

Dokumentation Toningenieurprojekt

# **Entwicklung eines audiometrischen Reimtestverfahrens für Kinder mit Hörbeeinträchtigung zur Detektion und Analyse von Fehlinterpretationen in der Sprachwahrnehmung**

Werner Moitzi

Uwe Neugebauer

**Betreuung:**

DI Dr. Alois Sontacchi

Institut für Elektronische Musik und Akustik (IEM)  
Universität für Musik und darstellende Kunst Graz

Graz, Jänner 2011

# Kurzfassung/Abstract

Im Rahmen dieser Projektarbeit soll ein neues, sprachaudiometrisches Testverfahren für binaural hörbeeinträchtigte Kindern im Alter von 4-7 Jahren entwickelt werden. Dieses Testverfahren soll als Ergänzung zur klassischen Tonaudiometrie verstanden werden und die Fähigkeit des Sprachverständnisses der Testperson widerspiegeln. Klassische sprachaudiometrische Testverfahren ermöglichen eine Aussage über das Sprachverständnis, erlauben es allerdings nicht, die Gründe für etwaige Schwächen im Sprachverständnis zu diagnostizieren. Das entwickelte Testverfahren geht hier einen Schritt weiter. Durch eine Analyse der verwendeten Sprachsignale werden Signalparameter definiert, mit Hilfe derer eine Erklärung für etwaiges fehlerhaftes Antwortverhalten gesucht wird. Eine weitere Besonderheit stellt die Auswahl des Bild- und Tonmaterials da, bei der besonders auf den altersabhängigen Wortschatz der Zielgruppe, sowie die für Österreich typische Aussprache der Testwörter Wert gelegt wird.

Als Ziel dieser Arbeit kann die Verbesserung der Hörfähigkeit und eine damit verbundene bessere Entwicklung des Kindes definiert werden.

This project aims at developing a new audio-metric rhyme test procedure for binaurally hearing-impaired children aged 4-7. The testing procedure is meant to be a supplement to traditional sound audiometry to reflect speech comprehension of test participants. Traditional sound audiometry testing procedures do allow for conclusions to be drawn on speech comprehension, but do not allow for determining the reasons of possible disturbance in speech comprehension. The test procedure developed as part of this project takes this one step further. Through analysis of speech signals applied, signal parameters are defined to help in the search for explanations for potential deficient response behaviour. The selection of visual and audio material is another special feature of this test procedure as a particular focus was put on age-related vocabulary of the target group and typical Austrian pronunciation of test words.

The objective of this project is to improve hearing capabilities and, thus, improve child development.

# Inhaltsverzeichnis

<b>Inhaltsverzeichnis</b>	<b>iii</b>
<b>Abbildungsverzeichnis</b>	<b>v</b>
<b>Tabellenverzeichnis</b>	<b>vi</b>
<b>1 Einleitung</b>	<b>1</b>
1.1 Sprachaudiometrie im deutschsprachigen Raum . . . . .	1
1.1.1 Göttinger Kindersprachverständnistest . . . . .	1
1.1.2 Heidelberger CVC-Audiometrie . . . . .	2
1.1.3 Mainzer Kindersprachtest . . . . .	2
1.1.4 Oldenburger Kinderreimtest (OLKI) . . . . .	2
1.1.5 Oldenburger Satztest (OLSA) . . . . .	3
<b>2 Der EWU - Test</b>	<b>4</b>
2.1 Die Grundidee . . . . .	4
2.2 Testsignale und Analyse . . . . .	5
2.2.1 Auswahl der Sprachsignale . . . . .	5
2.2.2 Analyse der Sprachsignale . . . . .	7
2.2.3 Datenaufbereitung . . . . .	12
2.2.4 Minimierung äußerer Einflüsse auf das Testergebnis . . . . .	14
<b>3 Die Testsoftware</b>	<b>16</b>
3.1 Aufbau und Funktionsweise . . . . .	16
3.1.1 Kalibrierung des Wiedergabesetups . . . . .	16
3.1.2 Eingabe der personenbezogenen Daten der Versuchsperson . . . . .	17
3.1.3 Durchführung des Hörtests . . . . .	18
3.1.4 Speicherung . . . . .	20
<b>4 Testablauf</b>	<b>22</b>
4.1 Vorbereitung der Probanden/Probandinnen . . . . .	22
4.2 Evaluierungstest mit hörenden Kindern . . . . .	22
4.2.1 Testdurchführung . . . . .	23
4.2.2 Ergebnisse . . . . .	23
4.3 Simulierter Testdurchlauf . . . . .	24
4.3.1 Alle Wörter richtig erkannt . . . . .	24
4.3.2 Ein spezieller Konsonant wird nicht erkannt . . . . .	24

---

4.3.3	Anzahl der gewählten Distraktoren zu hoch . . . . .	25
4.3.4	Allgemeiner Fall/mögliches reales Testszenario . . . . .	25
<b>5</b>	<b>Zusammenfassung und Ausblick</b>	<b>26</b>
	<b>Literaturverzeichnis</b>	<b>27</b>
<b>A</b>	<b>Randbedingungen für die Testdurchführung</b>	<b>29</b>
A.1	Anforderungen an der Prüfraum . . . . .	29
A.2	Empfehlungen zur Wiedergabehardware . . . . .	29
A.2.1	Tablet-PC . . . . .	29
A.2.2	Lautsprecher . . . . .	30
<b>B</b>	<b>Technische Dokumentation</b>	<b>32</b>
B.1	Kalibriersignal . . . . .	32
B.1.1	Pegelanpassung zwischen Kalibriersignal und Zielitem . . . . .	32
B.2	Technische Spezifikationen . . . . .	33
B.2.1	Spezifikationen der Audiosignale . . . . .	33
B.2.2	Spezifikationen der Entwicklungsumgebung . . . . .	33
B.3	Detailanalyse aller Audiofiles . . . . .	34

# Abbildungsverzeichnis

2.1	Ein mögliches Bildertriolett . . . . .	5
2.2	Schematische Darstellung Frame l=220 Samples . . . . .	8
2.3	Spektrogramm des Wortes Bach . . . . .	9
2.4	Spektrogramm des Anlautes (B) . . . . .	9
2.5	oben: Energiedichtespektrum Anlaut (K); unten: geglättetes Spektrum . . . . .	10
2.6	oben: Energiedichtespektrum Anlaut (Sch); unten: geglättetes Spektrum . . . . .	11
2.7	Berechnung eines Graustufenbildes mit einer unterschiedlichen Anzahl an Schwellwerten . . . . .	13
2.8	Schematischer Datenwürfel . . . . .	15
3.1	GUI - Startbildschirm . . . . .	17
3.2	GUI - Dateneingabe . . . . .	18
3.3	GUI - Hauptfenster . . . . .	19
3.4	GUI - Auswahl am Ende des Tests . . . . .	19
3.5	GUI - Flussdiagramm . . . . .	21
4.1	Detailanalyse; Szenario: alle Wörter mit „S“ nicht erkannt . . . . .	24
4.2	Detailanalyse; Szenario: Wörter mit Otsu-Wert 3 bei Zentroid nicht erkannt . . . . .	25
A.1	Beispielumgebung für die Testdurchführung . . . . .	30
A.2	HP EliteBook 2740p . . . . .	31
B.1	Leistungsdichtespektrum des Kalibriersignals . . . . .	32

# Tabellenverzeichnis

2.1	Worttriplets und deren Vorkommen in unterschiedlichen Erhebungen ( <sup>1</sup> = AWST-R, <sup>2</sup> = ELAN, <sup>3</sup> = ELFRA 1+2, <sup>4</sup> = SEKT-2) . . . . .	6
2.2	Unter- und Obergrenzen der mittels Otsu-Algorithmus bestimmten Gruppen . . . . .	13
2.3	Zielitems und Reimwörter nach Parametern sortiert (F = Anlautfrequenz, Z = Zentroid, FL = Flachheit; 1 = Otsu-Gruppe 1, 2 = Otsu-Gruppe 2, 3 = Otsu-Gruppe 3) . . . . .	14
4.1	Lösungshäufigkeiten der falsch erkannten Zielitems . . . . .	23
A.1	Auszug aus der Konfiguration des verwendeten HP EliteBook 2740p . . . . .	31
B.1	Spezifikationen der Audiosignale . . . . .	33
B.2	Spezifikationen der Entwicklungsumgebung . . . . .	33
B.3	Detailanalyse der Zielitems und Reimwörter . . . . .	35

# Kapitel 1

## Einleitung

Die Basis jeder audiometrischen Untersuchung stellt auch für Kinder das Tonaudiogramm dar. Dieses gibt Auskunft über die frequenzabhängige Hörschwelle der getesteten Person. Im Gegensatz zum Tonaudiogramm, bei dem als Testsignale einzelne Sinustöne verwendet werden, werden im Bereich der Sprachaudiometrie ganze Wörter oder Sätze als Testitems eingesetzt. Die Sprachaudiometrie stellt keine Alternative zur klassischen Tonaudiometrie dar, sondern dient als Ergänzung um beispielsweise Ergebnisse aus der Tonaudiometrie zu verifizieren oder das Sprachverständnis im Speziellen zu prüfen [Leh01].

### 1.1 Sprachaudiometrie im deutschsprachigen Raum

Im deutschsprachigen Raum häufig eingesetzte Testverfahren sind der Göttinger Kindersprachverständnistest, die Heidelberger CVC-Audiometrie, der Mainzer Kindersprachtest, der Oldenburger Kinderreimtest (OLKI), sowie der Oldenburger Satztest (OLSA). Diese Testverfahren sind auf unterschiedliche Altersgruppen und deren spezifischen Wortschatz ausgelegt und ermöglichen alle eine Aussage über das Sprachverstehen der Testperson.

Keiner der genannten Tests erlaubt es allerdings, Rückschlüsse auf die Ursachen problematischen Sprachverstehens zu ziehen. In [Bis05] wird ein Zusammenhang zwischen Sprachverständnis und Frequenzspektrum der Anlaute der verwendeten Testitems hergestellt, um mögliche Fehleinstellungen der apparativen Versorgung hörbeeinträchtigter Kinder zu erklären. Die Anwendbarkeit der erwähnten Tests in Österreich ist leider nur bedingt möglich, da sich der altersabhängige Wortschatz deutschsprachiger Kinder unterschiedlicher Nationalität stark unterscheidet. Es erscheint daher sinnvoll, ein Testverfahren zu entwickeln, das im Besonderen auf Testpersonen österreichischer Herkunft ausgelegt ist. Im folgenden Abschnitt, der den Ausführungen in [Kom04, HZ08, BWM05, Leh01, Eys05, MS06, Thi00, Hör00] folgt, sollen die oben angesprochenen, bereits existierenden Testverfahren kurz erläutert werden.

#### 1.1.1 Göttinger Kindersprachverständnistest

Der Göttinger Kindersprachverständnistest stammt aus dem Jahre 1976 und ist ein so genannter Bildertest. Das bedeutet, das Kind soll nach der Darbietung des Zielitems auf ein vorliegendes (mit dem Zielitem korrespondierendes) Bild zeigen. Die Testwörter bestehen ausschließlich aus Einsilbern. Das Ziel dieses Tests ist vor allem die Detektion von Phonemverwechslungen<sup>1</sup>. Der Göttinger Kindersprachverständnistest setzt sich aus zwei Subtests, welche für zwei verschiedene Zielgruppen konzipiert sind zusammen.

<sup>1</sup>Als Phonem bezeichnet man die kleinste bedeutungsunterscheidende Lauteinheit.

Subtest 1 wurde für Kinder im Alter zwischen 3 und 4 Jahren entwickelt und besteht aus 20 Wörtern, die dem initialen Wortschatz<sup>1</sup> der Zielgruppe entnommen wurden. Zusätzlich kommt Subtest 1 bei entwicklungsverzögerten Kindern zum Einsatz. Subtest 2 wurde für Vorschulkinder bis zum 6. Lebensjahr entwickelt und besteht aus 100 unterschiedlichen Wörtern. Subtest 2 kann weiters verwendet werden, um das Sprachverstehen von 8-Jährigen unter Einfluss von Störlärm zu ermitteln, indem zusätzlich zum Zielitem ein Störgeräusch eingespielt wird.

### 1.1.2 Heidelberger CVC-Audiometrie

Die Heidelberger CVC-Audiometrie, entwickelt im Jahr 1981, ist für die Zielgruppe von hörbeeinträchtigten Kindern im Grund- Hauptschulalter bestimmt. Dieser Test arbeitet mit Konsonant-Vokal-Konsonant (CVC) Verbindungen. Das Besondere an diesem Testverfahren ist die Auswertungsmethode, die nicht wie bei den meisten anderen Testverfahren mit „Richtig“ und „Falsch“ bewertet wird, sondern nach einem bestimmten Punkteschlüssel auch Teilantworten zulässt. Das Testschema selbst funktioniert nach dem Prinzip „Hören und Nachsprechen“, also mit offener Antwortmöglichkeit. Es werden Phonemverwechslungen der folgenden Fehlerkategorien analysiert:

- Stimmhaftigkeit
- Nasalität
- Artikulationsstellung
- Artikulationsmodus

Das Verfahren ist zusätzlich auch zur Beurteilung der Sprachwahrnehmungsleistungen sprachbeeinträchtigter Kinder geeignet.

### 1.1.3 Mainzer Kindersprachtest

Der Mainzer Kindersprachtest wurde 1974 entwickelt und spricht die Zielgruppe der 4- bis 8-jährigen Kinder an. Ähnlich der Heidelberger CVC-Audiometrie soll das Kind das Gehörte nachsprechen. Der Sprachtest ist nach Schwierigkeitsgrad, gemäß Alter und Hörbeeinträchtigung in drei Subtests unterteilt.

Subtest 1 besteht aus 10, dem Wortschatz eines 4-Jährigen entsprechenden, ein- und zweisilbigen Wörtern sowie Lautmalereien, welche in fünf verschiedenen Reihenfolgen präsentiert werden können. Subtest 2 besteht aus 25 der oben genannten Wortkonstruktionen und ist für Kinder bis zu einem Alter von 5 Jahren gedacht. Bei diesen beiden Testverfahren ist eine zusätzliche Unterstützung mittels Bildern die den jeweiligen Zielitems entsprechen zulässig.

Subtest 3 funktioniert nach dem selben Prinzip wie Subtest 1 und 2 und ist für Kinder im Alter von 5 bis 6 Jahren bestimmt.

