

---

VERMESSUNG DER  
GESAMTQUALITÄT VON SPRACH-  
UND AUDIOKODIERERN

BACHELOR ARBEIT

vorgelegt von

Falk Schiffner

STUDIENGANG "HÖRAKUSTIK"

25. JUNI 2010

## BETREUUNG

FH-LÜBECK: ..... PROF. DR. JÜRGEN TCHORZ  
T-LABS: ..... PROF. DR.-ING. ALEXANDER RAAKE  
T-LABS: ..... DIPL.-ING. MARCEL WÄLTERMANN

UNIVERSITY OF APPLIED SCIENCE LÜBECK  
Fachbereich „Angewandte Naturwissenschaften“  
Mönkhofer Weg 239  
D-23562 Lübeck

TECHNISCHE UNIVERSITÄT BERLIN  
An-Institut Deutsche Telekom Laboratories  
Fachgebiet „Assessment of IP-based Applications“  
Ernst-Reuter-Platz 7  
D-10587 Berlin

FALK SCHIFFNER  
Matrikel - Nr: 154972  
Dulonstr. 2  
D-16515 Oranienburg

# Inhaltsverzeichnis

<b>1 Kurzfassung</b>	<b>3</b>
<b>2 Einleitung</b>	<b>4</b>
2.1 Motivation und Zielsetzung . . . . .	4
2.2 Strukturierung der Arbeit . . . . .	6
<b>3 Grundlagen</b>	<b>7</b>
3.1 Qualitätsbewertung . . . . .	7
3.1.1 Subjektive Qualitätsbewertung . . . . .	7
3.1.2 Instrumentelle Qualitätsbewertung . . . . .	10
3.2 Testverfahren ACR und MUSHRA . . . . .	11
<b>4 Experimente</b>	<b>13</b>
4.1 Vorbereitung der Experimente . . . . .	13
4.1.1 Konditionen . . . . .	17
4.1.2 Generieren der Sprachdateien . . . . .	19
4.1.3 Zusammenfassung der Konditionen . . . . .	21
4.2 Durchführung . . . . .	23
<b>5 Ergebnisse</b>	<b>26</b>
5.1 Dokumentation der Ergebnisse ACR und MUSHRA . . . . .	26
5.2 Auswertung der Ergebnisse . . . . .	30
5.2.1 Betrachtung der Qualitätsbewertungen . . . . .	30
5.2.2 Qualitätsverbesserung und E-Modell . . . . .	33
5.3 Vergleich ACR und MUSHRA . . . . .	36
<b>6 Zusammenfassung</b>	<b>40</b>
<b>Abbildungsverzeichnis</b>	<b>43</b>
<b>Tabellenverzeichnis</b>	<b>44</b>

<b>Abkürzungsverzeichnis</b>	<b>45</b>
<b>Literaturverzeichnis</b>	<b>47</b>
<b>A Testeinweisungen</b>	<b>50</b>
A.1 ACR . . . . .	51
A.2 MUSHRA . . . . .	53
<b>B Batchfile</b>	<b>55</b>
<b>C Playlisten</b>	<b>59</b>
C.1 ACR . . . . .	60
C.2 MUSHRA . . . . .	62
<b>D Log - Files</b>	<b>66</b>
D.1 ACR . . . . .	67
D.2 MUSHRA . . . . .	69

# **Kapitel 1**

## **Kurzfassung**

Das Ziel dieser Arbeit ist das Erfassen der subjektiven Qualität von verschiedenen Sprach- und Audiokodierern, sowie der Vergleich zweier Testverfahren. Die verwendeten Kodierer kommen bei der Festnetztelephonie, beim Mobilfunk und bei Internetdiensten zur Anwendung. Der Schwerpunkt liegt dabei auf der Übertragung von Sprache.

Grundlagen zur Qualität und zum Bewertungsprozess werden erläutert. Ebenso werden die Beurteilungsverfahren und ihre Durchführung beschrieben.

Die Qualitätsurteile wurden in Hörversuchen ermittelt. Es wurden die zwei am häufigsten verwendeten Qualitätserfassungsmethoden genutzt. Das gewählte Testdesign zielte auf eine Vergleichbarkeit mit vorherigen Untersuchungen ab. Aus den Ergebnissen lassen sich die Einflüsse von auftretenden Übertragungsfehlern, wie Paketverlusten, Hintergrundrauschen oder Frequenzgangbeschränkungen ableiten. Zusätzlich werden die beiden Verfahren hinsichtlich ihrer Ergebnisse und ihrer Durchführbarkeit verglichen.

# **Kapitel 2**

## **Einleitung**

### **2.1 Motivation und Zielsetzung**

In den letzten zwei Jahrzehnten hat sich die Informationsgesellschaft verändert. Heutzutage sind Informationen immer und zu jeder Zeit verfügbar. Die dafür verwendete Technologie schreitet immer weiter voran. Als kleines Beispiel seien hier Smartphones erwähnt. Mit diesen ist es unter anderem möglich, jeder Zeit im Internet zu surfen oder Fotos direkt über das Internet zu verbreiten. Im Gegensatz zur Zeit der Erfindung des Telefons, in der es eine große technische Leistung war die menschliche Sprache zu übertragen, ist heute Kommunikationstechnologie im Alltag eingebunden. Mittlerweile besitzt fast Jeder ein Mobiltelefon. In großen Unternehmen sind Dinge wie Video- und Telefonkonferenzen Standard und kaum noch wegzudenken. So- gar in einer Live-Fernsehübertragung einer Landtagswahl wurde jüngst der Internetdienst „Skype“ als Direktschaltung benutzt.

Die Kommunikationstechnologie, von der hier gesprochen wird, besteht nicht nur aus rein physischen Produkten, sondern auch aus Kommunikationsdiensten. Bei der Fülle von Anbietern, die zum einen die Technologie bereitstellen, zum anderen die Dienste, spielt die Qualität eine gewichtige Rolle. Zu diesen Diensten gehört nicht nur das Telefonieren über den heimischen Festnetzanschluss oder mit dem Handy, sondern auch die Internettelefonie. Mittlerweile sind die Grenzen zwischen den Diensten nicht mehr genau auszumachen. So werden über das Mobiltelefon Videos von Internetplattformen geschaut und über den heimischen PC telefoniert.

Es befassen sich viele Forschungszentren und Studien mit der Frage der Signalqualität. So wird viel Energie und Geld in die Verbesserung dieser

Dienste investiert. Es klingt einleuchtend, dass eine gut klingende und fehlerfreie Übertragung vom Nutzer bevorzugt wird. Die Qualität einer Übertragung, unter anderen Aspekten, ist maßgebend für den Erfolg auf dem Markt verantwortlich. Die durchlässigen Grenzen in der modernen Kommunikationstechnologie führen auch zu einer Verschmelzung der unterschiedlichen Bereiche. Bei der Planung eines Dienstes ist es zum Beispiel nicht mehr nur von Interesse, über das Handy Sprache gut zu übertragen, sondern auch Musik und Videos. Die Wahl der Kodierungsart spielt dabei eine Rolle, denn je nach verwendetem Kodierer kann das Ergebnis eines Dienstes mehr oder weniger zufriedenstellend sein. Neben diversen technischen Aspekten, sind auch Nutzerbewertungen nötig, damit ein Entwickler die Übertragungsqualität besser einschätzen kann und dem Nutzer die bestmögliche Qualität zur Verfügung stellt. Gerade durch die Überschneidung der Bereiche Video-Audio- und Sprachübertragung, ist es notwendig, die Ergebnisse aus den unterschiedlichen „Lagern“ zu vergleichen.

Das Ziel dieser Arbeit ist das Erfassen der subjektiven Qualität von verschiedenen Sprach- und Audiokodierern. Dafür wird ein Vergleich zweier Methoden, die zur Ermittlung von Gesamtqualität eingesetzt werden, durchgeführt. Es handelt sich hierbei um einen ACR-Test und um einen MUSHRA-Test. Beide Verfahren werden im Abschnitt 3.2 genauer betrachtet. Der „5 Punkt ACR-Test“ mit seiner MOS-Skala wird typischerweise im Zusammenhang mit Sprachübertragungen eingesetzt, hingegen wird der MUSHRA-Test eher im Audio-Kontext eingesetzt. Es werden in dieser Studie verschiedene qualitätsmindernde Beeinträchtigungen untersucht. Dabei werden auch neuere Kodierungsverfahren, sowie eine Vergrößerung der Audiobandbreite berücksichtigt. Bei den neuen Kodierungsarten handelt es sich um Kodierer, die für Echtzeit-Sprachanwendungen entwickelt wurden. Es stellen sich Fragen:

- *Verändert sich die Beurteilung der Gesamtqualität, wenn das eine oder andere Verfahren verwendet wird?*
- *Eignet sich eines der Verfahren besser, um die Gesamtqualität zu ermitteln?*
- *Kann man die Spanne der Bedingungen überhaupt sinnvoll abbilden?*
- *Verbessert sich die Gesamtqualität mit einer Vergrößerung der Übertragungsbandbreite?*

Die Ergebnisse dieser Tests werden aus Hörversuchen ermittelt. Die Arbeit enthält die Planung und Durchführung, sowie die Analyse der oben erwähnten Methoden; ebenso eine Interpretation der Ergebnisse.

## 2.2 Strukturierung der Arbeit

Diese Arbeit unterteilt sich in 6 Kapitel.

Kapitel 1 enthält eine Kurzfassung der Arbeit, welche einen Abriss über die Arbeit darstellt. Kapitel 2 beinhaltet die Einleitung. Sie beschreibt die Motivation, sowie die Aufgabenstellung. In Kapitel 3 wird ein Einblick in die theoretischen Grundlagen gegeben. Dabei handelt es sich zum einen um die Prozesskette bei subjektiver Bewertung von Qualität, zum anderen wird dort das E-Modell vorgestellt. Dieses Modell dient zur Berechnung der Sprachqualität in einem Übertragungssystem. Es wird zur Planung und Simulation von Netzen verwendet und ermöglicht einen Vergleich verschiedenen Übertragungen. Außerdem werden in diesem Kapitel auch die verwendeten Methoden zur Qualitätsbeurteilung beschrieben. Abschließend wird noch ein Einblick in voran gegangene Studien und ihrer Ergebnisse gewährt. In Kapitel 4 werden die aktuellen Hörversuche beschrieben. Dabei wird auf die Versuchsplanung sowie auf deren Durchführung eingegangen. Es beschreibt die verwendeten Konditionen, Versuchspersonen, Umgebungsbedingungen und das verwendete Equipment. Kapitel 5 befasst sich mit der Dokumentation und Auswertung der Ergebnisse. Nach der Betrachtung der Ergebnisse, werden beide Versuche zusammengeführt und verglichen. Im letzten Kapitel folgt die Zufassung und Schlussfolgerung der Arbeit. Des Weiteren wird ein Ausblick auf zukünftige Aufgaben gegeben.

# Kapitel 3

## Grundlagen

### 3.1 Qualitätsbewertung

Möchte man sich, im ersten Schritt dem Begriff der Qualität nähern, ist die Definition für Qualität von Jekosch [19] behilflich. Diese besagt sinngemäß:

*Eine Qualitätsbewertung ist das Ergebnis eines Vergleichs zwischen den wahrgenommenen und den gewünschten Eigenschaften eines Objekts.*

Dabei muss das Objekt nicht zwangsläufig ein physisches Objekt sein. Ebenso wird ein Dienst / eine Dienstleistung zum Objekt, wenn man ihm Eigenschaften zuordnen kann. Diese Eigenschaften müssen qualitätsrelevante und erfassbare Merkmale des Objektes sein. Um eine Aussage zur Qualität zu machen, werden die Merkmale beurteilt. Dabei spielt die Erwartung eine entscheidende Rolle. Als Beispiel sollen an dieser Stelle Filmaufzeichnungen dienen. Schaut man heute eine VHS-Aufzeichnung, so empfindet man die Qualität des Videomaterials als schlecht, da heutzutage die Bild- und Tonqualität einer DVD erwartet wird. Hingegen wurde damals die Qualität der VHS als hervorragend empfunden, da sie dem technischen Stand entsprach.

#### 3.1.1 Subjektive Qualitätsbewertung

Die Bewertung der Qualität ist ein komplexer Vorgang. Um Aussagen über die Qualität zu erhalten, wird eine gewisse Anzahl von Proben einer Bewertungsinstanz zur Beurteilung gegeben. Im speziellen Fall der Qualität von Sprachsignalen, bewertet in erster Linie der Mensch. Die Bewertungen sind rein subjektiver Natur und erzielen die zuverlässigsten Ergebnisse. Es gibt

unterschiedliche Methoden um Qualitätsurteile zu erhalten z.B.: Verständlichkeitstest, Konversationstest, reine Hörtests (Listening only Test (LOT)). Die in der Studie verwendeten Methoden gehören zur Gruppe der LOT.

Bei der Qualitätsbeurteilung spielen eine Vielzahl von Faktoren eine Rolle. Der gesamte Beurteilungsprozess kann dabei gemäß Raake [24] und auf der Grundlage von Jekosch [19] in aufeinander folgende Schritte (1-4) aufgeteilt werden.

0. Schallereignis: Ist ein physikalisch-akustischer Vorgang, der durch seine physikalischen Eigenschaften bestimmt wird. Ein Schallereignis ist objektiv und existiert somit unabhängig von der subjektiven Wahrnehmungen.
1. Wahrnehmung: Ein Schallereignis wird über das Gehör aufgenommen (siehe Abbildung 3.1 auf Seite 9) und wird in diesem Moment zum Hörrereignis. Es enthält eine Fülle von physikalischen Informationen (z.B. Lautstärke, Frequenzgehalt) aber auch Informationen der Senderseite (z.B. Ort, Stimmung, Geschlecht des Sprechers o.ä.). Zusätzlich zur Aufnahme dieser Informationen spielen Stimmung, Erfahrung, Situation usw. des Empfängers eine Rolle bei der Wahrnehmung. So wird zum Beispiel ein identisches Signal je nach Verfassung unterschiedlich wahrgenommen<sup>1</sup>. All diese Faktoren passen die Wahrnehmung des Gehörten immer wieder neu an.
2. Betrachtung: Der Empfänger betrachtet alle Faktoren des Gehörten. Dabei werden die qualitätsrelevanten Merkmale betrachtet. Die Faktoren werden zu Dimensionen zusammengefasst (z.B. Rauschhaftigkeit, Klangfarbe)<sup>2</sup>. Dieser Prozess geschieht unterbewusst. Die bewerteten Eigenschaften  $\vec{m}$  kann man in einem multidimensionalen Raum abbilden. Gleichzeitig werden in diesem Raum auch die erwarteten Eigenschaften  $\vec{n}$  abgebildet. Jede Raumachse bildet eine Qualitätsdimension ab.
3. Vergleich: Nun werden die Werte jeder Dimension  $\vec{m}$  und  $\vec{n}$  intern

---

<sup>1</sup>Schulz von Thun [25] hat dem Sender einer Information „vier Ebenen“ und dem Empfänger „vier Ebenen“ zugeordnet (Sachebene, Selbstkundgabe, Beziehungsebene, Apellebene). Die Art des Gespräches hängt davon ab, in welcher Weise diese zusammenspielen.

<sup>2</sup>In anderen Untersuchungen wurden drei übergeordnete Qualitätsdimensionen von Wältermann [21] gefunden (Frequenzgehalt / Rauschhaftigkeit / Kontinuität). Jede dieser übergeordneten Dimensionen setzt sich aus einer Anzahl von Unterdimensionen zusammen. Wie im Beispiel der Dimension: Kontinuität [20] nachzulesen.

verglichen. Der Abstand der Werte bildet die Grundlage für den nachfolgenden Schritt.

4. Beurteilung: Der Empfänger nimmt die jeweiligen Abstände der einzelnen Dimensionen und fügt sie zu einer Bewertung zusammen. Die Abstände gehen je nach Größe unterschiedlich gewichtet in die Beurteilung ein und bilden die Gesamtqualität.

Die Gesamtqualität eines Telekommunikationsnetzwerkes ist von entscheidender Rolle, da der Nutzer im alltäglichen Umfeld nicht den Frequenzgehalt oder die Kontinuität einer Übertragung beurteilt, sondern ihre Gesamterscheinung. In dieser Arbeit wird sich nur mit der Gesamtqualität von Sprachsignalen beschäftigt und weiterhin als Qualität benannt. Der Einfluss verschiedener Qualitätsdimensionen wird nicht explizit betrachtet. Für die Sprachübertragung ist nur die akustische Beschaffenheit des Signals und nicht der Inhalt von Bedeutung, da nur sie auf dem Übertragungsweg von Störungen (Verzerrungen, Paketverlust, Rauschen, Klangveränderung) betroffen ist. Im Rahmen einer Sprachsignalübertragung kann auch nur darauf Einfluss genommen werden.

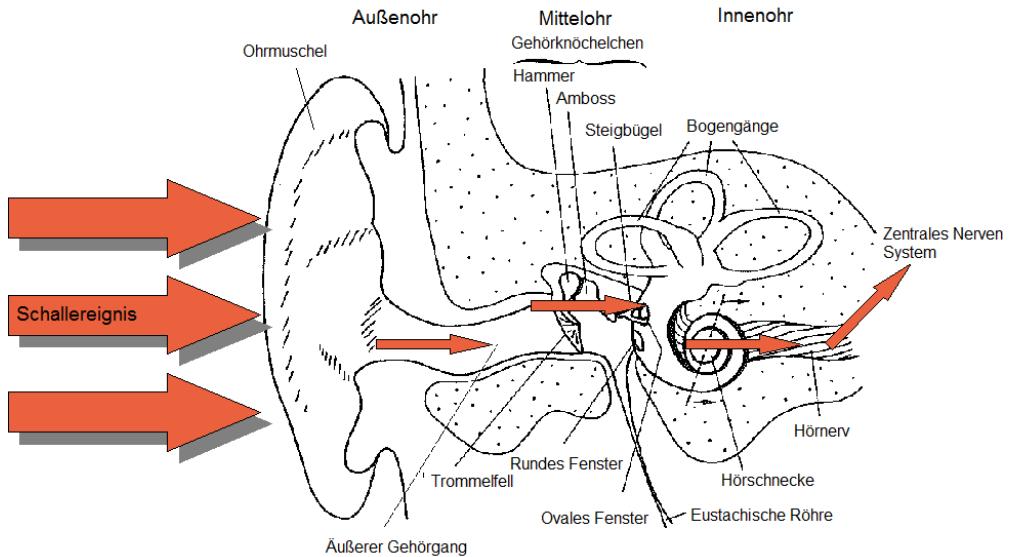


Abbildung 3.1: Weg der Schallaufnahme beim Menschen (Grundlage der Abbildung aus [31])

### 3.1.2 Instrumentelle Qualitätsbewertung

Subjektive Studien zur Qualitätsbewertung sind zeit- und kostintensiv. Es wurden deshalb Modelle entwickelt, die die wahrgenommene Qualität schätzen. Diese instrumentellen Verfahren ermitteln die Sprachqualität aus mess-technisch erfassbaren Werten. Dabei versuchen sie die menschliche Wahrnehmung zu simulieren. Eines dieser Verfahren ist das E-Modell.

Das E-Modell [9] ist ein Verfahren um die Gesamtqualität von Telefonnetzwerken zu beurteilen. Eine Beschreibung des E-Modells ist z.B. bei Möller [22] zu finden. Es wird primär als Planungswerkzeug eingesetzt und wurde von der ITU-T standardisiert. Im E-Modell werden qualitätsbeeinträchtigenden Faktoren addiert. Dabei erhält man einen einzahligen R-Wert, der die Qualität angibt. Gleichung 3.1 zeigt, wie sich der R-Wert zusammensetzt.  $R_0$  beschreibt dabei den Signal-Rausch-Abstand und beinhaltet Leitungs- sowie Hintergrundrauschen.  $I_i$  sind die Beeinträchtigungsfaktoren und A steht für einen Vorteilsfaktor. Dieser Vorteilsfaktor hängt mit der genutzten Technologie zusammen. Zum Beispiel kann in der Bewertung der Vorteil der Mobilität eines schnurlosen Telefons den Nachteil einer schlechteren Übertragung zum Teil aufheben.

$$R = R_0 - \sum_{i=1}^n I_i + A \quad (3.1)$$

Das E-Modell wurde ursprünglich für Narrowband- Anwendungen (NB: 300 Hz - 3,4 kHz) entwickelt, wobei die Qualität auf einer Skala von 0 (schlecht) bis 100 (ausgezeichnet) bewertet wurde. Inzwischen wurde das E-Modell auch für Wideband-Bedingungen (WB: 50 Hz - 7 kHz) und damit die Skala von 0 - 129 erweitert. Dies hat den Vorteil, dass die Codecs aus dem NB-Kontext direkt mit den Codecs aus dem WB-Kontext verglichen werden können. Der maximale Wert der R-Skala hängt direkt mit einem MOS-Wert von 4,5 zusammen. Es ist ebenso möglich die R-Werte in entsprechende MOS-Werte umzurechnen, wie aus ermittelten MOS-Werten R-Werte berechnet werden können. Für die Berechnung [24] des Letzteren verwendet man die Formel 3.2. Jedoch ist sie nur gültig für einen Bereich von  $6,5 \leq R \leq 100$

$$R = \frac{20}{3} \left( 8 - \sqrt{226} \cos\left(h + \frac{\pi}{3}\right) \right) \quad (3.2)$$

$$h = \frac{1}{3} \arctan 2(18566 - 6750MOS, \\ 15 \cdot (-903522 + 1113960MOS - 202500MOS^2)^{\frac{1}{2}}) \quad (3.3)$$

$$\arctan 2(x, y) = \begin{cases} \arctan\left(\frac{x}{y}\right) & \text{für } x \geq 0 \\ \pi - \arctan\left(\frac{y}{-x}\right) & \text{für } x < 0 \end{cases} \quad (3.4)$$

## 3.2 Testverfahren ACR und MUSHRA

Wie zuvor schon erwähnt sind Hörstudien sehr kostspielig und zeitintensiv, trotzdem sind sie die zuverlässigste Methode um Sprachqualität zu ermitteln. Es gibt verschiedene Verfahren, um von einer Testperson Qualitätsurteile zu erhalten. Zwei oft verwendete Verfahren sind der 5-Punkt ACR-Test und der MUSRHA-Test. Dieser Abschnitt enthält eine Beschreibung beider Testverfahren und einen kurzen Überblick über vorherige Studien zu diesem Thema.

