

TUM School of Computation, Information and Technology

**Entwicklung und Implementierung psycho-
akustischer Testverfahren zur Untersuchung
des Hörorgans von Kleinkindern**

Volker Schirkonyer

Vollständiger Abdruck der von der TUM School of Computation, Information and
Technology der Technischen Universität München zur Erlangung des akademischen Grades
eines

Doktors der Ingenieurwissenschaften (Dr.-Ing.)

genehmigten Dissertation.

Vorsitz: Prof. Dr.-Ing. Bernhard Seeber

Prüfer der Dissertation: 1. Prof. Dr.-Ing. W. Hemmert

2. apl. Prof. Dr.-Ing. Dr. med. habil. T. Janssen

Die Dissertation wurde am 16.11.2023 bei der Technischen Universität München eingereicht
und durch die TUM School of Computation, Information and Technology am 14.06.2024 an-
genommen.

Danksagung

An dieser Stelle möchte ich mich bei allen Personen bedanken, die mich bei der Durchführung dieser Dissertation unterstützt haben. Die wesentlichen Beiträge der vorliegenden Arbeit entstanden während meiner Tätigkeit als wissenschaftlicher Mitarbeiter in der Arbeitsgruppe Experimentelle Audiologie am Klinikum rechts der Isar der Technischen Universität München. Teile der Arbeit wurden von der Bayerischen Forschungstiftung und der Europäischen Kommission innerhalb des 7. EU-Forschungsrahmenprogramms in den Projekten *Automatisierte Hördiagnostik und Hörgeräteanpassung* und *AHEAD iii* gefördert.

Mein besonderer Dank gilt Prof. Dr. Dr. Thomas Janssen, der mir als Leiter der Experimentellen Audiologie am Klinikum rechts der Isar ein hervorragendes persönliches und fachliches Umfeld für die Durchführung meiner Arbeit bot und seiner stets ansteckenden Begeisterung für immer neue Facetten im Fachgebiet und darüber hinaus. Insbesondere danke ich ihm dafür, auch Jahre nach seinem Eintritt in den Ruhestand noch für die Betreuung und Begutachtung meiner Dissertation zur Verfügung zu stehen.

Ebenso danke ich Prof. Dr. Werner Hemmert, der durch die langjährige interdisziplinäre Zusammenarbeit zwischen dem Fachgebiet Bioanaloge Informationsverarbeitung und der Experimentellen Audiologie diese Arbeit ermöglicht hat. Das von ihm initiierte Forschungsnetz HöReN, dessen Teil ich sein darf, trägt in herausragender Weise zur interdisziplinären Vernetzung der Hörforschung bei. Zudem danke ich ihm für die administrative Unterstützung im Promotionsprozess.

Meinen Kollegen Dr. Florian Kandzia, Dr. Jörg Müller und Dr. Thomas Rosner danke ich für die langjährige, produktive und vor allem freundschaftliche Zusammenarbeit. Große Unterstützung habe ich auch durch Dr. Hans Oswald und Dr. Andre Lodwig erfahren, die mir die Mitarbeit bei der PATH MEDICAL GmbH ermöglichten, wodurch ich wertvolle Erfahrungen im kommerziellen Umfeld der Thematik sammeln konnte. Auch hat das großzügig gewährte Sabbatical entscheidend zur Fertigstellung meiner Dissertation beigetragen.

Außerdem danke ich – in alphabetischer Reihenfolge – Andrea Bohnert für die intensive Zusammenarbeit bei der Ton- und Sprachaudiometrie; Dr. Frank Böhnke für die Unterstützung mit dem Onlinesystem der TUM; Christine Harmuth und Sabine Lochmahr für die vielen Versuchspersonenmessungen; Dr. Ingo Lepper für Hinweise zu \LaTeX und Tippfehlern; Prof. Dr. Katrin Neumann für Rückmeldungen und Input zum BASD; allen Versuchspersonen, die großzügigerweise ihre Zeit und Mühe für die Teilnahme an den Messungen aufgewendet haben.