### 1.1.4 Oldenburger Kinderreimtest (OLKI)

Der OLKI, entwickelt 2000, wird für die Anwendung bei Kindern ab dem sechsten Lebensjahr empfohlen. Dieses Testverfahren ist ebenfalls ein Bildertest und verwendet Dreiergruppen von Wörtern, welche jeweils aus zwei Silben bestehen und sich nur im An- oder Inlaut unterscheiden. Das Kind muss während des Tests eines von drei möglichen Antwortbildern auswählen. Die Präsentation des Zielitems erfolgt nach einem einleitenden Aufforderungssatz.

---

<sup>1</sup>Als initialer Wortschatz wird der allererste sich bildende Wortschatz von Kindern bezeichnet.

### 1.1.5 Oldenburger Satztest (OLSA)

Vom Oldenburger Satztest gibt es verschiedene, leicht modifizierte Varianten, die sich vor allem in der Art der dargebotenen Sätze, aber auch in der Art der Auswertung unterscheiden. Der OLSA, der die allgemeine Altersklasse der Schulkinder anspricht, setzt sich, wie der Name schon sagt, aus ganzen Sätzen zusammen. Diese Sätze bestehen aus der einfachen Form Subjekt-Prädikat-Objekt. Die Reihenfolge der Wörter ist zufällig gewählt und ergibt somit nicht immer eine sinnvolle Aussage. Dadurch wird die Antizipation aber auch das mögliche Auswendiglernen der Satzlisten vermieden. Gemessen wird die Sprachverständlichkeitsschwelle mit und ohne Störgeräusch. Dies wird erreicht, indem die Lautstärke der dargebotenen Sätze adaptiv an das Antwortverhalten der Versuchsperson angepasst wird.

## Kapitel 2

# Der EWU - Test

In diesem Kapitel wird erläutert, welche Grundidee hinter dem neu entwickelten EWU-Sprachtest steckt und wie diese umgesetzt wurde. Weiters wird gezeigt, wie die entsprechenden Testitems selektiert und analysiert wurden.

### 2.1 Die Grundidee

Das im Rahmen dieser Arbeit entwickelte sprachaudiometrische Testverfahren soll für Kinder mit Hörbeeinträchtigung im Alter von 4-7 Jahren als Erweiterung zu einem bereits bestehenden Tonaudiogramm verstanden werden und die Fähigkeit des Sprachverständnisses widerspiegeln. Die Besonderheit an diesem Verfahren ist, dass Schwächen im Sprachverständnis nicht nur diagnostiziert werden, sondern das versucht wird, die Gründe für diese zu erörtern. Hierfür wird die Idee aus [Bis05] aufgegriffen und um weitere Signalanalysen ergänzt (siehe Kapitel 2.2).

Der Grundgedanke hinter dem EWU-Test sieht vor, dass die Versuchsperson ein Bild aus einem Bildertriolett richtig dem zugespielten Schallereignis zuordnen soll. Das Schallereignis ist von der Form „Zeig auf das Bild Haus“, wobei „Haus“ als Zielitem anzusehen ist und „Zeig auf das Bild“ als Ankerschall (Die Notwendigkeit dieses Ankerschalls wird in Kapitel 2.2.1 genauer erläutert). Das Bildertriolett besteht aus einem Minimalpaar und einem Distraktor. Ein Minimalpaar ist dadurch gekennzeichnet, dass sich die beiden Wörter nur in einem Phonem unterscheiden. In diesem speziellen Fall wurden Minimalpaare gewählt, die sich im Anlaut unterscheiden, d.h. es wurden Reimpaare gebildet. Das Distraktorbild dient lediglich dazu, die Validität der Antworten der Versuchsperson zu überprüfen. Die jeweiligen Distraktoren wurden so gewählt, dass keinerlei Zusammenhang zu den Minimalpaaren besteht. Das bedeutet, dass sich einerseits der Anlaut von jenen des Minimalpaares unterscheidet und andererseits die Bildinformation in keinerlei kausalem Zusammenhang mit jener des Ziel- bzw. Reimwortes steht. Es wurde nun die Hypothese aufgestellt, dass wenn das nicht zum Schallereignis passende Bild, also das Reimwort gewählt wird, die Ursache in den Signaleigenschaften des Anlautes zu suchen ist. Die Anzahl der gewählten Distraktorbilder soll Aufschluss über die Verwertbarkeit der Testergebnisse geben. Wird der Prozentsatz an gewählten Distraktoren zu hoch, wird der gesamte Test als ungültig gewertet.

Ein mögliches, schematisches Bildertriolett zeigt Abbildung 2.1. Hier stellt zum Beispiel der Hund das Zielitem, der Mund das Reimwort und das Auto den Distraktor dar.

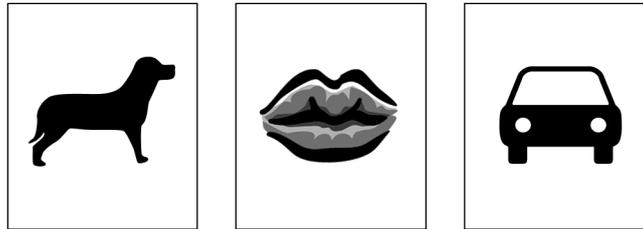


Abbildung 2.1: Ein mögliches Bildertriolett

## 2.2 Testsignale und Analyse

Die Liste der verwendeten Worttrioletts wurde von Frau Mag.<sup>a</sup> Eva-Maria Scheikl, die im Bereich Interdisziplinäre Frühförderung tätig ist entwickelt. Um zu gewährleisten, dass alle Worttrioletts durch den Wortschatz der Testpersonen abgedeckt werden, wurde die Liste der gewählten Wörter mit Ergebnissen aus folgenden Wortschatztests und Umfragen abgeglichen [KK05, BKH06, GD06, GAF00]:

- Aktiver Wortschatztest für 3- bis 5-jährige Kinder (AWST-R)
- Eltern Antworten - Elternfragebogen zur Wortschatzentwicklung im frühen Kindesalter (ELAN)
- Elternfragebögen für die Früherkennung von Risikokindern (ELFRA 1 + 2)
- Sprachentwicklungstest für zweijährige Kinder (SETK-2)

Tabelle 2.1 zeigt eine Aufstellung der ausgewählten Worttrioletts, sowie das Vorkommen der einzelnen Wörter in den entsprechenden Erhebungen.

Das Bildmaterial wurde ebenfalls im Vorfeld mit einer Kontrollgruppe hörender Kinder entsprechenden Alters auf dessen Erkennbarkeit und Eignung getestet. Eine Schwierigkeit stellt auch die notwendige Kürze des gesamten Hörtests dar, da die altersbedingte Aufmerksamkeitsspanne der Zielgruppe berücksichtigt werden muss. Es wurde daher entschieden, den Test auf eine Gesamtanzahl von 27 Trioletts zu beschränken, aus der eine Testdauer von ca. 10 Minuten resultieren soll. Statistisch gesehen würde ein Test mit besonders vielen Trioletts und Durchläufen natürlich eine gesteigerte Qualität der Ergebnisse zulassen.

### 2.2.1 Auswahl der Sprachsignale

Zu Beginn wurde überlegt, von einer Text-to-Speech-Software<sup>1</sup> generierte Wörter einzusetzen. Aufgrund des unnatürlichen, für die Zielgruppe ungewohnten Klangbildes generierter Sprache, musste allerdings nach einer alternativen Lösung gesucht werden. Die Entscheidung ist letztendlich auf Klangbeispiele aus der Österreichischen Aussprachedatenbank (ADABA) [Muh07] gefallen. Diese Datenbank beinhaltet alle für den Hörtest benötigten Klangbeispiele. Die Soundfiles der ADABA liegen in den drei Aussprachevarianten der Länder Österreich, Deutschland und Schweiz vor und wurden mit professionellen Sprecherinnen und Sprechern aufgezeichnet. Laut Scheikl ist für einen sprachaudiometrischen Test eine Sprecherin vorzuziehen, da die Zielgruppe in diesem Alter für eine weibliche Stimme eine höhere Vertrautheit empfindet. Diese Vertrautheit konnte durch die Wahl der Soundfiles mit österreichischer Aussprache weiter erhöht werden.

Die ausgewählten Sprachsignale wurden, wie bereits erwähnt, zusätzlich um einen einleitenden Satz erweitert. Dieser immer gleich bleibende Satz wird als „Ankerschall“ bezeichnet. Diese einleitende Phrase

<sup>1</sup>Eine Text-to-Speech-Software ist ein Computerprogramm, das zur synthetischen Spracherzeugung (Sprachsynthese) verwendet wird.

Triplet	Zielitem	Reimwort	Distraktor
1	Bauch	Rauch <sup>1</sup>	Keks <sup>2,3</sup>
2	Bein <sup>2</sup>	Stein <sup>2,3</sup>	Kopf <sup>1</sup>
3	Buch <sup>2,3,4</sup>	Tuch <sup>2</sup>	Schaf <sup>1,3</sup>
4	Fest	Nest <sup>1</sup>	Schiff <sup>2</sup>
5	Fisch <sup>2</sup>	Tisch <sup>2,3</sup>	Korb
6	Geld <sup>2</sup>	Zelt	Bus <sup>2,3,4</sup>
7	Hahn <sup>1</sup>	Zahn <sup>2</sup>	Mond <sup>1,2,3</sup>
8	Hand <sup>2,3</sup>	Sand <sup>2,3</sup>	Topf <sup>2</sup>
9	Haus <sup>2,3,4</sup>	Maus <sup>2</sup>	Zwerg
10	Herd	Pferd <sup>2,3</sup>	Zug <sup>2,3,4</sup>
11	Hose <sup>2,3</sup>	Dose <sup>3</sup>	Gabel <sup>2,3,4</sup>
12	Hund <sup>2,3</sup>	Mund <sup>2,3,4</sup>	Kamm <sup>3,4</sup>
13	Leiter	Reiter	Sonne <sup>2</sup>
14	Loch <sup>1</sup>	Koch	Saft <sup>2,3</sup>
15	Mauer	Bauer	Suppe <sup>3</sup>
16	Mutter <sup>2,3</sup>	Butter <sup>2</sup>	Stiege <sup>2,3</sup>
17	Nase <sup>2,3,4</sup>	Vase	Sessel <sup>2,3,4</sup>
18	Rad <sup>2</sup>	Bad	Kuss
19	Rock	Stock <sup>1,3</sup>	Bär <sup>1,2,3,4</sup>
20	Rüssel	Schüssel	Blume <sup>2,3</sup>
21	Schere <sup>3,4</sup>	Beere	Kübel <sup>2,3</sup>
22	Schiene	Biene	Löffel <sup>2,3</sup>
23	Schuh <sup>2,3</sup>	Kuh <sup>1,2,3,4</sup>	Bett <sup>2,3</sup>
24	Stern <sup>1,3,4</sup>	Kern	Ball <sup>2,3,4</sup>
25	Tee <sup>2</sup>	See <sup>1</sup>	Baum <sup>2,3,4</sup>
26	Teller <sup>2</sup>	Keller	Hase <sup>2,3,4</sup>
27	Wecker	Bäcker	Auto <sup>2,3,4</sup>

Tabelle 2.1: Worttriplets und deren Vorkommen in unterschiedlichen Erhebungen (<sup>1</sup> = AWST-R, <sup>2</sup> = ELAN, <sup>3</sup> = ELFRA 1+2, <sup>4</sup> = SEKT-2)

hat aus pädagogischer Sicht gesehen die Erhöhung der Aufmerksamkeit der Testperson vor dem Einspielen des Zielitems als Ziel. Da sich der gewünschte Ankerschall nicht im Umfang der ADABA befand, musste dieser separat aufgenommen werden. Für das Einsprechen wurde eine Sprecherin gewählt, deren Stimme nach der Aufzeichnung soweit verändert wurde, dass deren Klang, unter Bewahrung der notwendigen Natürlichkeit, der Sprecherinnenstimme der Audiofiles aus der ADABA möglichst ähnelt.

Sowohl bei der Aufzeichnung des Ankerschalls, als auch bei den ausgewählten Wörtern aus der ADABA, wurde darauf Wert gelegt, den Signalrauschabstand zwischen dem Hintergrundrauschen der Aufnahme und dem Kalibrierungspegel unter 40 dB zu halten. [Eur98]

### 2.2.2 Analyse der Sprachsignale

Sprache beinhaltet eine Vielzahl an Parametern, die für das Sprachverständnis des Menschen eine Rolle spielen. Eine Herausforderung stellt nun die geeignete Auswahl einiger weniger Signalparameter dar, von denen anzunehmen ist, dass diese für diesen Anwendungsfall von hoher Bedeutung sind. Es hat sich gezeigt, dass für die menschliche Sprachwahrnehmung niederfrequente Amplitudenmodulationen im Sprachsignal eine besondere Wichtigkeit haben. Die ausschlaggebenden Modulationsfrequenzen befinden sich im Bereich von ca. 4 Hz bis 16 Hz. Höhere Frequenzen scheinen nur noch geringfügigen Einfluss auf die Sprachwahrnehmung zu nehmen [CG05, Bal99]. Betrachtet man nun die zu analysierenden Daten, so muss man feststellen, dass eine Analyse der Modulationsfrequenzen keine geeigneten Ergebnisse liefern würde. Der Grund dafür ist in der Kürze der Anlaute zu suchen. Diese weisen eine durchschnittliche Länge von 0.08 - 0.14 Sekunden auf. In dieser kurzen Zeit ist es physikalisch nicht möglich, dass sich eine entsprechende Anzahl an Wellen der niederfrequenten Modulationsfrequenzen, aufgrund der zu langen Periodendauern ausbildet.

Für die Gestaltung eines möglichst effektiven Testverfahrens erscheint es zielführend, Signalparameter festzulegen, die einerseits im Anlaut detektiert werden können und auf die andererseits in der Anpassung der Hörbehelfe Einfluss genommen werden kann. Hierfür ist es allerdings unabdingbar, die Funktionsweise der Hörbehelfe genau zu kennen. Dies gestaltet sich in zweierlei Hinsicht problematisch. Der zu entwickelnde Test soll möglichst unabhängig von der Art der apparativen Hörversorgung der Testpersonen durchführbar sein, da eine Berücksichtigung der individuellen Hörbehelfe und der jeweiligen spezifischen Einstellungen nur sehr schwer möglich wäre. Hinzu kommt, dass alle Hersteller dieser Hörbehelfe deren genaue Funktionsweise offenlegen müssten. Verständlicher Weise war keiner der von uns kontaktierten Hersteller zu dieser Auskunft bereit. Es musste daher eine Annahme getroffen werden, welche Signalparameter generell und geräteunabhängig von Interesse sein könnten. Die Entscheidung viel letztendlich auf drei Parameter, die sich für eine Analyse der Anlaute anbieten. Diese sind die Anlautfrequenz, der spektrale Zentroid und die spektrale Flachheit, auf die in den folgenden Abschnitten näher eingegangen wird.

