteil ist, dass die Trainingsphase direkt in der Software durchgeführt werden kann und es keiner weiteren von außen bedarf. Allerdings wäre es hilfreich, wenn die genaue Dateienbeschriftung sichtbar wäre, damit im Training genau ersichtlich ist, welches File aktuell gehört wird. Zudem startet *MUSHRATest* im Training den Programmausschnitt jedes Mal von vorne. Im Hinblick auf den späteren Test, wäre das Training effektiver, wenn man auch hier während dem Abspielen eines Ausschnitts die einzelnen Stufen direkt vergleichen könnte.

Ferner wäre ein kurzer schriftlicher Teil hilfreich, damit man den Hörer mit einer Beschreibung auf den Test vorbereiten kann.

Das Einfügen der Testdateien könnte man weniger fehleranfällig gestalten, indem über eine GUI die Dateien direkt dem jeweiligen Stimulus zugeordnet werden können. Momentan kann es, hinsichtlich der genauen Ordner- und Namensstruktur der Software, zu Problemen kommen.

Die Speicherung der Daten für die statische Auswertung könnte ebenfalls etwas komfortabler dargestellt werden. Alle Daten werden für jede Testperson als Textdatei angelegt, die etwas unübersichtlich ist und für eine statistische Auswertung in ein weiteres Programm manuell übertragen werden muss. Hier wäre von Vorteil, wenn *MUSHRATest* die Daten direkt als Excel-Tabelle speichern könnte oder ein weiteres Tool zur Verfügung stellt, um die Daten direkt auswerten zu können.

9. Weitere Softwares zur Durchführung von Hörtests

Im Folgenden werden noch weitere softwaregestützte Hörversuche vorgestellt und kurz bewertet. Neben den oben genannten (*foobar2000*, *MUSHRA*Test und *STEP*) gibt es beispielsweise noch *Amorak*, welches sich ausschließlich für Linux-Systeme eignet. Für die Durchführung von MUSHRA-Tests kann auch die Software *Scale*, die im Juli 2013 entwickelt wurde, verwendet werden. Für ABC/HR (ITU-R BS.1666-1) eignen sich noch viele kleine Programme, die teilweise für Java- oder MatLap-Konsolen geschrieben wurden.

9.1 Das WinABX-Tool

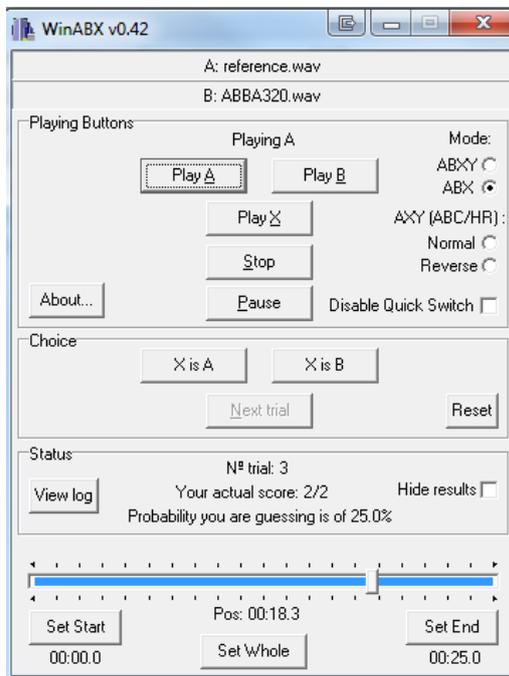


Abbildung CC: ABX-Test mit *WinABX* ⁹¹

WinABX ist eine Windowsapplikation, welche 2003 entwickelt wurde. Sie erlaubt es, einen ABX-Test durchzuführen. ⁹² Die Benutzeroberfläche ist sehr schlicht und einfach gehalten. Nach dem Öffnen fordert das Programm sofort dazu auf, zwei Testdateien einzufügen. Das Einfügen und Starten der Testfiles ist sehr einfach, allerdings weist die Software einige Schwachpunkte auf.

⁹¹ (Klein(Screenshot))

⁹² vgl. (<http://www.kikeg.arrakis.es>)

Im oberen Bereich der Applikation ist dauerhaft sichtbar, um welche Dateien es sich handelt. Dies kann bereits zu einer ersten Beeinflussung führen; besser wäre hier eine absolut blinde Versuchsreihe, bei der die Probanden nicht wissen, wo sich welches Dateiformat versteckt.

Des Weiteren stockt die Software kurz beim Umschalten der einzelnen Stimuli, dies kann zu Verwirrungen führen und ein genaues Ergebnis etwas verfälschen. Die Software zeigt ebenfalls wie *foobar2000* an, wie hoch die Wahrscheinlichkeit ist, dass das Ergebnis geraten wurde. Es ist noch positiv anzumerken, dass *WinABX* eine Textdatei mit den Ergebnissen des Probanden anlegt. Somit können die Daten für weitere statistische Auswertungen genutzt werden.

9.2 Das ABC Hidden Reference-Audio Comparison Tool (ABC/HR)

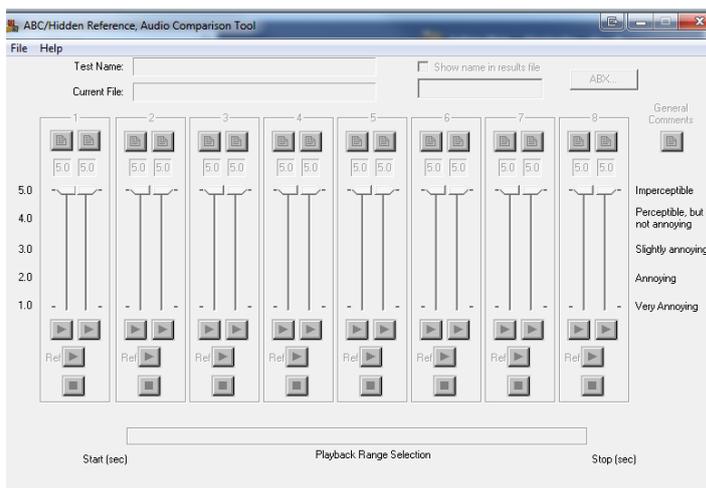


Abbildung DD: ABC/HR-Test mit ABC Hidden Reference-Audio Comparison Tool ⁹³

Die Windows Applikation *ABC Hidden Reference-Audio Comparison Tool* lässt sehr einfach einen ABC/HR Test erstellen. Im Setup-Fenster lassen sich acht Tests direkt durchführen.⁹⁴ Im Setup Fenster wird eine Referenzdatei geladen, für die anderen acht Tests kann eine Datei gewählt werden. Im Test werden dann die Referenz und die Testdatei gemischt und der Hörer muss für Test 1-8 die beiden Programmausschnitte bewerten, wobei ein Ausschnitt mit 5.0 bewertet werden muss. Das Ergebnis lässt sich dann über „File – Save Result“ speichern und ggf. auswerten. Leider bietet dieses Tool keine Trainingsphase, außerdem sollte es laut ITU nicht möglich sein, im Test

⁹³ (Klein(Screenshot))

⁹⁴ vgl. (Softpedia)

nochmal zu einem bereits bewerteten Programmausschnitt zurückzukehren. In diesem Tool hat der User aber die Möglichkeit, selbst wenn er den letzten Vergleich beurteilt, nochmals zum ersten zurückzuspringen. Leider bietet das *ABC Hidden Reference-Audio Comparison Tool* auch keine Auswertungsfunktion an. Neben dem ABC/HR Test lässt sich noch einen ABX-Test durchführen, welcher dem Testhörer bereits nach Beendigung von acht Durchgängen anzeigt, wie oft das Ergebnis richtig festgestellt wurde.

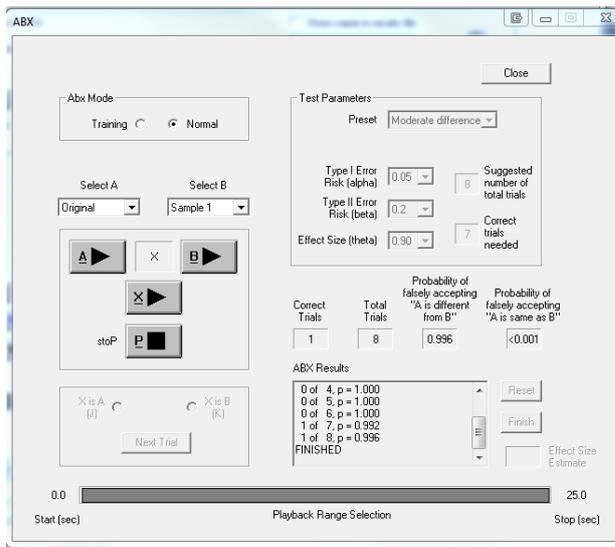


Abbildung EE: ABX-Test mit dem *ABC Hidden Reference Audio Comparison Tool* ⁹⁵

9.3 Die Software *scale* – Kurzbeschreibung zum MUSHRA-Test

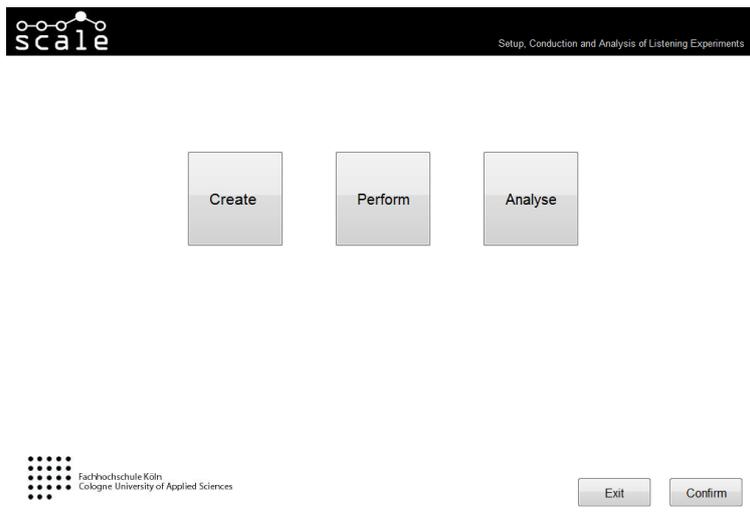


Abbildung FF: *scale* Startoberfläche ⁹⁶

⁹⁵ (Klein(Screenshot))

⁹⁶ (Klein(Screenshot))

Scale ist ein sehr umfangreiches Programm, mit welchem sich verschiedene Tests kreieren, durchführen und auswerten lassen. *Scale* wurde von der Universität für angewandte Wissenschaft in Köln entwickelt und am 15.07.2013 veröffentlicht. Das Programm wurde mit „MatLap“ geschrieben und lässt sich direkt auf der Homepage kostenlos downloaden und mittels eines „MCR - Matlab Compiler Runtime“ ausführen.⁹⁷

9.3.1 Die Testerstellung in *scale*

Mit *scale* lassen sich neben MUSHRA-Tests auch ABX, 2AFC,3AFC,4AFC oder YES/NO-Tests durchführen. Betrachtet man beispielsweise die Erstellung des MUSHRA-Tests, dann lassen sich eine Vielzahl von Einstellungen vornehmen: Die Anzahl der verschiedenen Vergleichsprogramme ist auf maximal 15 festgelegt, wie es in der ITU-R BS.1534-1 empfohlen wird. Zudem muss für jeden Test gewählt werden, wie viele Beispiele und Hörbeispiele es pro Szenario gibt. Über „add“ lassen sich Hörbeispiele einfügen und den jeweiligen Szenarien zuordnen. Zu jedem Szenario muss noch ein Arbeitsauftrag bzw. eine detaillierte Beschreibung des Testversuches geschrieben werden, welche dem Hörer angezeigt wird, sobald er den Test startet.

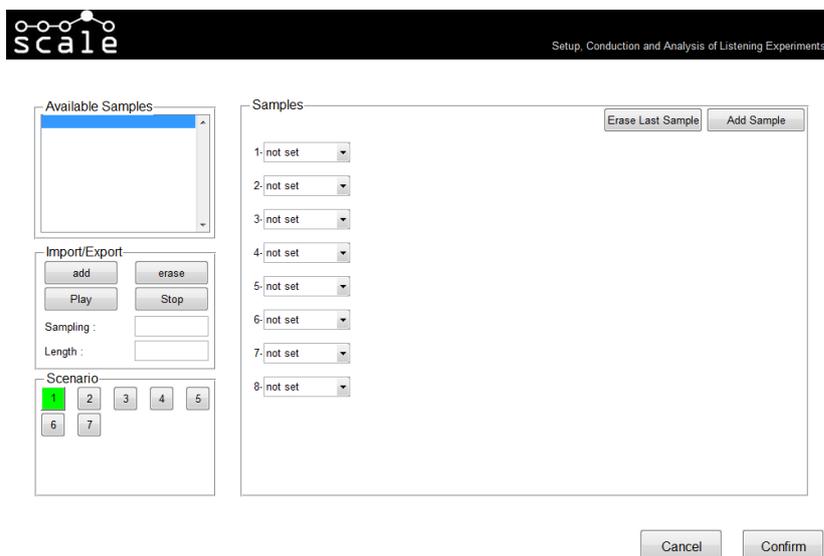


Abbildung GG: Testerstellung in *scale*⁹⁸

⁹⁷ (Pörschmann)

⁹⁸ (Klein(Screenshot))

9.3.2 Die Testdurchführung in *scale*

Der Test lässt sich in der Startoberfläche unter „Perform“ auswählen und starten. Für die Testdurchführung wird jeder Hörer im Programm mit seinem Namen und seinem Geburtstag registriert, somit fällt nachher die Auswertung leichter.

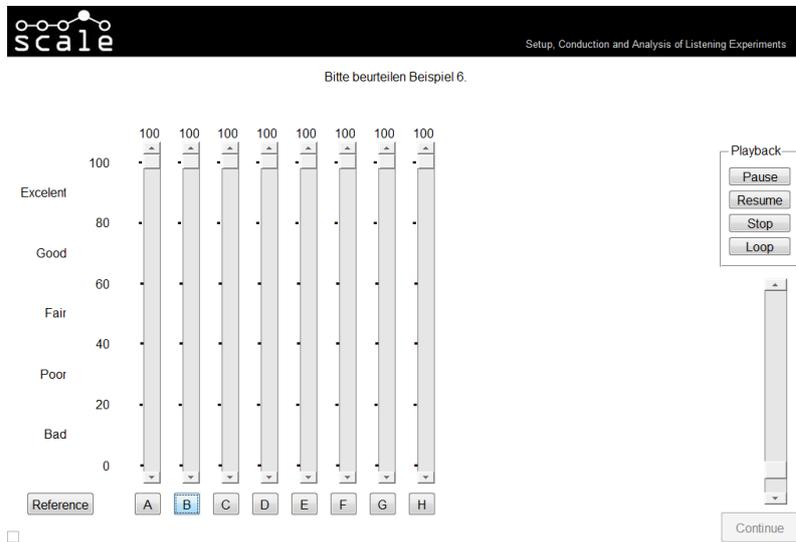


Abbildung HH: Testdurchführung in *scale* ⁹⁹

Die Testdateien können wie gewohnt abgehört und beurteilt werden. Der jeweilige Programmausschnitt wird, sobald ein anderer Stimulus gewählt wurde, von vorne gestartet.