**ACR - Test [17]:** ACR steht für „Absolute Category Rating“. Es handelt sich um einen Hörtest, der die subjektiv empfundene Qualität eines Sprachsignals ermitteln soll. Es wird die Gesamtqualität der Probe ermittelt. Die ITU-T empfiehlt dafür eine diskrete fünf Punkte Skala, bei der 1 für „schlecht“ steht und 5 für „ausgezeichnet“. Jedoch kann die Skala verschiedenen Untersuchungszielen angepasst werden. Diese Methode wird eingesetzt, um einen großen Bereich von Qualitäten zu untersuchen und ist weniger geeignet, kleine Schwankungen in der Signalqualität zu detektieren. Beim ACR werden die verschiedenen Signalproben einzeln vorgespielt und direkt nach jedem Beispiel bewertet. Unter den Teststimuli befindet sich ein Referenzsignal<sup>3</sup>, welches jedoch nicht explizit als solches deklariert ist.

**MUSHRA - Test [7]:** MUSHRA steht für „Multi Stimulus test with Hidden Reference and Anchor“. Es handelt sich um eine Testmethode für die subjektive Ermittlung von Audioqualität, mit der auch kleinere Unterschiede zwischen den Signalen erfasst werden können. Wie beim ACR wird beim MUSHRA die Gesamtqualität bewertet. Die Bewertung wird mittels einer kontinuierlichen Skala im Bereich von 0 „schlecht“ bis 100 „ausgezeichnet“

---

<sup>3</sup>Ein Referenzsignal stellt die optimale Signalqualität dar und dient somit als Richtwert, mit dem alle anderen Bedingungen verglichen werden.

durchgeführt. Im Test kann zwischen mehreren zeitsynchronen Testsignalen beliebig gewählt werden. Der Stimulus kann über Kopfhörer oder Lautsprecher dargeboten werden. Neben den eigentlichen Testsequenzen werden vier zusätzliche Signale dargeboten. Sie beinhalten die gekennzeichnete Referenz, sowie eine versteckte Referenz und ein oder mehrere versteckte Anker (bandbegrenztes Referenzsignal).

Weitere Informationen zur Durchführung beider Verfahren in dieser Studie befinden sich in Abschnitt 4.2.

Die Idee, den Einfluss verschiedener Testverfahren und Skalen zu vergleichen ist nicht neu. So wurden bereits mehrere Untersuchungen zum Thema durchgeführt. In der Studie von Q. Huynh-Thu et al. [23] wurde der Einfluss der Skaleneinteilung bei einem ACR-Test untersucht. Diese Studie wurde im Video-Kontext durchgeführt und verwendete vier verschiedene Skalen. Zwischen den verschiedenen Skalen wurde ein starker linearer Zusammenhang und keine signifikanten Unterschiede gefunden. Jedoch zeigte sich, dass sich die meisten Versuchspersonen bei ihrer Bewertung tendenziell an den Skalenmarkierungen orientierten. Dieser Effekt zeigt, dass ein Großteil der Versuchspersonen den Vorteil der kontinuierlichen Skala nicht nutzen.

Eine weitere Studie aus dem Video-Bereich, welche sich mit dem Vergleich verschiedener subjektiver Qualitätsbeurteilungsmethoden befasst, wurde in Japan durchgeführt [27]. Dabei wurden insgesamt sechs Methoden, darunter auch ein 5-Punkt ACR-Test, untersucht. Aus den sehr ähnlichen Ergebnissen schlussfolgerte die Gruppe, dass die Wahl der Bewertungsskala keinen erheblichen Einfluss auf das Ergebnis hat. In der Studie trat der 5-Punkt ACR-Test mit der schnellsten Bearbeitungszeit hervor und wird deshalb empfohlen.

Die Untersuchung, die dieser Arbeit voran gegangen ist, wurde von Feiten et al. [3] erarbeitet. In zwei durchgeführten ACR-Tests wurde eine große Spanne von Beeinträchtigungen bei IPTV-Diensten untersucht. Zusätzlich wurde ein MUSHRA-Test durchgeführt, in dem jedoch nur die Proben aus dem ACR-Test verwendet wurden, die im hochqualitativen Bereich zu finden sind. Die Zielsetzung in dieser Studie war die Untersuchung der oberen Qualitätsspanne und ob beide Testverfahren kombiniert werden können. Es wurde ermittelt, dass der ACR-Test die Qualitätsunterschiede bei hoher Qualität ähnlich gut ermittelt wie beim MUSHRA. Jedoch hat, wie erwartet der MUSHRA-Test eine höhere Auflösung bei kleinen Beeinträchtigungen. Im Allgemeinen wurde davon ausgegangen, ähnlich der Schlussfolgerung des japanischen Forscherteams, dass ein ACR-Test als Grundlage für ein Qualitätsmodell dienen kann.

# Kapitel 4

## Experimente

### 4.1 Vorbereitung der Experimente

Vor der Durchführung der Versuche stand die Planung und Vorbereitung. Beim Design beider Versuche wurde darauf geachtet, dass beide etwa gleich lang sind und die Versuchspersonen gleich fordern. Ebenso sollte die Anzahl der Bewertungen für die einzelnen Konditionen zwischen beiden Tests etwa gleich sein, um eine bessere Vergleichbarkeit zu gewährleisten. Weiterhin sollten die Konditionen sowie die verschiedenen Sprecher gleichhäufig vertreten sein. Es wurde ein zeitlicher Ablaufplan erstellt. Dadurch war es möglich abzuschätzen, wie viele Konditionen von einer Versuchsperson überhaupt bewertet werden können, ohne dass die Versuchsperson ermüdet. Am Ende soll für jede Testbedingung je eine Beurteilung zur Gesamtqualität aus dem ACR-Test und dem MUSHRA-Test ermittelt werden. Die Berechnung der Dauer beider Durchgänge ist in der Gleichung 4.1 zu sehen. Die Sprachdateien hatte durchschnittlich eine Länge von 8 Sekunden. Die Bewertungszeit für die abgespielten Dateien wurde auf 5 Sekunden geschätzt. In einem Durchgang sollten 39 Konditioen jeweils 2-mal vorkommen. Zusätzlich ist der Trainingsdurchgang mit 6 Dateien mit in die Berechnung eingegangen.

$$Dauer = \frac{2 \cdot (6 + (2 \cdot 39)) \cdot (8s + 5s)}{60s} \approx 37min \quad (4.1)$$

Die errechneten 37 Minuten wurden anschließend auf 40 Minuten aufgerundet. Zusammen mit der Einweisung, der Pause und der Audiometrie wird die komplette Versuchsdauer auf eine Stunde geschätzt (siehe Tabelle 4.1 auf Seite 14).

Jeder Proband hat somit genug Zeit, die verschiedenen Signalqualitäten 4-mal zu bewerten. In vorangegangenen Versuchen hat sich eine Testdauer von rund 60 Minuten als sinnvoll erwiesen. In dieser Zeit ist es möglich, viele Bewertungen zu erhalten und den Probanden nicht zu überfordern.

5 min	Einweisung
20 min	Durchgang 1
5 min	Pause
20 min	Durchgang 2
10 min	Audiometrie
<hr/>	
60 min	Komplette Versuchsdauer

Tabelle 4.1: Zeitplan Hörversuch

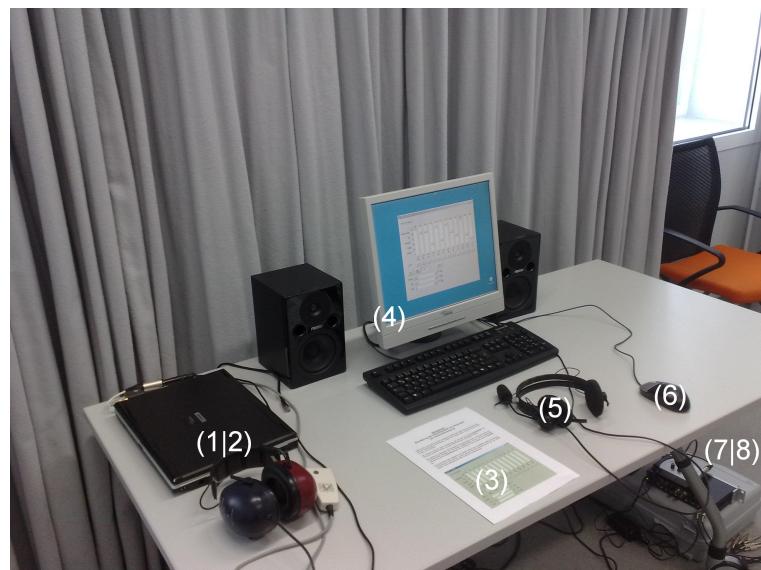


Abbildung 4.1: Versuchsaufbau

1. Digitales PC-Audiometer „ear1.7“
2. Sennheiser Kopfhörer HDA - 200 (Hörscreening)
3. Testeinweisung für die Probanden (hier für den MUSHRA - Test)

4. Bildschirm (auf dem die Probanden bewerteten)
5. Sennheiser Headset HMD - 410 (Hörversuch)
6. Eingabe der Bewertungen über optische Maus
7. Kopfhörerverstärker Millenium HA4
8. Soundkarte RME Fireface 400

**Laborraum:** Um für beide Hörversuche eine möglichst identische Umgebung zu schaffen, wurden beide Hörversuche innerhalb von zwei Wochen im T-Labs Labor „Calypso“ durchgeführt. Dieses Labor ist speziell für Versuche im Videobereich ausgestattet. Jedoch ist er durch seine hervorragenden Eigenschaften auch für Akustikexperimente sehr gut geeignet und erfüllt die Anforderungen aus der ITU-T Rec. P.800 [17]. Bei der Ermittlung des Ruhepegels wurde ein Wert von etwa 30 dB(A) erzielt. Die Nachhallzeit ( $RT_{60}$ )<sup>1</sup> ist ebenfalls sehr gering und liegt im Bereich von 125 Hz - 5000 Hz unter 0,25 s. Während des Versuchs befand sich kein PC im Raum, dessen Lüfterauschen stören konnte. Die Signale wurde vom benachbarten Raum in den Laborraum geschleift.

**Signalpfad:** Im benachbarten Raum „Belafonte“ befand sich der Testcomputer an dem eine RME HDSP Multiface II Soundkarte angeschlossen war. Mittels eines KVM-Schalters wurde das Monitorbild sowie die Eingabesteuerung der Maus und des Keyboards über ein Netzwerksystem in den Prüfraum geschleift. Der linke und rechte Ausgangskanal der Soundkarte wurde separat in den Prüfraum geschickt. Im Prüfraum wurden die beiden Monokanäle über die Soundkarte RME Fireface 400 zu einem Stereokanal zusammengefasst. Dieses Vorgehen war nötig, um über die Kontrollsoftware der Soundkarte die Kanäle einzeln auf den gewünschten Ausgangspegel zu bringen. Anschließend wurde das Stereosignal an den Kopfhörerverstärker Millenium HA4 geleitet, an dem das Sennheiser Headset HMD-410 angeschlossen war.

**Sprachmaterial:** Für die Hörtests wurden 15 deutschsprachige Sätze gewählt; jeweils 2 männliche und 2 weibliche Sprecher. Der Stimuluspool bestand somit aus insgesamt 30 Ursprungsdateien. Das Testmaterial wurde aus

---

<sup>1</sup>Die Nachhallzeit  $RT_{60}$  ist das Zeitintervall, in dem der Pegel nach dem Abschalten der Schallquelle um 60 dB fällt.

der Sprachdatenbank EUROM.1 [6] entnommen.

**Kopfhörer:** In beiden Versuchen wurde das Headset Sennheiser HMD-410 verwendet. Die Stimuluslautstärke war deutlich überschwellig und entsprach etwa lautgesprochener Sprache. Für den Hörtest wird ein Ausgangspegel von 79 dB SPL vorgeschlagen [18]. Der Ausgangspegel des Headsets wurde über einen Kunstkopf von „Head-Acoustics“ überprüft und justiert. Die Studie wurde binaural durchgeführt. Durch den Effekt der binauralen Lautheit über Kopfhörer [31] musste der Ausgangspegel der linken sowie der rechten Headsetseite auf 73 dB(A) verändert werden, denn bei einer Addition zweier kohärenter Schallsignale erhöht sich der Pegel von monauraler Lautheit zu binauraler Lautheit um 6 dB(A). Dieser Effekt lässt sich auch durch die Formel 4.2 für die Addition von kohärenten Schallquellen nachvollziehen [26].

$$L_{ges} = 20 \cdot \log(10^{\frac{L_l}{20}} + 10^{\frac{L_r}{20}}) \quad (4.2)$$

- $L_{ges}$ : Gesamtlautstärkepegel
- $L_r$ : Lautstärkepegel der Kopfhörerseite rechts
- $L_l$ : Lautstärkepegel der Kopfhörerseite links

**Versuchsteilnehmer:** Es wurden für die Hörversuche insgesamt 44 Testpersonen eingeladen. Dabei handelte es sich in fast allen Fällen um Studenten der nahegelegenen Universität, die in zwei Gruppen unterteilt wurden. Gruppe 1 nahm am ACR-Test teil und Gruppe 2 am MUSHRA-Test. Alle Versuchspersonen haben für ihre Teilnahme eine Aufwandsentschädigung erhalten. Als Teilnahmebedingung wurde das Beherrschung der deutschen Sprache und ein unbeeinträchtigtes Gehör vorausgesetzt.

Durchschnittsalter in Jahren	Geschlecht(w/m)	Test
25,67	12/12	ACR
25,75	11/9	MUSHRA

Tabelle 4.2: Zusammensetzung der Versuchspersonen

**Audiometer:** Dennoch wurden zusätzlich alle Teilnehmer einer kurzen Hörprüfung unterzogen. Dabei wurde nur die Luftleitung nach einem modifizierten Békésy-Audiometrieverfahren ermittelt. Für die Audiometrie wurde

das digitale PC-Audiometer „ear1.7“ verwendet. Keine Versuchsperson müsste aus der Studie ausgeschlossen werden musste, da keine Anzeichen für eine Hörstörung gefunden wurden.

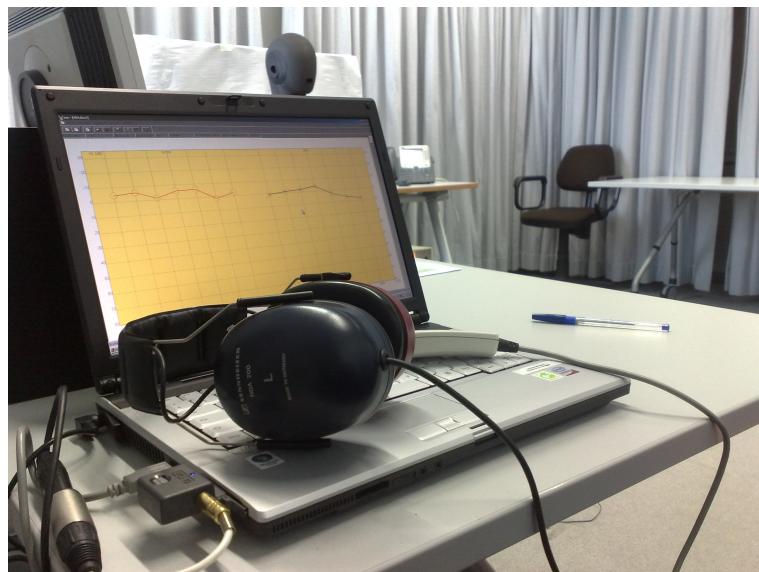


Abbildung 4.2: PC-Audiometer „ear1.7“

#### 4.1.1 Konditionen

In der Studie sollten die Testbedingungen einen großen Qualitätsbereich in Telefonnetzwerken abdecken. In vorherigen Studien wurde die Übertragungsbandbreite bereits bis auf Super-Wideband (SWB) vergrößert [28]. In dieser Studie wurde noch ein Schritt weiter gegangen und der komplette Frequenzbereich des menschlichen Hörfeldes (20 Hz - 20 kHz) gewählt (siehe auch Abbildung 5.4 auf Seite 33). Dieser Abschnitt beschreibt die verwendeten Codecs und Bandpassfilter.

- AAC - LD[5] - steht für „Advanced Audio Coding-Low Delay“ und wurde vom Fraunhofer Institut entwickelt. Er bietet eine, für alle Arten von Audiosignalen, hohe Klangqualität und besitzt eine sehr kurze Kodiererverzögerung von nur 20 ms. Deshalb kommt er unter anderem bei Videokonferenzen und in VoIP-Systemen zum Einsatz; bei diesen Anwendungen sind eine hohe Klangqualität und kurze Verzögerungen gefordert.

- CELT[29] - steht für „Constrained Energy Lapped Transform“. Er wurde von der Xiph.Org Foundation entwickelt und ist frei erhältlich. Die Kodierungsverzögerung ist sehr kurz, liegt typischerweise bei 3 ms bis 9 ms. Deshalb eignet er sich für das gleiche Anwendungsgebiet wie der AAC-LD. Er unterstützt Sprach- und Audiosignale und versucht, ähnlich wie der AAC-LD, die Lücke zwischen hoher Qualitätsforderung und kurzer Verzögerung zu schließen.
- SILK[30] - ist ein von der Firma Skype Limited entwickelter Audio-codec. Er soll nicht nur eine hohe Qualität bei niedriger Verzögerung liefern, sondern sich auch an die verschiedenen Netzwerke anpassen. Der SILK-Codec wurde für Echtzeit-Sprachanwendungen ausgelegt.
- AMR - WBplus[2] - Dabei handelt es sich um eine Erweiterung des AMR-WB Codecs. Er unterstützt Stereosignale und höhere Abtastraten und benutzt zusätzlich zur „Algebraic Code Excited Linear Prediction“ (ACELP) einen weiteren Algorithmus für die Kodierung. Dabei handelt es sich um die „Transform Coded Excitation“(TCX).
- G.711[10] - ist ein Codec, der nur den Schmalbandbereich (300 Hz - 3400 Hz) verarbeitet. Er wird hauptsächlich in ISDN Netzen benutzt.
- G.722[11] - ist ein Codec, der den Frequenzbereich von 50 Hz - 7000 Hz verarbeitet. Er arbeitet mit einer Sub-Band Adaptive Differential Pulse Code Modulation (SB-ADPCM) bei einer Bitrate von 64 kbit/s. Das Eingangssignal wird in zwei Bänder aufgeteilt, separat verarbeitet und später wieder zusammengefügt.
- G.722.2[12] - ist ebenfalls ein Codec der den Frequenzbereich von 50 Hz - 7000 Hz verarbeitet. Dabei benutzt er jedoch die „Algebraic Code Excited Linear Prediction“ (ACELP) für die Kodierung. Es handelt sich dabei um den gleichen Codec wie der AMR-WB.
- G.729A[14] - ist ein Codec, der bei der Kodierung keine Art der PCM, sondern einen speziellen Algorithmus „Conjugate-Structure Algebraic-Code-Excited Linear Prediction“(CS-ACELP) nutzt.
- G.722.1 Annex C [13] - ist ein Codec für Superbreitband-Kodierung. Er kann für Video- und Telefonkonferenzen eingesetzt werden oder in Internetanwendungen (Videostreaming).
- Bandpassfilter - Es wurden in der Studie eine ganze Reihe von Bandpässen verwendet. Diese Konditionen wurden keiner Kodierung unterzogen, sondern lediglich mit verschiedenen FIR-Bandpässen gefiltert.

Der Grund für die Verwendung von bandbegrenzten Sprachsignalen ist, dass eine Aussage über den Einfluss der Bandbreite auf die Qualität nur dann gemacht werden kann, wenn keine weiteren Störungen das Ergebnis beeinflussen können. Es wurden Bandpässe gewählt, die sowohl Schmalband-, Breitband- sowie Superbreitbandübertragungen abbilden. Zusätzlich wurden noch einige Zwischengrößen gewählt.