#### Spektrale Bandbegrenzung

Für die Separation der Anlaute und die anschließenden Analysen betrachteten wir ausschließlich den für Sprache besonders interessanten Frequenzbereich zwischen 0 Hz und 10 kHz. Bei Frequenzen über 10 kHz konnten keine relevanten Unterschiede zwischen den Anlauten festgestellt werden, mit Ausnahme von Zischlauten wie „S“ oder „SCH“. Da das Spektrum dieser Zischlaute aber schon ab ca. 2 kHz eine stark ausgeprägte Amplitude aufweist, sind die anlautspezifischen Unterschiede in dem beschränkten Frequenzbereich bereits zur Genüge abgedeckt.

Einen weiteren Grund für die Begrenzung des betrachteten Frequenzbereiches stellt der Frequenzgang der eingesetzten Lautsprecher dar. Das Standardlautsprecher-setup im Rahmen einer sprachaudiometrischen Untersuchung, ist normalerweise auf einen Frequenzbereich zwischen 20 Hz und 10 kHz beschränkt, da

hier bereits auf den Frequenzumfang der menschlichen Sprache Rücksicht genommen wird. Allerdings kommen meist Lautsprecher zum Einsatz, welche speziell im oben genannten Frequenzbereich ein sehr lineares Verhalten aufweisen (siehe Anhang A.2.2).

### Separation der Anlaute

Um eine korrekte Analyse der Anlaute vornehmen zu können, mussten diese vom Rest des Wortes getrennt werden. Aufgrund der Vielzahl an Wörtern und auch aus Gründen der Objektivität, haben wir uns dazu entschlossen, die Anlautseparation in Matlab zu implementieren. Die Separation der Anlaute erfolgte über die Berechnung der Energieänderung in Bezug auf den darauf folgenden Vokal. Energie eines Signals allgemein:

$$E = \int_{-\infty}^{\infty} |x(t)|^2 dt, \quad (2.1)$$

bzw. im diskreten Fall:

$$E = \sum_{n=0}^N |x(n)|^2. \quad (2.2)$$

Um dies zu realisieren, wurde ein Frame der Länge  $l$  erzeugt und schrittweise über das Wort geschoben (Abbildung 2.2). Eine Energiezunahme über einen bestimmten Schwellwert markiert das Ende des Anlautes. Diese Methode hat sich als einfachste und beste Lösung herausgestellt, da der Vokal nach dem Anlaut

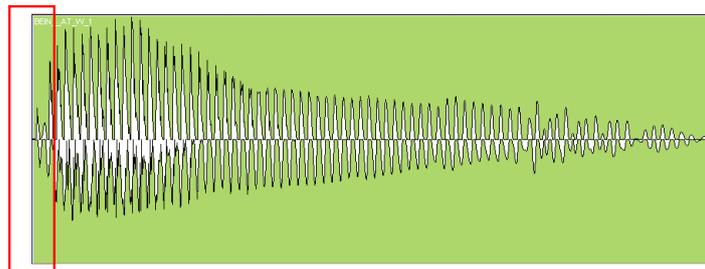


Abbildung 2.2: Schematische Darstellung Frame  $l=220$  Samples

immer einen höheren Energielevel aufweist als der Anlaut selbst. Das Rauschen vor dem Anlaut hat keinen Einfluss auf die Analyse und wurde somit auch nicht abgeschnitten.

Ein wichtiger Aspekt bei der Trennung der Anlaute war es, herauszufinden wo genau der Anlaut endet und der Vokal beginnt. Wir mussten uns in weiterer Folge überlegen, wie viel Vokalanteil bei dem separierten Anlaut verbleiben durfte, da eine hundertprozent genaue Trennung aufgrund der Kürze des Anlautes nicht möglich ist. (Anm.: Bei der Aussprache der meisten Anlaute spielt der folgende Vokal eine entscheidende Rolle, da dieser unterschiedliche Mundstellungen bei selben Anlauten bewirken kann.)

Der Prozess der Anlautseparation setzte sich schlussendlich aus der maschinellen Trennung mittels Matlab-File und einer zusätzlichen akustischen Überprüfung der resultierenden Anlaute zusammen. Gegebenenfalls wurde der Schwellwert leicht modifiziert, um ein sinnvolles Ergebnis der folgenden Analysen zu gewährleisten. In Abbildung 2.3 ist der Vokal „a“ durch die Periodizität zwischen 0.1 und 0.2 Sekunden gut erkennbar. In Abbildung 2.4 ist nur mehr der Anlaut zu sehen. Die Artefakte vor 0.09 Sekunden sind auf das Rauschen vor dem Vokal zurückzuführen.

### Die Anlautfrequenz

Als erster und naheliegendster Parameter wurde die „Anlautfrequenz“ untersucht. Sie soll Aufschluss über eine Grundeigenschaft des Anlautes geben - über dessen auf die Amplitude bezogen, dominanteste Frequenz.

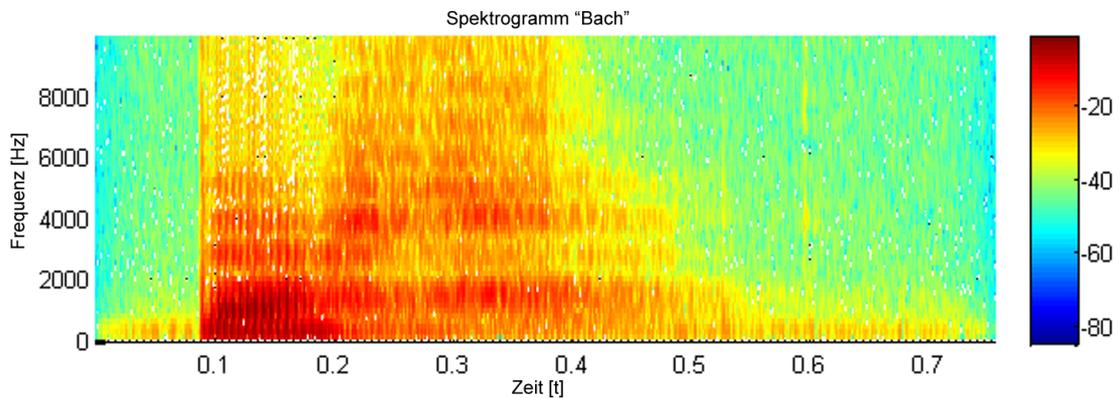


Abbildung 2.3: Spektrogramm des Wortes Bach

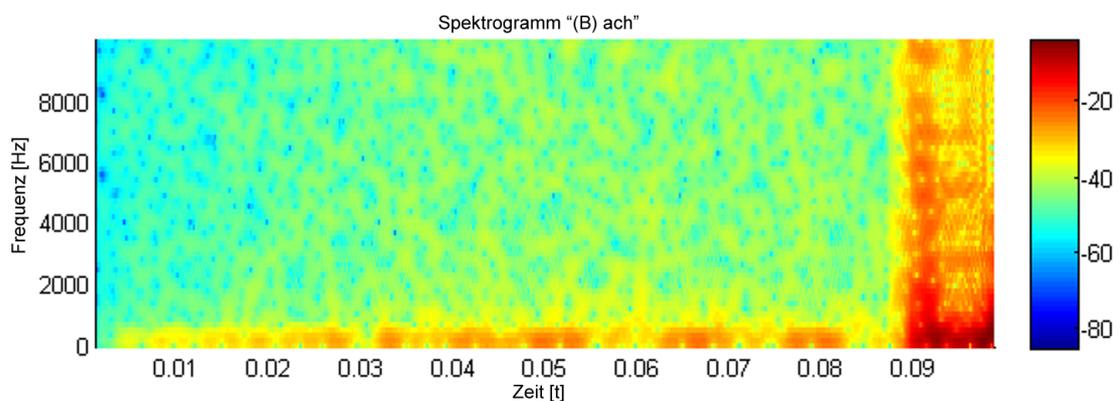


Abbildung 2.4: Spektrogramm des Anlautes (B)

Die besten Ergebnisse zur Bestimmung der Anlauf Frequenz lieferte die Analyse des Energiedichtespektrums. Da in den Spektren der einzelnen Anlaute oft sehr viele Einzelfrequenzen, die relativ gesehen sehr wenig Energie enthalten vorkommen, erschien eine Glättung des erhaltenen Spektrums sinnvoll. Diese Glättung und somit die schlussendliche Berechnung erfolgte mittels Welch - Algorithmus. Der Welch - Algorithmus berechnet das Energiedichtespektrum, indem er das Eingangssignal in 8 Segmente unterteilt (Overlap 50 Prozent) und diese mittels Hamming-Window fenstert. Man erhält somit eine gute Mittelung der spektralen Frequenzen. Die Fensterung mittels Hamming-Window ist für uns aus einem weiteren Grund wichtig. Wir mussten gewährleisten, dass die „scharf“ abgeschnittenen Anlautvektoren keine Artefakte in der FFT-Analyse hinterlassen.

Die Abbildungen 2.5 und 2.6 zeigen die Spektren der Anlaute „K“ des Wortes „Koch“ und „Sch“ des Wortes „Schiene“. Gut zu erkennen ist die größere Amplitude im oberen Frequenzbereich des Frikativs „Sch“ des Wortes „Schiene“. Im Gegensatz dazu steht der stimmlose Plosiv „K“ von „Koch“, welcher bei 1 kHz die größte Amplitude aufweist. In weiterer Folge wurde dieses Prozedere bei allen vorhandenen Wörtern durchgeführt. Die Frequenzen mit der jeweils größten Amplitude wurden dem Wort zugehörig in eine Parameterliste aufgenommen, um für die weiterführende Einteilung in die verschiedenen Otsu-Gruppen (siehe Kapitel 2.2.3) zur Verfügung zu stehen.

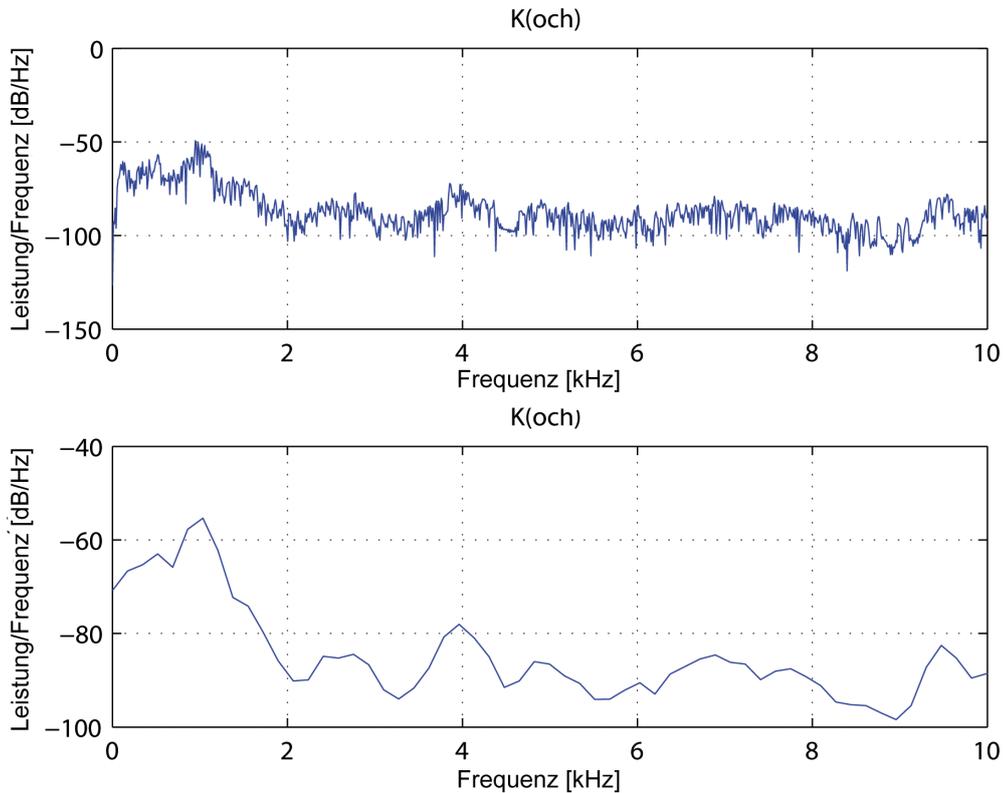


Abbildung 2.5: oben: Energiedichtespektrum Anlaut (K); unten: geglättetes Spektrum

### Der spektrale Zentroid

Der spektrale Zentroid ist eine Eigenschaft des Frequenzspektrums und beschreibt dessen Schwerpunkt bzw. Massenmittelpunkt. Dieser Parameter wurde gewählt, da dieser von hoher Wichtigkeit für die audiotische Szenenanalyse ist. Diese beschreibt die Mechanismen des menschlichen Gehörs, die es ermöglichen akustische Umgebungen zu analysieren [ABDL02]. In der Hörgeräteakustik wird unter anderem der Spektrale Schwerpunkt verwendet um die notwendige Geräuschklassifizierung durchzuführen, die es in der Folge ermöglicht, erwünschte Schallereignisse zu verstärken und Störschall (z.B. Hintergrundgeräusch) auszublenden. Menschen die keine Beeinträchtigung des Gehörs aufweisen, sind in der Lage diese Trennung zwischen Stör- und Nutzschall bewusst zu steuern und die Wahrnehmung auf ein Schallereignis zu fokussieren. Diese Fähigkeit des menschlichen Gehörs bzw. Gehirns wird als „Cocktail-Party-Effekt“ bezeichnet [Che53]. Im musikalischen Kontext wird dem spektralen Schwerpunkt außerdem ein Einfluss auf den „Glanz“ eines Tones zugesprochen [TSW04].