9.3.3 Die Testanalyse in *scale*

Die durchgeführten Tests lassen sich direkt in *scale* auswerten und beurteilen. Hierzu beinhaltet *scale* eine spezielle Analyseabteilung (siehe Abbildung FF: *scale* Startoberfläche auf S. - 53 -).

⁹⁹ (Klein(Screenshot))

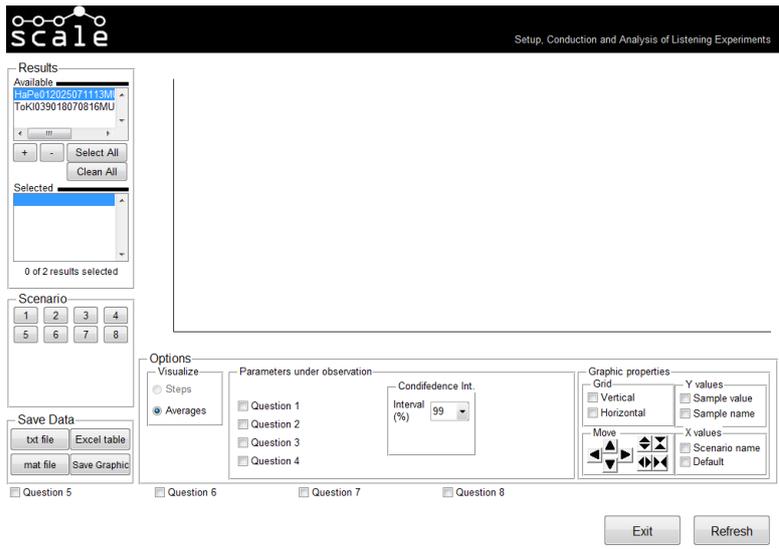


Abbildung II: Testanalyse in *scale* ¹⁰⁰

Mit der Analysesoftware von *scale* lassen sich die gewünschten Testhörer und Szenarien anwählen, zudem lassen sich alle Ergebnisse in Grafiken anzeigen und ggf. speichern. Die Daten können auch direkt an „MatLap“ übermittelt werden oder in Excel weiterverarbeitet werden.

9.3.4 Bewertung der Software *scale*

Scale arbeitet nach den Richtlinien der ITU. Die verschiedenen Szenarien und Programmausschnitte werden für jeden Test neu gemischt, damit keine Absprache zwischen den Probanden zustande kommen kann. Zudem bietet *scale* die Möglichkeit, einen Einführungstext zu schreiben und den Testhörer auf verschiedene Dinge aufmerksam zu machen.

Folgende Kritikpunkte sind an *scale* im MUSHRA-Test zu üben:

Es ist kein Springen während dem Hören der einzelnen Programmausschnitte möglich, ohne dass diese jedes Mal von vorne starten.

Das Schreiben der Einführung ist verpflichtend und muss für jedes Szenario erstellt werden, dies ist ein unnötiger Aufwand, eine schriftliche Einführung würde ausreichen.

Die Erstellung von verschiedenen Tests erweist sich als kompliziert und nicht unbedingt selbsterklärend, dies hätte einfacher gelöst werden können.

¹⁰⁰ (Klein(Screenshot))

Leider bietet *scale* gerade beim MUSHRA-Test, keine spezielle Trainingsphase an, mit der sich die Hörer in die Thematik und in die Software einarbeiten können.

Zudem wird bei der Durchführung des MUSHRA-Tests keine 100%-Bewertung bei mindestens einem Programmausschnitt gefordert, dies wäre hilfreich, da in jedem Szenario immer eine Kopie der Referenz vorhanden ist, die eigentlich mit 100% bewertet werden sollte.

10. Durchführung und Protokollierung eines MUSHRA-Hörversuches

Die Durchführung eines Hörtests fordert einige Vorbereitungen und eine Einarbeitung in eine Software. Der Hörtest wird nach den Richtlinien der ITU-R BS.1534-1 durchgeführt. Hierbei handelt es sich um die sogenannte MUSHRA-Methode, um mittlere bis grobe Qualitätsunterscheide aufzuzeigen. Der Test soll zeigen, ab wann ein Unterschied von CD-Qualität zu einer MP3-Datei für den Laien hörbar ist.

10.1 Die Grundlagen der MP3-Datei

Audiosignale sind Schallwellen, welche nicht aus einer einzigen Sinuswelle, sondern aus einer Überlagerung von unendlich vielen Sinuskurven bestehen. Die Schallwelle wird bestimmt durch die Amplitude A , die für die Lautstärke verantwortlich ist und die Frequenz f , die für die Tonhöhe verantwortlich ist. Jede einzelne Sinuskurve eines Signals besitzt diese Eigenschaften. Eine Schallwelle entsteht also aus der Addition, der mathematischen Überlagerung von unendlich vielen Sinuskurven.



Abbildung JJ: Analoges Eingangssignal ¹⁰¹

¹⁰¹vgl. (Kappes)

Für die digitale Speicherung eines Audiosignals ist es nötig, das Signal in kontinuierliche und diskrete Werte zu unterteilen. Hierzu misst man in bestimmten Abständen den Signalwert. Diesen Prozess bezeichnet man als Abtasten. Die Abtastwerte werden auch Samples genannt. Als Standard hat sich nach dem Abtasttheorem von Shannon eine Abtastung von 44100 Mal pro Sekunde etabliert. Dies ist auch die Qualität einer Audio CD (44,1 kHz mit 16 Bit).

Für die Audiokompression und das dadurch entstehende MP3-Format, macht sich MP3 das Wissen der Psychoakustik zunutze.

Unser Gehör besitzt eine Art Filterfunktion, bei einer Überlagerung von Geräuschen dominiert das lauteste Signal. Leisere Signale werden teilweise sogar überhaupt nicht gehört. Der zentrale Effekt, der zur Irrelevanz von bestimmten Daten in Audiosignalen führt, ist die Maskierung bzw. die Verdeckung. Diese bildet beim MPEG-Algorithmus das Fundament der Kompression. Der MPEG-Algorithmus macht sich den Verdeckungseffekt (siehe 2 Psychoakustik – der Verdeckungseffekt auf S. - 7 -) zunutze.

Um nun irrelevante Daten zu erkennen, muss man die Maskierungskurve für das Audiosignal ermitteln. Anhand dieser Kurve ist ersichtlich, welche Daten wirklich hörbar sind.¹⁰²

Beispiel: Eine Audiodatei hat einen Ton mit 1 kHz, ein weiterer Ton liegt bei 1,1 kHz und ist um zirka 18 dB leiser. Der Ton bei 1,1 kHz wird vollkommen vom ersten Ton verdeckt, da er unterhalb der Mithörschwelle liegt. Folglich wären auch andere schwächere Töne in unmittelbarer Nähe des ersten Tones maskiert. Kommt ein dritter Ton mit 2 kHz hinzu und ist dieser ebenfalls 18 dB leiser als der erste, wäre er jedoch hörbar, da er in einem anderen Frequenzband liegt. Für MPEG-Audio-Encoder bedeutet dies: Töne, die überdeckt sind, können weggelassen werden, da sie nicht hörbar sind.

Der eigentliche MPEG-1 Layer 3 Algorithmus ist in der Praxis viel komplexer und beinhaltet noch viele weitere Schritte, die an dieser Stelle nicht aufgeführt werden.

10.2 Auswahl der Audiobeispiele

Für einen MUSHRA-Test spielt die Auswahl der Hörbeispiele eine bedeutende Rolle, nur so können wirklich aussagekräftige Ergebnisse entstehen. Für den Test werden

¹⁰² Vgl. (Kappes)

Audiomaterialien verwendet, welche die EBU auf einer CD zusammengestellt hat. Es handelt sich hier um die spezielle SQAM-CD für Hörtests (siehe 3.7 Audiomaterial für Hörtests – SQAM-CD auf S. - 21 -). Hier werden speziell für den Anwendungsbereich der Bitratenreduktion verschiedene Audiotestfiles zur Verfügung gestellt. Das Testmaterial sollte laut ITU-R BS.1534 bei maximal 20-25 Sekunden liegen. Zudem sollten keine Lautstärkeunterschiede auftreten, da dies Einfluss auf das Ergebnis nehmen könnte. Für den durchzuführenden MUSHRA-Hörtest wurden folgende Audiomaterialien¹⁰³ aus der SQAM-CD ausgewählt:

| Track | Inhalt | Bemerkung |
|--------------|------------------------------------------------|---------------------------------------------------------------------------------------------------------|
| 32 | Triangles (Eznelnstrument) | Stereoaufnahme Triangelschläge |
| 35 | Glockenspiel (Eznelnstrument) | Stereoaufnahme Glockenspielschläge |
| 39 | Grand piano (Soloinstrument) | Stereoaufnahme eines Pianos |
| 60 | Piano – Schubert (Soloinstrument) | Stereoaufnahme Soloklavier von Schubert |
| 62 | Orchestra – Strauss (Orchestermusik) | Seij I Ozawa – Strauss: Also sprach Zarahuatra CD Philips No. 400 072-2, Track 1, 0:32 ... 0:58 |
| 66 | Wind Ensemble – Stravinsky (Orchestermusik) | Charles Dutoit – Stravinsky: Le Scare du Printemps CD Decca No. 414 202-2, Track 3, 5:20 ... 0:32 |
| 69 | ABBA (Popmusik) | ABBA – The Visitors (Head Over Heels) CD Polydor No. 800 011-2, Track 2, 0:10 ... 0:34 |

Für jedes Hörbeispiel stehen insgesamt acht verschiedene Codierungen zum Vergleich. Jedes Stück wurde in folgenden Bitraten als MP3 gespeichert:

16 kBit/s; 32 kBit/s; 64 kBit/s; 128 kBit/s; 256 kBit/s; 320 kBit/s

Zusätzlich gibt es noch die Originalreferenz im WAV-Format und die sogenannte Anchor-Referenz, welche ebenfalls als Vergleichsmaterial zu Verfügung steht. In der

¹⁰³ (3253, September 2008 S. 11f)

Anchor-Referenz wurde die Datei mit einem High-Pass bei 3,5 kHz bearbeitet. Die MP3-Dateien wurden in einem zweiten Schritt in WAV-Dateien umgewandelt. Alle Dateien wurden in folgendes Endformat konvertiert: WAV Stereo 32Bit float, 44,1 kHz (siehe 10.3.1 Cubase S. - 60 -). Siehe hierzu auch Anhang 13.4 auf S. XXII-.

Das Audiomaterial wird den Testhörern in unterschiedlicher Weise präsentiert. Die Probanden hören das Material in unterschiedlicher Reihenfolge, zudem werden die Qualitätsstufen bei jeder Testperson in anderer Abfolge abgespielt. Somit lässt sich eine Absprache bzw. ein Austausch zwischen den Testpersonen verhindern. Wie bereits in der ITU-R BS.1116-1 dargestellt, könnten bei Nichtmischung der Hörbeispiele beispielsweise folgende Gedanken der Testhörer auftreten: „*Das dritte Kodierverfahren klingt immer am besten.*“ Somit könnten sich Tendenzen herauskristallisieren, welche die Ergebnisse der Hörtests verfälschen.

10.3 Verwendete Softwares für den Hörtest

10.3.1 Cubase

Für die Umwandlung der Testdateien wurde die Software *Cubase* von Steinberg verwendet. Die WAV-Dateien der SQAM-CD wurden in *Cubase* geladen und auf die gewünschte Länge geschnitten und gekürzt. Danach wurde der jeweilige Ausschnitt als Referenzobjekt in WAV gespeichert und als Anchor-Referenz mit einem Highpass bei 3,5 kHz exportiert. Zudem wurde jedes Testsignal im MP3-Format mit 16,32,64, 128, 256 und 320 kBit/s gespeichert. Die Software *MUSHRATest* kann nur WAV-Dateien im 32Bit float-Format verarbeiten. Daher wurden alle Files in ein gesondertes Projekt geladen und als WAV mit 44,1 kHz und 32 Bit float gespeichert.

10.3.2 MUSHRATest

Es wurde die Software „MUSHRATest“ von Roman Kosobrodov verwendet. Es handelt sich hier um eine Freeware-Software. Eine genaue Beschreibung, Durchführung und Beurteilung der Software ist unter Absatz 8 Softwarebeschreibung: MUHSRATest auf Seite - 44 - nachzulesen.

10.4 Testpersonen

Als Testpersonen ist es nach ITU-R BS.1534-1 nicht unbedingt nötig mit Experten zu arbeiten, allerdings sollten die Hörer eine gewisse Erfahrung im kritischen Hören von Audiomaterial mitbringen, da diese Hörer ein schnelleres und glaubwürdigeres Ergebnis erzielen. Für den Test konnten Hörer aus verschiedenen Sparten gewonnen werden. Neben Studenten im Bereich der Audiovisuellen Medien absolvierten verschiedene Musiker und DJs den Test, welche ebenfalls eine gewisse Erfahrung bzw. ein geschultes Gehör mitbringen. Alle Probanden besitzen ein intaktes Gehör, dies wurde für den Hörtest vorausgesetzt. Es wurden insgesamt 21 Probanden getestet welche im Alter zwischen 20 und 40 Jahren waren.

10.5 Die Abhörsituation

Es gab verschiedene Ideen wie der Test durchgeführt werden sollte. Neben der Möglichkeit eines nach ITU-Richtlinien geeigneten Raumes mit zwei professionellen Abhörlautsprechern, bestand die Option, den Hörtest mit einem professionellen Kopfhörer durchzuführen. Da es nicht möglich war, alle Probanden an einen Ort zu bringen, wurde der Test mit einem Laptop von HP mit einer externen Firewire-Soundkarte von Focusrite und einem Kopfhörer AKG K 271 MK II durchgeführt. Alle Testhörer hatten denselben Kopfhörer, somit ergab sich die Möglichkeit, den Test ohne eine räumliche Bindung an verschiedenen Standorten durchzuführen.