Ein besonderer Dank gilt meiner Familie, ohne die dieser Weg nicht möglich gewesen wäre.

München, im November 2023

Inhaltsverzeichnis

Verzeichnis der verwendeten Symbole	vii
Abbildungsverzeichnis	ix
Tabellenverzeichnis	xi
Zusammenfassung	xiii
1 Einleitung	1
1.1 Motivation	1
1.2 Überblick über die vorliegende Arbeit	2
2 Grundlagen	5
2.1 Überblick über das Hörsystem	5
2.1.1 Schall	5
2.1.2 Anatomie des Ohres	5
2.1.3 Schallverarbeitung	7
2.2 Audiologische Testverfahren	9
2.2.1 Tonaudiometrie	10
2.2.2 Sprachaudiometrie	11
2.3 Psychoakustik	12
2.3.1 Methoden und Verfahren der Psychoakustik	13
2.3.2 Unterschiedsschwellen	14
2.4 Kalibrierung	15
2.5 Methoden	16
3 Messapparatur	19
3.1 Hardware	19
3.2 Firmwarestruktur	20
4 Testverfahren	25
4.1 Bildgestützte Tonschwellenaudiometrie (MAGIC)	25
4.1.1 Einleitung	25
4.1.2 Testdesign	26
4.1.3 Implementierung	30
4.2 Bildgestützte Sprachaudiometrie (MATCH)	34
4.2.1 Einleitung	34
4.2.2 Erstellung von Sprach- und Bildmaterial	35
4.2.3 Design und Implementierung	41
4.3 Test zur zentralen Verarbeitung (BASD)	48
4.3.1 Einleitung	48
4.3.2 Testdesign	48
4.3.3 Implementierung	54
5 Klinische Studien	57
5.1 MAGIC	57
5.1.1 Studiendesign	57
5.1.2 Ergebnisse	58

5.2	MATCH	64
5.2.1	Studiendesign	64
5.2.2	Ergebnisse	66
5.3	BASD	76
5.3.1	Studiendesign	76
5.3.2	Ergebnisse	78
6	Diskussion	87
6.1	MAGIC	87
6.2	MATCH	90
6.3	BASD	93
7	Fazit und Ausblick	101
	Literatur	105

Verzeichnis der verwendeten Symbole

ADC	Analog to Digital Conversion
AVWS	Auditive Verarbeitungs- und Wahrnehmungsstörungen
BASD	Bochum Auditory Speech Discrimination Test
BAS	Bayerisches Archiv für Sprachsignale
BASSS	BAS Services Schiel
ca.	circa
CODEC	Coder/Decoder
Δ	Differenz
DAC	Digital to Analog Conversion
dB	Dezibel
d.h.	das heißt
DMA	Direct Memory Access
DSP	Digitaler Signalprozessor
EBIU	External Bus Interface Unit
f	Frequenz [Hz, kHz]
FAEP	Frühe Akustisch Evozierte Potentiale
GUI	Graphical User Interface
HL	Hearing Level
HNO	Hals-Nasen-Ohren
HV	Hörverlust
I/O	Input/Output
IPA	Internationales Phonetisches Alphabet
ISR	Interrupt Service Routine
JND	Just Noticeable Differences
L	Schalldruckpegel [dB SPL, dB HL, dB SL]
L	Links
Li-Ion	Lithium-Ionen
MAGIC	Multiple-choice Auditory Graphical Interactive Check
MATCH	Mainzer Audiometric Test for Children
MAUS	Munich Automatic Segmentation
MDD	Medical Device Directive (93/42/EWG)
mHV	mittlerer Hörverlust
MW	Mittelwert
n	Anzahl
OAE	otoakustische Emission(en)
p	Schalldruck
p	p-Wert (Wahrscheinlichkeit, dass die Unterschiede zufällig sind)
PPI	Parallel Peripheral Interface