Mathematisch lässt sich der Zentroid  $\mu$  als durch die Amplituden gewichteter Mittelwert der Frequenzen darstellen [Pee04]:

$$\mu = \int xp(x)dx \quad (2.3)$$

mit

$$x = freq_v(x), \quad (2.4)$$

bzw.

$$p(x) = \frac{ampl_v(x)}{\sum_x ampl_v(x)}. \quad (2.5)$$

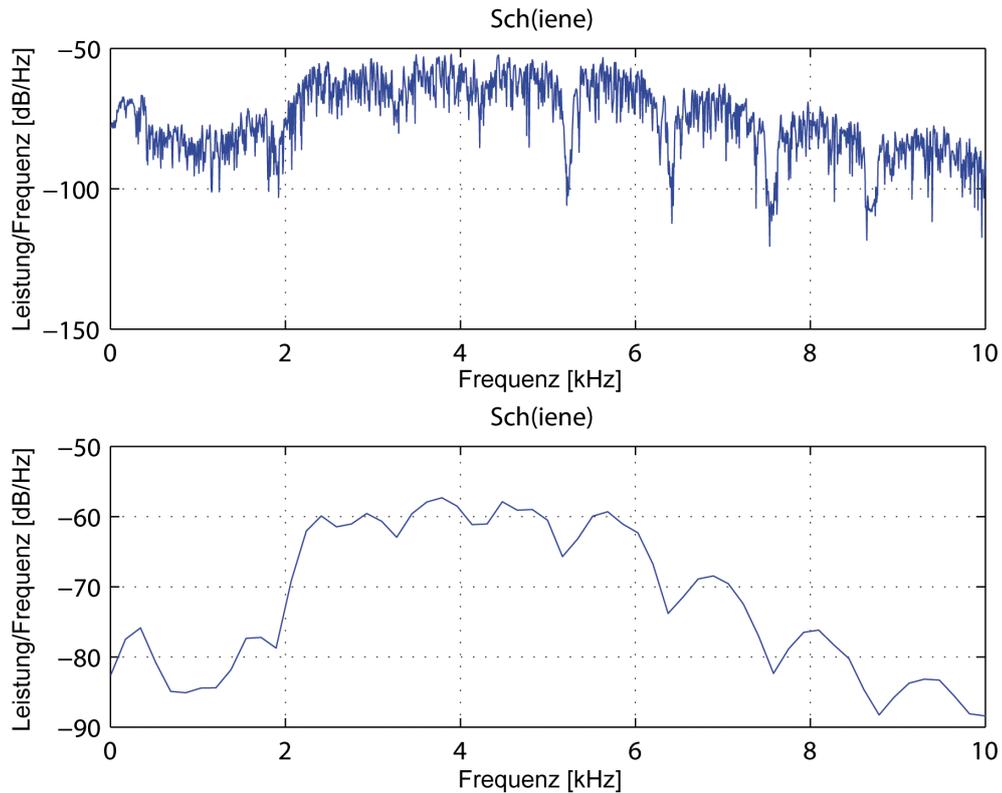


Abbildung 2.6: oben: Energiedichtespektrum Anlaut (Sch); unten: geglättetes Spektrum

Die Zentroid-Analyse der Anlaute erfolgte mit einer zusätzlichen Toolbox in Matlab, der so genannten „MIRtoolbox“ in der Version 1.2, die unter der GNU Lizenz publiziert wird [LT07]. Diese Toolbox besteht aus einer Fülle an in Matlab implementierten Funktionen, die zur Analyse von Audiosignalen verwendet werden können. Zur Bestimmung des spektralen Schwerpunktes wurde der Befehl „mircentroid“ herangezogen.

### Die spektrale Flachheit

Die spektrale Flachheit ist ein Maß für die „Rauschartigkeit“ eines Frequenzspektrums. Dieser Parameter beschreibt, in welchem Maß alle untersuchten Frequenzen gleichmäßig präsent im Spektrum vorhanden sind. Die Werte für die Flachheit liegen zwischen 0 und 1. Eine Flachheit von 1 würde z.B. das Spektrum eines gaußverteilten weißen Rauschen aufweisen, das durch eine konstante Rauschleistungsdichte im Leistungsdichtespektrum gekennzeichnet ist. Eine Flachheit von 0 würde dagegen ein reines Sinussignal charakterisieren. Mathematisch lässt sich die Flachheit als Verhältnis von geometrischem und arithmetischem Mittelwert folgendermaßen darstellen [Lar09, Pee04]:

$$SF = \frac{\sqrt[n]{\prod_{n=0}^{N-1} x(n)}}{\left(\frac{\sum_{n=0}^{N-1} x(n)}{N}\right)}. \quad (2.6)$$

$N$  steht hier für die FFT-Größe und  $x(n)$  für die Amplitude des Frequenzbins  $n$ . Dieser Parameter wurde von uns gewählt, da in ihm einige interessante Eigenschaften der zu analysierenden Sprache abgebildet werden können. So ist in der Sprachsignalverarbeitung die spektrale Flachheit eine nützliche Messgröße,

um zwischen stimmhaften und stimmlosen Lauten zu unterscheiden [PP09]. Außerdem lässt sich mit dieser Messgröße eine Klassifizierung von Sprache nach deren emotionalen Eigenschaften durchführen.

Für die Bestimmung der spektralen Flachheit der Anlaute wurde wiederum die MIRtoolbox verwendet. Der Befehl „mirflatness“ ermöglicht hier die entsprechende Berechnung.

### 2.2.3 Datenaufbereitung

Eine der Herausforderungen dieses Projektes ist die verlässliche Datenauswertung trotz der minimalen Anzahl an Probanden/Probandinnen. Ein weiteres Problem stellt die geringe Belastbarkeit der Versuchspersonen dar. Wie bereits in der Einleitung erwähnt, beträgt das Durchschnittsalter 4-7 Jahre. Daraus resultiert eine maximale Testzeit  $t = 10$  min pro Kind, um eine ausreichende Konzentration der Probanden/Probandinnen zu gewährleisten. Diese beiden Umstände erfordern eine spezielle Herangehensweise bezüglich der Datenauswertung. Um die erhaltenen Daten Anlautfrequenz, Zentroid und Flachheit der Reimwörter nun zur Weiterverwendung im Reimest zu klassifizieren, wurde der Otsu-Algorithmus [Ots79] für die Gruppierung der Worttriplets herangezogen.

#### Der Otsu-Algorithmus

Der Otsu-Algorithmus [Ots79], entwickelt 1979 von Nobuyuki Otsu, ist einer von vielen „Binärisierungsalgorithmen“. Er wurde ursprünglich entwickelt um Graustufenbilder in Binärbilder umzusetzen. Hierfür ist die Wahl des richtigen Schwellwertes entscheidend für das Ergebnis. Der Otsu-Algorithmus kommt hierbei sehr oft zum Einsatz, da er ein besseres Ergebnis liefert als die meisten anderen Verfahren. Der Unterschied zu anderen Verfahren besteht darin, dass der Otsu-Algorithmus statistische Hilfsmittel verwendet um den optimalen Schwellwert zu finden. Zur Berechnung wird die Varianz herangezogen, die als Maß für die Streuung der Grauwerte interpretiert werden kann.

Im Falle eines Binärbildes werden zwei Klassen  $K_0$  und  $K_1$  gebildet, der Hintergrund und das gewünschte Objekt. Nun wird ein Schwellwert  $t$  bestimmt, bei dem die Streuung der Daten innerhalb der Klassen möglichst klein, die Streuung zwischen den Klassen hingegen möglichst groß ist.

Die Summe der Auftrittswahrscheinlichkeiten  $p(g)$  der unterschiedlichen Grauwerte  $0 < g < G$  ( $G =$  maximaler Grauwert) in beiden Klassen lässt sich darstellen als

$$K_0 : P_0(t) = \sum_{g=0}^t p(g), \quad (2.7)$$

bzw.

$$K_1 : P_1(t) = \sum_{g=t+1}^G p(g). \quad (2.8)$$

Die Summe der Wahrscheinlichkeiten ergibt 1:

$$P_1(t) = 1 - P_0(t). \quad (2.9)$$

Die Varianzen innerhalb der Klassen ergeben sich zu:

$$\sigma_0^2(t) = \sum_{g=0}^t (g - g'_0)^2 p(g), \quad (2.10)$$

bzw.

$$\sigma_1^2(t) = \sum_{g=t+1}^G (g - g'_1)^2 p(g). \quad (2.11)$$

$g'_0$  und  $g'_1$  stehen hier für den arithmetischen Mittelwert der Grauwerte innerhalb der einzelnen Klassen und  $g'$  bezeichnet den arithmetischen Mittelwert der Grauwerte innerhalb des gesamten Bildes.

Die Varianz zwischen den Klassen lässt sich darstellen als

$$\sigma_{zw}^2(t) = P_0(t) \cdot (g'_0 - g')^2 + P_1(t) \cdot (g'_1 - g')^2 \tag{2.12}$$

und die Varianz in den Klassen als

$$\sigma_{in}^2(t) = P_0(t) \cdot \sigma_0^2(t) + P_1(t) \cdot \sigma_1^2(t). \tag{2.13}$$

Um die Varianz der Grauwerte in den Klassen möglichst klein, aber zwischen den Klassen möglichst groß zu halten, muss  $t$  so gewählt werden, dass der Quotient  $Q(t)$  maximal wird:

$$Q(t) = \frac{\sigma_{zw}^2(t)}{\sigma_{in}^2(t)}. \tag{2.14}$$

**Gruppierung der Daten**

Die Methode nach Otsu ist natürlich nicht nur auf grafisches Material anwendbar, sondern auf jegliche Art von Daten. In weiterer Folge ist es auch möglich, mehrere Klassen in die Berechnung mit einzubeziehen. Hierzu wurde ein bereits implementiertes Matlab-File verwendet. Mit Hilfe dieser auf dem Otsu-Algorithmus basierenden Berechnungen ist es möglich, eine beliebige Anzahl an Schwellwerten  $n$  von einem Input Array  $I$  zu bestimmen. Ein Beispiel hierfür zeigt Abbildung 2.7.



Abbildung 2.7: Berechnung eines Graustufenbildes mit einer unterschiedlichen Anzahl an Schwellwerten

In unserem Fall haben wir uns dazu entschlossen, jeden Signalparameter in drei Gruppen zu unterteilen ( $n=2$ ). Am Beispiel Anlautfrequenz bedeutet das, dass Wörter die eine niedrige Anlautfrequenz aufweisen, in eine Gruppe zusammengefasst werden. Selbiges gilt für Wörter mit mittlerer bzw. hoher Hauptfrequenz im Anlaut. Nach der selben Vorgehensweise werden auch die Werte für den Zentroid und die Flachheit gruppiert. Die Unter- und Obergrenzen der jeweiligen Parameter in den drei Otsu-Gruppen sind in Tabelle 2.2 ersichtlich.

Otsu-Wert	Anlautfrequenz	Zentroid	Flachheit
1	172 - 434 Hz	807 - 2341 Hz	0,06 - 0,14
2	775 - 2844 Hz	2481 - 4686 Hz	0,15 - 0,24
3	3530 - 9905 Hz	4848 - 9646 Hz	0,25 - 0,45

Tabelle 2.2: Unter- und Obergrenzen der mittels Otsu-Algorithmus bestimmten Gruppen

### Selektion der Testwörter

Um ein möglichst großes Spektrum in den zugrundeliegenden Analyseparametern gleichmäßig abdecken zu können, wurden die 27 Zielitems derart gewählt, dass jeder Analyseparameter mit jedem möglichen Otsu-Wert mindestens einmal existiert. Die endgültigen, im Test verwendeten Reimpaarkombinationen sind in Tabelle 2.3 dargestellt.

F 1		Z 1		FL 1	
Bauch	Rauch	Leiter	Reiter	Rad	Bad
Bein	Stein	Loch	Koch	Rock	Stock
Mutter	Butter	Nase	Vase	Rüssel	Schüssel
F 2		Z 2		FL 2	
Hand	Sand	Herd	Pferd	Mauer	Bauer
Haus	Maus	Hahn	Zahn	Hose	Dose
Hund	Mund	Schere	Beere	Buch	Tuch
F 3		Z 3		FL 3	
Stern	Kern	Fest	Nest	Fisch	Tisch
Schiene	Biene	Tee	See	Teller	Keller
Schuh	Kuh	Wecker	Bäcker	Geld	Zelt

Tabelle 2.3: Zielitems und Reimwörter nach Parametern sortiert (F = Anlautfrequenz, Z = Zentroid, FL = Flachheit; 1 = Otsu-Gruppe 1, 2 = Otsu-Gruppe 2, 3 = Otsu-Gruppe 3)

### Sortierung der Wörter

Die nun bereits gruppierten Wörter können in eine 3x3x3 Matrix einsortiert werden. Diese Matrix kann als Würfel verstanden werden, auf dessen räumlichen Achsen die drei Parameter Anlautfrequenz, Zentroid und Flachheit aufgetragen werden. Dadurch, dass jedem Wort eine der möglichen 27 Kombinationen aus drei Otsu-Werten zugeordnet wurde, kann nun jedes Wort eindeutig einem der 27 Würfel zugeordnet werden. Abbildung 2.8 zeigt schematisch den dreidimensionalen „Datenwürfel“. Alle Wörter die niedrige Otsu-Werte in allen drei Parametern aufweisen, werden zum Beispiel dem grün eingefärbten Würfel zugeordnet. Durch dieses Zuordnungssystem ist es möglich die Eigenschaften jedes Wortes auf einen Blick zu erfassen. Es bietet zusätzlich die Möglichkeit, nicht nur eine Eigenschaft jedes Wortes zu betrachten, sondern alle drei Eigenschaften jedes Wortes in Beziehung zueinander zu stellen. Eine Liste aller Wörter mit den zugehörigen Otsu-Werten zeigt Tabelle B.3.

#### 2.2.4 Minimierung äußerer Einflüsse auf das Testergebnis

Um einer Verfälschung der Testergebnisse durch äußere Einflüsse vorzubeugen, wurden einige vorbeugende Maßnahmen getroffen.