- NB, WB, SWB - Zusätzlich zu den zuvor genannten Bandpässen wurden noch drei Bandpässe erstellt. (NB: 300 - 3400 Hz, WB: 50 - 7000 Hz, SWB: 50 - 14000 Hz)
- FB\_Org - hinter der Bezeichnung FB\_Org verbergen sich die unbekränkigen Sprachdateien. Sie gelten für beide Versuche als optimale Qualität, da weder Störungen wie Paketverlust, Quantisierungsfehler o.ä. vorkommen, noch die Bandbreite beschränkt wurde.

#### 4.1.2 Generieren der Sprachdateien

Dieser Abschnitt befasst sich mit dem Erstellen der Sprachbeispiele. Es mussten nicht alle der zuvor genannten Konditionen erstellt werden, da sie schon in anderen Studien verwendet wurden [28] [8]. Zusätzlich dienen diese Dateien als Vergleichskonditionen, die es ermöglichen die Ergebnisse direkt mit Ergebnissen aus älteren Studien zu vergleichen. Für diese Studie wurden die folgenden Konditionen neu erstellt:

- AAC-LD, CELT, SILK, NB, WB, SWB, G.711, G.722, G.722.2
- Bandpässe 300 - 7k, 600 - 7k, 50 - 3.4k [Hz]

Das Erstellen aller Dateien ist in drei Abschnitte unterteilt. In Abbildung 4.3 ist der schematische Ablauf zu sehen. Er zeigt, welche Schritte für das Prozessieren nötig waren.

1. Vorverarbeitung
2. Kodierung / Dekodierung
3. Nachverarbeitung

Die Vor- und Nachverarbeitung wurde gemäß den Empfehlungen der ITU-T [18] [16] durchgeführt. Die Software, die für die Verarbeitung benutzt wurde, stammt aus der ITU-T Rec. G.191 [15]. Es handelt sich dabei um standardisierte Funktionen, um Audio- und Sprachdateien zu bearbeiten.

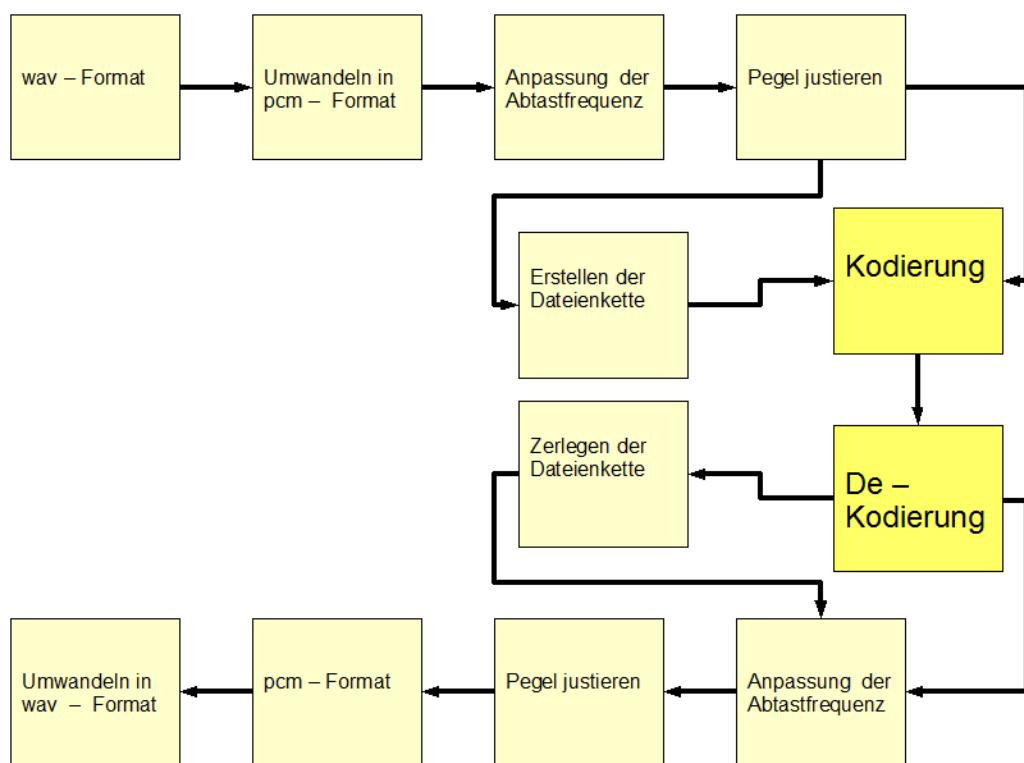


Abbildung 4.3: Schema zum Generieren der Sprachbeispiele

Die Ursprungsdateien lagen im wav-Format vor und mussten vor der Bearbeitung mit Hilfe eines Matlab-Skripts ins pcm-Format gebracht werden. Um die Verarbeitung der vielen Dateien möglichst automatisiert ablaufen zu lassen, wurden MS-DOS-Stapelverarbeitungsdateien geschrieben. Diese wurden mit Hilfe der DOS-Eingabeforderung aufgerufen. Dabei arbeitet das System die Befehle Zeilenweise ab. Es wurde für jede Kondition eine separate Datei geschrieben, die von einer Hauptdatei aufgerufen wurden. Ein Beispiel für eine Stapelverarbeitungsdatei (Batchfile) ist im Anhang B zu finden. Einzig die drei Bandpässe wurden mit Matlab gefiltert.

Die Ursprungsdateien hatten eine Abtastrate von 48 kHz. Für einige Filter und Codecs war es notwendig die Abtastfrequenz zu verändern. Es wurde immer eine Kombination aus zwei hochwertigen FIR-Tiefpassfiltern benutzt. Mit diesen Filtern konnte man im Verhältnis 1:2 und 1:3 entweder herauf- oder heruntertakten und so die Abtastfrequenz anpassen. Nach dem die Tastfrequenz den jeweiligen Bedürfnissen angepasst war, wurde der Pegel auf -26 dBov justiert<sup>2</sup>. Dadurch wurde verhindert, dass es zu Begrenzungseffekten kommt. Danach wurde die eigentliche Kodierung durchgeführt. Bei einigen Codecs wurden, alle vorverarbeiteten Sprachdateien zu einer Kette zusammenzufassen. Spezielle Funktionen (aus der ITU-T Rec. G.191 [15]) fügten alle Dateien aneinander und gaben zusätzlich ein Log-File aus, in dem die Länge und Positionen der Dateien stand.

Nach diesem Schritt wurden die Dateien wieder dekodiert. Bei Codecs, bei denen eine zusammengefasste Datei verarbeitet wurde, wurde die Dateienkette wieder in einzelne Dateien zerlegt. Anschließend wurden sie auf die ursprünglichen 48 kHz getaktet und der Ausgangspegel nochmals auf -26 dBov justiert. Im letzten Schritt mussten die Sprachbeispiele wieder aus dem pcm - Format zurück ins wav - Format gebracht werden. Hierfür wurde ebenfalls ein Matlab - Skript verwendet.

### 4.1.3 Zusammenfassung der Konditionen

Die in der Studie verwendeten Konditionen und ihre Eigenschaften sind in Tabelle 4.3 aufgelistet.

---

<sup>2</sup>Die Einheit dBov steht für Dezibel overload. Sie gibt die Amplitude eines Signals im Vergleich zum maximalen Wert, den ein Systems verarbeiten kann, an. Wird der maximale Wert überschritten, kommt es zu Begrenzungseffekten.

Kondition	Anzahl	Eigenschaften
AAC-LD	5	32, 40, 44, 48, 64 kbit/s
CELT	4	32, 40, 48, 64 kbit/s
SILK	2	12, 32 kbit/s
AMR-WB+	3	6, 13, 36 kbit/s
G.711	1	64 kbit/s
G.722	3	64 kbit/s, 2% und 8% Paketverlust
G.722.2	3	6.6, 12.65, 23.05 kbit/s
G.729A	2	8 kbit/s, 8% Paketverlust
G.722.1 Ann.C	3	24, 32, 48 kbit/s
Bandfilter	6	300-7k, 600-7k, 50-3.4k, 50-8k, 50-10k, 50-12k [Hz]
NB	1	50-3.4k [Hz]
WB	3	50-7k [Hz], Rauschen KFZ 55, 70 dB(A)
SWB	1	50-14k [Hz]
FB_Org	1	unbearbeitet
Tandem	2	G.722[64]xG.722.2[12.65][kbit/s], G.729AxG.729A
Weitere Eigenschaften:		
Anzahl der Sprecher		2 männliche und 2 weibliche
Anzahl der Sätze		15 pro Sprecher
Ausgangspegel		73dB(A)
Versuchspersonen		24 ACR und 20 MUSHRA (unerfahren)
Bewertungsskala		5-Punkt ACR und 0-100 Skala MUSHRA
Sprache		deutsch
Testumgebung		in Anlehnung an ITU-T Rec. P.800 [17]

Tabelle 4.3: Übersicht der verwendeten Konditionen

## 4.2 Durchführung

Die Hörversuche wurden nach der Empfehlungen der ITU-T [17] durchgeführt.

**Durchführung ACR - Test:** Jeder Proband bekam die Einweisung in schriftlicher Form (Anhang A), welche er zuerst lesen sollte. Anschließend wurden noch aufkommende Frage beantwortet. Der Versuch bestand aus zwei Durchgängen mit einer fünfminütigen Pause dazwischen. Die Pause wurde eingesetzt, um einer eventuellen Ermüdung des Probanden entgegenzuwirken. Jeder Durchgang startete mit einer kurzen Trainingsphase. Das Training war für alle Teilnehmer identisch und diente zur Eingewöhnung in die Testoberfläche. Die Beurteilung wurde auf einer diskreten 5-Punkte Skala durchgeführt. Um dem Probanden eine kleine Hilfestellung zu geben sind die Ziffern [1 - 5] zusätzlich mit Adjektiven gekennzeichnet worden.

- ausgezeichnet / 5
- gut / 4
- ordentlich / 3
- dürftig / 2
- schlecht / 1

Darüber hinaus sollte das Training als Überblick dienen, welche Arten von Störungen auftreten könnten. Die Ergebnisse aus dem Training wurden nicht mit in die Auswertung genommen. Das verwendete Programm erfasste die Bewertungen automatisch und schrieb sie in eine Log-Datei. Die Versuchspersonen hatten die Möglichkeit, durch drücken der „Play“-Taste das Beispiel noch einmal zu hören. Nach der Beurteilung wurde über die Taste „Weiter“ zum nächsten Beispiel gesprungen.

Es wurden im Test 39 unterschiedliche Bedingungen zur Bewertung dargeboten. Für jeden Durchgang wurde eine randomisierte Playliste erstellt. Ein Beispiel für eine Playliste ist im Anhang C zu sehen. Nach beiden Durchgängen wurde die zuvor erwähnte Audiometrie durchgeführt. Durchschnittlich benötigten alle Versuchspersonen im ACR-Test die geplanten 60 Minuten, um den kompletten Versuch zu absolvieren.

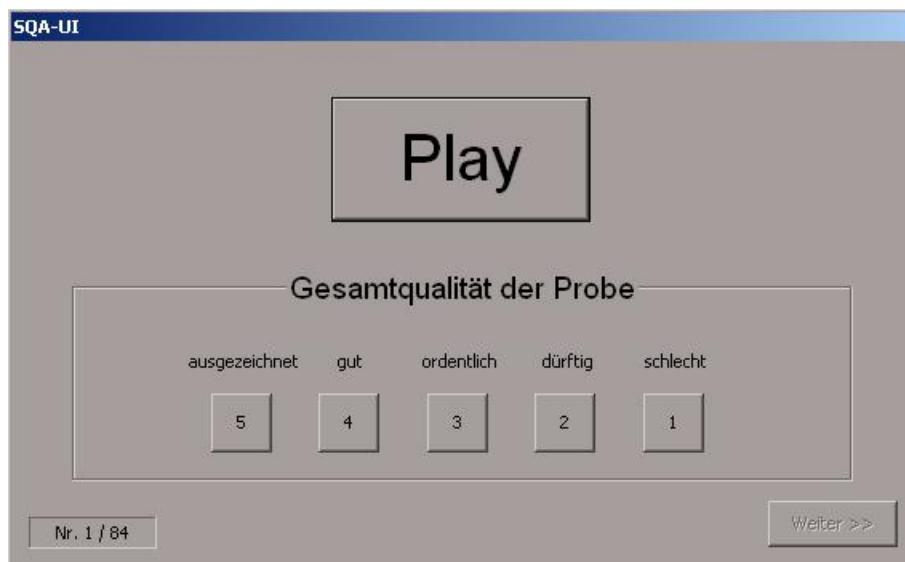


Abbildung 4.4: Eingabefenster ACR-Test

**Durchführung MUSHRA-Test:** Ähnlich dem Ablauf des ACR-Tests wurde auch beim MUSHRA-Test eine Einweisung in ähnlicher Form (Anhang A) vom Probanden gelesen. Anschließend wurde ein Trainingsdurchgang durchgeführt. Im Vergleich zum ACR-Test ist die Testoberfläche deutlich komplexer. Deshalb war der Versuchsleiter beim Training anwesend und konnte aufkommende Fragen direkt klären. Die Bewertung wurde hier auf einer kontinuierlichen Skala [0 - 100] vorgenommen. Zusätzlich zu den Zahlen waren an der Seite der Schieberegler auch noch dieselben Adjektive wie im ACR-Test abgebildet (siehe Abbildung 4.5). Die Trainingsphase war für alle Probanden wiederum gleich. Anschließend wurde mit dem eigentlichen Testdurchlauf begonnen. Der Test bestand aus 16 Blöcken, in denen jeweils 12 Sprachproben zu bewerten waren. Um auch im MUSHRA-Test aufkommenden Konzentrationsverlust zu vermeiden, wurde nach dem 8. Block eine fünfminütige Pause eingelegt. Die Bewertungen aus der Trainingsphase wurden nicht mit in die Auswertung übernommen. Die Versuchspersonen konnten jederzeit zwischen den Qualitätsbeispielen in einem Block wechseln und ihr Urteil anpassen. Nach der Bewertung aller Beispiele wurde über die „Next“-Schaltfläche die Beurteilung des Blockes gespeichert und zum nächsten Block gesprungen.

Es wurden im Test 39 unterschiedliche Bedingungen zur Bewertung dargeboten. Für jeden Durchgang wurde eine randomisierte Playliste erstellt. Ein Beispiel für eine Playliste ist im Anhang C zu finden. Zusätzlich spielt

das benutzte Programm „Step“ die 19 Durchgänge in zufälliger Reihenfolge ab und setzt die Reihenfolge der Sprachproben innerhalb eines Durchgangs auch zufällig. Das Programm erstellt von jedem Probanden eine Log-Datei für die spätere Auswertung. Nach beiden Durchgängen wurde wie zuvor die Audiometrie durchgeführt. Im Gegensatz zum ACR-Test benötigten hier die Versuchspersonen unterschiedlich lange (50 - 75 Minuten), um den Versuch zu absolvieren. Bei späterer Befragung gaben einige Probanden an, dass sie sehr oft zwischen der Referenz und dem verschlechterten Signal hin und her schalteten und dadurch verhältnismäßig lange brauchten.

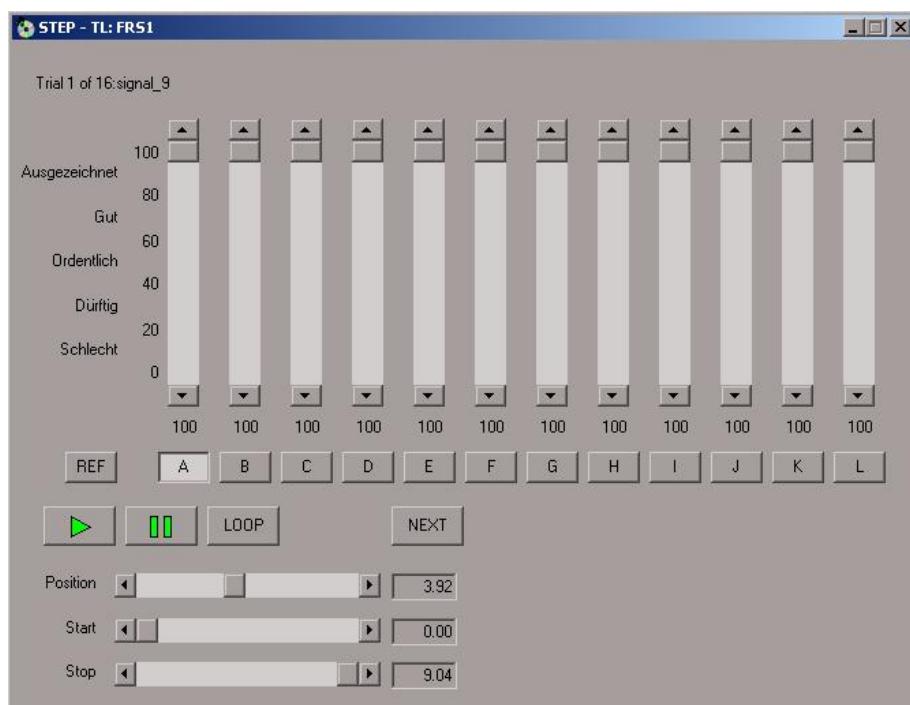


Abbildung 4.5: Eingabefenster MUSHRA-Test

# Kapitel 5

## Ergebnisse

### 5.1 Dokumentation der Ergebnisse ACR und MUSHRA

Die Rohdaten aus den Log-Dateien wurden mit Hilfe von MATLAB ausgelesen und in Exceldateien zur weiteren Bearbeitung gespeichert. Für jeden Test ist ein Beispiel für eine Log-Datei im Anhang D angefügt. Dabei wurden die Mittelwerte sowie die Standardabweichung und die 95% Konfidenzintervalle berechnet.

Die Ergebnisse beider Experimente sind in Tabelle 5.1 zu sehen. Die Tabelle listet die Konditionen, ihre mittleren Bewertungen, Standardabweichungen und Konfidenzintervalle aus beiden Tests auf. Im ACR-Test wurde über 96 Bewertungen pro Kondition, im MUSHRA-Test über 80 Bewertungen pro Kondition gemittelt. An dieser Stelle sollte erwähnt werden, dass im MUSHRA-Test FB\_Org, NB, WB über jeweils 320 Bewertungen gemittelt wurde. Diese drei Konditionen wurden in jedem Block bewertet und dienten als Referenz bzw. Anker. Es ist zu beachten, dass beide Tests in dieser Phase noch nicht verglichen werden können, da sie noch nicht auf einer gemeinsamen Skala abgebildet sind. Die Gegenüberstellung der Testergebnisse wird in Abschnitt 5.3 behandelt.

Anschließend ist für beide Tests jeweils ein Diagramm abgebildet (ACR: Diagramm 5.1 auf Seite 28; MUSHRA: Diagramm 5.2 auf Seite 29). Darin sind die gemittelten Bewertungen der einzelnen Konditionen und ihre 95% Konfidenzintervalle als Fehlerbalken abgebildet. Die Kondition FB\_Org, die in beiden Versuchen als optimale Qualität gilt, ist farblich hervorgehoben.