PTA	Pure Tone Audiometrie
R	Rechts
r	Korrelationskoeffizient nach Pearson
ρ	Korrelationskoeffizient nach Spearman
SAMPA	Speech Assessment Methods Phonetic Alphabet
σ	Standardabweichung
SD	Standardabweichung (standard deviation)
SDRAM	Synchronous Dynamic Random Access Memory
SL	Sensation Level
SPI	Serial Peripheral Interface
SPL	Sound Pressure Level (Schalldruckpegel)
SPORT	Serial Port
SVS	Sprachverständlichkeitsschwelle
T	Task
TFT	Thin-Film Transistor
UART	Universal Asynchronous Receiver/Transmitter
USB	Universal Serial Bus
z.B.	zum Beispiel

Abbildungsverzeichnis

Abb. 2.1	Übersicht über das periphere Hörorgan	5
Abb. 2.2	Schema von Mittelohr und Cochlea.....	6
Abb. 2.3	Schematische Darstellung der Hörbahn.....	7
Abb. 2.4	Isophone, Hörfläche und Hauptsprachbereich	8
Abb. 2.5	Empfindungsfunktion und Bestimmung des Grenzwertes.....	14
Abb. 2.6	Übersicht SAMPA-Zeichen.....	17
Abb. 3.1	Produktfoto des neuen Messsystems	20
Abb. 3.2	Hardwareschema des Tischgeräts	21
Abb. 3.3	Firmwarestruktur des Messgerätes	22
Abb. 3.4	Schema der Kommunikation zwischen DSP und Peripherie	23
Abb. 4.1	MAGIC adaptiver Algorithmus	28
Abb. 4.2	MAGIC Bildkarten.....	29
Abb. 4.3	Zustände der Tiere	30
Abb. 4.4	MAGIC Messverlauf.....	30
Abb. 4.5	MAGIC Einstelloptionen während des Tests	31
Abb. 4.6	MAGIC Testeinstellungen	32
Abb. 4.7	MAGIC Ergebnisdarstellung.....	33
Abb. 4.8	Phonemstatistik.....	36
Abb. 4.9	Langzeit-Sprachspektren	38
Abb. 4.10	Sprachverständlichkeit des Testmaterials bei Erwachsenen	39
Abb. 4.11	MATCH Bildkarten	40
Abb. 4.12	MATCH Ohr-Abbildung.....	41
Abb. 4.13	MATCH PC Darstellung	43
Abb. 4.14	MATCH Testverlauf	44
Abb. 4.15	MATCH Testeinstellungen.....	45
Abb. 4.16	MATCH Ergebnisdarstellung adaptive Messung.....	46
Abb. 4.17	MATCH Ergebnisdarstellung fester Messpegel	47
Abb. 4.18	BASD Signalpräsentation	49
Abb. 4.19	Spektrogramm und Zeitsignal der Sprachstimuli	50
Abb. 4.20	Simulation der BASD Schwellenbestimmung.....	53