##### Einfluss der Bildreihenfolge

Um den Einfluss der Bildreihenfolge auf das Testergebnis auszuschließen, wurde bei der Auswahl und Positionierung der Worttripletts auf zwei Faktoren besonderer Wert gelegt:

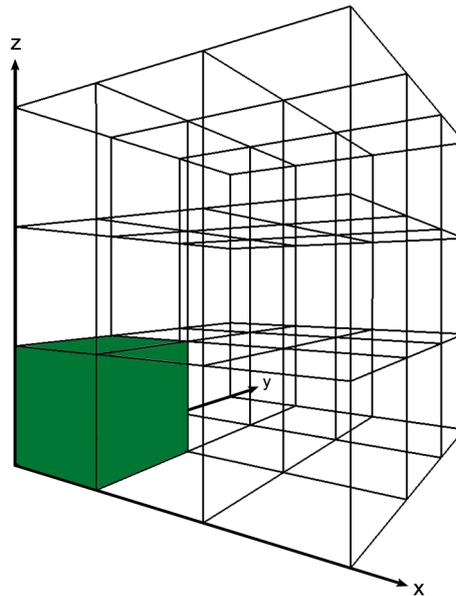


Abbildung 2.8: Schematischer Datenwürfel

- Reihenfolge der Triplets im Test
- Anordnung der Bilder untereinander

Die Reihenfolge der Zielitems wurde nun aus der in 2.2.3 vorgestellten Wortmatrix so gewählt, dass Wörter mit den selben oder ähnlichen Eigenschaften nicht unmittelbar hintereinander abgefragt werden. Das bringt den Vorteil mit sich, dass Konzentrationsschwächen, welche sich im Testablauf über mehrere Wörter erstrecken, keinen Einfluss auf einen bestimmten Analyseparameter ausüben können.

Um ein erkennbares Muster der Bildabfolge während des Versuches zu vermeiden, wurde die Anordnung der Bilder untereinander mit einem Zufallsgenerator festgelegt.

#### **Einfluss des/der Versuchleiters/Versuchsleiterin**

Um einen weiteren möglichen Einfluss auf das Testergebnis zu minimieren, sollte es der/die Versuchleiter/Versuchsleiterin vermeiden, während dem Test mit der Versuchsperson zu sprechen. Da dies bei der Zielgruppe der 4- bis 7-Jährigen wahrscheinlich nicht immer möglich sein wird, sollte die Kommunikation zwischen Versuchleiter/Versuchsleiterin und Testperson zumindest auf ein Minimum reduziert werden. Dies sollte nach einer entsprechenden Einweisung (siehe Anhang 4.1) durchführbar sein.

# Kapitel 3

## Die Testsoftware

Für die praktische Durchführung des Tests ist es sinnvoll, eine Umgebung zu schaffen, die ein möglichst einfaches Handling ermöglicht. Die Umsetzung des Tests in einer Software ermöglicht nicht nur einen definierten, immer gleich bleibenden Ablauf der Versuchsdurchläufe, sondern auch eine automatische Auswertung der von der Versuchsperson gegebenen Antworten. Positiv ist auch der geringe Aufwand für den/die Versuchsleiter/Versuchsleiterin und die einfache Bedienung für die Versuchsperson. Wir haben uns daher entschieden, den Test mit Hilfe der Software „Matlab“, die vom Unternehmen „The MathWorks“ entwickelt wird zu implementieren. Matlab ermöglicht das Erstellen einer grafischen Bedienoberfläche, eines so genannten „GUI“ (Graphical User Interface), über das alle Eingaben getätigt werden können. Zusätzlich wurde die Software so ausgelegt, dass eine Touch-Bedienung für die Versuchsperson ermöglicht wird. Das bedeutet, dass das Kind durch eine Berührung mit einem Finger das gewünschte Bild auswählen kann. Diese Art der Antwortmöglichkeit wurde implementiert, um dem Kind einen möglichst intuitiven Umgang mit dem Programm zu ermöglichen. Diese Form der Eingabe ist allerdings nur mit entsprechender Hardware umsetzbar. Hier haben wir uns für einen Tablet-PC mit kapazitivem Touchscreen und das Betriebssystem Windows 7 Professional der Firma Microsoft entschieden (siehe Anhang A.2). Diese Kombination ermöglicht problemlos die gewünschte Touchbedienung.

### 3.1 Aufbau und Funktionsweise

Die Software lässt sich grob in vier Teile gliedern:

1. Kalibrierung des Wiedergabesetups
2. Eingabe der personenbezogenen Daten der Versuchsperson
3. Durchführung des Hörtests
4. Auswertung und Speicherung

#### 3.1.1 Kalibrierung des Wiedergabesetups

Vor dem Beginn eines Testdurchlaufes ist es von essenzieller Bedeutung, den Schalldruckpegel  $L_p$  zu messen und das Wiedergabesetup zu kalibrieren. Dieser Kalibriervorgang kann mit Hilfe des Startbildschirms (siehe Abbildung 3.1) durchgeführt werden. An dieser Stelle ist es nun möglich ein Kalibriersignal auszugeben, das so lange abgespielt werden kann, bis der gewünschte Schalldruckpegel eingestellt ist. Das Kalibriersignal (genaue Spezifikation siehe Anhang B.1) wurde mittels Matlab generiert. Der Kalibrierpegel  $L_p$  soll mit

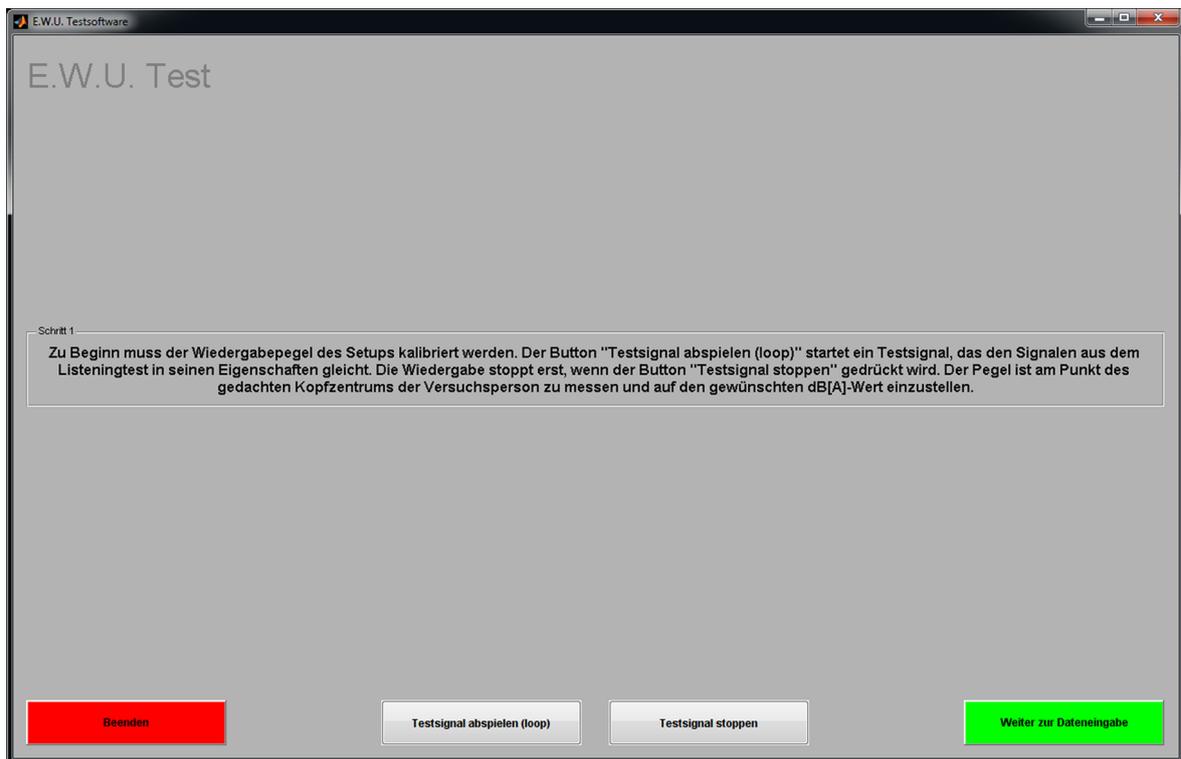


Abbildung 3.1: GUI - Startbildschirm

einem handelsüblichen dB-Meter (z.B. Minilyzer ML1 der Firma NTI), im RMS-Modus, an der Position des gedachten Kopfbereichs der Versuchsperson, in 1 Meter Abstand, frontal vor dem Lautsprecher gemessen werden und  $65 \text{ dB}(A)$  betragen [Eur98]. Die Messdauer soll mindestens 30 Sekunden betragen.

Es wurde überlegt, den Schalldruckpegel individuell, mit  $20 \text{ dB}$  über dem Reintonhörschwellenpegel des/der jeweiligen Probanden/Probandin anzubieten. Aufgrund der Frequenzabhängigkeit des Reintonhörschwellenpegels wird aber in dieser ersten Versuchsreihe davon Abstand genommen. (Anm.: Der Reintonhörschwellenpegel wird bei mit Hörbehelfen ausgestatteten Personen regelmäßig gemessen und festgehalten und entspricht der Ruhehörschwelle eines Normalhörenden.)

### 3.1.2 Eingabe der personenbezogenen Daten der Versuchsperson

Um wichtige Daten der Versuchsperson aufnehmen zu können und eine automatische Datenauswertung und -speicherung zu ermöglichen, sind an dieser Stelle der Vorbereitungen einige persönliche Daten der Versuchsperson einzugeben. Die Eingabe folgender Daten ist vorgesehen:

- Vor- und Nachname
- Geschlecht
- Erst- und gegebenenfalls Zweitsprache
- Hörstörungen in der Familie
- Technische Hörversorgung
- Höralter

- Letztes Tonaudiogramm erstellt am
- Hoch- Mittel- oder Tieftonschwerhörigkeit
- Wiedergabepegel in  $dB(A)$

Die Eingabe des vollständigen Namens ist hier zwingend erforderlich, die restlichen Parameter hingegen können optional eingegeben werden. Abbildung 3.2 zeigt die vorgestellte Eingabemaske. Durch Drücken der

Abbildung 3.2: GUI - Dateneingabe

Schaltfläche „Daten speichern und Test starten“ wird eine xls-Datei der Form „Vorname Nachname.xls“ erzeugt, in der tabellarisch die eingegebenen Daten gespeichert werden.

### 3.1.3 Durchführung des Hörtests

Im Anschluss an das Exportieren der eingegebenen Daten öffnet sich das eigentliche Hauptfenster, in dem der Test abläuft. Abbildung 3.3 zeigt eine schematische Variante dieses Fensters. An der Position, an der die drei hellblauen Rechtecke zu sehen sind, erscheinen die jeweiligen Grafiken des Ziel- und des Reimwortes sowie des Distraktors in zufälliger Reihenfolge. Der restliche Bildschirm bleibt in schwarz gehalten, damit die Versuchsperson möglichst auf die Testbilder fokussiert bleibt. Die ersten drei Triplets sind als Trainingstriplets gedacht und sollen der Versuchsperson eine Gewöhnung an die Bedienung ermöglichen. Ab dem vierten Triplet werden alle Antworten protokolliert und für die Auswertung herangezogen. Um den Fluss nicht zu unterbrechen, wird nach dem dritten Triplet nicht mehr explizit darauf hingewiesen.

Sobald ein Bild durch eine Fingerberührung gewählt wurde, werden alle drei Bilder ausgeblendet um der Versuchsperson ein Feedback über die erfolgreiche Eingabe zu geben. Der Test wird an dieser Stelle erst dann fortgesetzt, wenn der/die Versuchsleiter/Versuchsleiterin die Schaltfläche „W“ am rechten unteren

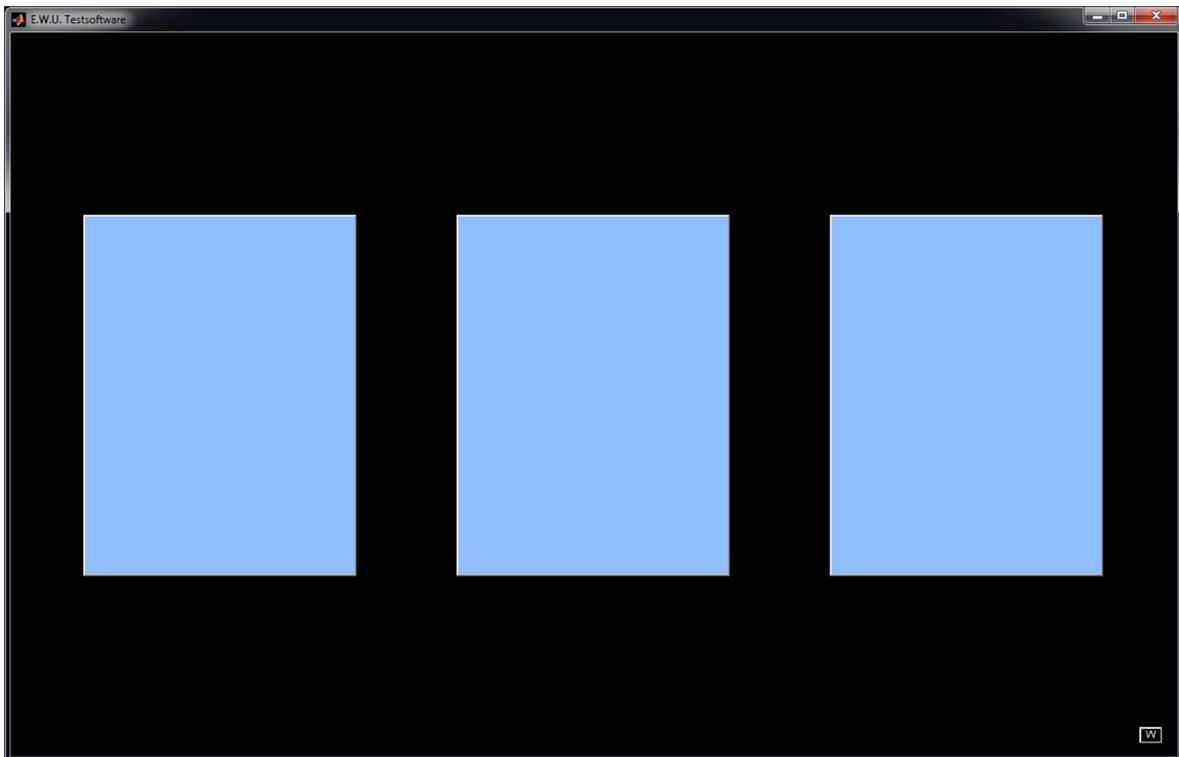


Abbildung 3.3: GUI - Hauptfenster

Bildschirmrand drückt. Erst durch diese Eingabe wird das nächste Soundfile abgespielt und das entsprechende Bildertrio geladen. Der Grund, warum die Entscheidung gegen ein vollautomatisches Laden der nächsten Files getroffen wurde ist der, dass der/die Versuchsleiter/Versuchsleiterin durch diese Form des Ablaufs jederzeit die Möglichkeit hat, auf unkonzentriertes Verhalten der Versuchsperson oder anderweitige Störungen einzugehen und ein Neustart des Tests dadurch vermieden werden kann.