Laptopeigenschaften HP 620¹⁰⁴

| | |
|------------------------|------------------------------------------|
| Prozessor | Core 2 Duo T6670 / 2.2 GHz – |
| Betriebssystem | Windows 7 Pro 64-Bit |
| Arbeitsspeicher | 2 GB RAM |
| Festplatte | SSD VERTEX 3 60 GB |
| Laufwerk | DVD SuperMulti DL – |
| Display | 39.6 cm (15.6") breit 1366 x 768 Scale |
| Soundkarte | Interne High-Definition Audiokarte |

¹⁰⁴ (HP)

Kopfhörereigenschaften AKG K271 MkII¹⁰⁵

- geschlossen
- ohrmschließend
- dynamisch
- Nennimpedanz: 55 Ohm
- Nennbelastbarkeit: 200 mW
- Übertragungsbereich: 16-28.000 Hz
- Kennschalldruck: 104dB/V
- einseitiger Kabelanschluss
- Kopfband-Abschaltautomatik
- Kabellänge gerades Kabel 3 m mit 3,5 mm Klinenstecker
- Gewicht ohne Kabel 240 g

Focusrite Saffire Pro 24 DSP¹⁰⁶

- mit Virtual Reference Monitoring Technologie (VRM - simuliert verschiedene Monitorszenarios) und thomann Echtzeiteffekte Equalizer
- Kompressor und 'Comfort' Reverb
- 24-Bit/ 96 kHz
- insgesamt 16 Ein- und 8 Ausgänge
- neuer Firewire-Chipsatz/DSP-Hardware
- 2 Mic-Preamps
- 4 analoge Inputs (2 Mic/Line/Inst Combo XLR
- 2 Line Klinke)
- 1x S/PDIF I/O coaxial
- 1x ADAT/optical S/PDIF Eingang
- 2x Kopfhörer-Ausgang
- MIDI In/Midi Out
- 1x Firewire-Anschluss
- zuschaltbare Phantomspeisung 48V
- LED-Anzeige
- Saffire Mix Control Zero-latency DSP Mixer/Router
- Abmaße: 215 x 45 x 220 mm
- Gewicht: 1.6kg

10.6 Einführungs- und Trainingsphase

Vor der eigentlichen Testdurchführung hatten die Probanden, wie in ITU-R BS.1534-1 beschrieben, die Möglichkeit die Software kennenzulernen. Den Testhörern wurde die Software in Ruhe erklärt und gezeigt. Zudem wurden die Testbeispiele in einem

¹⁰⁵ (Thomann)

¹⁰⁶ (Thomann)

Trainingsmodus aufgezeigt. Somit konnte sich jeder Hörer einen Eindruck davon machen, worauf im Test geachtet werden muss. Somit gliedert sich das Training in zwei Parts.

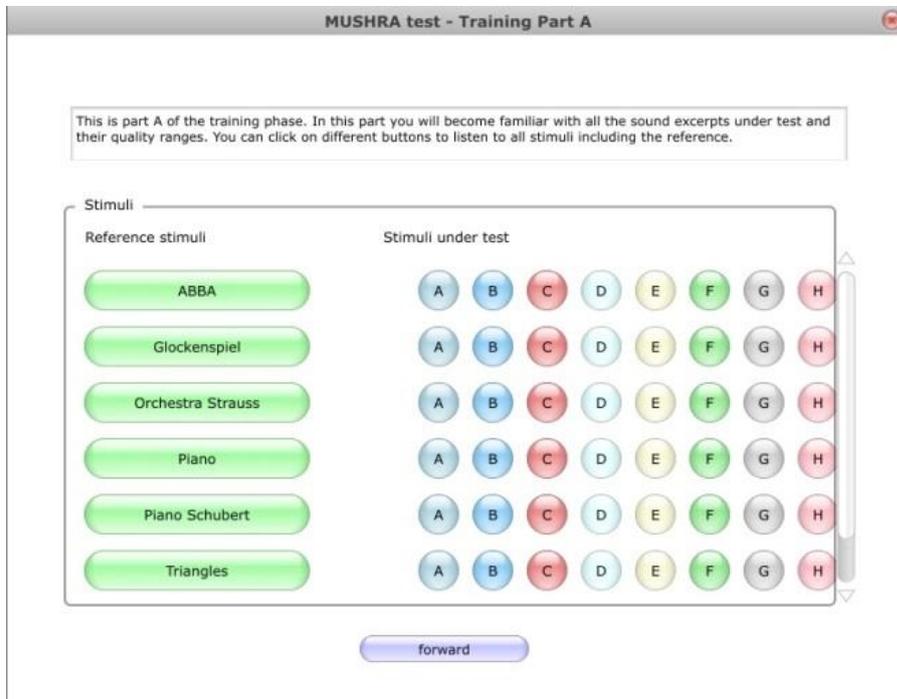


Abbildung KK: *MUSHRA*Test Training Part A ¹⁰⁷

In Part A hatten die Probanden die Möglichkeit sich alle Beispiele anzuhören, sowohl die Referenz als auch die veränderten Stimuli. Neben der Referenz und den verschiedenen MP3-Dateien stand auch die Hidden Reference Anchor zur Verfügung. Deren Bedeutung wurde den Testpersonen erklärt. In Part B erschien dasselbe Fenster, mit welchem später auch der Test durchgeführt wurde. Hier sollten die Testhörer sich mit der Softwarebedienung vertraut machen. Zudem wurde den Probanden erklärt, was zu tun ist und dass es die Aufgabe sei, die Qualitätsstufen der 8 Audiodateien in Bezug auf die Referenz zu bewerten, wobei ein Hörbeispiel identisch mit der Referenz festgestellt werden müsse, da eine versteckte Kopie der Referenz ebenfalls unter dem Vergleichsmaterial sei. Nach der Einführung, Erklärung und Durchführung des Trainings bzw. der Kennenlernphase wurden die Testpersonen alleine gelassen und konnten in Ruhe den Test durchführen.

¹⁰⁷ (Klein(Screenshot))

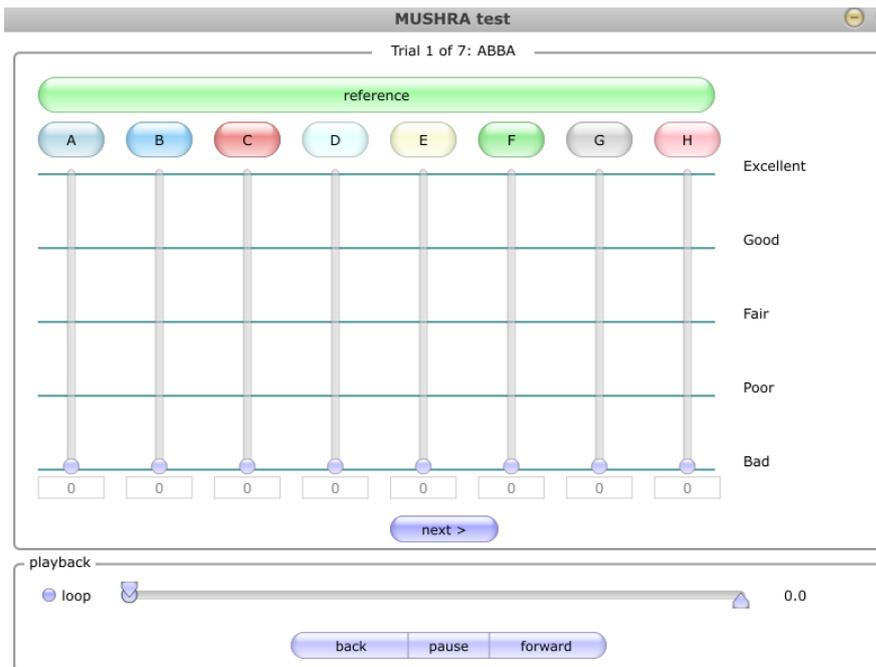


Abbildung LL: MUSHRA-Test-Hörtest ¹⁰⁸

10.7 Die Durchführung

Der Hörtest fand im Zeitraum vom 22.07.2013 bis 10.08.2013 statt. In dieser Zeit wurden 21 Hörer eingeladen, um den Test zu absolvieren. Die Absolventen hatten keinen Zeitdruck und konnten sich für den Test so lange Zeit lassen wie sie mochten. Die Anzahl der Beobachter und Auswerter belief sich auf eine Person. Die getesteten Programmausschnitte sind im Anhang 13.4 auf S. XXII- zu finden.

11. Auswertung des durchgeführten MUSHRA-Tests

Der letzte Aspekt des Hörtests ist die Beurteilung der Ergebnisse. Es geht um die statistische Auswertung der gesammelten Daten. Die ITU-R BS.1116-1 und ITU-R BS.1534-1 schreiben keine spezielle statistische Methode vor.

Für die Auswertung der vorliegenden Ergebnisse wird die Software Excel mit dem Add-on WinStat der Firma Microsoft verwendet. Hier werden alle Testergebnisse eingefügt und ausgewertet.

¹⁰⁸ (Klein(Screenshot))

11.1 Postscreening der Testhörer

Wie in der ITU empfohlen, wurde die Möglichkeit genutzt, Hörer nach Durchführung eines Hörtests aus der Bewertung auszuschließen. In diesem Fall wurden vor allem die Bewertungen der „reference“ und der „anchor-reference“ geprüft und bereinigt.

Insgesamt haben 21 Hörer am Test teilgenommen, die 7 Hörbeispiele mit je 8 Qualitätstufen evaluieren mussten. Sobald ein Hörer öfter als 4 Mal die Referenzen mit weniger als 70/100 bzw. mehr als 20/100 Punkte bewertete, wurde dieser vom Test ausgeschlossen, da davon ausgegangen werden kann, dass entweder ein Softwarefehler oder eine Fehleinschätzung durch zu kritische bzw. unkritische Bewertungen vorliegen.

Insgesamt wurde ein Testhörer aus der Versuchsreihe genommen, welcher mehr als viermal die Referenzen falsch erkannt hatte. Zum Schluss blieben 20 Testhörer mit je 7 Hörbeispielen à 8 Qualitätstufen übrig. Es wurden 1120 Zahlen ausgewertet.

11.2 Erklärung der Box-Plot-Übersicht

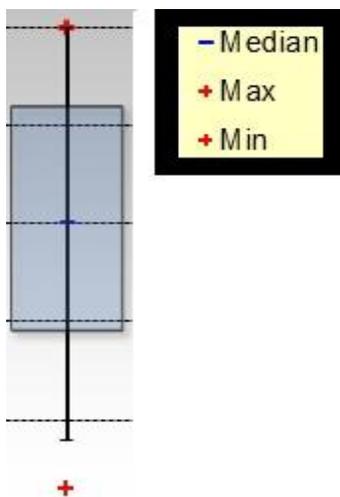


Abbildung MM: Grundschemata eines Box-Plots ¹⁰⁹

Die blaue Linie in der Box gibt den Mittelwert der jeweiligen Daten an, die Box beinhaltet 50 % der zentralen Daten um den Mittelwert. Die roten Kreuze zeigen die Extreme der Daten auf. Die schwarze Linie durch die Box markiert, wo sich 95% der Daten befinden. ¹¹⁰

¹⁰⁹ (Klein- Excelscreenshot, 2013)

¹¹⁰ Vgl. (Schlittgen, 2012 S. 32)

11.3 Auswertung der Qualitätsstufen über alle Hörbeispiele

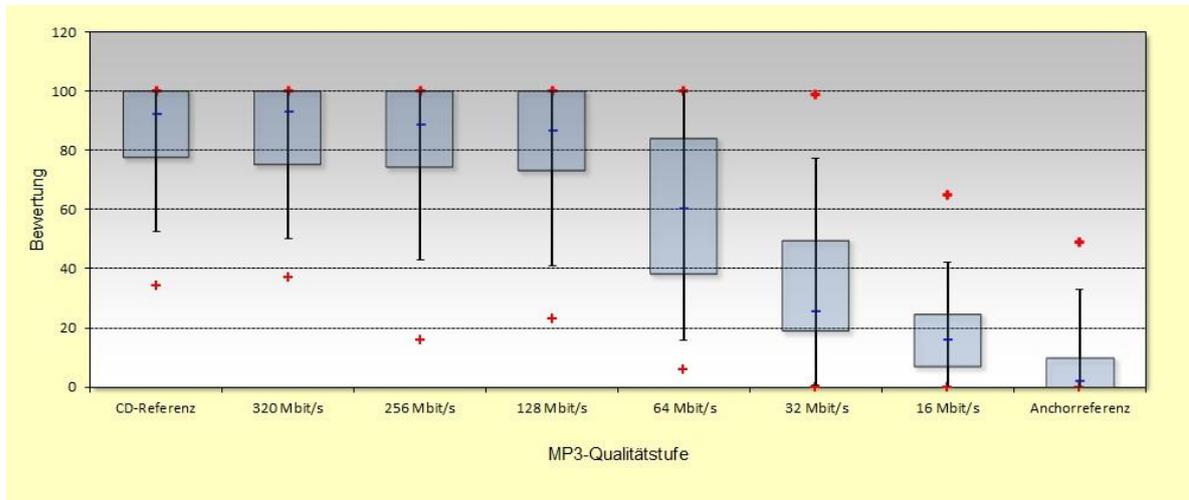


Abbildung NN: Box-Plot der Qualitätsstufen über alle Hörbeispiele ¹¹¹

Die Box-Plot-Ansicht liefert die Auswertung der verschiedenen Qualitätsstufen über alle Hörbeispiele. Die wenigsten Testhörer hatten die Referenz nicht erkannt und bewerteten diese unter 40 Punkte. Die meisten stuften das Referenz-Hörbeispiel im Bereich exzellent (80-100) ein. Ebenso verhielt es sich mit den drei darauffolgenden MP3-Qualitätsstufen. Die MP3-Stufen schlechter als 64 MBit/s wurden vom Hörer als hörbar schlechter eingestuft, auch wenn es teilweise Ausreißer nach oben gab. Die Anchor-Referenz wurde immer erkannt, somit kann davon ausgegangen werden, dass es sich um vertrauenswürdige Daten handelt.

Folgende Werte lassen sich aus dem Datenbestand berechnen:

| | N | Mittelwert | Konf. (±) | Std.Fehler | Std.Abw. |
|----------------|-----|-------------|-------------|-------------|-------------|
| CD-Referenz | 136 | 86,78676471 | 2,662442503 | 1,346236912 | 15,69968534 |
| 320 Mbit/s | 136 | 86,63235294 | 2,667350881 | 1,348718783 | 15,72862869 |
| 256 Mbit/s | 136 | 83,10294118 | 3,276920687 | 1,656941542 | 19,32309285 |
| 128 Mbit/s | 136 | 81,89705882 | 3,240836602 | 1,638695992 | 19,110315 |
| 64 Mbit/s | 136 | 60,72058824 | 4,502679453 | 2,276733967 | 26,55105248 |
| 32 Mbit/s | 136 | 32,50735294 | 3,717901517 | 1,879919004 | 21,92343455 |
| 16 Mbit/s | 136 | 17,63970588 | 2,279503613 | 1,152607766 | 13,44160087 |
| Anchorreferenz | 136 | 7,257352941 | 1,826726969 | 0,923665871 | 10,77170252 |

Abbildung OO: Auswertungstabelle aller Qualitätsstufen über alle Hörbeispiele ¹¹²

¹¹¹ (Klein- Excelscreenshot, 2013)

¹¹² (Klein- Excelscreenshot, 2013)

Die Tabelle liefert die Zahlen der Auswertung. Neben dem arithmetischen Mittelwert werden die Konfidenzintervalle, die Standardfehler und deren Abweichungen angezeigt.