Abb. 4.21	BASD Messverlauf.....	55
Abb. 4.22	BASD Testeinstellungen	55
Abb. 4.23	BASD Ergebnisdarstellung.....	56
Abb. 5.1	Beispiele für Audiogramme mit Ausreißern.....	58
Abb. 5.2	Boxplot der Differenzen	59
Abb. 5.3	Korrelation zwischen MAGIC und PTA.....	60
Abb. 5.4	Differenz zwischen MAGIC und PTA in Abhängigkeit des Alters.....	60
Abb. 5.5	Reliabilität der MAGIC Test-Retest-Messungen	61
Abb. 5.6	Differenz zwischen Test und Retest in Abhängigkeit des Alters.....	62
Abb. 5.7	Messdauer pro Schwelle in Abhängigkeit des Alters.....	63
Abb. 5.8	Sprachverständlichkeit des Testmaterials bei Kindern.....	66
Abb. 5.9	Leichte und schwere Worte des Testmaterials	67
Abb. 5.10	Psychometrische Funktion der Sprachverständlichkeit.....	68
Abb. 5.11	71,4 % Sprachverständnisschwelle	70
Abb. 5.12	Streudiagramm aller Sprachverständlichkeitsschwellen	70
Abb. 5.13	Korrelation zwischen MATCH und mHV PTA	71
Abb. 5.14	Reliabilität der MATCH Test-Retest-Messungen.....	72
Abb. 5.15	Testdauer aller Kinder für ein Ohr	73
Abb. 5.16	Testdauer in Abhängigkeit des Alters.....	74
Abb. 5.17	Reaktionszeiten Streudiagramm	75
Abb. 5.18	Reaktionszeiten Heatmap.....	75
Abb. 5.19	Frequenz-Differenz Unterschiedsschwellen.....	78
Abb. 5.20	Hörpegel-Differenz Unterschiedsschwellen	79
Abb. 5.21	Tondauer-Differenz Unterschiedsschwellen.....	80
Abb. 5.22	Kinderergebnisse für sprachhaltige Tests.....	84
Abb. 5.23	Kinderergebnisse für sprachfreie Tests.....	85

Tabellenverzeichnis

Tab. 4.1	Pegelkorrekturen	39
Tab. 5.1	Ausreißer in Abhängigkeit der Frequenz	59
Tab. 5.2	Kinderbefragung zu MAGIC und PTA	63
Tab. 5.3	Testkollektiv	65
Tab. 5.4	Leicht- und schwerverständliche Wörter	67
Tab. 5.5	Kennzahlen der angepassten Diskriminationsfunktionen	69
Tab. 5.6	Zusammenfassung der gemessenen 71,4%-Schwellenwerte	69
Tab. 5.7	MATCH Testdauer pro Ohr	73
Tab. 5.8	Frequenz-Differenz Unterschiedsschwellen	78
Tab. 5.9	Hörpegel-Differenz Unterschiedsschwellen	79
Tab. 5.10	Tondauer-Differenz Unterschiedsschwellen	80
Tab. 5.11	Altersabhängige Mediane der Diskriminationsschwellen	81
Tab. 5.12	Normwerte für das jeweils 80., 85., 90. und 95. Perzentil	82
Tab. 5.13	Kinderergebnisse für sprachhaltige Tests	84
Tab. 5.14	Kinderergebnisse für sprachfreie Tests	85

Zusammenfassung

In dieser Arbeit wurden drei neue Verfahren zur Untersuchung des Gehörs bei Kleinkindern und Vorschulkindern entwickelt und getestet, um eine verbesserte und frühzeitige Diagnose von Hörstörungen zu ermöglichen.

Der „Multiple-choice Auditory Graphical Interactive Check“ (MAGIC) ist ein neuer Selbsttest für Kinder ab 3 Jahren zur Bestimmung der Reintonschwelle. Die Kinder konnten die Testdurchführung selbst per Touchscreen steuern, was die Motivation erhöhte. Der Test wurde an 156 normal- und schwerhörenden Kindern im Alter von 3;6 bis 10;3 Jahren validiert. Die mittlere Abweichung der mit MAGIC und konventioneller Tonaudiometrie (PTA) erfassten Hörschwellen betrug 0,5 dB bei einer Standardabweichung (SD) von 7,6 dB. Es bestand eine starke positive Korrelation von 0,79 ($p < 0,001$) zwischen den beiden Methoden, wie auch zwischen Test und Retest ($r = 0,81$; $p < 0,001$). Test und Retest wichen im Mittel um 1,4 dB (SD = 8,0 dB) ab. Die Kinder gaben an, dass ihnen der MAGIC-Test mehr Spaß machte und weniger anstrengend war als die herkömmliche Methode. Der MAGIC stellt somit eine gute Ergänzung zu den etablierten Testverfahren in der Audiologie dar.