Da es in der Verantwortung der/des jeweiligen Versuchsleiterin/Versuchsleiters liegt, die einzelnen Tripletts manuell weiterzuschalten, sollte darauf geachtet werden, die Zeitintervalle zwischen den aufeinander folgenden Tripletts innerhalb von  $\pm 10\%$  konstant zu halten. Es wird weiter empfohlen, ein Mindestzeitintervall von nicht weniger als 4 Sekunden einzuhalten [Eur98].

Am Ende des Hörtests kann der/die Versuchsleiter/Versuchsleiterin entscheiden, ob der Test ohne eine Speicherung der Ergebnisse beendet werden soll (Schaltfläche „Beenden“), ob der Test beendet werden soll und die Ergebnisse gespeichert werden sollen (Schaltfläche „Speichern und Beenden“), oder ob alle Ergebnisse gespeichert werden sollen und der Test im Anschluss daran neu gestartet werden soll (Schaltfläche „Speichern und Neustarten“) (siehe Abbildung 3.4). Wird die Variante gewählt, bei der der Test wieder neu gestartet wird, so wird der Bildschirm zur Dateneingabe aufgerufen, da nun eine erneute Kalibrierung in den meisten Fällen nicht benötigt werden würde.

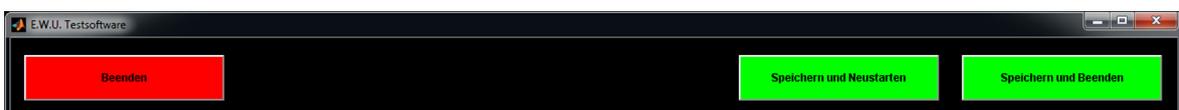


Abbildung 3.4: GUI - Auswahl am Ende des Tests

### 3.1.4 Speicherung

Wird in der letzten Auswahl eine Schaltfläche gedrückt, die das Speichern der Daten nach sich zieht, so wird die zuvor erzeugte xls-Datei um folgende, in dem gerade erfolgten Testdurchlauf erhobenen Daten erweitert:

- Anzahl der richtig gewählten Wörter
- Anzahl der falsch gewählten Wörter
- Anzahl der gewählten Distraktoren
- Prozent der richtig gewählten Wörter
- Dezidierte Angabe der richtig gewählten Wörter
- Dezidierte Angabe der falsch gewählten Wörter
- Dezidierte Angabe der Reimwörter zu falsch gewählten Wörter

Zusätzlich wird eine jpg-Grafik ausgegeben, welche den Namen der Versuchsperson trägt. Diese Grafik zeigt die Verteilung laut Otsu (Wertebereich: 1 - 3) aller richtig und aller falsch gewählten Wörter in Abhängigkeit der 3 untersuchten Parameter Anlautfrequenz, Zentroid und Flachheit. Weiters wird eine Grafik ausgegeben, welche die selbe Analyse für die zugehörigen Reimwörter der falsch gewählten Bilder zeigt. Das ermöglicht eine detailliertere Analyse der Ergebnisse. Insbesondere für den Fall, dass ein bestimmter Anlaut, also eine Konsonant, mehrmals verwechselt wird.

Abbildung 3.5 zeigt abschließend ein vereinfachtes Flussdiagramm der Testsoftware.

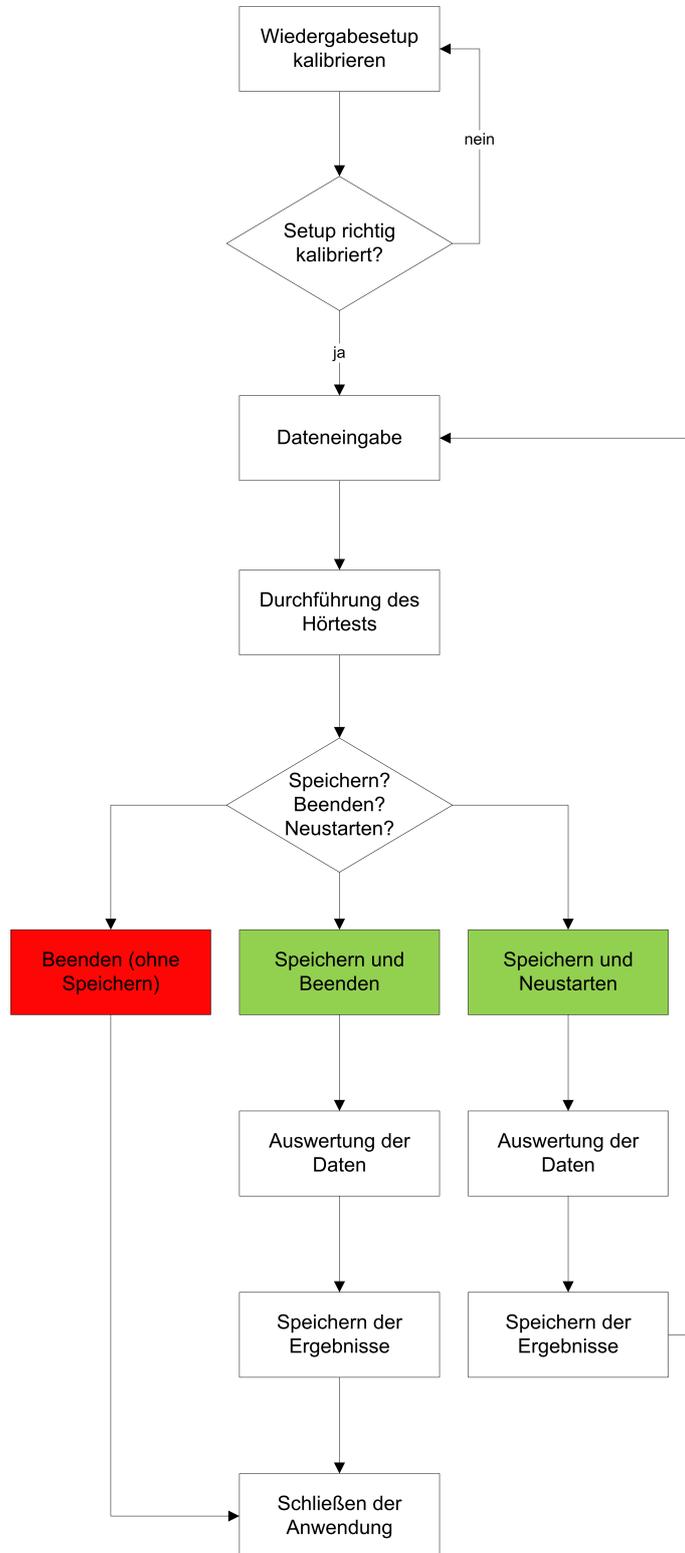


Abbildung 3.5: GUI - Flussdiagramm

# Kapitel 4

## Testablauf

### 4.1 Vorbereitung der Probanden/Probandinnen

Es ist anzustreben, den Testdurchlauf ohne Unterbrechungen und ohne zusätzliche Kommunikation mit den Testpersonen durchzuführen. Das erfordert, speziell für die Zielgruppe dieses Test, eine gründliche Einweisung der Versuchsteilnehmer/Versuchsteilnehmerinnen durch den/die Versuchsleiter/Versuchsleiterin.

Um ein weitestgehend unverfälschtes Testergebnis zu erhalten ist es wichtig, möglichst alle äußeren Einflüsse die den Testablauf negativ beeinflussen könnten, zu beseitigen (siehe auch Anhang A.1). Es sollte versucht werden, die Versuchsperson mindestens fünf Minuten vor dem Test in den Raum zu bringen um sie mit der Umgebung vertraut zu machen. Weiters wird so sichergestellt, dass der Hörschwellenpegel zu Testbeginn jener ist, welcher sich für die jeweilige Person als normal auszeichnet. Sollte es vor dem Test zu signifikanter Lärmeinwirkung gekommen sein, sollte dies vermerkt werden [Eur98].

Im Idealfall geht dem Reimtest eine otoskopische Untersuchung voraus, die von einer qualifizierten Person durchgeführt wird. Da sich aber die Zielgruppe für diesen Test, aufgrund der ständigen Versorgung mit Hörbehelfen, in regelmäßigen Abständen solchen Untersuchungen unterziehen muss, kann in der Regel von der Voruntersuchung abgesehen werden.

Die Fähigkeit der Probanden/Probandinnen das Testmaterial zu verstehen muss unbedingt gewährleistet sein. Dies wurde zwar weitestgehend bei der Zielitemauswahl durch Fr. Mag.<sup>a</sup> Scheikl sichergestellt, sollte aber aufgrund der erheblichen Wichtigkeit hier nochmals erwähnt sein.

### 4.2 Evaluierungstest mit hörenden Kindern

Um die Eignung des EWU-Sprachtests in der Praxis zu überprüfen, wurde ein Vortest mit hörenden Kindern unter realen Testbedingungen durchgeführt. Das Hauptziel dieser Testreihe war die Überprüfung folgender Eigenschaften:

- Deckung des passiven Wortschatzes der Zielgruppe mit den verwendeten Worttriplets
- Eignung des Bildmaterials für die Zielgruppe
- Eignung der Präsentationsform

### 4.2.1 Testdurchführung

Der Test wurde unter den in Anhang A empfohlenen Randbedingungen am Förderzentrum des Landes Steiermark für Hör- und Sprachbildung durchgeführt. Das Testverfahren wurde bei 51 hörenden Kindern (darunter 6 zweisprachig aufwachsende Kinder) des Heilpädagogischen Kindergartens für Hör- und Sprachbildung des Landes Steiermark erprobt. Die erste Teilstichprobe setzte sich aus 22 Mädchen und 29 Buben im Alter zwischen 3,3 und 6,1 Jahren zusammen. Der Mittelwert von 4,4 Jahren liegt im unteren Durchschnitt des für den EWU-Test vorgesehenen Altersbereiches und lässt somit Rückschlüsse auf die geplanten Testreihen mit hörbeeinträchtigten Kindern zu.

Im Zuge des Sprachtests wurde von jedem Kind der Normakusis (normale Hörfähigkeit) festgestellt. Anschließend wurde der EWU-Test mit einem Wiedergabepegel von 65 dB(A) (gemessen in 1 m Entfernung) durchgeführt. Laut Literatur entspricht dieser Schalldruckpegel dem Schalldruckpegel für „normale Umgangssprache“ [Leh01].

### 4.2.2 Ergebnisse

Von den 27 Zielitems wurden 12 Zielitems von allen hörenden Kindern korrekt erkannt. Die verbleibenden 15 Zielitems wurden mit folgender Lösungshäufigkeit gewählt:

Nr.	Zielitem	Häufigkeit	Reimwort	Häufigkeit	Distraktor	Häufigkeit
1	Herd	74,5 %	Pferd	21,6 %	Zug	3,9 %
2	Rock	90,2 %	Stock	7,8 %	Bär	2,0 %
3	Bauch	92,2 %	Rauch	2,0 %	Keks	5,9 %
4	Fest	92,2 %	Nest	3,9 %	Schiff	3,9 %
5	Mutter	94,1 %	Butter	5,9 %	Stiege	-
6	Rüssel	94,1 %	Schüssel	5,9 %	Blume	-
7	Nase	96,1 %	Vase	2,0 %	Sessel	2,0 %
8	Rad	96,1 %	Bad	-	Kuss	3,9 %
9	Wecker	96,1 %	Bäcker	3,9 %	Auto	-
10	Geld	98,0 %	Zelt	2,0 %	Bus	-
11	Haus	98,0 %	Maus	2,0 %	Zwerg	-
12	Mauer	98,0 %	Bauer	-	Suppe	2,0 %
13	Schere	98,0 %	Beere	-	Kübel	2,0 %
14	Stern	98,0 %	Kern	2,0 %	Ball	-
15	Teller	98,0 %	Keller	2,0 %	Hase	-

Tabelle 4.1: Lösungshäufigkeiten der falsch erkannten Zielitems

Nach diesem Testdurchgang wurden, unter Berücksichtigung der Ergebnisse, Änderungen an jenen Bildern vorgenommen, bei denen - nach subjektiver Einschätzung der Versuchsleiterin - undeutliche Bild Darstellungen das Lösungsverhalten negativ beeinflusst haben könnten. Der Bildaustausch fand bei den Wörtern **Rad**, **Butter**, **Stock**, **Rock**, **Vase** und **Keks** statt.

Diese abgeänderte Zusammenstellung an Bildern und Items stellt nun das endgültige Testsetup dar und wird für alle weiteren Versuche verwendet werden.

### 4.3 Simulierter Testdurchlauf

Da es bis dato leider nicht möglich war eine Testreihe mit der Zielgruppe durchzuführen, folgen an dieser Stelle einige simulierte Testszenarien und deren Interpretation.

#### 4.3.1 Alle Wörter richtig erkannt

In diesem Fall liefert der Test keine nützlichen Informationen.

Achtung: Da dieser Reimtest als zusätzliches Instrumentarium zur gezielten Detektion von Hörschwächen entwickelt wurde, gibt er keine (!) allgemeine Auskunft über das Hörvermögen der jeweiligen Versuchspersonen. Das bedeutet, aus den Diagrammen können keine spezifischen, frequenzabhängigen Hörkurven für die Versuchspersonen abgeleitet werden.

#### 4.3.2 Ein spezieller Konsonant wird nicht erkannt

Tritt dieser Fall ein, ist das alleine aus der xls-Datei erkennbar, da es hier eine dezidierte Angabe der Reimwörter zu den jeweiligen falsch gewählten Wörtern gibt. Hätte ein Kind allerdings eine Hörschwäche in einem bestimmten Frequenzbereich, so würde sich das wahrscheinlich in der Fehldetektion von bestimmten, sich ähnelnden Anlauten (also Konsonanten) widerspiegeln (siehe auch 4.3.4). Abbildung 4.1 stellt ein Ergebnis dar, welches auf eine Fehldetektion des Konsonanten „S“ zurückzuführen ist. Drei von vier Zielitems der falsch erkannten Wörter weisen hier den für einen Zischlaut typischen hohen Otsu-Wert 3 (Viel Energie im hochfrequenten Spektralbereich) bei der Anlautanalyse auf.