Anhand dieser Werte und Grafiken können viele Erkenntnisse gewonnen werden: Die Referenz wurde von den Testhörer erkannt, jedoch zeigen die besten drei MP3-Stufen ebenfalls eine exzellente Bewertung auf. Die Mittelwerte liegen so eng beieinander das subjektiv für den Normalhörer kein Unterscheid vorliegt. Auch die Konfidenzintervalle verstärken diese Erkenntnis.

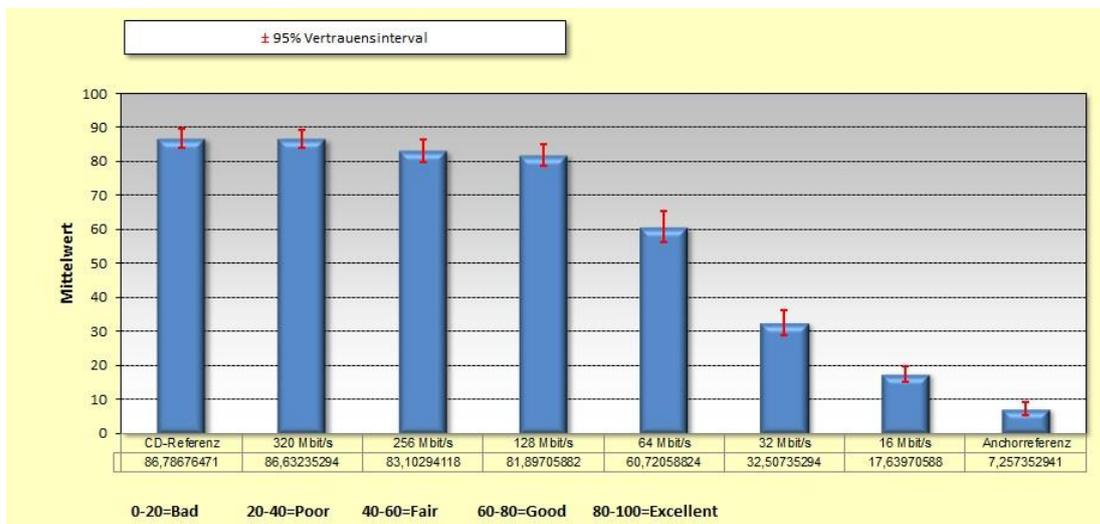


Abbildung PP: Mittelwerte der Qualitätsstufen über alle Hörbeispiele

Im Diagramm der Abbildung PP sind der arithmetische Mittelwert und das Konfidenzintervall dargestellt. Erst ab einer MP3-Qualitätsstufe von 64 MBit/s ist ein subjektiver Unterscheid für die meisten Testhörer erkennbar. Der t-Test hierzu ist unter 11.4 Auswertung der Hörbeispiele mittels eines t-Tests auf S- 67 - nachlesbar.

11.4 Auswertung der Hörbeispiele mittels eines t-Tests

Mittels eines t-Tests lassen sich Angaben machen, inwiefern sich Ergebnisse signifikant unterscheiden. Betrachtet man die Qualitätsstufen über alle Hörbeispiele (vgl. Abbildung PP: Mittelwerte der Qualitätsstufen über alle Hörbeispiele) lässt sich eine gewisse Tendenz erkennen, welche vermuten lässt, dass sich MP3-Qualitätsstufen mit 320 MBit/s, 256 MBit/s und 128MBit/s nicht signifikant von der Referenz unterscheiden. Um diese Annahme zu bestätigen wird der p-Wert durch einen t-Test ermittelt:

| | N | Mittelwert | Vertrauen (±) | Std. Fehler | Std. Abw. |
|--------------------|-----|-------------|---------------|-------------|-------------|
| CD-Referenz | 136 | 86,78676471 | 2,662442503 | 1,346236912 | 15,69968534 |
| 320 Mbit/s | 136 | 86,63235294 | 2,667350881 | 1,348718783 | 15,72862869 |
| gesamte Stichprobe | 272 | 86,70955882 | 1,872411511 | 0,951063112 | 15,68533467 |

t-Test:

| Varianzschätzung | T | Freiheits- grade | P |
|------------------|-------------|---------------------|-------------|
| Homogen | 0,081029577 | 270 | 0,935478462 |
| Heterogen | 0,081029577 | 269,999084 | 0,935478462 |

Abbildung QQ: t-Test CD-Referenz und MP3 mit 320 MBit/s ¹¹³

Der t-Test liefert einen p-Wert von > 0,05, somit unterscheiden sich die Referenz und das 320 MBit/s-MP3 nicht signifikant voneinander.

| | N | Mittelwert | Vertrauen (±) | Std. Fehler | Std. Abw. |
|--------------------|-----|-------------|---------------|-------------|-------------|
| CD-Referenz | 136 | 86,78676471 | 2,662442503 | 1,346236912 | 15,69968534 |
| 256 Mbit/s | 136 | 83,10294118 | 3,276920687 | 1,656941542 | 19,32309285 |
| gesamte Stichprobe | 272 | 84,94485294 | 2,10920484 | 1,071338703 | 17,66897053 |

t-Test:

| Varianzschätzung | T | Freiheits- grade | P |
|------------------|-------------|---------------------|-------------|
| Homogen | 1,725522998 | 270 | 0,085577412 |
| Heterogen | 1,725522998 | 259,1387976 | 0,085625348 |

Abbildung RR: t-Test CD-Referenz und MP3 mit 256 MBit/s ¹¹⁴

Der t-Test liefert einen p-Wert von > 0,05, somit unterscheiden sich die Referenz und das 256 MBit/s-MP3 nicht signifikant voneinander.

¹¹³ (Klein- Excelscreenshot, 2013)

¹¹⁴ (Klein- Excelscreenshot, 2013)

| | N | Mittelwert | Vertrauen (±) | Std. Fehler | Std. Abw. |
|--------------------|-----|-------------|---------------|-------------|-------------|
| CD-Referenz | 136 | 86,78676471 | 2,662442503 | 1,346236912 | 15,69968534 |
| 128 Mbit/s | 136 | 81,89705882 | 3,240836602 | 1,638695992 | 19,110315 |
| gesamte Stichprobe | 272 | 84,34191176 | 2,104200751 | 1,068796952 | 17,6270509 |

t-Test:

| Varianzschätzung | T | Freiheits- grade | P |
|------------------|------------|---------------------|-------------|
| Homogen | 2,30562429 | 270 | 0,021889769 |
| Heterogen | 2,30562429 | 260,1977427 | 0,021918487 |

Abbildung SS: t-Test CD-Referenz und MP3 mit 128 MBit/s ¹¹⁵

Der t-Test liefert einen p-Wert von <0,05, somit unterscheiden sich die Referenz und das 128 MBit/s-MP3 signifikant voneinander.

Abschließend lässt sich sagen, dass MP3-Dateien mit einer Bitrate von 320 oder 256 MBit/s sich vom Normalhörer nicht von der Originalreferenz unterscheiden lassen. Für den Konsument ist kein Unterschied zu erkennen. Ab einer Qualitätsstufe von 128 MBit/s oder schlechter werden Unterschiede zur Referenzdatei wahrgenommen.

11.5 Auswertung der Hörbeispiele über die verschiedenen Qualitätsstufen

Im Folgenden werden die Mittelwerte und Konfidenzintervalle der einzelnen Qualitätsstufen betrachtet. Es wird vermutet, dass für spezielle Instrumentengruppierungen der aus der SQAM-CD vorgeschlagenen Programmausschnitte, niederere Qualitätsstufen ausreichen, um keine signifikanten Unterschiede zu erzielen.

¹¹⁵ (Klein- Excelscreenshot, 2013)

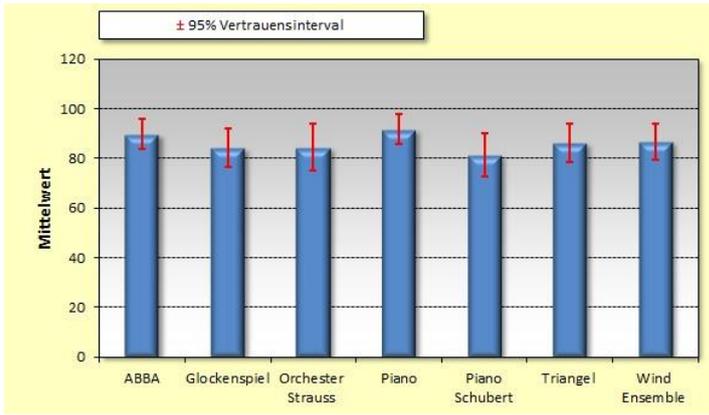


Abbildung TT: Mittelwerte bei MP3-Dateien mit 320MBit/s ¹¹⁶

Im Bereich der hochwertigeren MP3-Dateien liegen alle Hörbeispiele im exzellenten Bereich zwischen 80 und 100 Punkten. Es ist kein Trend erkennbar, dass bestimmte Programmausschnitte eine hochwertigere Kodierung benötigen.

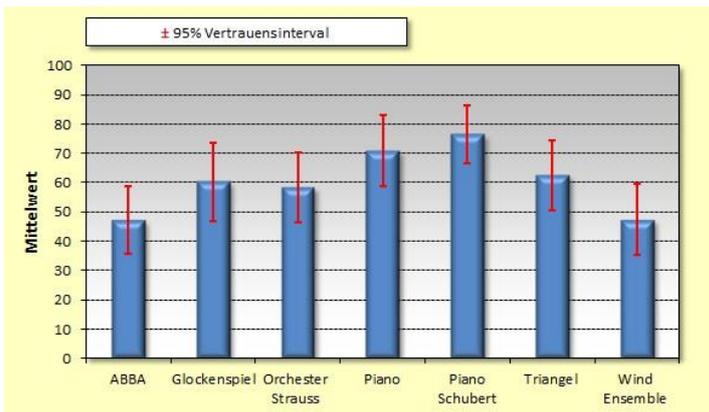


Abbildung UU: Mittelwerte bei MP3-Dateien mit 64 MBit/s ¹¹⁷

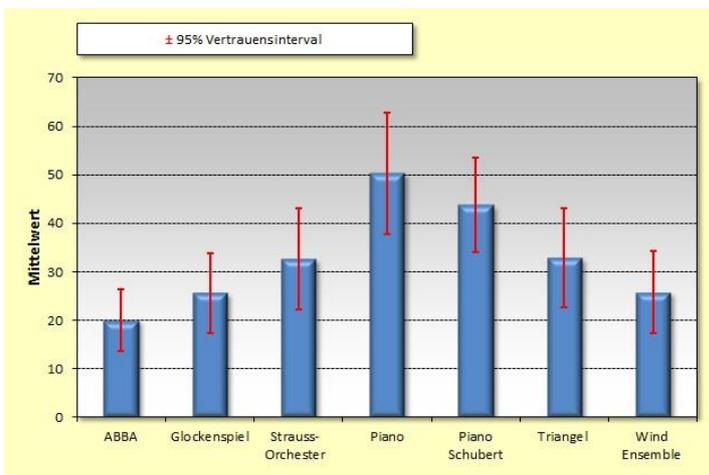


Abbildung VV: Mittelwerte bei MP3-Dateien mit 32MBit/s ¹¹⁸

¹¹⁶ (Klein- Excelscreenshot, 2013)

¹¹⁷ (Klein- Excelscreenshot, 2013)

Betrachtet man die verschiedenen Programmausschnitte in den Qualitätsstufen von 32 MBit/s und 64 MBit/s lässt die Tendenz erkennen, dass Soloinstrumente (Glockenspiel, Solo-Piano, Triangel) bereits in niedriger Qualitätsstufe als sehr gut eingeschätzt werden. Deutliche Unterschiede der Stufen lassen sich vor allem in den Orchester- und Popwerken hören. Hier ist erkennbar, dass niedrige MP3-Stufen eine subjektive Einschränkung der Qualität mit sich bringen. Somit könnte man annehmen, dass gewissen Stil- und Musikrichtungen oder Besetzungen eine niedrigere bzw. höhere Qualitätsstufe voraussetzen.

11.6 Abschlussbericht des MUSHRA-Hörtests

Grundsätzlich lässt sich sagen, dass der MPEG-Layer-3-Codec sehr gut arbeitet. Die hochwertigen MP3-Bitraten lassen kaum einen Spielraum zum Original offen. Lediglich extrem niedrige Bitraten (16, 32, 64, 128 MBit/s) wurden von den Normalhörern (keine Experten) als schlechte Qualität dotiert. Bitraten höher als 128 MBit/s liefern bereits eine sehr starke und stabile Qualität und unterscheiden sich nicht signifikant voneinander.

Ferner kann vermutet werden, dass Einzelinstrumente, wie beispielsweise ein Soloklavier, eine Triangel oder ein Glockenspiel in der Bitratereduktion deutlich unproblematischer sind. Hier lassen sich auch mit kleinen Bitrates teilweise geringe Unterscheide zur Referenz erkennen.

Durch Gespräche mit den Testpersonen nach dem Hörvorgang wurde ebenfalls deutlich, dass auch leise Passagen oft als unproblematischer beschrieben werden als laute.

12. Fazit

Die wohl wichtigsten Dokumente, auf welche sich auch diese Arbeit stützt, sind die ITU-Empfehlungen BS.1116-1 und BS.1534-1. Diese geben Anstöße bzw. Empfehlungen, wie Hörtests durchgeführt werden sollen. Dabei werden für verschiedene Zwecke unterschiedliche Möglichkeiten vorgestellt: Für feine Beurteilungen steht der sogenannte ABC/HR-Test (BS.1116-1) und für grobe und mittlere Qualitätsunterschiede steht der MUSHRA-Test (BS.1534-1) zur Verfügung. Beide Testmethoden sind sehr gut und detailliert beschrieben.

Softwaregestützte Hörversuche erleichtern die Durchführung von Hörtests im Vergleich zur analogen Methoden. Durch die Möglichkeiten der digitalen Verarbeitung können Daten schneller und leichter verglichen werden, als mit Zettel und Stift. Trotzdem bringen verschiedenen Softwares unterschiedliche Vor- und Nachteile mit sich. Viele Anwendungen und Programme sind frei im Internet erhältlich. Leider arbeiten diese nicht alle nach den Empfehlungen der ITU. Oft sind es Kleinigkeiten, die den Benutzer letztendlich zu einem verfälschten Ergebnis führen. Bei den getesteten kommerziellen Softwares konnten kaum Mängel diesbezüglich festgestellt werden, alle arbeitet sehr exakt nach den Anforderungen.