Kondition	ACR Score	Sdv	CI95%	MUSHRA Score	Sdv	CI95%
<i>AAC_LD[32k]</i>	3,57	1,01	0,20	65,33	20,81	4,56
<i>AAC_LD[40k]</i>	---	---	---	78,30	16,84	3,69
<i>AAC_LD[44k]</i>	4,35	0,68	0,14	---	---	---
<i>AAC_LD[48k]</i>	4,34	0,75	0,15	90,20	12,23	2,68
<i>AAC_LD[64k]</i>	4,55	0,65	0,13	93,24	9,99	2,19
<i>AMR-WBplus@13</i>	3,47	0,91	0,18	70,71	24,30	5,33
<i>AMR-WBplus@36</i>	4,28	0,72	0,14	88,45	14,18	3,11
<i>AMR-WBplus@6</i>	1,57	0,82	0,16	17,96	18,40	4,03
<i>bp_filter_300 – 7000Hz</i>	2,84	0,94	0,19	51,66	21,43	4,70
<i>bp_filter_50 – 10000Hz</i>	4,26	0,74	0,15	84,15	14,57	3,19
<i>bp_filter_50 – 12000Hz</i>	4,48	0,70	0,14	94,26	7,87	1,72
<i>bp_filter_50 – 3400Hz</i>	2,23	1,01	0,20	33,29	20,83	4,56
<i>bp_filter_50 – 8000Hz</i>	3,93	0,80	0,16	74,61	17,57	3,85
<i>bp_filter_600 – 7000Hz</i>	2,28	0,95	0,19	32,30	18,89	4,14
<i>CELT[32k]</i>	3,01	1,17	0,23	60,16	29,61	6,49
<i>CELT[40k]</i>	3,49	1,09	0,22	69,31	27,21	5,96
<i>CELT[48k]</i>	3,97	1,00	0,20	83,05	19,59	4,29
<i>CELT[64k]</i>	4,49	0,68	0,14	90,15	11,57	2,54
<i>direct_wb_car55dB(A)</i>	2,90	0,85	0,17	48,78	23,00	5,04
<i>direct_wb_car70dB(A)</i>	2,09	0,81	0,16	18,65	14,75	3,23
<i>FB_Org</i>	4,29	0,75	0,15	97,77	6,47	0,71
<i>G.711@64</i>	2,31	0,97	0,19	36,09	18,01	3,95
<i>G.722.2@12.65</i>	2,83	0,84	0,17	51,86	18,79	4,12
<i>G.722.2@23.05</i>	3,20	0,96	0,19	59,11	18,28	4,01
<i>G.722.2@6.6</i>	1,63	0,67	0,13	24,39	19,52	4,28
<i>G.722@64</i>	3,27	0,84	0,17	66,59	16,96	3,72
<i>G.722@64_2%FER</i>	2,57	0,88	0,18	53,01	21,76	4,77
<i>G.722@64_8%FER</i>	1,51	0,60	0,12	27,05	23,65	5,18
<i>G.722@64xG.722.2@12.65</i>	2,69	0,85	0,17	49,01	20,51	4,49
<i>G.729A@8</i>	2,11	0,87	0,17	30,48	19,94	4,37
<i>G.729A_8%FER</i>	1,40	0,61	0,12	16,94	14,59	3,20
<i>G.729AxG.729A</i>	1,78	0,76	0,15	25,39	16,20	3,55
<i>G.722_1Ann.C@24</i>	2,91	0,97	0,19	50,48	24,82	5,44
<i>G.722_1Ann.C@32</i>	3,36	1,02	0,20	64,79	24,55	5,38
<i>G.722_1Ann.C@48</i>	3,93	0,95	0,19	76,39	21,58	4,73
<i>NB</i>	2,43	1,03	0,21	37,97	19,69	2,16
<i>SILK[12k]</i>	3,10	0,93	0,19	63,95	21,95	4,81
<i>SILK[32k]</i>	4,15	0,74	0,15	87,88	12,53	2,75
<i>SWB</i>	4,50	0,66	0,13	91,31	13,78	3,02
<i>WB</i>	3,64	0,91	0,18	64,17	18,06	1,98

Tabelle 5.1: Übersicht der Ergebnisse des ACR- sowie des MUSHRA-Tests. Score steht für den Mittelwerte der Bewertungen, Sdv für die Standardabweichung und CI95% für das 95% Konfidenzintervall

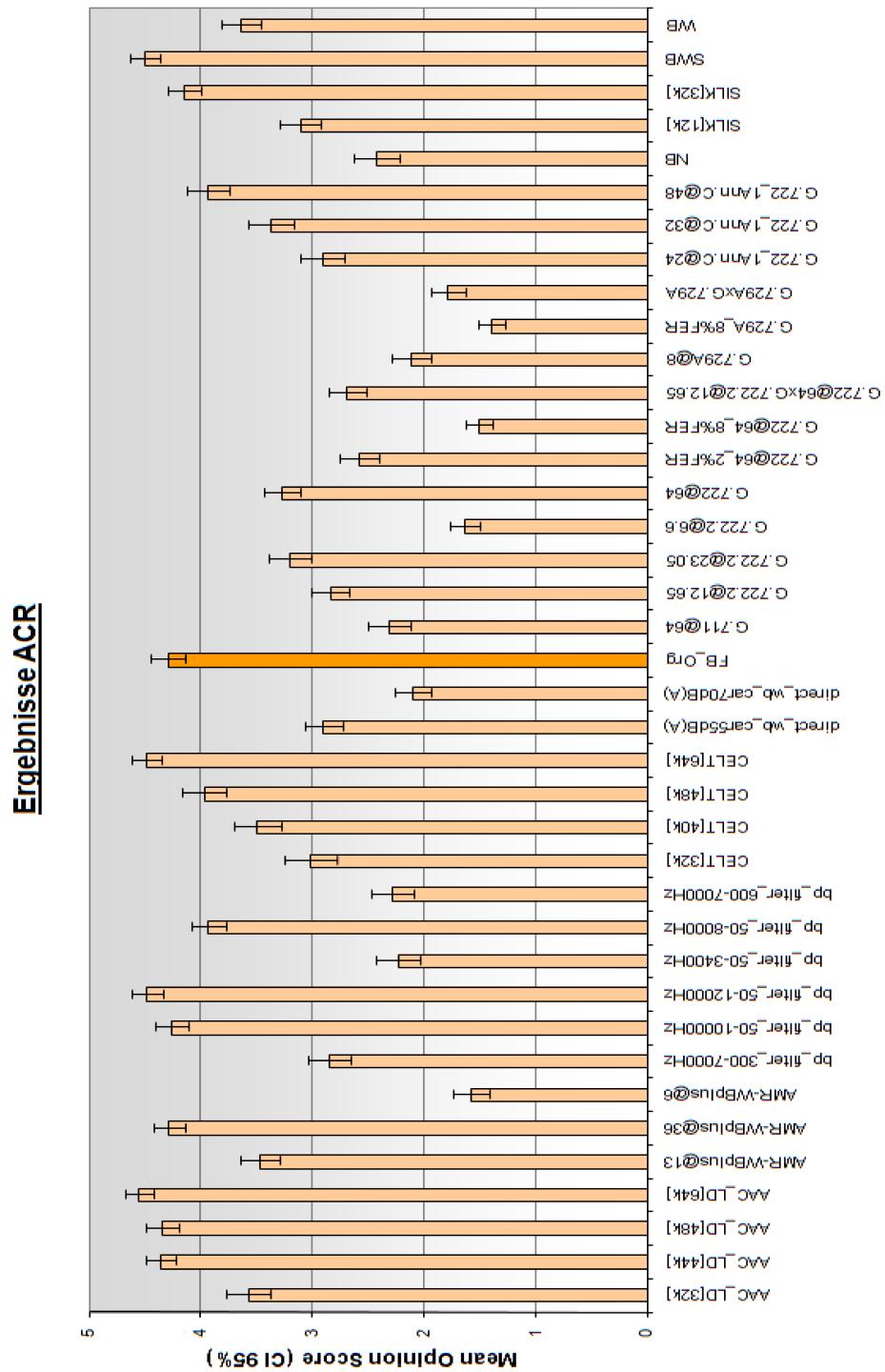


Abbildung 5.1: Mittelwerte der Versuchskonditionen, 95% Konfidenzintervalle als Fehlerbalken

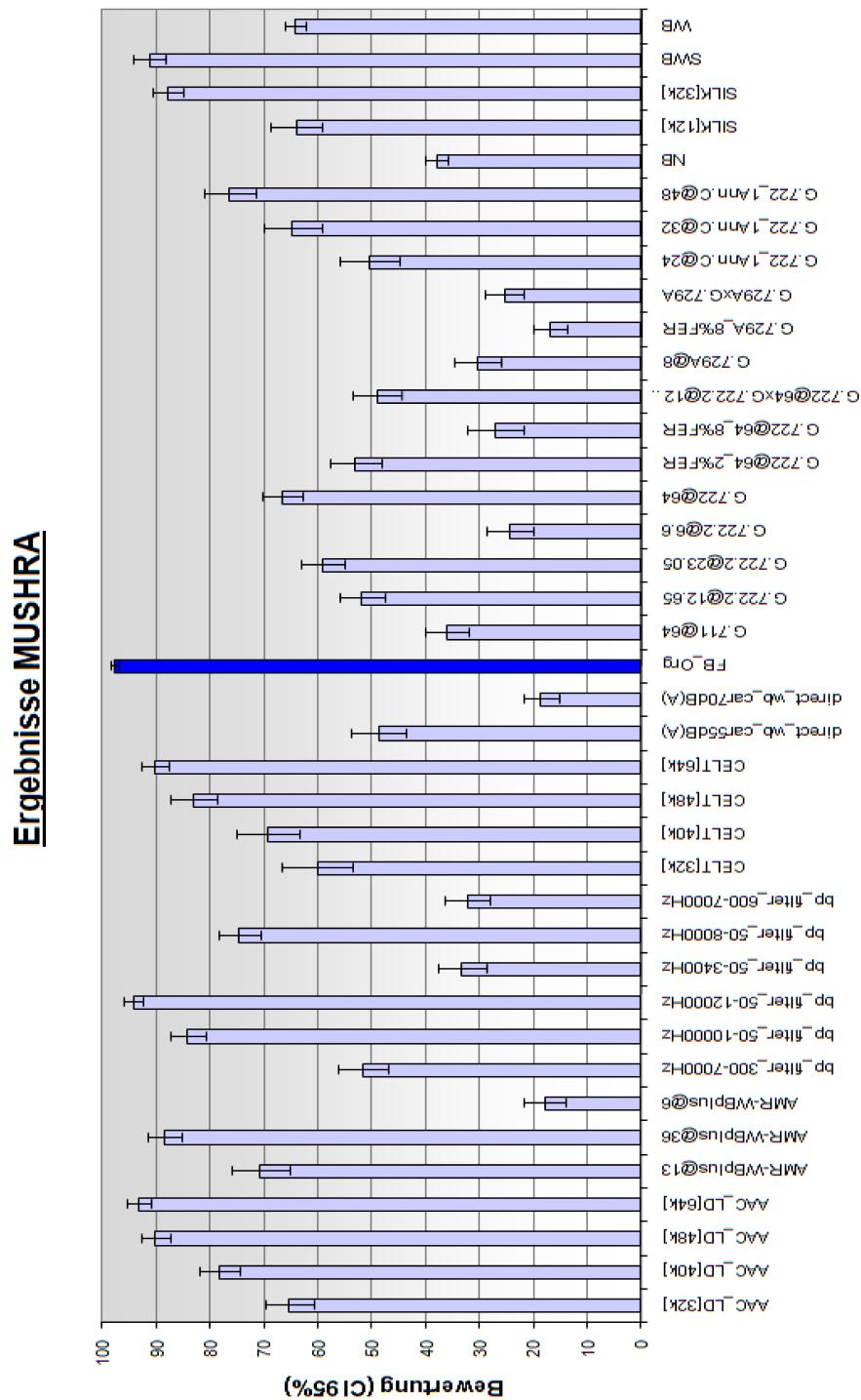


Abbildung 5.2: Mittelwerte der Versuchskonditionen, 95% Konfidenzintervalle als Fehlerbalken

## 5.2 Auswertung der Ergebnisse

### 5.2.1 Betrachtung der Qualitätsbewertungen

Im Diagramm 5.1 auf Seite 28 sind deutlich die Unterschiede zwischen den Qualitäten aus dem ACR-Test zu sehen.

Wie erwartet steigt bei höherer Bitrate auch die Qualität. Als Beispiel sei dafür der CELT Codec genannt. Ebenso deutlich ist eine Verschlechterung der Qualitätsbewertung zu erkennen, je größer die Störeinflüsse in der Übertragung sind, z.B. die Konditionen mit Hintergrundrauschen (`direct_wb_car55dB(A)` und `direct_wb_car70dB(A)`). Im Diagramm 5.2 auf Seite 29 sind die Testergebnisse der einzelnen Bedingungen des MUSHRA-Tests abgebildet. Auch hier sind die gleichen Effekte zu beobachten wie beim ACR-Test. Bei einer Erhöhung der Bitrate steigt die Gesamtqualität und bei einer Vergrößerung der Störeinflüsse sinkt sie. Die Bewertungen der Bandpässe zeigen dieselbe Ausprägung. Generell decken sich die Ergebnisse aus beiden Tests sehr stark. In der gesamten Arbeit wurden in der Summe aller Urteile mehr „gut“ und „ausgezeichnet“ abgegeben, als „dürftig“ und „schlecht“.

Beim Vergleich der Kondition „WB“ mit beiden Rauschkonditionen, ist der Störeinfluss des Rauschens deutlich auszumachen. So verschlechtert sich die Bewertung um etwa einen halben MOS-Wert beim Auftreten von Rauschen (55 dB(A)). Die Qualität wird durch das lautere Rauschen noch weiter verschlechtert; von durchschnittlich „ordentlich“ (bei 55 dB(A)) auf durchschnittlich „dürftig“ (bei 70 dB(A)), also rund einen ganzen MOS-Wert. Es ist also anzunehmen, wenn die Lautstärke des Rauschens weiterhin zunimmt, stinkt die Qualität weiter. Der Grund dafür ist, dass trotz der hervorragenden Fähigkeiten des Gehörs, Nutzschall von Störschall zu trennen, das Rauschen Anteile der Sprache überdeckt. Somit wird die Verständlichkeit schlechter und der Konzentrationsaufwand größer. Beide Faktoren wirken sich negativ auf die Bewertung aus.

Mehrere Konditionen, darunter der AAC-LD, CELT, SILK, AMR-WB+ und einige Bandpässe haben einen Wert über 4 erreicht. Der SILK-Codec mit einer Bitrate von 32 kbit/s, hat einen Wert von 4,15 erzielt. Die drei Codecs AAC-LD, CELT und SILK sind für ähnliche Anwendungen gedacht (z.B. Echtzeit-Sprachübertragung). Vergleicht man diese bei einer Bitrate von 32 kbit/s, erkennt man, dass der SILK eine deutlich bessere Bewertung erhält.

Entgegen der Erwartung wurde die Ursprungsdatei („FB\_Org“), die unverändert in den Test übernommen wurde, beim ACR-Test nicht am besten

bewertet. Eine eindeutige Erklärung wurde dafür nicht gefunden. Die größere Bandbreite der originalen Sprachdatei wird nicht mehr als Verbesserung wahrgenommen. Der ermittelte Wert liegt jedoch nur leicht unter den besten Bandpässen, auf dem Niveau der Bandpassfilter 50 Hz - 10 kHz, 50 Hz - 12 kHz und SWB. Der Fehler könnte in der schlechteren Auflösung der MOS-Skala liegen. Da eventuell die Versuchspersonen bei einigen Konditionen sich schwer entscheiden konnten („Die Kondition ist zwar nicht ausgezeichnet, aber besser als gut“). Diesen Ansatz könnte man durch ein Wiederholen des Tests mit einem stufenlosen Schieber überprüfen. Sollte die Vermutung stimmen, müsste dann „FB\_Org“ am besten abschneiden.

Ein weiterer Aspekt ist, dass die Codecs AAC-LD (44, 48, 64 kbit/s), Celt (64 kbit/s), SILK (32 kbit/s) besser im Test abschnitten als die Quellendatei. Ein zufälliges Auftreten eines Reihenfolgeneffektes wurde ausgeschlossen, nachdem die Playlisten überprüft wurden. Es ist auch denkbar, dass die Bewertung eher von klanglicher Präferenz der Versuchspersonen herrührt. Etwa die Hälfte aller Probanden hatte einen mp3-Player dabei. Mit sehr großer Wahrscheinlichkeit besitzt auch ein Großteil der anderen Hälfte einen mp3-Player. Das Klangempfinden könnte sich durch das ständige Hören von digitalisierten Audiosignalen verändert haben. Die kodierten Sprachdateien somit als klanglich besser bewertet wurden. Diese Vermutung müsste jedoch in weiteren Untersuchungen überprüft werden.

Ein Unterschied in der Bewertung zwischen ACR und MUSHRA liegt bei der ungestörten Kondition „FB\_Org“. Im MUSHRA-Test wurde sie wie erwartet als beste Kondition beurteilt. Der Grund für diese Abweichung zum ACR-Test liegt darin, dass das Referenzsignal explizit als „beste“ Qualität angegeben ist. Die Referenz wird im Testablauf durch ständiges Vergleichen leichter heraus gehört. Die Standardabweichungen sind im MUSHRA-Test für 36 der 38 Konditionen kleiner, was belegt, dass der stufenlos einstellbare Schieberegler und die im MUSHRA-Test verwendete Skala [0 - 100] besser auflösen.

**Auswertung der Bandpasskonditionen:** Um den Einfluss der Bandbreite auf die Signalqualität zu untersuchen, ist es notwendig nur Signale zu betrachten, die keine weiteren Störungen aufweisen. Deshalb wurde eine Vielzahl von Bandpässen verwendet. Dabei ist zu erkennen, dass mit steigender Bandbreite auch die Qualitätsbewertung steigt. Schaut man sich den Zuwachs an Qualität genauer an, ist deutlich ein Sättigungsverhalten erkennbar; Abbildung 5.3 Seite 32.

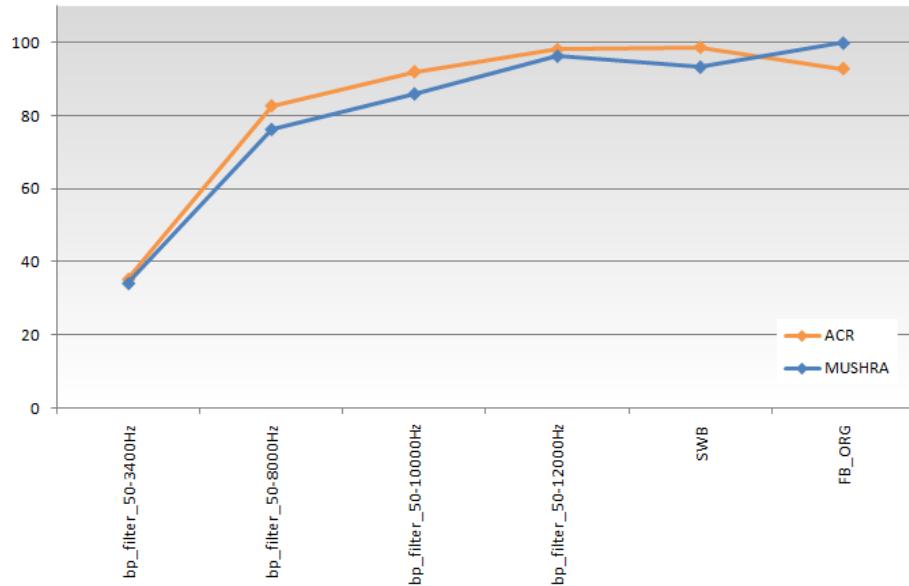


Abbildung 5.3: Sättigungsverhalten in der Beurteilung der Bandpässe mit unterer Grenzfrequenz  $f_u = 50\text{Hz}$  und FB\_Org, Werte für den Vergleich von ACR und MUSHRA auf 0 - 100 normalisiert

Daraus lässt sich schließen, dass ab einer Übertragungsbandbreite größer als SWB sich kein weiterer Qualitätsgewinn erzielen lässt. Kein Proband hatte im Hörscreening ein Anzeichen für eine Hörstörung. Bei einem Durchschnittsalter von 26 Jahren ist nicht davon auszugehen, dass der hörbare Bereich der Versuchspersonen eingeschränkt ist und nicht mehr bei ca. 20 Hz - 20 kHz liegt. Der Frequenzbereich jenseits der 8 kHz wurde in der Audiometrie nicht überprüft.

Der ermittelte Sättigungsbereich bei der Bewertung der Bandpasskonditionen erklärt sich durch die Wahl des verwendeten Testmaterials. In Abbildung 5.4 auf Seite 33 ist das Hörfeld des Menschen abgebildet. Der Frequenzbereich der vom Menschen gesprochen Sprache liegt im Bereich von etwa 100 Hz - 8 kHz. Schaut man sich das Diagramm der Bandpässe Abbildung 5.3 im Vergleich zum Hörfeld an, zeigt sich der Vorteil von WB- zu NB-Übertragung sehr deutlich im steilen Anstieg. Beim weiteren Blick wird daraus auch deutlich, dass jenseits der WB-Übertragung es keinen weiteren Gewinn an Sprachinformationen gibt. Dadurch ist es nicht verwunderlich, dass die großen Bandbreiten nur noch eine geringere Verbesserung der Qualität erzielen.

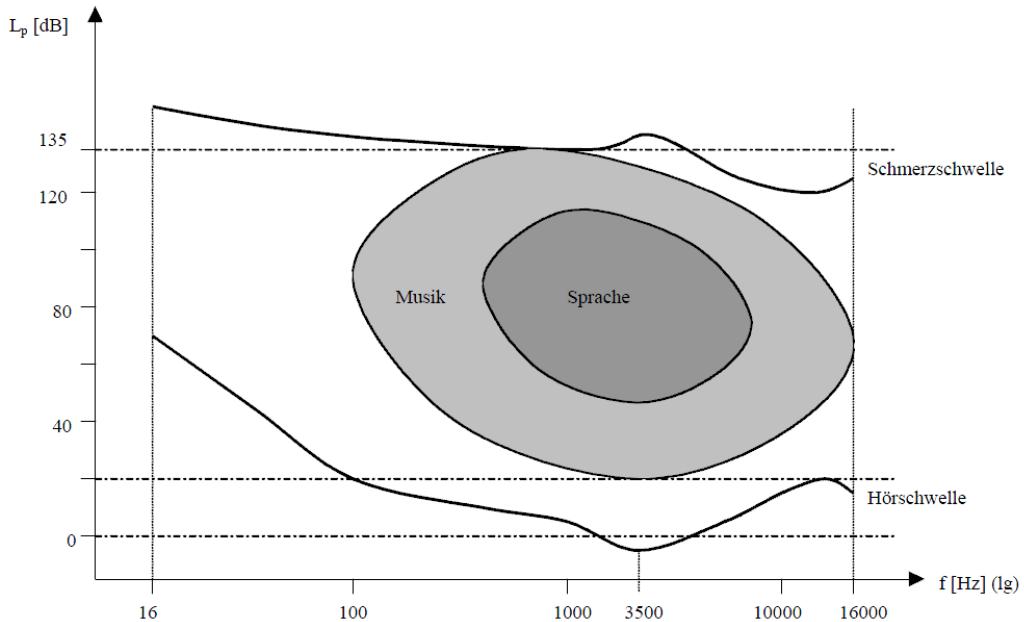


Abbildung 5.4: Hörfeld des Menschen [4]

### 5.2.2 Qualitätsverbesserung und E-Modell

Die Verwendung der großen Bandbreite von Qualitäten in dieser Studie hat einen Einfluss auf die Bewertung. Führt man eine Studie im WB-Kontext durch, werden die WB-Konditionen tendenziell sehr gut bewertet, da sie die obere Grenze der Qualität widerspiegeln. Wenn jedoch in einem Test zusätzlich SWB- und FB-Konditionen vorhanden sind, werden die selben WB-Konditionen schlechter bewertet. Denn im direkten Vergleich sind sie schlechter als die SWB-/FB-Konditionen. Die MOS-Werte der WB-Konditionen im SWB-/FB-Kontext sind somit geringer als die MOS-Werte der WB-Konditionen im WB-Kontext.