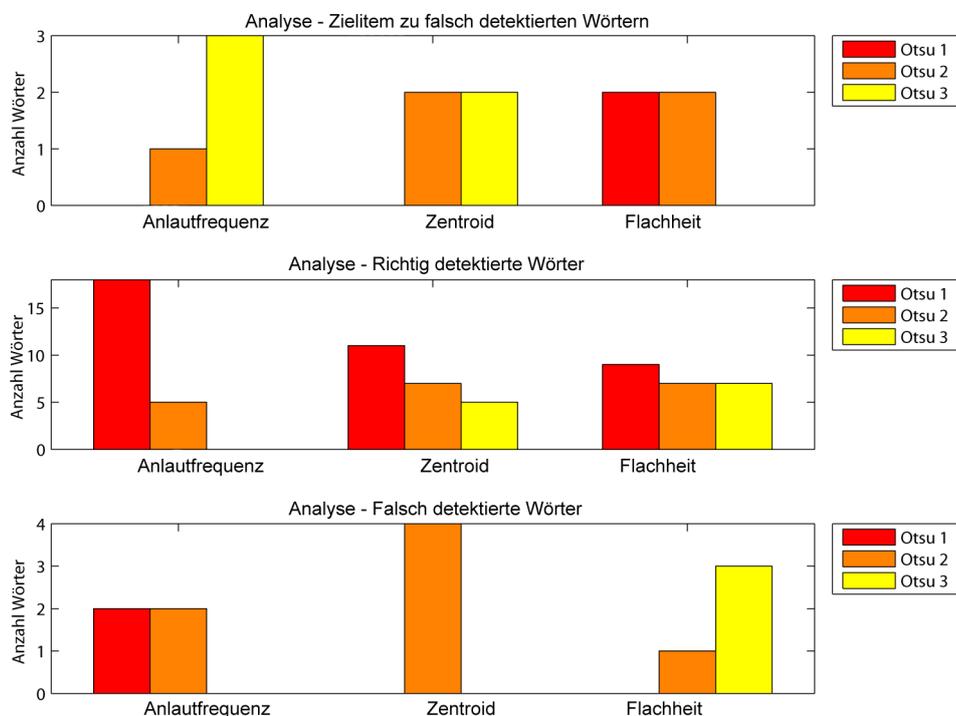


Abbildung 4.1: Detailanalyse; Szenario: alle Wörter mit „S“ nicht erkannt

### 4.3.3 Anzahl der gewählten Distraktoren zu hoch

Wird eine zu hohe Anzahl an Distraktoren gewählt, wird das Testergebnis als ungünstig deklariert, da anzunehmen ist, dass die Versuchsperson nicht konzentriert war, oder den Wortschatz, welcher nötig ist um ein sinnvolles Testergebnis zu erhalten, nicht zur Gänze beherrscht. Dies wird im Analysefile vermerkt. Die erhaltenen Ergebnisse werden jedoch trotzdem gespeichert.

### 4.3.4 Allgemeiner Fall/mögliches reales Testszenario

Sehr viel wahrscheinlicher ist es aber, dass sich mögliche Hördefizite nicht über ein und denselben Anlaut, bzw. dessen spezifische Anlautfrequenz detektieren lassen. In diesem Fall muss die Fehlhörigkeit anhand der Parameter Zentroid bzw. Flachheit analysiert werden. Hätte die getestete Person zum Beispiel eine Hörschwäche, welche sich im oberen Frequenzbereich des Zentroids widerspiegelt (Otsu 3), wäre das in dem zugehörigen Diagramm eindeutig nachweisbar (siehe Abbildung 4.2). Hier weisen alle sieben zugehörigen Zielitems, welche nicht erkannt wurden, einen Otsu-Wert 3 auf.

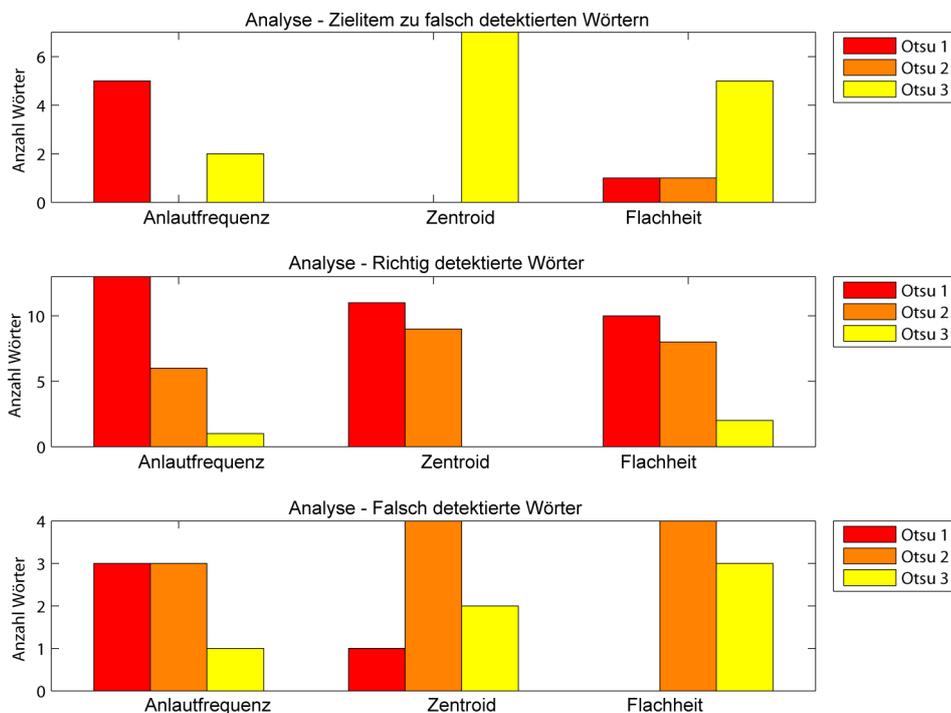


Abbildung 4.2: Detailanalyse; Szenario: Wörter mit Otsu-Wert 3 bei Zentroid nicht erkannt

Mit derselben Vorgehensweise können Defizite auch in anderen Bereichen, wie der Flachheit oder den Anlautfrequenzen, festgestellt werden. Eine wie schon in Punkt 4.3.2 angesprochene Fehlhörigkeit in einem spezifischen Frequenzbereich sollte sich mit einer erhöhten Anzahl an falsch detektierten Wörtern eines bestimmten Otsu-Wertes im Anlautfrequenzspektrum niederschlagen.

Um die Ergebnisse der Auswertung weiter zu festigen, sollten sich die Analysewerte zwischen richtig detektierten Wörtern und den Zielitems zu den falsch detektierten Wörtern invers verhalten. Das wäre die Bestätigung für ein Hördefizit in einem Bereich innerhalb einer der drei analysierten Parameter.

## Kapitel 5

# Zusammenfassung und Ausblick

Vergleicht man das in dieser Arbeit entwickelte Testverfahren mit im deutschsprachigen Raum bereits etablierten Tests, so eröffnet der EWU-Test die Möglichkeit, dem zu Beginn dieser Arbeit definierten Ziel der Verbesserung der Hörfähigkeit und einer damit verbundenen besseren Entwicklung des Kindes mit Hörbeeinträchtigung, ein Stück weit näher zu kommen. Das Abwenden vom Ansatz der reinen Diagnostik hin zu einer Kombination aus Diagnostik und Ursachenforschung, scheint hier einen besonders reizvollen Ansatz darzustellen.

In künftigen Arbeiten auf Basis des EWU-Tests, könnte eine immense Qualitätssteigerung dieses Testverfahrens vor allem durch eine Weiterentwicklung in enger Zusammenarbeit mit den Herstellern der Hörbehelfe erreicht werden. Die genaue Kenntnis der digitalen Signalverarbeitung in den unterschiedlichsten Hörbehelfen würde es ermöglichen, zusätzliche bzw. alternative Signalparameter zu definieren, mit Hilfe derer es unter Umständen möglich wäre, einen direkten Zusammenhang zwischen den Signaleigenschaften der Testwörter und deren Verarbeitung in der jeweiligen apparativen Versorgung herzustellen.

Im Zuge der ersten Evaluierungstests mit hörenden Kindern im Alterszielbereich von 4-7 Jahren hat sich gezeigt, dass sogar die bereits sehr stark begrenzte Testlaufzeit manchmal ein Problem darstellt. Besonders am unteren Ende der definierten Altersgrenze weisen die Probanden/Probandinnen teilweise Konzentrationsschwächen auf, die mit der aktuellen Umsetzung des Testverfahrens schwer kompensierbar sind. Hier wäre es notwendig, weitere Überlegungen zur Präsentation der Testwörter anzustellen. Der Mittelweg zwischen einem automatisierten, objektiven Testablauf und einer möglicherweise notwendigen Unterstützung des/der Versuchsleiters/Versuchsleiterin, sollte hier genau untersucht werden.

Aufgrund der definierten Zielgruppe war es bis dato leider nicht möglich, eine ausreichend große Anzahl an Testpersonen zu rekrutieren. Eine ausgedehnte Versuchsreihe mit einer ausreichend hohen Anzahl an Probanden/Probandinnen, wird aller Voraussicht nach im Rahmen der Dissertation von Frau Mag.<sup>a</sup> Scheikl durchgeführt werden.

# Literaturverzeichnis

- [ABDL02] S. Allegro, M. Büchler, N. Diller, and S. Launer, "Algorithmen für die Geräuschklassifizierung in Hörgeräten," in *Proceedings zur DGA-Jahrestagung*, Zürich, CH, 2002.
- [Bal99] P. Balabko, "Speech and Music Discrimination based on Signal Modulation Spectrum," Idiap Research Institute, Martigny, CH, Tech. Rep., 1999.
- [Bis05] S. Bischoff, "O H R Olsztyner Hör Reime," Ph.D. dissertation, Hochschule Heidelberg, 2005.
- [BKH06] A. Bockmann and C. Kiese Himmel, *ELAN - Eltern Antworten - Elternfragebogen zur Wortschatzentwicklung im frühen Kindesalter*. Göttingen, DE: Beltz Test, 2006.
- [BWM05] G. Böhme and K. Welzl Müller, *Audiometrie: Hörprüfungen im Erwachsenen- und Kindesalter*. Bern, CH: Huber, 2005.
- [CG05] T. U. Christiansen and S. Greenberg, "Frequency Selective Filtering of the Modulation Spectrum and its impact on consonant identification," Centre for Applied Hearing Research, Technical University of Denmark, Lyngby, DK, Tech. Rep., 2005.
- [Che53] E. C. Cherry, "Some experiments on the recognition of speech, with one and with two ears," *The Journal of the Acoustical Society of America*, vol. 25, no. 5, pp. 975–979, 1953.
- [Deu97] Deutsches Institut für Normung e.V., "DIN EN 60645-2:1997 Audiometer Teil 2: Geräte für die Sprachaudiometrie," 1997.
- [Eur92] Europäisches Komitee für Normung, "EN ISO 8253-2:1992 Audiometrische Prüfverfahren, Schallfeldaudiometrie mit reinen Tönen und schmalbandigen Prüfsignalen," 1992.
- [Eur98] Europäisches Komitee für Normung, "EN ISO 8253-3:1998 Audiometrische Prüfverfahren, Sprachaudiometrie," 1998.
- [Eys05] U. Eysholdt, "Pädaudiologische Diagnostik," in *Lehrbuch der Phoniatrie und Pädaudiologie*, J. Wendler, W. Seidner, and U. Eysholdt, Eds. Stuttgart, DE: Thieme, 2005.
- [GAF00] H. Grimm, M. Aktas, and S. Frevert, *SETK-2 - Sprachentwicklungstest für zweijährige Kinder - Diagnose rezeptiver und produktiver Sprachverarbeitungsfähigkeiten*. Göttingen, DE: Hogrefe, 2000.
- [GD06] H. Grimm and H. Doil, *ELFRA - Elternfragebögen für die Früherkennung von Risikokindern*. Göttingen, DE: Hogrefe, 2006.
- [Hör00] HörTech GmbH, *Oldenburger Satztest, Handbuch und Hintergrundwissen*, Oldenburg, DE, 2000.

- [HZ08] H. Höfler and P. Zorowka, "Grundlagen V: Ohr und Gehör," in *Phoniatrie und Pädaudiologie*, G. Friedrich, W. Bigenzahn, and P. Zorowka, Eds. Bern, CH: Huber, 2008.
- [KK05] C. Kiese and P. Kozielski, *AWST-R Aktiver Wortschatztest für 3- bis 5-jährige Kinder*. Göttingen, DE: Beltz Test GmbH, 2005.
- [Kom04] M. Kompis, *Audiologie*. Bern, CH: Huber, 2004.
- [Lar09] O. Lartillot, *MIRtoolbox 1.2 Users Manual*, Jyväskylä, FIN, 2009.
- [Leh01] E. Lehnhardt, "Hörprüfungen beim Säugling und Kleinkind," in *Praxis der Audiometrie*, E. Lehnhardt and R. Laszig, Eds. Stuttgart, DE: Thieme, 2001.
- [LT07] O. Lartillot and P. Toiviainen, "A Matlab Toolbox for Musical Feature Extraction From Audio," in *10th International Conference on Digital Audio Effects Proceedings*, Bordeaux, FR, 2007.
- [MS06] D. Mrowinski and G. Scholz, *Audiometrie, Eine Anleitung für die praktische Hörprüfung*. Stuttgart, DE: Thieme, 2006.
- [Muh07] R. Muhr, *Österreichisches Aussprachewörterbuch Österreichische Aussprachedatenbank*. Wien, AUT: Lang, 2007.
- [Ots79] N. Otsu, "A threshold selection method from Grey-Level Histograms," in *IEEE Trans. on Systems, Man and Cybernetics*, vol. SMC-9, No. 1, 1979, pp. 62–66.
- [Pee04] G. Peeters, "A large set of audio features for sound description (similarity and classification) in the CUIDADO project," Institut de Recherche et Coordination Acoustique/Musique (IRCAM), Paris, FR, Tech. Rep., 2004.
- [PP09] J. Přibíl and A. Přibílová, "Spectral flatness analysis for emotional speech synthesis and transformation," in *Cross-Modal Analysis of Speech, Gestures, Gaze and Facial Expressions*, A. Esposito and R. Vich, Eds. Berlin, Heidelberg: Springer, 2009.
- [Thi00] M. M. Thiel, *Logopädie bei kindlichen Hörstörungen: Ein mehrdimensionales Konzept für Therapie und Beratung*. Berlin, Heidelberg, DE: Springer, 2000.
- [TSW04] A. Tarnopolski, E. Schubert, and J. Wolfe, "Spectral centroid and timbre in complex, multiple instrumental textures," in *Proceedings of the 8th International Conference on Music Perception and Cognition*, Evanston, IL, USA, 2004.