Trotzdem können auch für kommerzielle Produkte im Bereich der GUI und dem Workflow gewisse Verbesserungen beanstandet werden. Sowohl bei Freeware-Produkten als auch kommerziellen Softwares erscheinen die Erstellung und vor allem die Auswertung von Tests teilweise sehr kompliziert und umständlich. In den meisten Fällen erhält man lediglich eine unübersichtliche Datei mit den Ergebnissen (oft als Textdatei), die dann manuell in eine zusätzliche Auswertungssoftware übertragen werden müssen. Die wenigsten bieten die Möglichkeit, zusammengetragene Daten direkt auszuwerten. Selbst wenn dies gegeben ist, erscheint zum Teil die Bedienung recht umständlich und bietet auch oft nicht alle Ergebnisse, die gerne ersichtlich wären. Wenige Softwares arbeiten mit einer direkten Verbindung zu Excel oder anderen Programmen zur Auswertung der Daten, was eine dann aber erhebliche Erleichterung beim Auswertungsprozess mit sich bringt.

Auch in der Testdurchführung, wie es sich die ITU vorstellt, gibt es noch bedeutende Verbesserungsmöglichkeiten; oft werden, selbst in kommerziellen Programmen, keine Trainingseinheiten angeboten und wenn, dann sind diese nicht nach den

Vorstellungen der ITU gestrickt. Die Trainingsphase ist ein elementarer Bestandteil bei der Durchführung der Hörtests und sollte auf jeden Fall gewährleistet sein.

Eine intensive Einarbeitungszeit ist auch definitiv notwendig, da der Proband sonst recht schnell überfordert wird und ihm nicht ganz klar ist, worum es in den Hörtests gehen soll.

Um die Einarbeitungszeit zu vereinfachen, wäre es möglich und hilfreich Videotutorials zur Verfügung zu stellen, die den Proband in die Thematik einführen und die Dokumente der ITU bzw. die Softwares verständlich erklären und beschreiben. Zudem könnte man mit Präsentationen und kurzen schriftlichen Statements schneller die Benutzer erreichen und eine Einarbeitung in die Softwarebedienung erleichtern.

Um die Funktionen der getesteten Softwares noch etwas zu verdeutlichen, befindet sich im Folgenden eine Tabelle, welche die jeweiligen Eigenschaften einiger der untersuchten Programme beschreibt:

| | foobar2000 | STEP | MUSHRA _{Test} | WinABX | ABC HRACT* | scale |
|------------------------------------------------------------|------------------|-----------------------------|---------------------------------------|-----------------------------------------|---------------------|-------|
| ABC/HR -Methode | | x | | x | x | x |
| MUSHRA -Methode | | x | x | | | x |
| AB-Methode | | x | | | | x |
| ABX -Methode | x | x | | | | x |
| Weitere Methoden | | | | x | | x |
| kontinuierliche Abhörmöglichkeit ohne stoppen beim Wechsel | x | | x | x | x | |
| schriftliche Einführung | | | | | | x |
| ITU-gerechte Skalen | | x | x | x | x | x |
| ITU-gerechte Trainingsphase | | teilweise | x | | | |
| ITU-gerechte Testphase | | x | x | x | | x |
| interne statistische Auswertungen | | | | | | x |
| Daten als Text- oder Excel-Datei | | x | x | x | x | x |
| kostenlos | x | | x | x | x | x |
| Sonstige Vorteile | | Verknüpfung mit Excel | | | | |
| Sonstige Nachteile | ausschließl. ABX | umständliche Testerstellung | umständliche Stimulieinbindung | ausschließl. ABX | ausschließl. ABC/HR | |
| | | | keine Stimulibeschreibung im Training | Strörung beim Abspielen der Audiotdaten | | |

* ABC Hidden Reference Audio Comparison Tool

Ich habe mich mit weitaus mehr Programmen beschäftigt, jedoch erschienen mir *foobar2000*, *STEP*, *MUSHRA_{Test}*, *WinABX*, *ABC HRACT* und *scale* die Wichtigsten, da viele andere gar nicht funktionieren oder extreme Schwächen aufzeigen, sodass diese nicht erwähnenswert sind.

Für mich persönlich bieten die Softwares *scale* und *STEP* die umfangreichsten und besten Eigenschaften, allerdings ist die Testerstellung in beiden Programmen sehr umständlich. Zudem liegen die Kosten der Software *STEP* bei 299\$.

Die Datenanalyse in *scale* weist eine sehr komplexe Struktur auf, welche eine längere Einarbeitungszeit in Anspruch nimmt; in *STEP* dagegen lassen sich über eine Verknüpfung mit Excel alle Daten schnell aufrufen und anzeigen.

Ansonsten bieten beide Softwares alle nötigen Optionen, die benötigt werden.

Für die Durchführung eines MUSHRA-Tests würde ich außerdem die von mir verwendete Software *MUSHRA*Test befürworten. Sie ist sehr ansprechend gestaltet und beinhaltet nur die notwendigsten Funktionen, welche aber absolut ausreichend und normgerecht nach der ITU programmiert sind. Leider fehlt auch hier der Teil für die statische Auswertung. Ferner ist auch hier die Einbindung der Stimuli etwas aufwändiger, zumal die Software nur mit WAV 44,1 kHz 32 Bit-float-Dateien arbeiten kann. Ein großer Vorteil jedoch ist: *MUSHRA*Test spielt die Programmausschnitte kontinuierlich ab und startet nicht beim Wechsel auf ein neues Stimuli den Ausschnitt von vorne.

Obwohl es in Zukunft voraussichtlich noch mehr Programme für die Durchführung von Hörtests geben wird, sollte man diese genau untersuchen, um eventuelle Schwachpunkte oder Fehler aufzudecken. Nicht jedes Programm arbeitet exakt nach den Empfehlungen der ITU, daher ist ein genaueres Hinsehen wichtig, um repräsentative Ergebnisse zu erzielen.

13. Anhang

13.1 Audio-CD 1: Psychoakustische Effekte ¹¹⁹

Track 1: Verdeckungseffekt durch Impulslänge

Track 2: Verdeckungseffekt durch Pegel und Tonhöhe

¹¹⁹ Vgl. Audio-CD (Zwicker, 1990, 1999, 2007) Track 32, Track 9

13.2 CD und Trackinformationen zur SQAM-CD

| Track information | | | | | | | | | Suggested Applications (see p. 12) | | | | | | | | | |
|-------------------|--------------|----------|---------------|--------|--|-----------------|---------------------|------------------|---------------------------------------|----------------------|---------------------|------------|--------------------|---------------|--------------------|---------------------------|---------------------------|--------------------|
| Track number | Index Number | Duration | Contents | | | Number of Takes | Mono (M)/Stereo (S) | Attenuation (dB) | Notes (See Page 10) | A/D & D/A conversion | Aliasing distortion | Bit errors | Bit-rate reduction | Dynamic range | Frequency response | Overload after processing | Programme modulated noise | Stereophonic image |
| 53 | 01 | 0:21 | Female speech | German | | 1 | M | 25 | 7 | . | . | . | . | . | . | . | . | . |
| 54 | 01 | 0:21 | Male speech | German | | 1 | M | 25 | 7 | . | . | . | . | . | . | . | . | . |

Solo instruments

| | | | | | | | | | | | | | | | | | | |
|----|----|------|---------|----------|--|---|---|---|----|---|---|---|---|---|---|---|---|---|
| 55 | 01 | 0:32 | Trumpet | Haydn | | 1 | S | | 9 | . | . | X | . | . | . | . | . | . |
| 56 | 01 | 0:34 | Organ | Handel | | 1 | S | | 10 | . | . | . | . | . | . | . | . | . |
| 57 | 01 | 0:20 | Organ | Bach | | 1 | S | | 11 | . | X | . | . | . | . | . | . | . |
| 58 | 01 | 0:16 | Guitar | Sarasate | | 1 | S | | 12 | . | X | . | . | . | . | X | X | . |
| 59 | 01 | 0:29 | Violin | Ravel | | 1 | S | | 13 | . | . | . | . | . | . | . | . | . |
| 60 | 01 | 1:32 | Piano | Schubert | | 1 | S | 9 | 4 | X | . | X | X | X | . | . | X | . |

Vocal & Orchestra

| | | | | | | | | | | | | | | | | | | |
|----|----|------|----------|-----------|--|---|---|--|----|---|---|---|---|---|---|---|---|---|
| 61 | 01 | 2:59 | Soprano | Mozart | | 1 | S | | 14 | . | . | . | . | . | X | X | . | . |
| 62 | 01 | 0:31 | Soprano | Spiritual | | 1 | S | | 15 | . | . | . | . | . | . | . | . | . |
| 63 | 01 | 0:57 | Soloists | Verdi | | 1 | S | | 16 | . | . | . | . | . | X | . | . | X |
| 64 | 01 | 0:31 | Choir | Orff | | 1 | S | | 17 | . | . | . | . | . | X | . | . | . |

Orchestra

| | | | | | | | | | | | | | | | | | | |
|----|----|------|---------------|------------|--|---|---|--|-----|---|---|---|---|---|---|---|---|---|
| 62 | 01 | 1:52 | Orchestra | R. Strauss | | 1 | S | | 18 | . | . | . | X | X | X | . | . | . |
| 66 | 01 | 0:18 | Wind ensemble | Stravinsky | | 1 | S | | 19 | . | X | . | X | X | . | . | . | . |
| 67 | 01 | 1:22 | Wind ensemble | Mozart | | 1 | S | | 20 | . | . | . | . | . | . | . | . | . |
| 68 | 01 | 2:44 | Orchestra | Baird | | 1 | S | | 2&8 | . | . | . | . | X | . | . | . | . |

Pop Music

| | | | | | | | | | | | | | | | | | | |
|----|----|------|---------------|--|--|---|---|--|----|---|---|---|---|---|---|---|---|---|
| 69 | 01 | 0:33 | ABBA | | | 1 | S | | 21 | . | . | . | X | . | . | . | . | X |
| 70 | 01 | 0:21 | Eddie Rabbitt | | | 1 | S | | 22 | . | . | . | . | . | . | X | X | . |

Abbildung WW: Track Index SQAM CD ¹²⁰

13.3 Softwaredokumentation von foobar2000

Die folgende Dokumentation beschreibt die wichtigsten Eigenschaften und Funktionen von foobar2000. ¹²¹

13.3.1 Systemvoraussetzungen und Installation

Die Software *foobar2000* läuft laut Hersteller auf allen Windows-Systemen ab Windows XP-SP2 oder höher.

Die Installation von *foobar2000* erweist sich als recht einfach. Die Software lässt sich auf der offiziellen *foobar2000* Website direkt herunterladen: <http://www.foobar2000.org/download> ¹²²

Mittlerweile gibt es die *foobar2000 v1.2.8* Version. Nachdem der Download abgeschlossen ist, lässt sich die *foobar2000_v1.2.8.exe* -Datei mittels Doppelklick öffnen. Der Installationsassistent öffnet sich und führt den Benutzer Schritt für Schritt durch die Installation. Nachdem Akzeptieren der Lizenzbedingungen kann man eine Standard-Installation auf dem ausführenden Datenträger oder einem externen bzw. mobilen Datenträger ausführen.

Im Regelfall entscheidet man sich für den internen Datenträger, auf dem auch die Systemdateien von Windows lagern. Nach dem Bestätigen der Installationswahl fragt *foobar2000* nach dem Speicherplatz für die Installation. Auch hier bietet sich der Standardplatz und *C:\Programme* an.

Folgende Installationsoptionen stehen zur Verfügung:

| Auswahl | Funktion |
|--------------------------------|--------------------------------------------------------|
| Audio CD support | Audio-CDs wiedergeben bzw. auf die Festplatte rippen |
| Album Lis | Eine Listenstruktur der Musiksammlung anzeigen lassen. |
| Standard DSPs/Equalizer DSP | Digitalem Signalprozessor (DSP) z.B. Equalizer |
| ReplayGain Scanner | Anpassung der Lautstärke aller Musiktitel |

¹²¹ Quelle dieser Informationen stammen aus der Homepage von foobar direkt (www.foobar.org) und durch ausprobieren und testen der Software

¹²² Stand 05.06.2013

| | |
|-----------------------------------------------------------------------------------------------------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Converter | Dateiformatkonvertierungen |
| File Operations | Audiodateien direkt aus <i>foobar2000</i> heraus auf der Festplatte löschen oder verschieben. |
| Archive Reader | Hier lassen sich ZIP- oder RAR-Archive ohne zu entpacken lesen und abspielen. |
| Online Tagger | Es können Meta-Informationen zu Musikdateien aus einer zentralen Datenbank aus dem Internet gelesen und in die Metadaten z.B. der mp3 eingetragen werden. |
| Shortcut-Icons <ul style="list-style-type: none"> • Desktop Icon • Start Menü Icon • Quick Launch Icon | Wo soll eine Icon-Verknüpfung entstehen: <ul style="list-style-type: none"> • Auf dem Desktop • Im Startmenü • Auf der Schnellstartleiste |

123

13.3.2 Oberflächenaufbau

Die Bedienung von *foobar2000* ist recht simpel und gestaltet sich als benutzerfreundlich und selbsterklärend. Am oberen Rand der Benutzeroberfläche ist auf der linken Seite das Menü von *foobar2000*, rechts davon befindet sich ein Bedienfeld mit: Stop, Start, Pause, Titel vor, Titel zurück und einer Titel random-Funktion. In der rechten Hälfte ist ein kleines Spektrometer vorhanden und ein Zeitverlauf des aktuellen Titels. Zudem lässt sich die Lautstärke über eine Volumenfunktion steuern. Jedes einzelne Feld lässt sich in der Größe variieren und anpassen. Diese Abspielfunktionen tauchen ebenfalls im Menüpunkt „Playback“ auf.



Abbildung XX: Menüleiste *foobar2000* ¹²⁴

Im unteren Bereich von *foobar2000* gibt es die sogenannten Playlists. Diese werden benötigt, damit überhaupt Musik abgespielt werden kann. Die Playlists sind in

¹²³ (foobar2000)

¹²⁴ (Klein(Screenshot))

Spalten und Zeilen aufgeteilt, aus denen verschiedene Informationen zu den einzelnen Titeln abgelesen werden können. Neben dem Album lassen sich Tracknummer, Titel und die Spielzeit für jeden Titel ablesen.

Durch einen Rechtsklick in einen leeren Bereich lassen sich in zwei Fenstern verschiedene Informationen der Titel detailliert anzeigen. Neben den Metadaten kann auch man die Location und allgemeine Informationen der Datei abrufen.