Um eine Qualitätsgewinn der FB-Übertragung gegenüber der WB-Übertragung zu ermitteln, werden die Ergebnisse aus dieser Studie mit den Ergebnissen aus der HiQ-Voice-Studie [8] mit Hilfe des E-Modells (siehe auch Abschnitt 3.1.2) verglichen. Im E-Modell erhält man aus den MOS-Werten ein einzigartiges Maß, den *R*-Wert. Dieser Wert gibt die Qualität an. Die *R*-Werte lassen sich direkt mit *R*-Werten aus anderen Studien vergleichen.

Vor der Berechnung der *R*-Werte müssen die, für jeden Kondition gemittelten, MOS-Werte normalisiert werden. Dafür wurde die nachstehende Formel 5.1 (entnommen aus [24], Seite 253) verwendet. Die Normalisierung setzt den, im Test erreichten, maximalen MOS-Wert auf den Skalenwert  $MOS_r =$

4.5. Somit wird die gesamte obere Skalenweite ausgenutzt.

$$MOS_{norm} = \frac{MOS_{test} - 1}{MOS_{test(maximal)} - 1} \cdot (MOS_r - 1) + 1 \quad (5.1)$$

Danach werden die normalisierten MOS-Werte auf die R-Skala [0 - 100] für Narrowband ( $R_{NB}$ ) transformiert. Die Formel 3.2 für die Berechnung steht in Abschnitt 3.1.2. Anschließend werden die  $R_{NB}$ -Werte durch eine lineare Expansion 5.2 auf die R-Skala des Wideband E-Modells ( $R_{WB}$ ) abgebildet. Dieser Vorgang ist notwendig, da die HiQ-Voice-Studie im WB-Kontext stattfand.

$$R_{WB} = 1.29 \cdot R_{NB} \quad (5.2)$$

Um die Verbesserung der Qualität zu ermitteln, wurden die Ergebnisse aus diesem Test mit den Ergebnissen des HiQ-Voice-Tests [8] verglichen. Dabei wurden nur die Konditionen verwendet, die in beiden Versuchen vorkamen. Die Ergebnisse sind in Tabelle 5.2 aufgelistet. Die Werte wurden in das Diagramm 5.5 übertragen. Im Diagramm sind die  $R$ -Werte aus dem HiQ-Voice-Test auf der Y-Achse aufgetragen, die  $R$ -Werte aus dem FB-Versuch auf der X-Achse. Anschließend wurden mittels Regression zwei Geraden errechnet. Die maximalen  $R$ -Werte wurden durch lineare Interpolation errechnet, indem in die Gleichungen der Regressionsgeraden der maximale Skalenwert von 129 aus der WB-E-Modell eingesetzt wurde. Aus dieser Rechnung ist ein  $R_{max1}$  durch den Koordinatenursprung und ein  $R_{max2}$  mit zwei Parametern hervorgegangen.

$$R_{max1} : f(x) = mx + 0 = 1,57x = 203$$

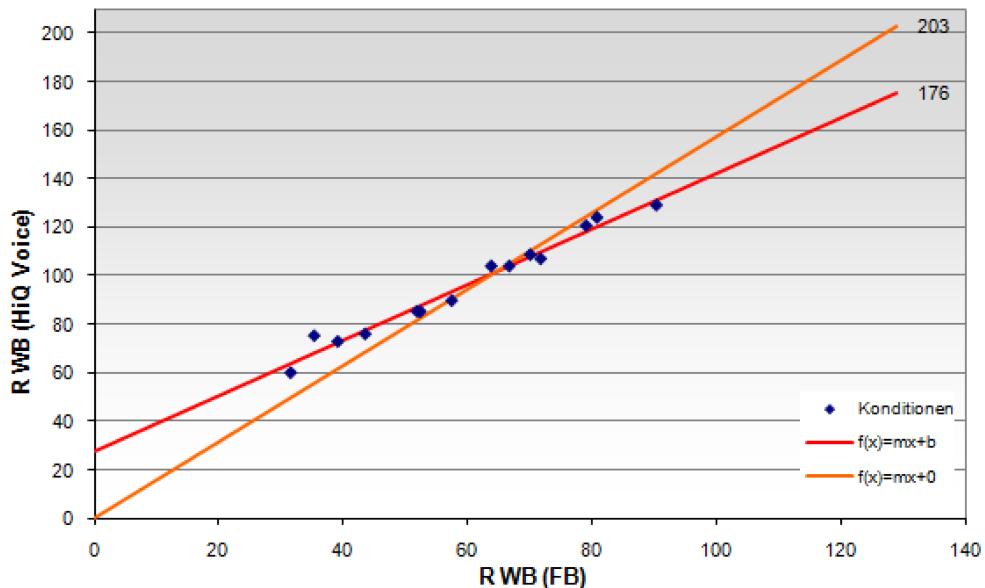
$$R_{max2} : f(x) = mx + b = 1,15x + 27,55 = 176$$

Dies führte zu einem  $R_{max}$ -Wert von 203 bzw. 176 und somit zu einem Qualitätsvorteil von 74 Punkten bei einem Skalenmaximum von 203 und 47 Punkten bei einem Maximum von 176. D.h., dass die R-Skala des E-Modells für den FB-Kontext auf ein Maximum von 203 bzw. 176 erweitert werden

müsste, um die Konditionen aus dem NB-, WB-, SWB- und FB-Kontext auf einer Skala abzubilden. Ein ähnliches Ergebnis wurde in einer anderen Studie ermittelt. In dieser Studie [28] wurde untersucht, welche Qualitätsverbesserung eine SWB-Übertragung im Vergleich zu einer WB-Übertragung ergibt und wie sich dieser Zugewinn auf das E-Modell auswirkt. Dort wurden  $R_{max,SWB}$  von 180 bzw. 174 ermittelt. Vergleicht man diese Ergebnisse, kommt man auch hier zu der Vermutung, dass trotz einer nochmaligen Erweiterung der Bandbreite es zu keiner weiteren Verbesserung kommt. Die Unterschiede im hochqualitativen Bereich sind, bezüglich der Qualität, so gering, dass sich keine Verbesserung der Gesamtqualität durch Vergrößerung der Bandbreite mehr erzielen lässt. Dennoch wird der Zugewinn an Qualität einer SWB-/FB-Übertragung im Vergleich zur WB-Übertragung deutlich.

<b>Kondition (bit-rate)</b>	$R_{WB}$ ( <b>HiQVoice</b> )	$R_{WB}$ ( <b>FB</b> )
G.711(64)	89,80	57,46
G.729A(8)	85,30	52,40
G.729A(8)8%FER	60,20	31,48
G.729A(8) x G.729A (8)	76,10	43,50
WB	129,00	90,46
direct wb car: 55 dB(A)	107,00	71,80
direct wb car: 70 dB(A)	85,40	51,88
G.722.2(23.05)	120,60	79,13
G.722.2(12.65)	108,70	70,11
G.722.2(6.6)	73,00	39,08
G.722(64)	124,00	80,87
G.722(64)2%FER	104,00	63,83
G.722(64)8%FER	75,40	35,27
G.722(64)xG.722.2(12.65)	104,00	66,74

Tabelle 5.2:  $R_{WB}$ -Werte aus dem HiQ Voice und dem FB-Test

Abbildung 5.5: Vergleich der  $R_{WB}$ -Werte (HiQ-Voice mit FB)

### 5.3 Vergleich ACR und MUSHRA

In diesem Abschnitt werden die Bewertungen direkt gegenübergestellt. Für die Auswertung wurden 38 Konditionen verwendet, da die AAC-LD[40k, 44k]-Konditionen nicht in beiden Hörtests vorkamen. Die einzelnen Urteile wurden vor der Mittelwertbildung auf die Skala [0 - 100] normalisiert. Dafür wurde ebenfalls die Formel 5.1 verwendet. Anschließend wurde für jede Kondition, für jeden Test separat, der Mittelwert und die Standardabweichung ( $Sdv$ ), sowie das Konfidenzintervall (CI 95%) berechnet. Die Werte sind in Tabelle 5.3 aufgelistet.  $Score_{norm}$  steht in der Tabelle für die Mittelwerte der normalisierten Bewertungen. Für eine bessere Vergleichbarkeit wurden die Werte zusammen mit ihren Konfidenzintervallen in das Diagramm 5.6 übertragen.

<b>Kondition</b>	<b>ACR</b>			<b>MUSHRA</b>		
	<i>Score<sub>norm</sub></i>	Sdv	CI95%	<i>Score<sub>norm</sub></i>	Sdv	CI95%
<i>AAC_LD[32k]</i>	72,75	28,25	5,65	66,81	21,28	4,66
<i>AAC_LD[48k]</i>	94,25	20,96	4,19	92,26	12,51	2,74
<i>AAC_LD[64k]</i>	100,06	18,04	3,61	95,36	10,22	2,24
<i>AMR-WBplus@13</i>	69,85	25,25	5,05	72,32	24,86	5,45
<i>AMR-WBplus@36</i>	92,51	20,11	4,02	90,47	14,51	3,18
<i>AMR-WBplus@6</i>	16,98	22,80	4,56	18,35	18,82	4,12
<i>bp_filter_300 – 7000Hz</i>	52,42	26,33	5,27	52,83	21,92	4,80
<i>bp_filter_50 – 10000Hz</i>	91,92	20,73	4,15	86,07	14,91	3,27
<i>bp_filter_50 – 12000Hz</i>	98,02	19,40	3,88	96,41	8,05	1,76
<i>bp_filter_50 – 3400Hz</i>	35,28	28,17	5,64	34,03	21,31	4,67
<i>bp_filter_50 – 8000Hz</i>	82,63	22,25	4,45	76,31	17,98	3,94
<i>bp_filter_600 – 7000Hz</i>	36,73	26,44	5,29	33,02	19,33	4,24
<i>CELT[32k]</i>	57,07	32,50	6,50	61,53	30,29	6,64
<i>CELT[40k]</i>	70,43	30,28	6,06	70,89	27,83	6,10
<i>CELT[48k]</i>	83,79	27,87	5,58	84,94	20,04	4,39
<i>CELT[64k]</i>	98,32	18,98	3,80	92,20	11,84	2,59
<i>direct_wb_car55dB(A)</i>	53,87	23,76	4,75	49,88	23,53	5,16
<i>direct_wb_car70dB(A)</i>	31,50	22,56	4,51	19,06	15,09	3,31
<i>FB_Org</i>	92,80	20,99	4,20	100,00	6,62	0,72
<i>G.711@64</i>	37,60	26,94	5,39	36,90	18,43	4,04
<i>G.722.2@12.65</i>	52,13	23,48	4,70	53,03	19,22	4,21
<i>G.722.2@23.05</i>	62,29	26,72	5,35	60,45	18,71	4,10
<i>G.722.2@6.6</i>	18,43	18,65	3,73	24,93	19,97	4,38
<i>G.722@64</i>	64,33	23,41	4,68	68,10	17,35	3,80
<i>G.722@64_2%FER</i>	44,86	24,53	4,91	54,21	22,26	4,88
<i>G.722@64_8%FER</i>	15,23	16,68	3,34	27,65	24,20	5,30
<i>G.722@64xG.722.2@12.65</i>	48,06	23,70	4,74	50,12	20,98	4,60
<i>G.729A@8</i>	32,08	24,23	4,85	31,15	20,40	4,47
<i>G.729A_8%FER</i>	12,04	16,92	3,38	17,30	14,93	3,27
<i>G.729AxG.729A</i>	22,79	21,10	4,22	25,95	16,58	3,63
<i>G.722_1Ann.C@24</i>	54,16	27,17	5,43	51,62	25,40	5,57
<i>G.722_1Ann.C@32</i>	66,94	28,36	5,67	66,26	25,11	5,50
<i>G.722_1Ann.C@48</i>	82,63	26,61	5,32	78,12	22,08	4,84
<i>NB</i>	40,80	28,82	5,77	38,82	20,15	2,21
<i>SILK[12k]</i>	59,68	26,06	5,21	65,40	22,45	4,92
<i>SILK[32k]</i>	88,73	20,62	4,13	89,88	12,82	2,81
<i>SWB</i>	98,61	18,54	3,71	93,39	14,10	3,09
<i>WB</i>	74,49	25,31	5,06	65,62	18,48	2,02

Tabelle 5.3: Vergleich der Ergebnisse des ACR- sowie des MUSHRA-Tests, beide Versuchsreihen wurden auf die Skala [0 - 100] normalisiert

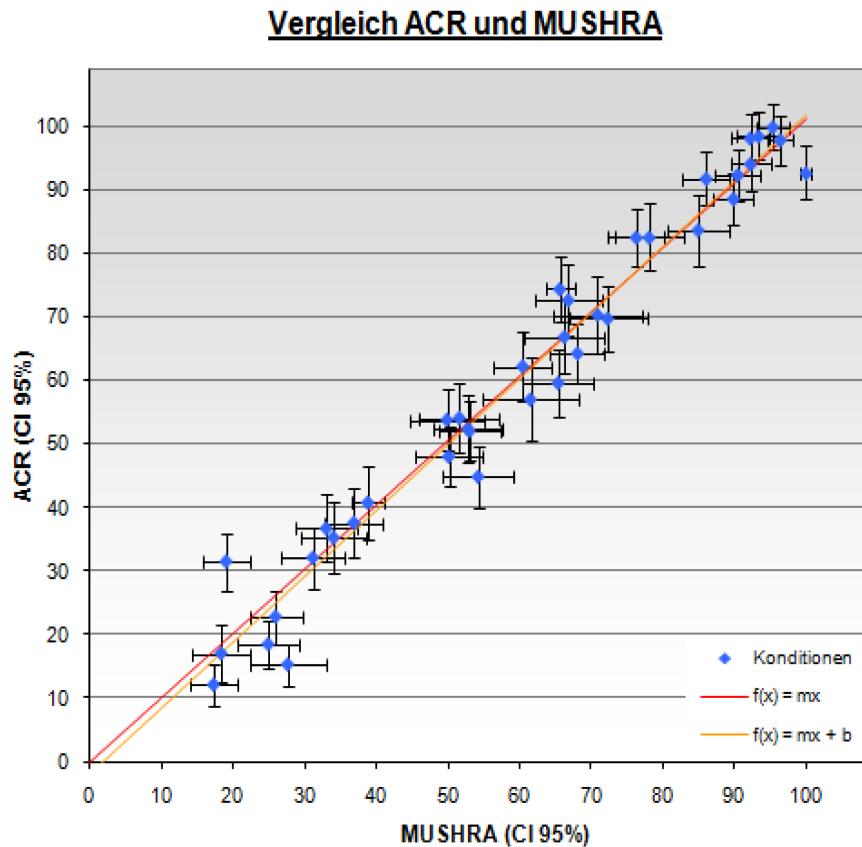


Abbildung 5.6: Gegenüberstellung der Konditionen aus ACR und MUSHRA und ihrer Konfidenzintervalle (CI 95%)

Aus dem Diagramm 5.6 ist deutlich ein linearer Zusammenhang zwischen der Beurteilung im ACR-Test und der Beurteilung aus dem MUSHRA-Test zu erkennen.

$$\text{Regressionskurve}_{\text{rot}} : f(x) = mx + 0 = 1,01x$$

$$\text{Regressionskurve}_{\text{orange}} : f(x) = mx + b = 1,04x - 2,15$$

Das Bestimmtheitsmaß  $R^2$  liegt bei 0,99 (rote Gerade) und bei 0,97 (orange Gerade). Schaut man sich die zugehörigen Konfidenzintervalle an, liegen bei 31 von 38 Konditionen die CI 95% aus dem MUSHRA-Versuch unterhalb der aus dem ACR-Versuch. Das ist wiederum ein Beleg für die bessere Auflö-

sung der Skala, sowie der Bewertungsmethode. Die Differenzen der Konfidenzintervalle sind jedoch sehr gering und liegen im Durchschnitt bei 0,77. Die größten Differenzen zwischen den Konfidenzintervallen findet man bei den Konditionen FB\_Org, NB und WB ( $\Delta FB_Org_{CI95\%} = 3,47$ ,  $\Delta NB_{CI95\%} = 3,56$ ,  $\Delta WB_{CI95\%} = 3,04$ ). Der Grund dafür ist, dass diese drei Konditionen 4-mal so häufig im MUSHRA-Test bewertet wurden, da sie als Referenz bzw. Anker in jedem Block vorkamen. Trotz der kleinen Unterschiede kann man generell die Vertrauensbereiche beider Test als vergleichbar bezeichnen.

Die Ergebnisse zeigen klar, dass ACR und MUSHRA annährend identische Bewertungen für Sprachsignale liefern. Vergleicht man die Methoden nicht nur an Hand der Qualitätsurteile der Probanden, sondern nimmt auch den Aufwand für eine Durchführung mit in Betracht, so besteht bei der Vorbereitung (Erstellen der Playlisten) kein Unterschied. Das MUSHRA-Verfahren benötigt eine intensivere Einweisung und Trainingsphase als ein ACR-Test. Demnach ist das ACR-Verfahren leichter und schneller durchführbar. Gerade hinsichtlich des Kostenfaktors, den ein Hörtest darstellt, ist ein schnell durchführbarer Test, der zeitgleich valide Ergebnisse produziert von Vorteil.

# Kapitel 6

## Zusammenfassung

In dieser Arbeit sind zwei subjektive Qualitätsbeurteilungsmethoden verwendet worden. Bei den zur Beurteilung stehenden Signalen handelte es sich um Sprachproben, die mit Hilfe verschiedener Sprach- sowie Audiokodierern bearbeitet wurden. Ein Überblick über die verwendeten Proben, sowie ihre Erstellung ist in der Arbeit zu finden, ebenso die Dokumentation der Durchführung und der Ergebnisse. Dabei wurden die Ergebnisse ausgewertet und mit einander in Verbindung gebracht. Außerdem wurden für den ACR-Test die entsprechenden R-Werte des E-Modells berechnet, um eine eventuelle Verbesserung der Qualitätsbeurteilung im Vergleich zum WB- und SWB-Kontext zu ermitteln.

Ein Vergleich der MOS-Werte (FB-Kontext) mit MOS-Werten aus älteren Studien [8] (WB-Kontext) bestätigt den Vorteil einer Bandbreitenvergrößerung von WB zu SWB [28] in einer Größenordnung von rund 35%. Außerdem ist zu erkennen, dass eine weitere Vergrößerung (größer als SWB) keine nennenswerte Verbesserung mehr hervorbringt. Jedoch bedarf es weiterer gezielter Untersuchungen auf dem Gebiet der Übertragungsbandbreiten (größer als WB), um die ermittelte Erweiterung des E-Modells in dieser Arbeit zu stützen.

Die beobachteten Effekte im ACR-Versuch entsprachen im Großen und Ganzen denen im MUSHRA-Versuch. Der Vergleich beider Verfahren zeigt einen starken linearen Zusammenhang. Das stellt den Beweis dar, dass der ACR die Unterschiede über die gesamte Qualitätsbreite, im Vergleich zum MUSHRA, annährend gleich ermitteln kann. Diese Schlussfolgerung deckt sich mit den Ergebnissen aus der Studie von Feiten [3] und der Studie der NTT [27], in der jeweils der ACR-Test als „Sieger“ hervorgegangen ist. In der aktuellen Arbeit zeigte sich auch der Vorteil, dass sich der ACR-Test zügig

und problemlos durchführen lässt. Damit verbunden besitzt der ACR-Test auch ein besseres Kosten-Nutzen-Verhältnis, so dass der ACR-Test auch hier als Methode empfohlen wird. Dennoch zeigt die Studie die bessere Auflösung des MUSHRA-Tests. Ein weiterer Vorteil des MUSHRA-Tests ist die größere Interaktion der Probanden mit der Test-GUI. Daraus resultiert eine größere Aufmerksamkeit der Probanden und verhindert die im ACR-Test eher bestehende Gefahr „des Durchklickens“ bei Ermüdung.

Beantwortung der anfangs gestellten Fragen:

- *Verändert sich die Beurteilung zur Gesamtqualität, wenn das eine oder andere Verfahren verwendet wird?*

Die Beurteilungen sind annährend identisch.

- *Eignet sich eines der Verfahren besser um die Gesamtqualität zu ermitteln?*

Beide Verfahren eignen sich hervorragend zur Ermittlung der Gesamtqualität. Liegt ein Augenmerk auf dem Zeitfaktor, so empfiehlt sich der ACR-Test.

- *Kann man die Spanne der Bedingungen überhaupt sinnvoll abbilden?*

Ja, trotz zwei verschiedener Methoden wurde ein annährend identisches Ergebnis erzielt.