## Anhang A

# Randbedingungen für die Testdurchführung

### A.1 Anforderungen an der Prüfraum

Da dieser Test in erster Linie für eine Zielgruppe im Altersbereich von 4-7 Jahren ausgelegt ist und diese Zielgruppe aufgrund der körperlichen Beeinträchtigungen ständig mit verschiedensten Testverfahren belastet wird, ist es unser Ziel den jungen Probanden/Probandinnen ein möglichst angenehmes Umfeld zu schaffen, sodass wir uns entschieden haben, den Test am jeweiligen Institut für Hör- und Frühförderung stattfinden zu lassen. Aus diesem Grund werden die raumakustischen Anforderungen auf ein notwendiges Minimum reduziert. Zusätzlich zu den akustischen Anforderungen, spielt in unserem Fall der visuelle Aspekt eine nicht zu vernachlässigende Rolle. Es hat sich herausgestellt, dass sich eine kindgerecht gestaltete Umgebung positiv auf das Testverhalten der Probanden/Probandinnen auswirkt. Es empfiehlt sich deshalb eine Umgebung, ähnlich der in Abbildung A.1 zu schaffen.

Es ist weiters anzustreben, einen Testraum mit einer Nachhallzeit  $t_h < 0.5$  s zu wählen [Eur98]. Idealerweise soll der Raum einen Grundgeräuschpegel von nicht mehr als 30  $dB(A)$  aufweisen, um einen Mindestsignalrauschabstand in Bezug auf das Testsignal von 35  $dB$  zu gewährleisten. Es sollten sich auch keine reflektierenden Flächen in unmittelbarer Nähe des Lautsprechers befinden.

### A.2 Empfehlungen zur Wiedergabehardware

Grundsätzlich erscheint es sinnvoll, eine rein elektronische Durchführung des Tests anzustreben. Dies setzt voraus, dass sowohl die gesamte Darbietung, als auch die Ausgabe der Ergebnisse rechnerunterstützt erfolgt. Die für den Test benötigte Hardware besteht aus einem einzigen Gerät, das sowohl die Bilddarstellung (Touchscreen) als auch die Audioausgabe (Mindestens 1 Ausgabekanal, Line Signal oder Kopfhörerausgang) sowie die Speicherung der Ergebnisse übernimmt und einem Lautsprecher der die empfangenen elektrischen Signale in Schall umwandelt und den Anforderungen in Anhang A.2.2 entspricht.

#### A.2.1 Tablet-PC

Für die erfolgreiche Durchführung des Tests wurden vorab folgende Anforderungen an die Wiedergabehardware gestellt:



Abbildung A.1: Beispielumgebung für die Testdurchführung

- Portabilität des Systems
- intuitive/kindgerechte Bedienung
- Wiedergabe und Ergebnisausgabe in einem Gerät
- sehr leises/unhörbares System
- keine Abschirmung des einfallenden Schalls durch das System

Alle diese Anforderungen werden durch den in Grafik A.2 bzw. Tabelle A.1 angeführten, speziell für diese Anwendung konfigurierten Tablet-PC erfüllt.

Herauszustreichen ist hierbei der kapazitive Touchscreen, der eine Bedienung mit den Fingern, ohne weitere Hilfsmittel ermöglicht, sowie die Verwendung eines Solid-State-Laufwerks, das aufgrund nicht vorhandener beweglicher Teile und einer geringen Wärmeentwicklung einen besonders geräuscharmen Betrieb begünstigt.

### A.2.2 Lautsprecher

Da die Durchführung des Tests in möglichst variabler Umgebung ermöglicht werden soll und in den meisten für Sprachaudiometrie geeigneten Prüfräumen bereits Lautsprecher vorhanden sind, wird an dieser Stelle eine Mindestanforderung für die Lautsprecher formuliert.



Abbildung A.2: HP EliteBook 2740p

OS	Microsoft Windows 7 Prof.
CPU	Intel Core i5
Display	12,1" WXGA-AntiGlare (Touch, kapazitiv)
Internes Laufwerk	SSD

Tabelle A.1: Auszug aus der Konfiguration des verwendeten HP EliteBook 2740p

### Empfehlungen zum Lautsprechersetup

Das Testverfahren wird ohne Störschall umgesetzt. Es wird empfohlen, den Lautsprecher frontal zum/zur Probanden/Probandin mit einem Einfallswinkel von  $0^\circ$  zu positionieren [Eur98]. Die Darbietung aus der  $0^\circ$ -Richtung mit nur einem Lautsprecher wurde empfohlen, da dieser Messaufbau einer realen Gesprächssituation am nächsten kommt. Die Entfernung des Lautsprechers vom gedachten Kopfbereich der Versuchsperson sollte mindestens 1 m betragen und der Lautsprecher sollte sich in Ohrenhöhe befinden [Eur92]. Sind in einem Testraum zwei Lautsprecher in Standardaufstellung (Stereodreieck) vorhanden, so kann einer der beiden Lautsprecher in die frontale Position gebracht werden. Der zweite Lautsprecher muss abgeschaltet bzw. stumm geschaltet werden. Die präsentierten Signale sind Monofiles, d.h. es kann der linke oder der rechte Lautsprecher verwendet werden. Für die Kalibrierung und den Testdurchlauf muss unbedingt der selbe Lautsprecher verwendet werden.

### Kenndaten des Lautsprechers

Bei der Auswahl des Lautsprechers sollte insbesondere darauf Wert gelegt werden, dass dieser einen möglichst glatten Frequenzgang zwischen 20 Hz und 10 kHz aufweist. Nur so kann gewährleistet werden, dass die Signale unverfälscht abgespielt werden.

Die ersten Durchgänge unseres Tests wurden am Institut für Hör- und Frühförderung durchgeführt. Als Abhörlautsprecher kam hier das Modell GB100 der Firma Gravenstein zum Einsatz, welches bis dato standardmäßig von der Firma Neuroth an diesem Standort für Hörversuche verschiedenster Art verwendet wird. Datenblatt zu diesem Lautsprecher ist leider keines mehr vorhanden, da die Firma Gravenstein schon seit längerem nicht mehr existiert. Der Frequenzgang wurde allerdings von einem Techniker der Firma Neuroth vermessen und ist zwischen 20 Hz und 12 kHz sehr linear. Weitere Eckdaten die der Lautsprecherplakette entnommen wurden:

- Modell: Gravenstein GB100
- Impedanz: 8 Ohm
- Leistung: 32 Watt

## Anhang B

# Technische Dokumentation

### B.1 Kalibriersignal

Als Kalibriersignal wird ein frequenzbewertetes Rauschen mit konstantem Spektraldichtepegel innerhalb des Frequenzbereiches von 125 Hz bis 1 kHz und oberhalb von 1 kHz um 12 dB/Okt. abfallend (Grenzabweichungen = +/- 5 dB) verwendet [Deu97, Eur98]. Abbildung B.1 zeigt das Leistungsdichtespektrum des Kalibriersignals.

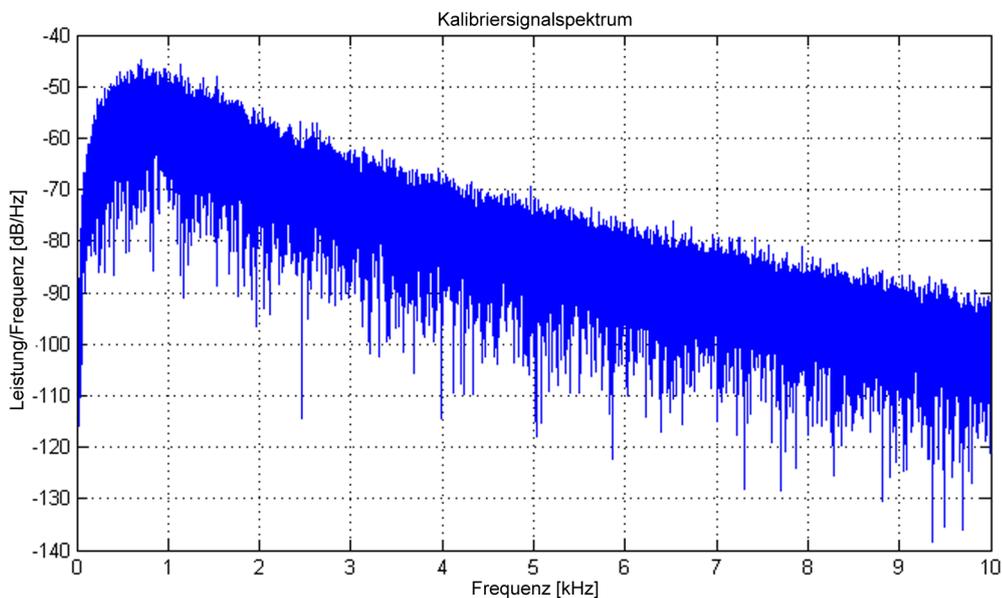


Abbildung B.1: Leistungsdichtespektrum des Kalibriersignals

#### B.1.1 Pegelanpassung zwischen Kalibriersignal und Zielitem

Um das korrekte Schalldruckpegelverhältnis zwischen Kalibriersignal und Zielitem in der Wiedergabe zu erreichen, wurde der RMS-Wert des Rauschsignals mit einem Proportionalitätsfaktor an den durchschnittlichen RMS-Wert der Zielitems angepasst. Hierzu wurden sowohl das Rauschsignal, als auch die Zielitems vor der RMS-Berechnung zusätzlich mit einer dB(A)-bewerteten Filterkurve behandelt, um so die Relation

zur natürlichen Filterkurve des menschlichen Gehörs beizubehalten. Um das laut DIN 60645-2 [Deu97] geforderte Frequenzspektrum des Kalibriersignals zu erstellen, wurde das Rauschsignal mit einem Butterworth Filter 2. Ordnung (Cut Off Frequency  $f_c = 1$  kHz) behandelt.

## B.2 Technische Spezifikationen

### B.2.1 Spezifikationen der Audiosignale

Alle Audiosignale, sowie das Kalibriersignal liegen in folgendem Format vor:

Audioformat	Wave (PCM/uncompressed)
Bitrate	16 Bit
Samplingfrequenz	44100 Hz

Tabelle B.1: Spezifikationen der Audiosignale

### B.2.2 Spezifikationen der Entwicklungsumgebung

Der gesamte EWU-Test wurde auf folgenden Plattformen entwickelt und/oder erfolgreich getestet:

PC	HP EliteBook 8510p
OS	Windows Vista Business SP1 inkl. Sysupdates bis 01.2011
Entwicklungssoftware	Matlab R2008a
PC	HP EliteBook 8530w
OS	Windows 7 Ultimate inkl. Sysupdates bis 01.2011
Entwicklungssoftware	Matlab R2010a

Tabelle B.2: Spezifikationen der Entwicklungsumgebung

### B.3 Detailanalyse aller Audiofiles

Wort	Anlautfrequenz	Zentroid	Flachheit	Otsu Frequenz	Otsu Zentroid	Otsu Flachheit
Bäcker	258	3817	0,42	1	2	3
Bad	258	2993	0,31	1	2	3
Bauch	172	1774	0,16	1	1	2
Bauer	257	2879	0,31	1	2	3
Beere	172	4188	0,45	1	2	3
Bein	258	2033	0,15	1	1	2
Biene	344	3084	0,25	1	2	3
Buch	258	3469	0,24	1	2	2
Butter	258	3934	0,33	1	2	3
Dose	172	2341	0,23	1	1	2
Fest	344	8013	0,29	1	3	3
Fisch	345	8399	0,43	1	3	3
Geld	2670	4060	0,3	2	2	3
Hahn	1378	3222	0,11	2	2	1
Hand	1720	2491	0,11	2	2	1
Haus	1120	2660	0,09	2	2	1
Herd	432	3625	0,17	1	2	2
Hose	434	2337	0,16	1	1	2
Hund	775	1506	0,09	2	1	1
Keller	2412	4173	0,25	2	2	3
Kern	2498	4617	0,25	2	2	3
Koch	960	3632	0,19	2	2	2
Kuh	1034	3866	0,19	2	2	2
Leiter	172	1827	0,18	1	1	2
Loch	172	1079	0,09	1	1	1
Mauer	172	1951	0,21	1	1	2
Maus	172	1535	0,16	1	1	2
Mund	172	1513	0,16	1	1	2
Mutter	172	2481	0,26	1	2	3
Nase	172	1405	0,14	1	1	1
Nest	172	1971	0,22	1	1	2
Pferd	345	7190	0,36	1	3	3
Rad	172	1280	0,07	1	1	1
Rauch	172	1506	0,08	1	1	1
Reiter	172	1233	0,07	1	1	1
Rock	172	1189	0,06	1	1	1
Rüssel	172	1032	0,07	1	1	1
Sand	9905	8819	0,29	3	3	3
Schere	2844	4686	0,15	2	2	2
Schiene	3790	5000	0,14	3	3	1
Schuh	5082	5278	0,16	3	3	2

Wort	Anlautfrequenz	Zentroid	Flachheit	Otsu Frequenz	Otsu Zentroid	Otsu Flachheit
Schüssel	5770	5079	0,1	3	3	1
See	5149	8493	0,23	3	3	2
Stein	3530	5015	0,17	3	3	2
Stern	3531	4604	0,13	3	2	1
Stock	4910	4848	0,16	3	3	2
Tee	345	6059	0,26	1	3	3
Teller	172	5356	0,35	1	3	3
Tisch	2325	5806	0,36	2	3	3
Tuch	1119	4894	0,2	2	3	2
Vase	172	1613	0,1	1	1	1
Wecker	172	5419	0,29	1	3	3
Zahn	9647	9646	0,31	3	3	3
Zelt	9819	8933	0,28	3	3	3

Tabelle B.3: Detailanalyse der Zielitems und Reimwörter