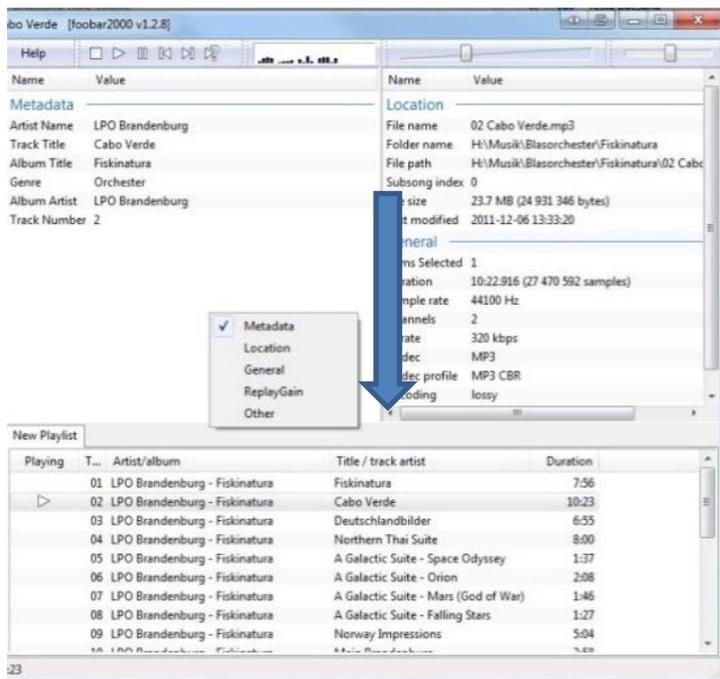


Abbildung YY: Ansicht anpassen foobar2000¹²⁵

13.3.3 Die Playlists

Sobald die Titel in die Playlist geladen wurden, lassen sich diese in ihrer Abspielreihenfolge durch die gedrückte linke Maustaste an die gewünschte Position ziehen und über die „Entf“-Taste aus der Playlist löschen. Neue Playlists lassen sich entweder über File/New Playlist, mit dem Tastenkürzel „Strg N“ oder einem Doppelklick auf einen grauen Bereich neben der aktuellen Playlist erzeugen. Über denselben Weg lassen sich auch die Playlists umbenennen, speichern oder löschen.

¹²⁵ (Klein(Screenshot))

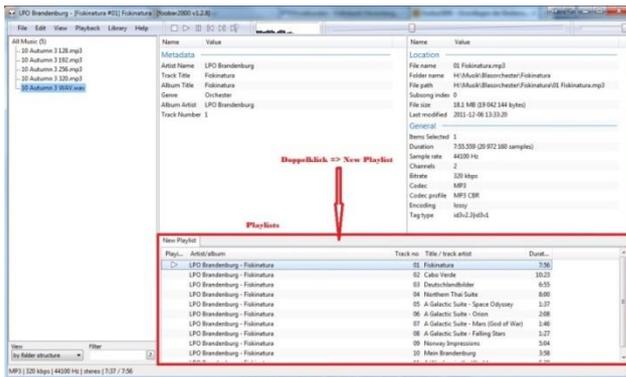


Abbildung ZZ: Playlists in *foobar2000* ¹²⁶

Playlists enthalten lokale Bezüge zu den Speicherorten der jeweiligen Dateien auf dem eigenen Rechner und sind somit nicht für die Weitergabe an andere Computer geeignet, solange sich die Dateien nicht in einem Netzwerk oder im Internet befinden. In letzterem Fall bietet sich *foobar2000s* Funktion an, die Software über einen Musikserver fernzusteuern.

13.3.4 Menüpunkt File

Die Titel lassen sich über verschiedene Wege der Playlist hinzufügen:

Über „Drag and Drop“ können Files direkt in eine Playlist gezogen werden, zudem können Songs über das Menüfeld File/Open direkt ausgewählt werden. Um eine Audio-CD einzubinden, öffnet man diese über File/Open CD. Hier muss entschieden werden, ob die Audio-CD nur abgespielt oder auch gerippt werden soll. *Foobar2000* kann dann mittels Datenbanken, die sich online befinden, die Metadaten und Informationen zum Album abspeichern. Nachdem die Audio-CD auf den PC gerippt wurde, können aus verschiedenen Quellen die Metadaten abgerufen werden. *Foobar2000* arbeitet hier primär mit der Datenbank „freedb“, eine öffentliche Datenbank, auf die jeder zugreifen kann. Man kann auch bereits gerippte Daten (wenn es sich um ein komplettes Album handelt) via Rechtsklick/Tagging auf Metadaten von „freedb“ überprüfen lassen.

¹²⁶ (Klein(Screenshot))

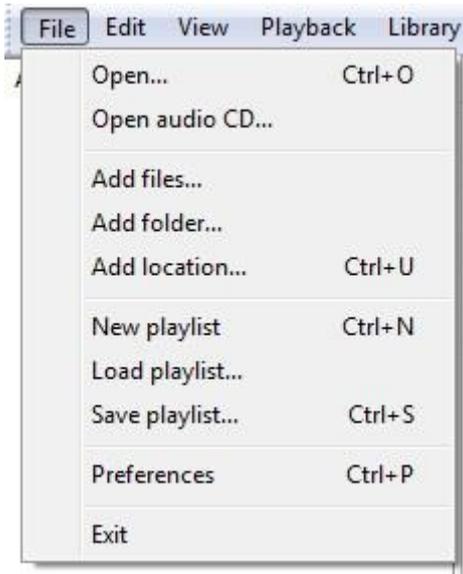


Abbildung AAA: Menüpunkt File in foobar2000 ¹²⁷

Unter „File“ tauchen die Möglichkeiten auf, eine neue Playlist zu erstellen, eine alte zu laden oder eine aktuelle zu speichern. Unter „Preferences“ bietet foobar2000 viele Optionen und Einstellungsmöglichkeiten, auf die später eingegangen wird. Über Exit kann das Programm geschlossen werden.

13.3.5 Menüpunkt Edit

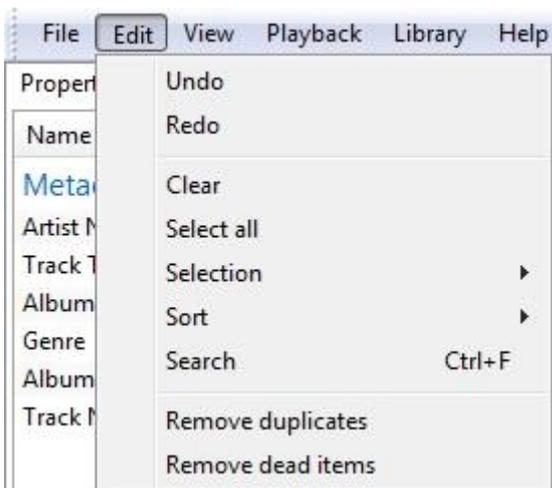


Abbildung BBB: Menüpunkt Edit in foobar2000 ¹²⁸

¹²⁷ (Klein(Screenshot))

¹²⁸ (Klein(Screenshot))

Im Menüpunkt „Edit“ lässt sich neben der Funktion, den letzten Vorgang wiederherzustellen oder rückgängig zu machen, die aktuelle Playlist bearbeiten. Folgende Funktionen stehen zur Verfügung:

| | |
|-------------------|------------------------------------------------------------------------------------------------------------------------------|
| Clear | Playlistinhalte löschen |
| Select all | Alle Inhalte der Playlist markieren. |
| Selection | Einzelne Titel entfernen bzw. alle außer einem Titel entfernen. |
| Sort | Hier gibt es verschiedene Möglichkeiten die aktuellen Songs der Playlist zu sortieren z.B. nach Länge, Album, Interpret usw. |
| Remove duplicates | Lieder, die doppelt in der Playlist auftauchen, werden gelöscht. |
| Remove dead items | Löscht Songs aus der Playlist die es auf dem Datenträger nicht mehr gibt. |

13.3.6 Menüpunkt View



Abbildung CCC: Menüpunkt View in foobar2000

Hier lassen sich verschiedene visuelle Eigenschaften ändern. Mit der Funktion „Always on top“ bleibt das foobar2000 Fenster immer an der Oberfläche und verschwindet nie hinter anderen Programmen. Unter Visualizations lassen sich optische Darstellungen der Audiospuren anzeigen:

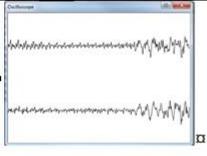
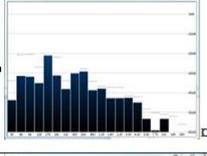
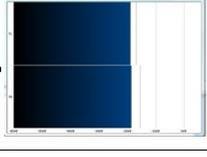
| | |
|-----------------------------------------------------------------------------------|--------------|
|  | Oszilloskop |
|  | Peak-Meter |
|  | Spektrogramm |
|  | Spektrum |
|  | VU-Meter |

Abbildung DDD: Tabelle der Visulationen von foobar2000 ¹²⁹

Im Untermenüpunkt „Console“ lässt sich der Verlauf nachvollziehen. Hier wird dokumentiert, was geöffnet, gelöscht, verschoben oder geändert wurde. Über die „Clear“-Funktion kann man den Verlauf löschen.

Mit der Equalizerfunktion kann ein EQ aktiviert werden, mit welchem man klangliche Veränderungen erreichen kann. Es handelt sich hier um einen 18-Band-EQ mit einer Steuerung von +/- 20 dB pro Band.

Der Playlistmanager ist standardmäßig in der Oberfläche eingebettet. Hier lassen sich verschiedene Playlists erstellen, umbenennen oder erzeugen.

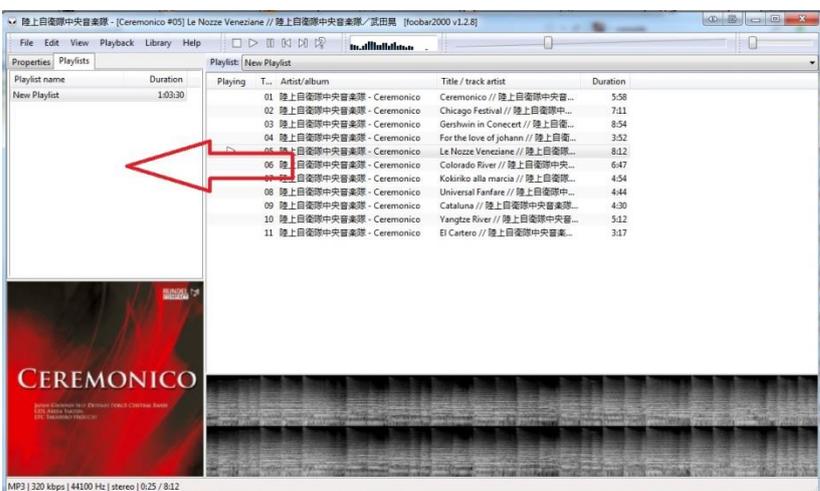


Abbildung EEE: Playlistmanager in foobar2000 ¹³⁰

¹²⁹ (Klein(Screenshot))

In der Unterfunktion Layout lässt sich die Benutzeroberfläche individuell anpassen. Es gibt vorgefertigte Layouts, die bestimmte Fenster bereits anzeigen.

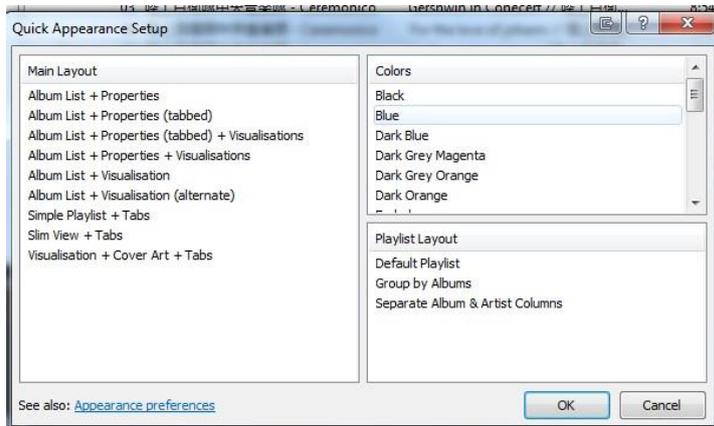


Abbildung FFF: Vorgefertigte Layouts in foobar2000 ¹³¹

Zudem lassen sich verschiedene Ansichten der Playlists einstellen und die Hintergrundfarbe ändern.

Mit der Funktion „Enable Layout editing mode“ lassen sich mit Rechtsklick optische Anpassungen vornehmen, beispielsweise lässt sich in der Playlist bei eingeschaltetem editing mode auswählen, welche Informationen angezeigt werden sollen (Titellänge, Titelname, Albumname).

Mit der Scratchbox können eigene Layouts entwickelt werden. Hier stehen die verschiedenen Bedienelemente wie z.B. Equalizer, Bibliothek oder Playlists einfügen, die nach eigenem Geschmack angelegt werden können oder auch als Standardlayout übernommen werden können, zur Verfügung. Diese Layouts lassen sich speichern und ggf. auch in ein anderes System importieren.

¹³⁰ (Klein(Screenshot))

¹³¹ (Klein(Screenshot))

13.3.7 Menüpunkt Playback



Abbildung GGG: Menüpunkt Playback in *foobar2000* ¹³²

Hier sind die Abspielmöglichkeiten der Songs aufgelistet, die sich ebenfalls in der Menüleiste widerspiegeln (siehe Abbildung XX: Menüleiste *foobar2000*). Zusätzlich lassen sich unter „Order“ weitere Einstellungen zum zufälligen Abspielen von Titeln vornehmen:

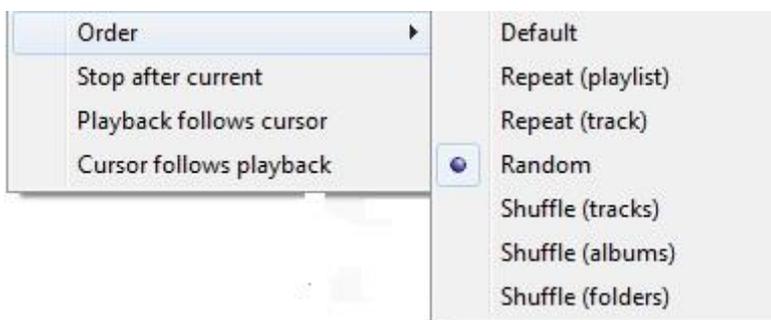


Abbildung HHH: Randomfunktionen in *foobar2000* ¹³³

Bei der Auswahlmöglichkeit „Shuffle (track)“ werden alle Tracks in der Playlist zufällig abgespielt. Mit „Shuffle (album)“ werden die verschiedenen Alben einer Playlist zufällig abgespielt, allerdings wird ein Album immer komplett durchgespielt. Das gleiche Prinzip gilt für den Menüpunkt „Shuffle (folders)“, allerdings verwendet hier *foobar2000* die verschiedenen Ordner aus der Bibliothek. Random hingegen spielt immer in einer neuen zufälligen Reihenfolge – es wird nach jedem Schleifendurchgang neu gewürfelt.