- *Verbessert sich die Gesamtqualität weiter mit der nochmaligen Vergrößerung der Übertragungsbandbreite?*

Die Ergebnisse aus den Versuchen lassen den Schluss zu, dass eine Übertragung jenseits von SWB keinen nennenswerten Qualitätsgewinn erzielt.

Um die Ergebnisse aus dieser Studie zu bestätigen, sind weiteren Untersuchungen nötig und geplant. Dabei sollten veränderte Versuchsbedingungen berücksichtigt werden (wie z.B.: Laborraum, Versuchspersonen, Kodierer, Sprache des Testmaterials). In folgenden Studien wäre es interessant, verschiedene Audiosignale zu integrieren, da in dieser Arbeit nur reine Sprachsignale verwendet wurden. Dabei könnte es sich zum Beispiel um Musik, Sprache vor Hintergrundmusik, Audioaufnahmen von Filmsequenzen etc. handeln. Die Vermutung liegt nahe, dass auch im Zusammenhang von unterschiedlichen Audiosignalen die Qualitätsresultate des ACR sowie des MUSHRA sehr ähnlich sind. Diese Vermutung wird zusätzlich durch die zuvor erwähnten Studien aus dem Audio- Videobereich gestützt. Weiterhin sollte der beobachtete Sättigungsbereich der Bewertungen der Bandbreitenkonditionen

gesondert untersucht werden. Es ist vorstellbar, dass dieser Effekt nur bei reinen Sprachsignalen auftritt und im Falle von Musikübertragungen die Qualitätsbewertung mit vergrößerter Bandbreite steigt. In diesem Zusammenhang kam die Frage auf, wie sich verschiedene Hörminderungen auf die Qualitätsbeurteilung auswirken. Es lässt sich vermuten, dass diese „wachsende“ Bevölkerungsgruppe der Hörbeeinträchtigten keinen wirklichen Vorteil einer erweiterten Übertragungsbandbreite hat. Weiter interessant ist auch, ob sich die Klangvorliebe mit dem Einzug der digitalen Audiokodierung in den Alltag verändert hat. Ein Indiz dafür wurde im ACR-Test gefunden, jedoch im MUSHRA-Test nicht bestätigt.

# Abbildungsverzeichnis

1	Titelbild: Corti - Organ, siehe Literaturverzeichnis [1] . . . . .	1
3.1	Weg der Schallaufnahme beim Menschen (Grundlage der Abbildung aus [31]), erstellt von Falk Schiffner . . . . .	9
4.1	Versuchsaufbau (Raum: „Calypso“), Foto by F.Schiffner 2010-04-01 . . . . .	14
4.2	Audiometer „ear1.7“, Foto by F.Schiffner 2010-04-01 . . . . .	17
4.3	Schema zur Generierung der Sprachbeispiele . . . . .	20
4.4	Screenshot, Eingabefenster ACR-Test . . . . .	24
4.5	Screenshot, Eingabefenster MUSHRA-Test . . . . .	25
5.1	ACR Resultate der Versuchskonditionen . . . . .	28
5.2	MUSHRA Resultate der Versuchskonditionen . . . . .	29
5.3	Sättigungsverhalten der Bandpässe mit unterer Grenzfrequenz $f_u = 50\text{Hz}$ . . . . .	32
5.4	Hörfeld des Menschen entnommen aus Skript [4] Seite 37 . . . . .	33
5.5	Vergleich der $R_{WB}$ -Werte aus HiQ-Voice und FB . . . . .	36
5.6	Vergleich der Ergebnisse des ACR- sowie des MUSHRA-Tests . . . . .	38

# Tabellenverzeichnis

4.1	Zeitplan Hörversuch . . . . .	14
4.2	Zusammensetzung der Versuchspersonen . . . . .	16
4.3	Übersicht der verwendeten Konditionen . . . . .	22
5.1	Übersicht der Ergebnisse des ACR- sowie des MUSHRA-Tests	27
5.2	$R_{WB}$ - Werte aus dem HiQ-Voice und dem FB-Test . . . . .	35
5.3	Vergleich der Ergebnisse des ACR- sowie des MUSHRA-Tests	37

# Abkürzungsverzeichnis

$f_u$ .....	Untere Grenzfrequenz
$I_i$ .....	Beeinträchtigungsfaktoren
$R_0$ .....	Basisqualität
A .....	E - Modell Vorteilsfaktor
AAC-LD .....	Advanced Audio Coding - Low Delay
ACELP .....	Algebraic Code Excited Linear Prediction
ACR .....	Absolute Category Rating
AMR-WB .....	Adaptive Multi-Rate Wideband
CELT .....	Constrained Energy Lapped Transform
CI 95% .....	95% Konfidezintervall
CS - ACELP .....	Conjugate-Structure Algebraic-Code-Excited Linear Prediction
dB .....	Dezibel
dB(A) .....	Dezibel A - bewertet
dBov .....	Dezibel overload
DVD .....	Digital Versatile Disc
FB .....	Fullband
FER .....	Frame Erasure Rate
FIR .....	Finite Impulse Response (Filter)
GUI .....	Graphical User Interface
HiQ-Voice .....	High Quality - Voice (Projekt)
IPTV .....	Internet Protokoll Television
ISDN .....	Integrated Services Digital Network
ITU - R .....	International Telecommunication Union Radiocommunications Sector
ITU - T .....	International Telecommunication Union Telecommunications Standardisation Sector
KVM .....	Keyboard Video Mouse
LOT .....	Listening Only Test
Mean .....	Mittelwert
MOS .....	Mean Opinion Score

MS - DOS .....	Microsoft - Disc Operating System
MUSHRA .....	MULTi Stimulus with Hidden Reference and Anchor
NB .....	Narrowband, Schmalband
NTT .....	Nippon Telegraph and Telephone Corporation
PCM .....	Puls Code Modulation
RT .....	Reverberation Time, Nachhallzeit
SB - ADPCM ....	Sub-Band Adaptive Differential Pulse Code Modulation
Sdv .....	Standard Deviation, Standardabweichung
SPL .....	Sound Pressure Level
SWB .....	Super-Wideband, Super-Breitband
TCX .....	Transform Coded Excitation
VHS .....	Video Home System
VoIP .....	Voice over Internet Protokoll
WB .....	Wideband, Breitband

# Literaturverzeichnis

- [1] *Corti - Organ.* Internet, URL: [www.teknat.uu.se/forskuing/uu/beskrivning.php?vetenskapsik=&forskomr=62&id=379&lang=en](http://www.teknat.uu.se/forskuing/uu/beskrivning.php?vetenskapsik=&forskomr=62&id=379&lang=en). letzter Besuch 05/2010.
- [2] 3RD GENERATION PARTNERSHIP PROJEKT: *ANSI-C code for the Fixed-point Extended AMR Wideband codec*, Mai 2004.
- [3] BERNHARD FEITEN, ALEXANDER RAAKE, MARIE-NEIGE GRACIA, ULF WÜSTENHAGEN AND JENS KROLL: *Subjective Quality Evaluation of Audio Streaming Applications on Absolute and Paired Rating Scales*. Convention Paper 7787, 126th AES Convention, GER-Munich, 2009.
- [4] FISCHER, J.: *Audiotechnik Skript*. pdf - Skript, 1999. Skript basiert auf den Mitschriften der Vorlesung „Audiodarstellungstechnik“ von Prof. Stollenwerk.
- [5] GERHÄUSER, H.: *MPEG-4 AAC-LD Low-Delay High-Quality Audio Coding*. Fraunhofer Institute for Integrated Circuits IIS, 2007.
- [6] GIBBON, D.: *German Speech Database*. ESPRIT Projekt 2589 Report (SAM, Multi-Lingual Speech Input/Output Assessment, Methodology and Standardization, University of Bielefeld, D-Bielefeld, 1992).
- [7] ITU-R REC. BS.1534: *Method for the subjective assessment of intermediate quality level of coding systems*. International Telecommunication Union, CH-Geneva, 2001.
- [8] ITU-T CONTRIBUTION COM 12 - C23 -E: *Provisional planning values for the wideband Equipment Impairment Factor*. International Telecommunication Union, CH-Geneva, 2009.
- [9] ITU-T REC. G.107: *The E-model, a computational model for use in transmission planning*. International Telecommunication Union, CH-Geneva, 2005.

- [10] ITU-T REC. G.711: *Pulse Code Modulation of Voice Frequencies*. International Telecommunication Union, CH-Geneva, 1988.
- [11] ITU-T REC. G.722: *7 kHz Audio-Coding Within 64 kbit/s*. International Telecommunication Union, CH-Geneva, 1988.
- [12] ITU-T REC. G.722.2: *Wideband coding of speech at around 16 kbit/s using Adaptive Multi-Rate Wideband (AMR-WB)*. International Telecommunication Union, CH-Geneva, 2003.
- [13] ITU-T REC. G.722.2 ANNEX C: *Wideband coding of speech at around 16 kbit/s using Adaptive Multi-Rate Wideband (AMR-WB)*. International Telecommunication Union, CH-Geneva, 2004.
- [14] ITU-T REC. G.729 ANNEX A: *Coding of speech at 8 kbit/s using conjugate structure algebraic-code-excited linear-prediction (CS-ACELP)*. International Telecommunication Union, CH-Geneva, 1996.
- [15] ITU-T REC. P.191: *Software Tools for Speech and Audio Coding Standardization*. International Telecommunication Union, CH-Geneva, 2000.
- [16] ITU-T REC. P.341: *Transmission Characteristics for Wideband (150-7000 Hz) digital hands-free Telephony Terminal*. International Telecommunication Union, CH-Geneva, 1996.
- [17] ITU-T REC. P.800: *Methods for Subjective Determination of Transmission Quality*. International Telecommunication Union, CH-Geneva, 1996.
- [18] ITU-T REC. P.830: *Subjective Performance Assessment of Telephone-Band and Wideband digital Codecs*. International Telecommunication Union, CH-Geneva, 1996.
- [19] JEKOSCH, U.: *Voice and Speech Quality Perception: Assessment and Evaluation*. Springer-Verlag, D-Berlin, 2005.
- [20] LU HUO, MARCEL WÄLTERMANN, ULRICH HEUTE AND SEBASTIAN MÖLLER: *Estimation of Speech Quality Dimension „Discontinuity“*. ITG-Fachtagung Sprachkommunikation, D-Aachen, 2008.
- [21] MARCEL WÄLTERMANN, KRISTIN SCHOLZ, ALEXANDER RAAKE, ULRICH HEUTE SEBASTIAN MÖLLER: *Underlying Quality Dimensions of Modern Telephone Connections*. in Proc. 9th International Conference on Spoken Language Processing (ICSLP), USA-Pittsburgh, 2006.

- [22] MÖLLER, S.: *Assessment and Prediction of Speech Quality in Telecommunications*. Kluwer Academic Publisher, NL-Dordrecht, 2000.
- [23] Q. HUJNH-THU, MARIE-NEIGE GARCIA, F. SPERANZA, P.-J. CORRIVEAU AND ALEXANDER RAAKE: *Study of Rating Scales for Subjective Quality Assessment of High-Definition Video*. IEEE Transactions on Broadcasting, 2010.
- [24] RAAKE, A.: *Speech Quality of VoIP: Assessment and Prediction*. John Wiley and Sons, UK-Chichester, 2006.
- [25] SCHULZ VON THUN: *Das Kommunikationsquadrat*. URL: <http://www.schulz-von-thun.de/mod-komquad.html>. letzter Besuch 05/2010.
- [26] SENGPIEL, D. E.: *Berechnen der Summen-Pegel von zwei kohärenten Quellen*. URL: <http://www.sengpielaudio.com/Rechnerkohquellen.htm>. letzter Besuch 05/2010.
- [27] T. TOMINAGA, T. HAYASHI, J. OKAMOTO, AND A. TAKAHASHI: *Performance comparisons of subjective quality assessment methods*. Nippon Telegraph and Telephone Corporation (NTT), 2010.
- [28] TUCKER, I.: *E-Model Scale Extension for Super-Wideband Speech Quality*. Bachelor Thesis, 2009. Technische Universität, D-Berlin.
- [29] VALIN, J.-M.: *CELT: A low-latency, High-quality Audio Codec*. Xiph.Org Foundation, 2009.
- [30] Vos, K.: *SILK Speech Codec*. Internet-Draft, URL: <http://tools.ietf.org/html/draft-vos-silk-01>, März 2010. letzter Besuch 06/2010.
- [31] ZWICKER, EBERHARD AND FASTL, HUGO: *Psychoacoustics: Facts and Models*. Springer-Verlag, 2007.



# Anhang A

## Testeinweisungen

### A.1 ACR

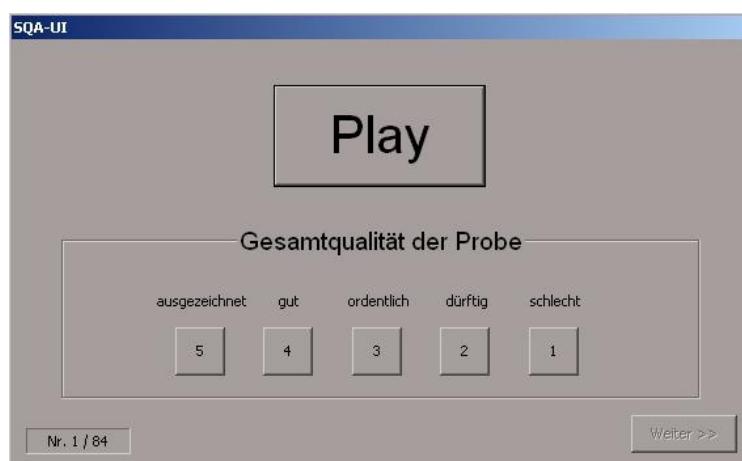
**Hörversuch:**  
**Beurteilung der Gesamtqualität von Hörproben**  
**bei Telefonübertragung**

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

Sie nehmen heute an einem Hörversuch teil, bei dem Sie die **Gesamtqualität** von Hörproben bewerten sollen, die über Telefonstrecken übertragen wurden.

Sie werden im Folgenden unterschiedliche Sprachproben hören, die hintereinander abgespielt werden. Jeder Durchgang enthält unterschiedliche Sätze, die von vier verschiedenen Sprechern gesprochen werden. In fünf verschiedenen Abstufungen geben Sie an, wie Ihr persönlich wahrgenommener **Gesamteindruck** bezogen auf die **Qualität** ist.

Die beschreibenden Adjektive und Zahlen auf den Schalttafeln sollen Sie bei Ihrer Urteilsfindung unterstützen. Das untenstehende Bild zeigt das Eingabefenster:



Vor Beginn des eigentlichen Versuchs findet eine Trainingsphase statt, die Ihnen einen Überblick über die im Versuch vorkommenden Störungen und Qualitäten verschafft. Zudem erlaubt Ihnen die Trainingsphase, sich mit der

Beurteilungsaufgabe und dem Computerprogramm vertraut zu machen. Durch einen Klick mit der Maus auf den mit „Play“ gekennzeichneten Knopf können Sie die Wiedergabe einer Hörprobe jederzeit wiederholen, bevor Sie Ihr Urteil abgeben.

Nehmen Sie die Beurteilung bitte intuitiv vor. Bei dieser Untersuchung, die rein subjektiver Natur ist, gibt es weder richtige noch falsche Antworten. Ausschließlich Ihr persönlicher Eindruck ist für die Untersuchung wichtig.

Um nach jeder Sprachprobe die nächste anzuhören, müssen Sie zunächst Ihre Bewertung abgeben. Wenn Sie auf „Weiter“ gedrückt haben (unten rechts), beginnt sofort das Abspielen der nächsten Sprachprobe.

Sollten Sie während des Versuchs Schwierigkeiten haben, melden Sie sich bitte beim Versuchsleiter.

**Ablauf:**

- Einweisung
- Training + Durchgang 1
- 5min Pause
- Training + Durchgang 2
- Audiometrie

## A.2 MUSHRA

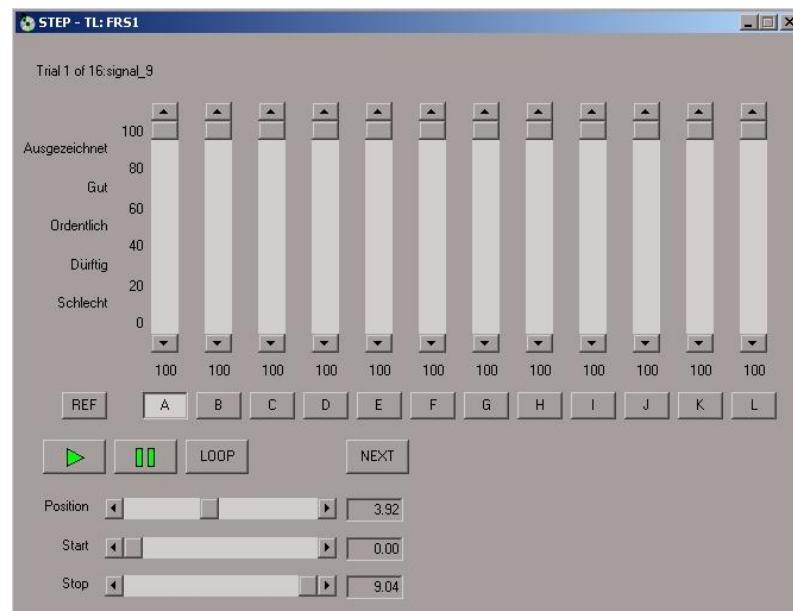
**Hörversuch:**  
**Beurteilung der Gesamtqualität von Hörproben**  
**bei Telefonübertragung**

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

Sie nehmen heute an einem Hörversuch teil, bei dem Sie die **Gesamtqualität** von Hörproben bewerten sollen, die über Telefonstrecken übertragen wurden.

Sie werden im Folgenden unterschiedliche Sprachproben hören. Auf einer Skala von 0 (schlechteste Qualität) bis 100 (beste Qualität) geben Sie an, wie Ihr persönlich wahrgenommener **Gesamteindruck** der einzelnen Probe bezogen auf die **Qualität** ist.

Die beschreibenden Adjektive und Zahlen neben den Schaltnäpfen sollen Sie bei Ihrer Urteilsfindung unterstützen, jedoch können Sie Ihre Bewertung frei auf der Skala setzen. Das untenstehende Bild zeigt das Eingabefenster:



Durch Anklicken der mit den Buchstaben (A - L) gekennzeichneten Schaltflächen, starten Sie die jeweilige Sprachprobe. Sie können zwischen den unterschiedlichen Qualitäten wechseln, während ein Satz abgespielt wird. Die Schaltfläche, die mit „REF“ gekennzeichnet ist, dient als „Referenz“, also die Sprachprobe in optimaler Qualität. Sie können die Referenz und alle weiteren Sprachproben in beliebiger Reihenfolge und so oft anhören, wie Sie zur Beurteilung benötigen. Es ist durchaus möglich, dass Sie eine oder mehrere Probe mit maximaler Punktzahl (100) bewerten, wenn Sie keinen Qualitätsunterschied zur Referenz hören.

Über den Schalttafeln befinden sich die zugehörigen Schieberegler, mit denen Sie Ihre Bewertung vornehmen können. Sie haben weiterhin die Möglichkeit, über den Knopf „LOOP“ die Hörprobe als Endlosschleife anzuhören. Mit dem Knopf, der das Symbol für Pause zeigt, können Sie die Endlosschleife anhalten. Der Knopf mit dem Symbol für Abspielen startet die Schleife wieder. Haben Sie alle Hörproben in einem Durchgang bewertet, kommen Sie über die Knopf „NEXT“ zum nächsten Durchgang.

Vor Beginn des eigentlichen Versuchs findet eine Trainingsphase statt, die Ihnen einen Überblick über die im Versuch vorkommenden Störungen und Qualitäten verschafft. Zudem erlaubt Ihnen die Trainingsphase, sich mit der Beurteilungsaufgabe und dem Computerprogramm vertraut zu machen. Dabei bewerten Sie nur zwei kürzere Durchgänge.

Der Test ist in zwei Blöcke mit jeweils 8 Durchgängen unterteilt. Nach der Bewertung des 8. Durchgangs (Zähler in der linken oberen Ecke zeigt „Trial 9 of 16“ an), ist eine kurze Pause vorgesehen. Bitte melden Sie sich dann beim Versuchsleiter.

Nehmen Sie die Beurteilung bitte intuitiv vor. Bei dieser Untersuchung, die rein subjektiver Natur ist, gibt es weder richtige noch falsche Antworten. Ausschließlich Ihr persönlicher Eindruck ist für die Untersuchung wichtig.

Sollten Sie während des Versuchs Schwierigkeiten haben, melden Sie sich bitte beim Versuchsleiter.

**Ablauf:**

- Einweisung
- Training
- Block 1 (Trial 01 - 08)
- 5 min. Pause
- Block 2 (Trial 09 - 19)
- Audiometrie

Dauer: ca. 1 Stunde



# Anhang B

## Batchfile

```

echo *
echo *
echo ***** End of preprocessing *****
echo *
echo *
echo *
echo ****
echo *
echo ***** Start Coding SILK *****

REM Coding with SILK Codec
Call :SILK

echo ****
echo **** End of Coding ****
echo *
echo *
echo *
echo **** Post-Processing ****
cd SILK_Fin-Dateien
IF NOT EXIST .\PostprocessSILK\ ( MKDIR .\PostprocessSILK\ )
FOR %%s IN ( * ) DO (
    REM HQ2 up filtering
    filter -q -up HQ2 %%s _tmp1 640
    REM Level adjustment to -26dBov
    sv56demo -q -lev -26 -sf 48000 _tmp1 tmp 1920
    copy tmp .\PostprocessSILK\%%s
    echo o | del tmp
    echo o | del _tmp*
    cd PostprocessSILK
    REN %%s SILK[32]_%%s
    cd..
)
cd.. cd.. cd..

echo ****
echo **** End of post-processing ****
echo ****

goto :eof

REM Coding Condition SILK
:SILK
    cd SILK-Dateien
    IF NOT EXIST .\SILK_Fin-Dateien\ (MD.\SILK_Fin-Dateien\ )
FOR %%s IN ( * ) Do (
    silk_encoder %%s tmpA.bit -rate 32000

```

```
silk_decoder tmpA.bit tmpB  
copy tmpB \SILK_Fin-Dateien\%%$  
echo o | del tmp*  
)  
goto :eof  
:end
```



# Anhang C

## Playlisten

### C.1 ACR

```
:TB
Soundfiles\AMR-WBplus_36_m_ar02.wav
Soundfiles\c108_w_cd09.wav
Soundfiles\NB_G711_m_ss017.wav
Soundfiles\WB_w_nsb025.wav
Soundfiles\c110_m_ar03.wav
Soundfiles\SWB_w_cd012.wav
:TE
Soundfiles\AAC_LD[32k]_w_nsb031.wav
Soundfiles\CELT[48k]_m_ar010.wav
Soundfiles\c110_w_nsb039.wav
Soundfiles\AMR-WBplus_6_m_ss023.wav
Soundfiles\WB_m_ar09.wav
Soundfiles\AAC_LD[48k]_w_nsb036.wav
Soundfiles\bp_300_7k_w_cd035.wav
Soundfiles\c043_w_cd012.wav
Soundfiles\CELT[32k]_w_nsb033.wav
Soundfiles\WB_G722m_ss017.wav
Soundfiles\AAC_LD[44k]_m_ss026.wav
Soundfiles\CELT[40k]_w_nsb032.wav
Soundfiles\CELT[32k]_m_ar05.wav
Soundfiles\AAC_LD[44k]_w_nsb027.wav
Soundfiles\HHH_m_ar013.wav
Soundfiles\WB_G722_2[2305]m_ss030.wav
Soundfiles\SWB_w_nsb035.wav
Soundfiles\c001_w_nsb035.wav
Soundfiles\bp_50_8000_m_ar011.wav
Soundfiles\c114_w_nsb040.wav
Soundfiles\bp_50_10000_w_cd015.wav
Soundfiles\AMR-WBplus_13_m_ar015.wav
Soundfiles\bp_50_8000_w_cd022.wav
Soundfiles\G7222_1_32_m_ss032.wav
Soundfiles\SILK[12]_w_cd012.wav
Soundfiles\WB_w_nsb030.wav
Soundfiles\NB_w_cd033.wav
Soundfiles\bp_50_3400_w_nsb038.wav
Soundfiles\HHH_w_cd036.wav
Soundfiles\SWB_m_ss022.wav
Soundfiles\WB_G722_2[1265]w_nsb039.wav
Soundfiles\c014_m_ar07.wav
Soundfiles\KKK_m_ar012.wav
Soundfiles\G7222_1_24_m_ar06.wav
Soundfiles\AMR-WBplus_13_w_nsb038.wav
Soundfiles\AAC_LD[48k]_m_ss020.wav
Soundfiles\SILK[12]_m_ar03.wav
Soundfiles\WB_G722_2[2305]w_cd036.wav
Soundfiles\II_m_ar08.wav
Soundfiles\AAC_LD[44k]_w_nsb025.wav
Soundfiles\c043_m_ss018.wav
Soundfiles\NB_m_ss019.wav
Soundfiles\c108_m_ss028.wav
Soundfiles\G7222_1_48_w_nsb039.wav
Soundfiles\AMR-WBplus_36_m_ss024.wav
Soundfiles\CELT[48k]_w_nsb028.wav
Soundfiles\c014_w_cd014.wav
Soundfiles\bp_300_7k_m_ss024.wav
Soundfiles\c001_m_ar015.wav
Soundfiles\AAC_LD[32k]_m_ss021.wav
Soundfiles\AAC_LD[64k]_m_ss021.wav
Soundfiles\bp_50_10000_m_ar013.wav
```

Soundfiles\c110\_m\_ar04.wav  
Soundfiles\WB\_G722\_2[1265]m\_ar016.wav  
Soundfiles\c108\_w\_cd033.wav  
Soundfiles\G722\_1\_48\_m\_ss017.wav  
Soundfiles\SILK[32]\_w\_cd014.wav  
Soundfiles\CELT[64k]\_w\_nsb030.wav  
Soundfiles\NB\_G711\_m\_ar016.wav  
Soundfiles\SILK[32]\_m\_ss018.wav  
Soundfiles\NB\_G711\_w\_cd021.wav  
Soundfiles\G722\_1\_32\_w\_cd035.wav  
Soundfiles\bp\_50\_3400\_m\_ar05.wav  
Soundfiles\AMR-WBplus\_6\_w\_nsb031.wav  
Soundfiles\c114\_m\_ss025.wav  
Soundfiles\KKK\_w\_nsb040.wav  
Soundfiles\bp\_50\_12000\_m\_ss027.wav  
Soundfiles\WB\_G722w\_cd033.wav  
Soundfiles\bp\_600\_7k\_w\_cd016.wav  
Soundfiles\CELT[64k]\_m\_ar011.wav  
Soundfiles\bp\_600\_7k\_m\_ss019.wav  
Soundfiles\FB\_w\_nsb040.wav  
Soundfiles\AMR-WBplus\_36\_w\_nsb026.wav  
Soundfiles\FB\_m\_ss027.wav  
Soundfiles\II1\_w\_nsb033.wav  
Soundfiles\bp\_50\_12000\_w\_nsb025.wav  
Soundfiles\G722\_1\_24\_w\_nsb025.wav  
Soundfiles\CELT[40k]\_m\_ar014.wav

## C.2 MUSHRA

```

session=mushra
#signal_1
./Soundfiles/FB_w_nsb025.wav REF
./Soundfiles/FB_w_nsb025.wav HR
./Soundfiles/NB_W_nsb025.wav LP35
./Soundfiles/WB_W_nsb025.wav LP70
./Soundfiles/WB_G722_2[1265]W_nsb025.wav
./Soundfiles/bp_50_8000_w_nsb025.wav
./Soundfiles/bp_50_12000_w_nsb025.wav
./Soundfiles/bp_300_7k_w_nsb025.wav
./Soundfiles/CELT[48k]_W_nsb025.wav
./Soundfiles/AAC_LD[40k]_W_nsb025.wav
./Soundfiles/WB_G722_2[2305]w_nsb025.wav
./Soundfiles/WB_G722W_nsb025.wav
./Soundfiles/c014_w_nsb025.wav

#signal_2
./Soundfiles/FB_m_ss017.wav REF
./Soundfiles/FB_m_ss017.wav HR
./Soundfiles/NB_m_ss017.wav LP35
./Soundfiles/WB_m_ss017.wav LP70
./Soundfiles/c114_m_ss017.wav
./Soundfiles/c110_m_ss017.wav
./Soundfiles/c014_m_ss017.wav
./Soundfiles/WB_G722_2[1265]m_ss017.wav
./Soundfiles/SILK[12]_m_ss017.wav
./Soundfiles/AMR-WBplus_36_m_ss017.wav
./Soundfiles/AMR-WBplus_6_m_ss017.wav
./Soundfiles/CELT[32k]_m_ss017.wav
./Soundfiles/G7222_1_48_m_ss017.wav

#signal_3
./Soundfiles/FB_m_ar02.wav REF
./Soundfiles/FB_m_ar02.wav HR
./Soundfiles/NB_m_ar02.wav LP35
./Soundfiles/WB_m_ar02.wav LP70
./Soundfiles/AAC_LD[48k]_m_ar02.wav
./Soundfiles/G7222_1_48_m_ar02.wav
./Soundfiles/c001_m_ar02.wav
./Soundfiles/AAC_LD[64k]_m_ar02.wav
./Soundfiles/SILK[32]_m_ar02.wav
./Soundfiles/bp_300_7k_m_ar02.wav
./Soundfiles/AMR-WBplus_13_m_ar02.wav
./Soundfiles/II_m_ar02.wav
./Soundfiles/WB_G722_2[2305]m_ar02.wav

#signal_4
./Soundfiles/FB_w_cd09.wav REF
./Soundfiles/FB_w_cd09.wav HR
./Soundfiles/NB_w_cd09.wav LP35
./Soundfiles/WB_w_cd09.wav LP70
./Soundfiles/G7222_1_32_w_cd09.wav
./Soundfiles/c110_w_cd09.wav
./Soundfiles/c043_w_cd09.wav
./Soundfiles/CELT[64k]_w_cd09.wav
./Soundfiles/bp_50_10000_w_cd09.wav
./Soundfiles/AAC_LD[64k]_w_cd09.wav
./Soundfiles/c043_w_cd09.wav
./Soundfiles/III_w_cd09.wav
./Soundfiles/c114_w_cd09.wav

```

```

#signal_5
./Soundfiles/FB_m_ar03.wav REF
./Soundfiles/FB_m_ar03.wav HR
./Soundfiles/NB_m_ar03.wav LP35
./Soundfiles/WB_m_ar03.wav LP70
./Soundfiles/CELT[40k]_m_ar03.wav
./Soundfiles/CELT[32k]_m_ar03.wav
./Soundfiles/AAC_LD[48k]_m_ar03.wav
./Soundfiles/KKK_m_ar03.wav
./Soundfiles/SWB_m_ar03.wav
./Soundfiles/c001_m_ar03.wav
./Soundfiles/G7222_1_24_m_ar03.wav
./Soundfiles/bp_600_7k_m_ar03.wav
./Soundfiles/AAC_LD[40k]_m_ar03.wav

#signal_6
./Soundfiles/FB_m_ss018.wav REF
./Soundfiles/FB_m_ss018.wav HR
./Soundfiles/NB_m_ss018.wav LP35
./Soundfiles/WB_m_ss018.wav LP70
./Soundfiles/bp_50_3400_m_ss018.wav
./Soundfiles/SILK[12]_m_ss018.wav
./Soundfiles/bp_50_8000_m_ss018.wav
./Soundfiles/SWB_m_ss018.wav
./Soundfiles/SILK[32]_m_ss018.wav
./Soundfiles/NB_G711_m_ss018.wav
./Soundfiles/HHH_m_ss018.wav
./Soundfiles/G7222_1_32_m_ss018.wav
./Soundfiles/AAC_LD[32k]_m_ss018.wav

#signal_7
./Soundfiles/FB_w_cd012.wav REF
./Soundfiles/FB_w_cd012.wav HR
./Soundfiles/NB_w_cd012.wav LP35
./Soundfiles/WB_w_cd012.wav LP70
./Soundfiles/G7222_1_24_w_cd012.wav
./Soundfiles/c108_w_cd012.wav
./Soundfiles/AMR-WBplus_13_w_cd012.wav
./Soundfiles/HHH_w_cd012.wav
./Soundfiles/AMR-WBplus_36_w_cd012.wav
./Soundfiles/c108_w_cd012.wav
./Soundfiles/bp_50_12000_w_cd012.wav
./Soundfiles/NB_G711_w_cd012.wav
./Soundfiles/AMR-WBplus_6_w_cd012.wav

#signal_8
./Soundfiles/FB_w_nsb026.wav REF
./Soundfiles/FB_w_nsb026.wav HR
./Soundfiles/NB_w_nsb026.wav LP35
./Soundfiles/WB_w_nsb026.wav LP70
./Soundfiles/KKK_w_nsb026.wav
./Soundfiles/WB_G722w_nsb026.wav
./Soundfiles/CELT[40k]_w_nsb026.wav
./Soundfiles/bp_50_3400_w_nsb026.wav
./Soundfiles/CELT[64k]_w_nsb026.wav
./Soundfiles/bp_600_7k_w_nsb026.wav
./Soundfiles/bp_50_10000_w_nsb026.wav
./Soundfiles/AAC_LD[32k]_w_nsb026.wav
./Soundfiles/CELT[48k]_w_nsb026.wav

```

```

#signal_9
./Soundfiles/FB_m_ar04.wav          REF
./Soundfiles/FB_m_ar04.wav          HR
./Soundfiles/NB_m_ar04.wav          LP35
./Soundfiles/WB_m_ar04.wav          LP70
./Soundfiles/c014_m_ar04.wav
./Soundfiles/SWB_m_ar04.wav
./Soundfiles/AAC_LD[32k]_m_ar04.wav
./Soundfiles/G7222_1_24_m_ar04.wav
./Soundfiles/AMR-WBplus_36_m_ar04.wav
./Soundfiles/AMR-WBplus_6_m_ar04.wav
./Soundfiles/NB_G711_m_ar04.wav
./Soundfiles/SILK[32]_m_ar04.wav
./Soundfiles/AMR-WBplus_13_m_ar04.wav

#signal_10
./Soundfiles/FB_m_ss019.wav          REF
./Soundfiles/FB_m_ss019.wav          HR
./Soundfiles/NB_m_ss019.wav          LP35
./Soundfiles/WB_m_ss019.wav          LP70
./Soundfiles/AMR-WBplus_36_m_ss019.wav
./Soundfiles/WB_G722_2[1265]m_ss019.wav
./Soundfiles/AMR-WBplus_13_m_ss019.wav
./Soundfiles/II_m_ss019.wav
./Soundfiles/SILK[12]_m_ss019.wav
./Soundfiles/WB_G722_2[2305]m_ss019.wav
./Soundfiles/CELT[48k]_m_ss019.wav
./Soundfiles/CELT[32k]_m_ss019.wav
./Soundfiles/bp_600_7k_m_ss019.wav

#signal_11
./Soundfiles/FB_w_cd014.wav          REF
./Soundfiles/FB_w_cd014.wav          HR
./Soundfiles/NB_w_cd014.wav          LP35
./Soundfiles/WB_w_cd014.wav          LP70
./Soundfiles/AMR-WBplus_6_w_cd014.wav
./Soundfiles/G7222_1_48_w_cd014.wav
./Soundfiles/CELT[64k]_w_cd014.wav
./Soundfiles/bp_50_8000_w_cd014.wav
./Soundfiles/G7222_1_24_w_cd014.wav
./Soundfiles/bp_50_3400_w_cd014.wav
./Soundfiles/bp_50_10000_w_cd014.wav
./Soundfiles/WB_G722_2[1265]w_cd014.wav
./Soundfiles/c110_w_cd014.wav

#signal_12
./Soundfiles/FB_w_nsb027.wav         REF
./Soundfiles/FB_w_nsb027.wav         HR
./Soundfiles/NB_w_nsb027.wav         LP35
./Soundfiles/WB_w_nsb027.wav         LP70
./Soundfiles/WB_G722w_nsb027.wav
./Soundfiles/c110_w_nsb027.wav
./Soundfiles/c108_w_nsb027.wav
./Soundfiles/c043_w_nsb027.wav
./Soundfiles/c001_w_nsb027.wav
./Soundfiles/c014_w_nsb027.wav
./Soundfiles/SILK[12]_w_nsb027.wav
./Soundfiles/WB_G722w_nsb027.wav
./Soundfiles/AAC_LD[40k]_w_nsb027.wav

```

```

#signal_13
./Soundfiles/FB_m_ar05.wav          REF
./Soundfiles/FB_m_ar05.wav          HR
./Soundfiles/NB_m_ar05.wav          LP35
./Soundfiles/WB_m_ar05.wav          LP70
./Soundfiles/bp_300_7k_m_ar05.wav
./Soundfiles/CELT[40k]_m_ar05.wav
./Soundfiles/HHH_m_ar05.wav
./Soundfiles/AAC_LD[40k]_m_ar05.wav
./Soundfiles/c001_m_ar05.wav
./Soundfiles/AAC_LD[48k]_m_ar05.wav
./Soundfiles/bp_50_12000_m_ar05.wav
./Soundfiles/KKK_m_ar05.wav
./Soundfiles/bp_50_3400_m_ar05.wav

#signal_14
./Soundfiles/FB_m_ss020.wav          REF
./Soundfiles/FB_m_ss020.wav          HR
./Soundfiles/NB_m_ss020.wav          LP35
./Soundfiles/WB_m_ss020.wav          LP70
./Soundfiles/bp_50_10000_m_ss020.wav
./Soundfiles/G7222_1_48_m_ss020.wav
./Soundfiles/NB_G711_m_ss020.wav
./Soundfiles/KKK_m_ss020.wav
./Soundfiles/SWB_m_ss020.wav
./Soundfiles/bp_50_8000_m_ss020.wav
./Soundfiles/AAC_LD[64k]_m_ss020.wav
./Soundfiles/c108_m_ss020.wav
./Soundfiles/AAC_LD[48k]_m_ss020.wav

#signal_15
./Soundfiles/FB_w_cd015.wav          REF
./Soundfiles/FB_w_cd015.wav          HR
./Soundfiles/NB_w_cd015.wav          LP35
./Soundfiles/WB_w_cd015.wav          LP70
./Soundfiles/G7222_1_32_w_cd015.wav
./Soundfiles/AAC_LD[64k]_w_cd015.wav
./Soundfiles/c114_w_cd015.wav
./Soundfiles/bp_50_12000_w_cd015.wav
./Soundfiles/SILK[32]_w_cd015.wav
./Soundfiles/CELT[48k]_w_cd015.wav
./Soundfiles/WB_G722_2[2305]w_cd015.wav
./Soundfiles/c043_w_cd015.wav
./Soundfiles/CELT[64k]_w_cd015.wav

#signal_16
./Soundfiles/FB_w_nsb028.wav         REF
./Soundfiles/FB_w_nsb028.wav         HR
./Soundfiles/NB_w_nsb028.wav         LP35
./Soundfiles/WB_w_nsb028.wav         LP70
./Soundfiles/bp_600_7k_w_nsb028.wav
./Soundfiles/CELT[32k]_w_nsb028.wav
./Soundfiles/bp_300_7k_w_nsb028.wav
./Soundfiles/CELT[40k]_w_nsb028.wav
./Soundfiles/G7222_1_32_w_nsb028.wav
./Soundfiles/III_w_nsb028.wav
./Soundfiles/c114_w_nsb028.wav
./Soundfiles/HHH_w_nsb028.wav
./Soundfiles/AAC_LD[32k]_w_nsb028.wav