¹³² (Klein(Screenshot))

¹³³ (Klein(Screenshot))

13.3.8 Menüpunkt Library

Wenn nicht bereits im Layout integriert, kann die Datenbank bzw. Bibliothek hier aufgerufen werden. Über den Menüpunkt Library/„Album List“ erreicht man alle importierten Ordner bzw. Alben.

In der „Album List“-Funktion kann die Datenbank nach der Ordnerstruktur auf dem PC, nach Album, nach Interpret, nach Genre, Jahr oder Album mit Interpret sortiert werden.

Über die Funktion „Config“ im Menüpunkt Library wird der User direkt zur Datenbankverwaltung in den Einstellungen weitergeleitet.

Zudem kann die Suchfunktion aufgerufen werden, mit welcher es möglich ist, die Bibliothek nach einzelnen Titeln zu durchsuchen. Die Suchfunktion bietet noch eine weitere Möglichkeit: Der „...“-Button mit den drei Punkten neben dem Suchfeld eröffnet die Möglichkeit, die gefundenen Titel in eine normale Playlist zu integrieren oder eine sogenannte Autoplaylist zu generieren. Eine Autoplaylist ist immer in Kontakt mit den realen Datenbeständen auf dem PC. Das bedeutet, erstellt man eine Autoplaylist und löscht bestimmte Dateien oder Musikdateien auf der Festplatte, dann werden diese auch automatisch in der Autoplaylist gelöscht. Wenn eine Autoplaylist von einem Künstler erzeugt wird, dann werden neue Titel in der Datenbank dieses Künstlers direkt in die Autoplaylist geschrieben.

13.3.9 Menüpunkt Help

Über die Menüfunktion Help kann *foobar2000* nach aktuellen Updates suchen oder leitet den User online auf die offizielle Homepage weiter, auf der aktuelle Fragen bzw. häufig gestellte Fragen in den FAQ geklärt werden.

13.3.10 Preferences

In den „Preferences“ lassen sich die Haupteinstellungen von *foobar2000* beeinflussen. Je nachdem, ob zusätzliche Plug-Ins installiert wurden, können hier die Veränderungen vorgenommen werden. Die Hauptbereiche (ohne zusätzliche Plug-Ins) sind recht übersichtlich und verständlich:

- In den Components erscheint eine Übersicht über alle installierten Plug-Ins und der vorhandenen Softwareversion.

- Im Punkt Display lassen sich alle Einstellungen zur Benutzeroberfläche kontrollieren. Treten neue Plug-Ins im Bereich der Oberfläche auf, dann tauchen diese ebenfalls hier auf. Zudem lassen sich hier verschiedene Themen und Layouteinstellungen speichern, ex- und importieren.
- Unter Keyboard Shortcuts können individuelle Tastenkürzel erstellt und gespeichert werden.
- Unter Media Library lässt sich die Musikdatenbank bzw. die Bibliothek verwalten.
- Im Bereich Networking lassen sich die Einstellungen für die Netzwerkübertragung auswählen und die „Buffer Size“ einstellen, damit eine störfreie Wiedergabe (z.B. bei WLAN Netzen) gewährleistet wird.
- Im Unterpunkt „Playback“ der Einstellungen lässt sich angeben, welche Soundtreiber *foobar2000* benutzen soll, normalerweise wählt *foobar2000* hier den Standard aus. Neben verschiedenen Treibern kann hier auch die Puffergröße eingestellt werden. Die Puffergröße ist ein kleiner Bereich des Arbeitsspeichers, in welcher die Audiofiles geladen werden, bevor diese abgespielt werden. Man spricht hier davon, dass die Daten (das Audio in diesem Fall) gepuffert werden. Eine Änderung des Puffers bewirkt eine Änderung der Latenz. Des Weiteren kann man verschiedene Attribute der Fade-in- bzw. Fade-out-Einstellungen bei Pause, Stopp usw. ändern. Neben den Output-Optionen können hier die DSP Einstellungen der Playbacks vorgenommen werden, das bedeutet dass *foobar2000* während der Wiedergabe den Klang mittels dieser Funktionen anpasst.

13.3.11 Metadaten

Unter Metadaten oder auch Tags genannt versteht man die Beschreibungen von Informationen zu einer Audiodatei. Die Metadaten sind für *foobar2000* wichtig, um viele hilfreiche Funktionen der Software zu nutzen. Metadaten werden u.a. für die Sortierung der Alben und Titel benötigt. Metadaten sind allerdings nur in bestimmten Formaten möglich. In einer WAV Datei lassen sich keine Metadaten abspeichern, da WAV ursprünglich nicht für die Archivierung großer Musiksammlungen gedacht war. Die in der Library auftauchende Funktion der „Albums List“ setzt eingetragene Metadaten voraus. Arbeitet man hier mit WAV Dateien sollte man im „View“-Bereich

die Sortierung nach der „Folder Structure“ anwenden, da keine Metadaten zu Interpretieren, Alben u.ä. verfügbar sind.

Arbeitet man mit Audiodaten wie MP3, die mit Metadaten umgehen können, kann man diese direkt in *foobar2000* ändern. Über Rechtsklick auf den Titel und auf „Proporities“ erscheinen die Metadaten, welche direkt per Tastatureingabe geändert werden können. Hier kann auch mit externen Programmen z.B. „mp3-Tag“ gearbeitet werden, bei denen dies teilweise noch komfortabler gelöst ist. Bei den Tags muss nicht jedes Feld ausgefüllt sein, es ist durchaus möglich, nur wenige Tags einzutragen. Leider gibt es keinen einheitlichen Standard für Tags, somit verwendet jedes Format seine individuellen Tags.

Foobar2000 erstellt gemeinsame Tags, die nur intern verwaltet werden. Dies bedeutet: Bei womöglich unterschiedlichen Tags im Metadatenfeld „Album“ wertet *Foobar2000* diese Informationen beim Lesen einer Datei intern aus und stellt sie immer auf die gleiche Weise als Feld „Album“ dar. Der Nutzer muss sich nicht darum kümmern, welches Audioformat und damit welcher Tag-Standard gerade genutzt wird. Diese Funktion erleichtert den Umgang mit unterschiedlichen Formaten erheblich.

Foobar2000 bietet die Möglichkeit fehlende Tags in die Daten einzutragen, ohne dies für jeden einzelnen Titel tun zu müssen: Der einfachste Weg ist wohl, die Tags direkt beim Rippen der CD aus einer Onlinedatenbank abzurufen und in die Audiodaten zu schreiben; natürlich lassen sich die Tags auch manuell eintragen. Eine weitere Möglichkeit ist es, die Tags aus den Dateinamen zu lesen, z.B. bei einer Beschriftung mit Titel-Album- Interpret-Jahr. Dies ist allerdings recht mühselig.

Neben den schriftlichen Metadaten lassen sich auch CD-Cover und Booklets hinzufügen. Diese werden eventuell beim Rippen direkt aus der Datenbank übernommen oder können über die Funktion „Tagging“ manuell eingefügt und verwaltet werden.

13.3.12 Dateiformatkonvertierung

Eine verlustfreie Speicherung der Musik von einer Audio-CD mit Metadaten lässt sich nur mit dem Dateiformat FLAC realisieren. Vielleicht möchte man auch seine komplette Musiksammlung in einem einheitlichen Format speichern. Eines der

beliebtesten Formate ist das alte und technisch nicht mehr aktuelle MP3-Format. Hier bietet *foobar2000* eine komfortable Lösung an: Voraussetzung für eine Konvertierung in ein anders File ist, dass alle Dateien sich in einer Playlist befinden. Es werden alle Titel in der Playlist markiert, die konvertiert werden sollen. Danach ruft man über Rechtsklick „Convert“ den passenden Menüpunkt dazu auf. Über „Quick Convert“ lassen sich sehr schnell die beliebtesten Formate auswählen und umwandeln. Zu jedem Format gibt es spezielle Einstellmöglichkeiten, die über die Funktion „Edit“ aufgerufen werden können. Hier lassen sich beispielsweise beim MP3-Format die Qualitätsstufe oder auch verlustfreie Kompressionslevel einstellen (z.B. bei FLAC). Zudem lassen sich Ditherwerte und eine Bittiefe festlegen. Mit einem Klick auf „Convert“ lässt sich der gewünschte Speicherplatz auswählen.

Es ist auch möglich sich hier sogenannte „Presets“ abzuspeichern. Klickt man nicht auf „Quick Convert“, sondern auf „...“, lassen sich Format, Zielordner, gewünschte DSP Effekte (z.B. EQ-Einstellungen oder Normalisierung) und weiteres festlegen und als „Preset“ wieder abrufen, die dann ggf. auch beim CD-Rippen direkt zur Verfügung stehen.

Mit den DSP-Effekten können bestimmte Konfigurationsmöglichkeiten genutzt werden, beispielsweise das Einbauen eines Limiters oder einer Mono-Stereo-Konvertierung.

Will man nun in das MP3-File konvertieren, fragt *foobar2000* nach dem sogenannten Lame-Code. Dies ist der benötigte Codec, um eine Umwandlung durchzuführen. Dieser lässt sich als exe-Datei im Internet herunterladen.

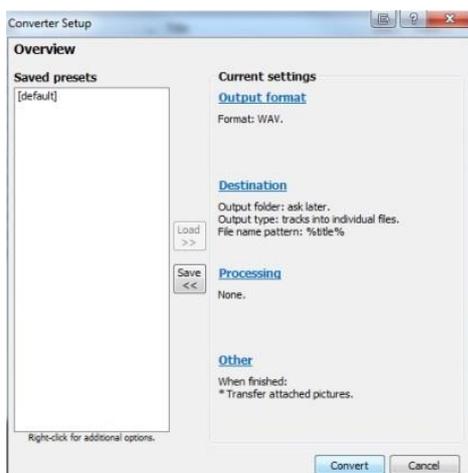


Abbildung III: Konvertieren in *foobar2000* ¹³⁴

¹³⁴ (Klein(Screenshot))

13.3.13 „ReplayGain“-Funktion

Eine hilfreiche Funktion von *foobar2000* ist die ReplayGain-Funktion. Diese bietet die Möglichkeit, die Lautstärke verschiedener Titel anzupassen, um Musik in einer einheitlichen Lautstärke genießen zu können. Alle Titel müssen eine Scanfunktion durchlaufen. Die Lautstärkeergebnisse werden in die Metadaten eingetragen.

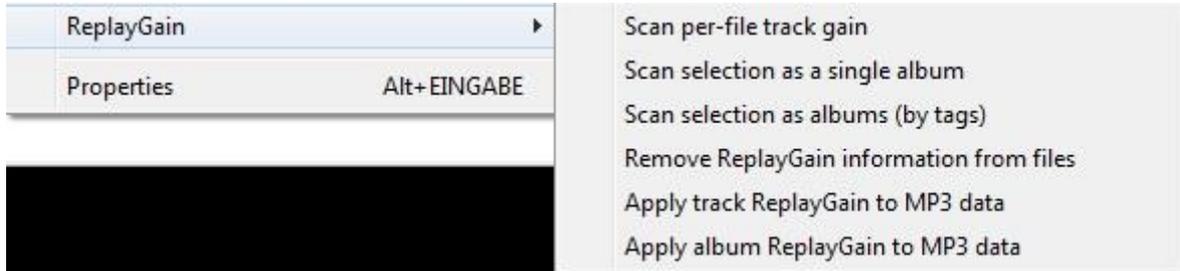


Abbildung JJJ: ReplayGain in *foobar2000* ¹³⁵

Die Replay-Funktion lässt sich nur in der Playlist aufrufen, somit müssen alle Daten, die miteinbezogen werden sollen, in der Playlist vorhanden sein. Mit Rechtsklick auf den Menüpunkt „ReplayGain“ lassen sich verschiedene Optionen auswählen:

| | |
|------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Scan per file track | Alle Daten werden einzeln gescannt. |
| Scan selection as single album | Alle Tracks der Playlist werden wie ein Album betrachtet. |
| Scan selection as albums (tags) | Es werden die track- und albumbasierten Informationen berechnet. Die Tracks erhalten alle den gleichen Pegel. Die Alben können sich voneinander unterscheiden. |
| Remove ReplayGain information from files | Alle berechneten Daten werden aus den Metadaten entfernt. |
| Apply track ReplayGain to mp3 data | Hier lassen sich die berechneten Track-Daten den Metadaten zufügen. |
| Apply album ReplayGain to mp3 data | Hier lassen sich die berechneten Album-Daten den Metadaten zufügen. |

¹³⁵ (Klein(Screenshot))

Allerdings muss beachtet werden, dass foobar2000 sich hier nicht mit der „Loudness“ beschäftigt, sondern ausschließlich mit den Pegeln zueinander.

13.3.14 Zusätzliche Plug-Ins in *foobar2000*

Plug-Ins dienen zur Erweiterung für eine komfortablere Bedienung von *foobar2000*. Plug-Ins werden sowohl auf der offiziellen Homepage (Components) von *foobar2000* angeboten, als auch von Drittanbietern im Internet. Die Plug-Ins werden als dll-Dateien im Unterordner von *foobar2000* (standardmäßig C:\Programme\Foobar2000\components) abgelegt.

Unter File/Preferences/Components sieht der User, welche Plug-Ins installiert bzw. in Verwendung sind. Allein auf der offiziellen Seite von *foobar2000* sind hunderte von Plug-Ins unter anderem verschiedene Oberflächeneinstellungen, CD-Brennfunktionen usw. zu finden, daher sollte jeder selbst entscheiden, was sinnvoll und nötig ist.