```



# Anhang D

## Log - Files

### D.1 ACR

```

Experiment initialized..
Playlist 'D:\experiments\Falk_BA\ACR_TEST\Test_Tool\Playlists mit SILK\VP_01_pl_1.lst' loaded
successfully. Number of conditions: 84
All needed audio files available
Name: 01
ID: 01
Training begins
22-03-10 09:59:36      48  5 Soundfiles\AMR-WBplus_36_m_ar02.wav
22-03-10 09:59:46      10  3 Soundfiles\c108_w_cd09.wav
22-03-10 09:59:55      9   4 Soundfiles\NB_G711_m_ss017.wav
22-03-10 10:00:06     11  4 Soundfiles\WB_w_nsb025.wav
22-03-10 10:00:14      8   2 Soundfiles\c110_m_ar03.wav
22-03-10 10:00:24     10  5 Soundfiles\SWB_w_cd012.wav
22-03-10 10:00:31          7 Training finished.
22-03-10 10:00:42      11  4 Soundfiles\AAC_LD[32k]_w_nsb031.wav
22-03-10 10:00:50      8   5 Soundfiles\CELT[48k]_m_ar010.wav
22-03-10 10:01:03     13  2 Soundfiles\c110_w_nsb039.wav
22-03-10 10:01:13     10  2 Soundfiles\AMR-WBplus_6_m_ss023.wav
22-03-10 10:01:23     10  5 Soundfiles\WB_m_ar09.wav
22-03-10 10:01:32      9   4 Soundfiles\AAC_LD[48k]_w_nsb036.wav
22-03-10 10:01:43     11  2 Soundfiles\bp_300_7k_w_cd035.wav
22-03-10 10:01:53     10  3 Soundfiles\c043_w_cd012.wav
22-03-10 10:02:05     12  2 Soundfiles\CELT[32k]_w_nsb033.wav
22-03-10 10:02:14      9   4 Soundfiles\WB_G722m_ss017.wav
22-03-10 10:02:24     10  4 Soundfiles\AAC_LD[44k]_m_ss026.wav
22-03-10 10:02:39     15  3 Soundfiles\CELT[40k]_w_nsb032.wav
22-03-10 10:02:48      9   2 Soundfiles\CELT[32k]_m_ar05.wav
22-03-10 10:02:57      9   4 Soundfiles\AAC_LD[44k]_w_nsb027.wav
22-03-10 10:03:06      9   2 Soundfiles\HHH_m_ar013.wav
22-03-10 10:03:17     11  3 Soundfiles\WB_G722_2[2305]m_ss030.wav
22-03-10 10:03:29     12  5 Soundfiles\SWB_w_nsb035.wav
22-03-10 10:03:41      12  1 Soundfiles\c001_w_nsb035.wav
22-03-10 10:03:51     10  5 Soundfiles\bp_50_8000_m_ar011.wav
22-03-10 10:04:03     12  3 Soundfiles\c114_w_nsb040.wav
22-03-10 10:04:16     13  4 Soundfiles\bp_50_10000_w_cd015.wav
22-03-10 10:04:27     11  4 Soundfiles\AMR-WBplus_13_m_ar015.wav
22-03-10 10:04:40     13  3 Soundfiles\bp_50_8000_w_cd022.wav
22-03-10 10:04:52     12  2 Soundfiles\G7222_1_32_m_ss032.wav
22-03-10 10:05:01      9   3 Soundfiles\SILK[12]_w_cd012.wav
22-03-10 10:05:11     10  3 Soundfiles\WB_w_nsb030.wav
22-03-10 10:05:21     10  2 Soundfiles\NB_w_cd033.wav
22-03-10 10:05:32      11  4 Soundfiles\bp_50_3400_w_nsb038.wav
22-03-10 10:05:41      9   2 Soundfiles\HHH_w_cd036.wav
22-03-10 10:05:51     10  5 Soundfiles\SWB_m_ss022.wav
22-03-10 10:06:03     12  4 Soundfiles\WB_G722_2[1265]w_nsb039.wav
22-03-10 10:06:13     10  4 Soundfiles\c014_m_ar07.wav
22-03-10 10:06:23      10  1 Soundfiles\KKK_m_ar012.wav
22-03-10 10:06:32      9   4 Soundfiles\G7222_1_24_m_ar06.wav
22-03-10 10:06:43     11  3 Soundfiles\AMR-WBplus_13_w_nsb038.wav
22-03-10 10:06:54     11  5 Soundfiles\AAC_LD[48k]_m_ss020.wav
22-03-10 10:07:02      8   4 Soundfiles\SILK[12]_m_ar03.wav
22-03-10 10:07:11      9   4 Soundfiles\WB_G722_2[2305]w_cd036.wav
22-03-10 10:07:21     10  2 Soundfiles\III_m_ar08.wav
22-03-10 10:07:31      10  4 Soundfiles\AAC_LD[64k]_w_nsb025.wav
22-03-10 10:07:41     10  2 Soundfiles\c043_m_ss018.wav
22-03-10 10:07:50      9   3 Soundfiles\NB_m_ss019.wav
22-03-10 10:08:00     10  2 Soundfiles\c108_m_ss028.wav
22-03-10 10:08:12     12  4 Soundfiles\G7222_1_48_w_nsb039.wav
22-03-10 10:08:22     10  5 Soundfiles\AMR-WBplus_36_m_ss024.wav
22-03-10 10:08:32     10  2 Soundfiles\CELT[48k]_w_nsb028.wav

```

22-03-10 10:08:41 9 2 Soundfiles\c014\_w\_cd014.wav  
22-03-10 10:08:51 10 3 Soundfiles\bp\_300\_7k\_m\_ss024.wav  
22-03-10 10:09:02 11 2 Soundfiles\c001\_m\_ar015.wav  
22-03-10 10:09:11 9 3 Soundfiles\AAC\_LD[32k]\_m\_ss021.wav  
22-03-10 10:09:21 10 5 Soundfiles\AAC\_LD[64k]\_m\_ss021.wav  
22-03-10 10:09:30 9 5 Soundfiles\bp\_50\_10000\_m\_ar013.wav  
22-03-10 10:09:41 11 2 Soundfiles\c110\_m\_ar04.wav  
22-03-10 10:09:52 11 4 Soundfiles\WB\_G722\_2[1265]m\_ar016.wav  
22-03-10 10:10:01 9 2 Soundfiles\c108\_w\_cd033.wav  
22-03-10 10:10:11 10 4 Soundfiles\G7222\_1\_48\_m\_ss017.wav  
22-03-10 10:10:20 9 4 Soundfiles\SILK[32]\_w\_cd014.wav  
22-03-10 10:10:30 10 3 Soundfiles\CELT[64k]\_w\_nsb030.wav  
22-03-10 10:10:42 12 2 Soundfiles\NB\_G711\_m\_ar016.wav  
22-03-10 10:10:52 10 4 Soundfiles\SILK[32]\_m\_ss018.wav  
22-03-10 10:11:02 10 2 Soundfiles\NB\_G711\_w\_cd021.wav  
22-03-10 10:11:13 11 4 Soundfiles\G7222\_1\_32\_w\_cd035.wav  
22-03-10 10:11:22 9 2 Soundfiles\bp\_50\_3400\_m\_ar05.wav  
22-03-10 10:11:32 10 2 Soundfiles\AMR-WBplus\_6\_w\_nsb031.wav  
22-03-10 10:11:43 11 3 Soundfiles\c114\_m\_ss025.wav  
22-03-10 10:11:54 11 2 Soundfiles\KKK\_w\_nsb040.wav  
22-03-10 10:12:04 10 5 Soundfiles\bp\_50\_12000\_m\_ss027.wav  
22-03-10 10:12:13 9 3 Soundfiles\WB\_G722w\_cd033.wav  
22-03-10 10:12:25 12 1 Soundfiles\bp\_600\_7k\_w\_cd016.wav  
22-03-10 10:12:35 10 5 Soundfiles\CELT[64k]\_m\_ar011.wav  
22-03-10 10:12:45 10 2 Soundfiles\bp\_600\_7k\_m\_ss019.wav  
22-03-10 10:12:57 12 4 Soundfiles\FB\_nsb040.wav  
22-03-10 10:13:07 10 4 Soundfiles\AMR-WBplus\_36\_w\_nsb026.wav  
22-03-10 10:13:17 10 5 Soundfiles\FB\_n\_ss027.wav  
22-03-10 10:13:29 12 2 Soundfiles\II\_w\_nsb033.wav  
22-03-10 10:13:40 11 5 Soundfiles\bp\_50\_12000\_w\_nsb025.wav  
22-03-10 10:13:51 11 2 Soundfiles\G7222\_1\_24\_w\_nsb025.wav  
22-03-10 10:14:01 10 3 Soundfiles\CELT[40k]\_m\_ar014.wav  
22-03-10 10:14:03 2 Total logging time = 931 seconds.

## D.2 MUSHRA

Lab	Listener	File	Score
TL	VP01	./Soundfiles/FB_w_nsb025.wav	100
TL	VP01	./Soundfiles/NB_W_nsb025.wav	4
TL	VP01	./Soundfiles/WB_W_nsb025.wav	65
TL	VP01	./Soundfiles/WB_G722_2[1265]W_nsb025.wav	45
TL	VP01	./Soundfiles/bp_50_8000_w_nsb025.wav	63
TL	VP01	./Soundfiles/bp_50_12000_w_nsb025.wav	100
TL	VP01	./Soundfiles/bp_300_7k_w_nsb025.wav	7
TL	VP01	./Soundfiles/CELT[48k]_W_nsb025.wav	100
TL	VP01	./Soundfiles/AAC_LD[40k]_w_nsb025.wav	93
TL	VP01	./Soundfiles/WB_G722_2[2305]w_nsb025.wav	52
TL	VP01	./Soundfiles/WB_G722W_nsb025.wav	15
TL	VP01	./Soundfiles/c014_w_nsb025.wav	17
TL	VP01	./Soundfiles/FB_m_ss017.wav	100
TL	VP01	./Soundfiles/NB_m_ss017.wav	4
TL	VP01	./Soundfiles/WB_m_ss017.wav	60
TL	VP01	./Soundfiles/c114_m_ss017.wav	50
TL	VP01	./Soundfiles/c110_m_ss017.wav	50
TL	VP01	./Soundfiles/c014_m_ss017.wav	50
TL	VP01	./Soundfiles/WB_G722_2[1265]m_ss017.wav	43
TL	VP01	./Soundfiles/SILK[12]_m_ss017.wav	51
TL	VP01	./Soundfiles/AMR-WBplus_36_m_ss017.wav	100
TL	VP01	./Soundfiles/AMR-WBplus_6_m_ss017.wav	48
TL	VP01	./Soundfiles/CELT[32k]_m_ss017.wav	63
TL	VP01	./Soundfiles/G7222_1_48_m_ss017.wav	100
TL	VP01	./Soundfiles/FB_m_ar02.wav	100
TL	VP01	./Soundfiles/NB_m_ar02.wav	11
TL	VP01	./Soundfiles/WB_m_ar02.wav	42
TL	VP01	./Soundfiles/AAC_LD[48k]_m_ar02.wav	100
TL	VP01	./Soundfiles/G7222_1_48_m_ar02.wav	91
TL	VP01	./Soundfiles/c001_m_ar02.wav	0
TL	VP01	./Soundfiles/AAC_LD[64k]_m_ar02.wav	100
TL	VP01	./Soundfiles/SILK[32]_m_ar02.wav	99
TL	VP01	./Soundfiles/bp_300_7k_m_ar02.wav	20
TL	VP01	./Soundfiles/AMR-WBplus_13_m_ar02.wav	96
TL	VP01	./Soundfiles/III_m_ar02.wav	0
TL	VP01	./Soundfiles/WB_G722_2[2305]m_ar02.wav	44
TL	VP01	./Soundfiles/FB_w_cd09.wav	100
TL	VP01	./Soundfiles/NB_w_cd09.wav	4
TL	VP01	./Soundfiles/WB_w_cd09.wav	61
TL	VP01	./Soundfiles/G7222_1_32_w_cd09.wav	67
TL	VP01	./Soundfiles/c110_w_cd09.wav	48
TL	VP01	./Soundfiles/c043_w_cd09.wav	49
TL	VP01	./Soundfiles/CELT[64k]_w_cd09.wav	100
TL	VP01	./Soundfiles/bp_50_10000_w_cd09.wav	100
TL	VP01	./Soundfiles/AAC_LD[64k]_w_cd09.wav	100
TL	VP01	./Soundfiles/c043_w_cd09.wav	53
TL	VP01	./Soundfiles/III_w_cd09.wav	2
TL	VP01	./Soundfiles/c114_w_cd09.wav	44
TL	VP01	./Soundfiles/FB_m_ar03.wav	100
TL	VP01	./Soundfiles/NB_m_ar03.wav	25
TL	VP01	./Soundfiles/WB_m_ar03.wav	56
TL	VP01	./Soundfiles/CELT[40k]_m_ar03.wav	75
TL	VP01	./Soundfiles/CELT[32k]_m_ar03.wav	69
TL	VP01	./Soundfiles/AAC_LD[48k]_m_ar03.wav	100
TL	VP01	./Soundfiles/KKK_m_ar03.wav	5
TL	VP01	./Soundfiles/SWB_m_ar03.wav	100
TL	VP01	./Soundfiles/c001_m_ar03.wav	5
TL	VP01	./Soundfiles/G7222_1_24_m_ar03.wav	47
TL	VP01	./Soundfiles/bp_600_7k_m_ar03.wav	4

```

TL  VP01  ./Soundfiles/AAC_LD[40k]_m_ar03.wav  96
TL  VP01  ./Soundfiles/FB_m_ss018.wav  100
TL  VP01  ./Soundfiles/NB_m_ss018.wav  0
TL  VP01  ./Soundfiles/WB_m_ss018.wav  56
TL  VP01  ./Soundfiles/bp_50_3400_m_ss018.wav  2
TL  VP01  ./Soundfiles/SILK[12]_m_ss018.wav  62
TL  VP01  ./Soundfiles/bp_50_8000_m_ss018.wav  67
TL  VP01  ./Soundfiles/SWB_m_ss018.wav  100
TL  VP01  ./Soundfiles/SILK[32]_m_ss018.wav  100
TL  VP01  ./Soundfiles/NB_G711_m_ss018.wav  0
TL  VP01  ./Soundfiles/HHH_m_ss018.wav  0
TL  VP01  ./Soundfiles/G7222_1_32_m_ss018.wav  67
TL  VP01  ./Soundfiles/AAC_LD[32k]_m_ss018.wav  76
TL  VP01  ./Soundfiles/FB_w_cd012.wav  100
TL  VP01  ./Soundfiles/NB_w_cd012.wav  28
TL  VP01  ./Soundfiles/WB_w_cd012.wav  72
TL  VP01  ./Soundfiles/G7222_1_24_w_cd012.wav  69
TL  VP01  ./Soundfiles/c108_w_cd012.wav  2
TL  VP01  ./Soundfiles/AMR-WBplus_13_w_cd012.wav  91
TL  VP01  ./Soundfiles/AMR-WBplus_13_w_cd012.wav  100
TL  VP01  ./Soundfiles/AMR-WBplus_36_w_cd012.wav  100
TL  VP01  ./Soundfiles/c108_w_cd012.wav  2
TL  VP01  ./Soundfiles/bp_50_12000_w_cd012.wav  100
TL  VP01  ./Soundfiles/NB_G711_w_cd012.wav  25
TL  VP01  ./Soundfiles/AMR-WBplus_6_w_cd012.wav  3
TL  VP01  ./Soundfiles/FB_w_nsb026.wav  80
TL  VP01  ./Soundfiles/NB_w_nsb026.wav  10
TL  VP01  ./Soundfiles/WB_w_nsb026.wav  67
TL  VP01  ./Soundfiles/KKK_w_nsb026.wav  5
TL  VP01  ./Soundfiles/WB_G722w_nsb026.wav  50
TL  VP01  ./Soundfiles/CELT[40k]_w_nsb026.wav  74
TL  VP01  ./Soundfiles/bp_50_3400_w_nsb026.wav  7
TL  VP01  ./Soundfiles/CELT[64k]_w_nsb026.wav  92
TL  VP01  ./Soundfiles/bp_600_7k_w_nsb026.wav  0
TL  VP01  ./Soundfiles/bp_50_10000_w_nsb026.wav  94
TL  VP01  ./Soundfiles/AAC_LD[32k]_w_nsb026.wav  72
TL  VP01  ./Soundfiles/CELT[48k]_w_nsb026.wav  93
TL  VP01  ./Soundfiles/FB_m_ar04.wav  100
TL  VP01  ./Soundfiles/NB_m_ar04.wav  7
TL  VP01  ./Soundfiles/WB_m_ar04.wav  66
TL  VP01  ./Soundfiles/c014_m_ar04.wav  37
TL  VP01  ./Soundfiles/SWB_m_ar04.wav  100
TL  VP01  ./Soundfiles/AAC_LD[32k]_m_ar04.wav  69
TL  VP01  ./Soundfiles/G7222_1_24_m_ar04.wav  40
TL  VP01  ./Soundfiles/AMR-WBplus_36_m_ar04.wav  100
TL  VP01  ./Soundfiles/AMR-WBplus_6_m_ar04.wav  4
TL  VP01  ./Soundfiles/NB_G711_m_ar04.wav  3
TL  VP01  ./Soundfiles/SILK[32]_m_ar04.wav  96
TL  VP01  ./Soundfiles/AMR-WBplus_13_m_ar04.wav  66
TL  VP01  ./Soundfiles/FB_m_ss019.wav  100
TL  VP01  ./Soundfiles/NB_m_ss019.wav  0
TL  VP01  ./Soundfiles/WB_m_ss019.wav  48
TL  VP01  ./Soundfiles/AMR-WBplus_36_m_ss019.wav  75
TL  VP01  ./Soundfiles/WB_G722_2[1265]m_ss019.wav  43
TL  VP01  ./Soundfiles/AMR-WBplus_13_m_ss019.wav  71
TL  VP01  ./Soundfiles/III_m_ss019.wav  2
TL  VP01  ./Soundfiles/SILK[12]_m_ss019.wav  100
TL  VP01  ./Soundfiles/WB_G722_2[2305]m_ss019.wav  36
TL  VP01  ./Soundfiles/CELT[48k]_m_ss019.wav  100
TL  VP01  ./Soundfiles/CELT[32k]_m_ss019.wav  35

```

```

TL  VP01  ./Soundfiles/bp_600_7k_m_ss019.wav      0
TL  VP01  ./Soundfiles/FB_w_cd014.wav      34
TL  VP01  ./Soundfiles/NB_w_cd014.wav      0
TL  VP01  ./Soundfiles/WB_w_cd014.wav      40
TL  VP01  ./Soundfiles/AMR-WBplus_6_w_cd014.wav      0
TL  VP01  ./Soundfiles/G7222_1_48_w_cd014.wav      63
TL  VP01  ./Soundfiles/CELT[64k]_w_cd014.wav      96
TL  VP01  ./Soundfiles/bp_50_8000_w_cd014.wav      91
TL  VP01  ./Soundfiles/G7222_1_24_w_cd014.wav      15
TL  VP01  ./Soundfiles/bp_50_3400_w_cd014.wav      35
TL  VP01  ./Soundfiles/bp_50_10000_w_cd014.wav      70
TL  VP01  ./Soundfiles/WB_G722_2[1265]w_cd014.wav      30
TL  VP01  ./Soundfiles/c110_w_cd014.wav      17
TL  VP01  ./Soundfiles/FB_w_nsb027.wav      100
TL  VP01  ./Soundfiles/NB_w_nsb027.wav      5
TL  VP01  ./Soundfiles/WB_w_nsb027.wav      48
TL  VP01  ./Soundfiles/WB_G722w_nsb027.wav      66
TL  VP01  ./Soundfiles/c110_w_nsb027.wav      45
TL  VP01  ./Soundfiles/c108_w_nsb027.wav      0
TL  VP01  ./Soundfiles/c043_w_nsb027.wav      38
TL  VP01  ./Soundfiles/c001_w_nsb027.wav      0
TL  VP01  ./Soundfiles/c014_w_nsb027.wav      37
TL  VP01  ./Soundfiles/SILK[12]_w_nsb027.wav      48
TL  VP01  ./Soundfiles/WB_G722w_nsb027.wav      74
TL  VP01  ./Soundfiles/AAC_LD[40k]_w_nsb027.wav      82
TL  VP01  ./Soundfiles/FB_m_ar05.wav      100
TL  VP01  ./Soundfiles/NB_m_ar05.wav      12
TL  VP01  ./Soundfiles/WB_m_ar05.wav      5
TL  VP01  ./Soundfiles/bp_300_7k_m_ar05.wav      17
TL  VP01  ./Soundfiles/CELT[40k]_m_ar05.wav      74
TL  VP01  ./Soundfiles/HHH_m_ar05.wav      4
TL  VP01  ./Soundfiles/AAC_LD[40k]_m_ar05.wav      96
TL  VP01  ./Soundfiles/c001_m_ar05.wav      0
TL  VP01  ./Soundfiles/AAC_LD[48k]_m_ar05.wav      100
TL  VP01  ./Soundfiles/bp_50_12000_m_ar05.wav      98
TL  VP01  ./Soundfiles/KKK_m_ar05.wav      4
TL  VP01  ./Soundfiles/bp_50_3400_m_ar05.wav      9
TL  VP01  ./Soundfiles/FB_m_ss020.wav      100
TL  VP01  ./Soundfiles/NB_m_ss020.wav      7
TL  VP01  ./Soundfiles/WB_m_ss020.wav      49
TL  VP01  ./Soundfiles/bp_50_10000_m_ss020.wav      91
TL  VP01  ./Soundfiles/G7222_1_48_m_ss020.wav      68
TL  VP01  ./Soundfiles/NB_G711_m_ss020.wav      5
TL  VP01  ./Soundfiles/KKK_m_ss020.wav      0
TL  VP01  ./Soundfiles/SWB_m_ss020.wav      90
TL  VP01  ./Soundfiles/bp_50_8000_m_ss020.wav      50
TL  VP01  ./Soundfiles/AAC_LD[64k]_m_ss020.wav      67
TL  VP01  ./Soundfiles/c108_m_ss020.wav      0
TL  VP01  ./Soundfiles/AAC_LD[48k]_m_ss020.wav      66
TL  VP01  ./Soundfiles/FB_w_cd015.wav      100
TL  VP01  ./Soundfiles/NB_w_cd015.wav      1
TL  VP01  ./Soundfiles/WB_w_cd015.wav      70
TL  VP01  ./Soundfiles/G7222_1_32_w_cd015.wav      83
TL  VP01  ./Soundfiles/AAC_LD[64k]_w_cd015.wav      100
TL  VP01  ./Soundfiles/c114_w_cd015.wav      12
TL  VP01  ./Soundfiles/bp_50_12000_w_cd015.wav      100
TL  VP01  ./Soundfiles/SILK[32]_w_cd015.wav      100
TL  VP01  ./Soundfiles/CELT[48k]_w_cd015.wav      100
TL  VP01  ./Soundfiles/WB_G722_2[2305]w_cd015.wav      100
TL  VP01  ./Soundfiles/c043_w_cd015.wav      67

```

## **Erklärung der Urheberschaft**

Hiermit erkläre ich an Eides statt, dass ich die vorliegende Arbeit mit dem Titel:

*„Vermessung der Gesamtqualität von Sprach- und Audiokodierern“*

selbstständig und ausschließlich unter Verwendung der angegebenen Quellen verfasst habe. Die Arbeit wurde bisher in gleicher oder ähnlicher Form in keiner anderen Prüfungsbehörde zum Erlangen eines akademischen Grades vorgelegt.

.....  
Ort, Datum

.....  
Unterschrift