13.4 Audioausschnitte im MUSHRA-Hörversuch

| Track | Inhalt | Bemerkung | Referenz | Stimuli 1 | Stimuli 2 | Stimuli 3 | Stimuli 4 | Stimuli 5 | Stimuli 6 | Stimuli 7 |
|-------------|----------------------------|--------------------------|--------------|--------------|--------------|--------------|---------------|---------------|---------------|------------|
| Original-CD | | | | | | | | | | **** |
| 32 | Triangles | Einzelinstrument Stereo | WAV-original | 16 Bit - MP3 | 32 Bit - MP3 | 64 Bit - MP3 | 128 Bit - MP3 | 256 Bit - MP3 | 320 Bit - MP3 | WAV-anchor |
| 35 | Glockenspiel | Einzelinstrument Stereo | WAV-original | 16 Bit - MP3 | 32 Bit - MP3 | 64 Bit - MP3 | 128 Bit - MP3 | 256 Bit - MP3 | 320 Bit - MP3 | WAV-anchor |
| 39 | Grand Piano | Einzelinstrument Stereo | WAV-original | 16 Bit - MP3 | 32 Bit - MP3 | 64 Bit - MP3 | 128 Bit - MP3 | 256 Bit - MP3 | 320 Bit - MP3 | WAV-anchor |
| 60 | Piano - Schubert | Einzelinstrument Stereo | WAV-original | 16 Bit - MP3 | 32 Bit - MP3 | 64 Bit - MP3 | 128 Bit - MP3 | 256 Bit - MP3 | 320 Bit - MP3 | WAV-anchor |
| 62 | Orchestra - Strauss | Also sprach Zarathustra* | WAV-original | 16 Bit - MP3 | 32 Bit - MP3 | 64 Bit - MP3 | 128 Bit - MP3 | 256 Bit - MP3 | 320 Bit - MP3 | WAV-anchor |
| 66 | Wind Ensemble - Stravinsky | Le Scare du Printemps** | WAV-original | 16 Bit - MP3 | 32 Bit - MP3 | 64 Bit - MP3 | 128 Bit - MP3 | 256 Bit - MP3 | 320 Bit - MP3 | WAV-anchor |
| 69 | ABBA | The Visitors*** | WAV-original | 16 Bit - MP3 | 32 Bit - MP3 | 64 Bit - MP3 | 128 Bit - MP3 | 256 Bit - MP3 | 320 Bit - MP3 | WAV-anchor |

* Seiji Ozawa - Strauss: Also sprach Zarathustra; CD Philips No. 400 072-2, Track 1, 0:32 ... 0:58

** Charles Dutoit - Stravinsky: Le Scare du Printemps; CD Decca No. 414 202-2, Track 3, 5:20 ... 0:32

*** ABBA - The Visitors (Head Over Heels); CD Polydor No. 800 011-2, Track 2, 0:10 ... 0:34

**** Highpass bei 3.5 kHz

14. Abkürzungsverzeichnis

| | | |
|--------|---|---------------------------------------------------------|
| dB | = | Dezibel |
| mp3 | = | MPEG-1 / Layer 3 |
| z.B. | = | zum Beispiel |
| u.a. | = | unter anderem |
| PC | = | Personal Computer |
| DSP | = | Digitaler Signalprozessor |
| EQ | = | Equalizer |
| FAQ | = | Frequently Asked Questions |
| STEP | = | Subjective Test and Evaluation Program |
| ITU | = | International Telecommunication Union |
| s | = | Sekunden |
| EBU | = | European Broadcasting Union |
| CD | = | Compact Disc |
| AD | = | Analog => Digital |
| DA | = | Digital => Analog |
| MUSHRA | = | Multi Stimulus test with Hidden Reference and Anchor |
| MAC | = | Macintosh-Computer |
| GUI | = | Graphical User Interface |

15. **Abbildungsverzeichnis**

Die Quellenangaben der Abbildungen sind jeweils direkt bei den Abbildungen dokumentiert.

| | |
|-----------------------------------------------------------------------------|--------|
| Abbildung A: Das menschliche Hörorgan..... | - 3 - |
| Abbildung B: Das Innenohr | - 5 - |
| Abbildung C: Kurven gleicher Lautstärke | - 6 - |
| Abbildung D: Lautstärkeempfinden verschiedener Impulslängen | - 8 - |
| Abbildung E: Testtöne im schmalbandigen Rauschen | - 9 - |
| Abbildung F: MUSHRA Oberfläche | - 17 - |
| Abbildung G: Bewertungsparameter für die Absolutbewertung | - 25 - |
| Abbildung H: ABX-Test mit <i>foobar2000</i> | - 26 - |
| Abbildung I: Verteilung bei kleinen Stichproben | - 31 - |
| Abbildung J: Binomialverteilung | - 32 - |
| Abbildung K: Summenverteilung | - 33 - |
| Abbildung L: ABX-Test in <i>foobar2000</i> aufrufen | - 35 - |
| Abbildung M: ABX-Komperator in <i>foobar2000</i> mit Erklärungen | - 35 - |
| Abbildung N: <i>STEP</i> -Oberfläche | - 37 - |
| Abbildung O: Optionenauswahl in verschiedenen Sessions in <i>STEP</i> | - 39 - |
| Abbildung P: MUSHRA-Sessionfile in <i>STEP</i> | - 39 - |
| Abbildung Q: AB-Test in <i>STEP</i> | - 40 - |
| Abbildung R: ABX-Test in <i>STEP</i> | - 40 - |
| Abbildung S: ABC/HR (BS. 1116-1) in <i>STEP</i> | - 41 - |
| Abbildung T: MUSHRA in <i>STEP</i> | - 42 - |
| Abbildung U: <i>MUSHRATest</i> – Dateien | - 44 - |
| Abbildung V: Tests erstellen in <i>MUSHRATest</i> | - 45 - |
| Abbildung W: Ordnerstruktur <i>MUSHRATest</i> | - 45 - |
| Abbildung X: Ordnerstruktur 2 in <i>MUSHRATest</i> | - 46 - |
| Abbildung Y: Training Part A in <i>MUSHRATest</i> | - 46 - |
| Abbildung Z: <i>MUSHRATest</i> Training Part B | - 47 - |
| Abbildung AA: Hörtest mit <i>MUSHRATest</i> | - 48 - |
| Abbildung BB: Datendokumentation in <i>MUSHRATest</i> | - 49 - |
| Abbildung CC: ABX-Test mit <i>WinABX</i> | - 51 - |

| | |
|--------------------------------------------------------------------------------------|--------|
| Abbildung DD: ABC/HR-Test mit <i>ABC Hidden Reference-Audio Comparison Tool</i> ... | - 52 - |
| Abbildung EE: ABX-Test mit dem <i>ABC Hidden Reference Audio Comparsion Tool</i> ... | - 53 - |
| Abbildung FF: <i>scale</i> Startoberfläche | - 53 - |
| Abbildung GG: Testerstellung in <i>scale</i> | - 54 - |
| Abbildung HH: Testdurchführung in <i>scale</i> | - 55 - |
| Abbildung II: Testanalyse in <i>scale</i> | - 56 - |
| Abbildung JJ: Analoges Eingangssignal | - 57 - |
| Abbildung KK: <i>MUSHRATest</i> Training Part A | - 63 - |
| Abbildung LL: <i>MUSHRATest</i> -Hörtest | - 64 - |
| Abbildung MM: Grundschemata eines Box-Plots | - 65 - |
| Abbildung NN: Box-Plot der Qualitätsstufen über alle Hörbeispiele | - 66 - |
| Abbildung OO: Auswertungstabelle aller Qualitätsstufen über alle Hörbeispiele .. | - 66 - |
| Abbildung PP: Mittelwerte der Qualitätsstufen über alle Hörbeispiele | - 67 - |
| Abbildung QQ: t-Test CD-Referenz und MP3 mit 320 MBit/s | - 68 - |
| Abbildung RR: t-Test CD-Referenz und MP3 mit 256 MBit/s | - 68 - |
| Abbildung SS: t-Test CD-Referenz und MP3 mit 128 MBit/s | - 69 - |
| Abbildung TT: Mittelwerte bei MP3-Dateien mit 320MBit/s | - 70 - |
| Abbildung UU: Mittelwerte bei MP3-Dateien mit 64 MBit/s | - 70 - |
| Abbildung VV: Mittelwerte bei MP3-Dateien mit 32MBit/s | - 70 - |
| Abbildung WW: Track Index SQAM CD | VI |
| Abbildung XX: Menüleiste <i>foobar2000</i> | VIII |
| Abbildung YY: Ansicht anpassen <i>foobar2000</i> | IX |
| Abbildung ZZ: Playlists in <i>foobar2000</i> | X |
| Abbildung AAA: Menüpunkt File in <i>foobar2000</i> | XI |
| Abbildung BBB: Menüpunkt Edit in <i>foobar2000</i> | XI |
| Abbildung CCC: Menüpunkt View in <i>foobar2000</i> | XII |
| Abbildung DDD: Tabelle der Visulationen von <i>foobar2000</i> | XIII |
| Abbildung EEE: Playlistmanager in <i>foobar2000</i> | XIII |
| Abbildung FFF: Vorgefertigte Layouts in <i>foobar2000</i> | XIV |
| Abbildung GGG: Menüpunkt Playback in <i>foobar2000</i> | XV |
| Abbildung HHH: Randomfunktionen in <i>foobar2000</i> | XV |
| Abbildung III: Konvertieren in <i>foobar2000</i> | XIX |
| Abbildung JJJ: ReplayGain in <i>foobar2000</i> | XX |

16. Literaturverzeichnis und Webhinweise

3253, EBU-TECH. September 2008. *Sound Quality Assessment Material recordings for subjective tests.* Geneva : EBU, September 2008.

Audio Research Labs, Manual. September 2010. *STEP Manual.* s.l. : STEP, September September 2010.

Bosch, Karl. 2002. *Statistik Wahrheit und Lüge.* München : Oldenburg Wissenschaftsverlag, 2002. 3-486-25986-5.

Curdt, Prof. Oliver. SS12. *Menschliches Gehör. HdM Sounddesign Präsentation.* Stuttgart : s.n., SS12.

Dickreiter, Michael. 1997. *Handbuch der Tonstudioteknik Band 1.* München : Saur, 1997.

— **2008.** *Handbuch der Tonstudioteknik Band 2; 7.Auflage.* München : K G Saur, 2008. 978-3-598-11765-7.

Euerle, Wolfgang. hören heute online. *hören heute online.* [Online] [Zitat vom: 04. Juni 2013.] <http://www hoeren-heute.de/ gehoer.htm>.

Feiten, Bernhard und Hoeg, Wolfgang. 1998. *Internationaler Hörtest der Mehrkanal-Systeme MPEG-2 Layer II und Dolby AC-3.* München : K.G. Sauer, 1998. 3-598--20361-6.

Fitch, Robert K. 2012. *WinSTAT.* [Bedienungsanleitung] Microsoft Excel : s.n., 2012.

foobar2000. foobar2000. *foobar2000.* [Online] [Zitat vom: 28. Juni 2013.] <http://www.foobar2000.org/>.

Geers. Geers Hörgeräte. [Online] [Zitat vom: 04. Juni 2013.] <http://www.geers.de/hoertest/so-hoert-der-mensch/>.

Gontard, Dres. Andrea & Maximilian. Hoersturz.de. *Hoersturz.de*. [Online] [Zitat vom: 04. Juni 2013.] <http://www.hoersturz.de/ohr/aeusseres-ohr.htm>.

Heß, Dr. med Heike. HNO Praxisteam. *HNO Praxisteam*. [Online] [Zitat vom: 04. Juni 2013.] <http://www.hno-praxisteam.de/html/horvorgang.html>.

HP. <http://www.hp.com>. *Hewlett Packard*. [Online] [Zitat vom: 07. Juli 2013.] http://www8.hp.com/de/de/hp-news/press-release.html?id=588635&jumpid=reg_r1002_dede_c-001_title_r0001#.Ue0cSqwzKvM.

<http://www.kikeg.arrakis.es>. KIKEG. [Online] [Zitat vom: 22. Juni 2013.] <http://www.kikeg.arrakis.es/winabx/winabx.zip>.

ITU. ITU. *ITU*. [Online] [Zitat vom: 04. Juli 2013.] <http://www.itu.int/en/wtisd/Pages/about.aspx>.

ITU-R BS.1116-1. 1994-1997. *Methods for the subjective assessment of small impairments in audio systems including multichannel sound systems*. <http://www.itu.int/rec/R-REC-BS.1116/en> : ITU, 1994-1997.

ITU-R BS.1283-1. 1997-2003. *A guide to ITU-R Recommendations for subjective assessment of sound quality*. <http://www.itu.int/rec/R-REC-BS.1283/en> : ITU, 1997-2003.

ITU-R BS.1284. 1997-2003. *General methods for the subjective assessment of sound quality*. <http://www.itu.int/rec/R-REC-BS.1284/en> : ITU, 1997-2003.

ITU-R BS.1534. 2001-2003. *Method for the subjective assessment of intermediate quality level of coding systems*. <http://www.itu.int/rec/R-REC-BS.1234/en> : ITU, 2001-2003.

Kappes, Andre. Die Audiokosierung mp3. *Die Audiokosierung mp3*. [Online] [Zitat vom: 13. Juni 2013.] <http://www.mp3encoding.de/.3772313604>.

Klein- Excelscreenshot, Tobias. 2013. 2013.

Klein(Screenshot), Tobias. *Screenshot.* PC Tobias Klein : s.n.

Kuttruff, Heinrich. 2004. *Akustik - Eine Einführung.* Stuttgart : s. Hirzel Verlag, 2004.
ISBN 3-7776-1244-8.

Oestreich, Markus und Romberg, Oliver. 2012. *Keine Panik vor Statistik.*
Wiesbaden : Springer Fachmedien, 2012. 987-3-8348-1946-8.

Pörschmann, Christoph. Fachhochschule Köln. *Cologne University of Applied
Acienes.* [Online] [Zitat vom: 25. Juli 2013.] <http://www.audiogroup.web.fh-koeln.de/>.

Rudolph, Thomas-Michael. TMR Elektronik GmbH . *TMR Elektronik GmbH.* [Online]
tmr. [Zitat vom: 06. Juni 2013.] [http://www.tmr-
audio.de/homemain/wissenswertes-ueber-hifi/faq/das-hoeren/145-
verdeckungseffekt](http://www.tmr-audio.de/homemain/wissenswertes-ueber-hifi/faq/das-hoeren/145-verdeckungseffekt).

Rumsey, Deborah. 2010. *Statistik für Dummies.* Weinheim : WILEY-VCH Verlag
GmbH & Co. KGaA, 2010. 987-3-527-70594-8.

Schlittgen, Rainer. 2012. *Einführung in die Statistik.* München : Oldenbourg
Wissenschaftsverlag GmbH, 2012. 978-3-486-71524-8.

Softpedia. Softpedia. <http://www.softpedia.com/>. [Online] [Zitat vom: 28. Juli 2013.]
[http://www.softpedia.com/get/Multimedia/Audio/Other-AUDIO-Tools/ABC-
Hidden-Reference.shtml](http://www.softpedia.com/get/Multimedia/Audio/Other-AUDIO-Tools/ABC-
Hidden-Reference.shtml).

Statistical Analysis of ABX Results Using Signal Detetion Theory. **Boley , Jon und
Lester, Michael. 2009.** New York : Audio AES, 2009.

Thomann. www.thomann.de. *Thomann.* [Online] [Zitat vom: 22. Juli 2013.]
http://www.thomann.de/de/akg_k271_mkii.htm.

Webers, Johannes. 2007. *Handbuch der Tonstudioteknik.* Poing : Franzis, 2007.

Zwicker, Hugo Fastl und Ebherhard. 1990, 1999, 2007. *Psychoacoustics.* Berlin Heidelberg : Springer-Verlag, 1990, 1999, 2007. 3-540-23159-5.

17. Erklärung

Hiermit erkläre ich, dass ich die vorliegende Bachelorarbeit selbständig angefertigt habe. Es wurden nur die in der Arbeit ausdrücklich benannten Quellen und Hilfsmittel benutzt. Wörtlich oder sinngemäß übernommenes Gedankengut habe ich als solches kenntlich gemacht.

Ort, Datum

Unterschrift