Der neu entwickelte „Mainzer Audiometric Test for Children“ (MATCH) ist ein Sprachtest, der auf 26 Nomen basiert, die im Wortschatz von 2-jährigen Kindern vorkommen und eine bessere Übereinstimmung mit der Alltagssprache haben als ältere Tests. Er wurde an 179 Kindern im Altersbereich von 2;11 bis 6;9 Jahren, die in drei Altersgruppen eingeteilt wurden, normiert und validiert. Der verwendete adaptive Algorithmus führte zu einer Sprachverständlichkeitsschwelle bei 71,4 % Sprachverständlichkeit. Die zugehörigen Werte für die 3 Altersgruppen sind 38,3 dB, 34,0 dB und 30,0 dB. Es bestand eine starke positive Korrelation von $r = 0,84$ ($p < 0,001$) zwischen den MATCH-Ergebnissen und dem mittleren tonaudiometrischen Hörverlust, wie auch zwischen Test und Retest ($r = 0,89$; $p < 0,001$). Test und Retest wichen im Mittel um 1,7 dB (SD = 4,6 dB) ab.

Der „Bochum Auditory Speech Discrimination Test“ (BASD) ist eine Testbatterie zur Untersuchung der zentralen Schallverarbeitung. Sie beinhaltet sowohl sprachfreie psychoakustische Verfahren als auch sprachhaltige Untertests. Der BASD wurde an 30 Erwachsenen im Alter zwischen 18;9 und 26;9 Jahren erprobt, an 293 Kindern im Alter zwischen 4;0 und 5;11 Jahren normiert und an 158 Kindern im Altersbereich von 4;0 bis 4;5 Jahren validiert. Der Test zeigte signifikante Unterschiede zwischen Kindern mit veränderter Sprachentwicklung oder auditiven Verarbeitungs- und Wahrnehmungsstörungen und Kindern der Kontrollgruppe.

Abstract

In this thesis, three new methods were developed and tested to assess the hearing of young and preschool children to improve the early diagnosis of hearing loss.

The “Multiple-choice Auditory Graphical Interactive Check” (MAGIC) is a novel self-test for children aged 3 years and older to determine their puretone thresholds. Children were able to independently control the test using a touchscreen, thus enhancing motivation to complete the task. The test was validated in 156 children with normal hearing and children with hearing loss, aged 3 yr 6 mo to 10 yr 3 mo. The mean difference between MAGIC and conventional puretone audiometry (PTA) was 0.5 dB with a standard deviation (SD) of 7.6 dB. There was a strong positive correlation of 0.79 ($p < 0.001$) between the two methods, as well as between test and retest ($r = 0.81$; $p < 0.001$). The mean difference between test and retest was 1.4 dB (SD = 8.0 dB). Children reported that the MAGIC test was more enjoyable and less strenuous than the conventional method. Thus, MAGIC is a valuable adjunct to established audiological testing procedures.

The newly developed “Mainzer Audiometric Test for Children” (MATCH) is a speech test based on 26 nouns commonly found in the vocabulary of 2-year-old children, offering a better alignment with everyday language compared to older tests. It was standardized and validated in 179 children aged 2 yr 11 mo to 6 yr 9 mo, divided into three age groups. The adaptive algorithm resulted in a speech intelligibility threshold of 71.4 % intelligibility. The corresponding values for the three age groups are 38.3 dB, 34.0 dB, and 30.0 dB. A strong positive correlation of $r = 0.84$ ($p < 0.001$) was observed between MATCH results and the average puretone hearing level, as well as between test and retest ($r = 0.89$; $p < 0.001$). The mean difference between test and retest was 1.7 dB (SD = 4.6 dB).

The “Bochum Auditory Speech Discrimination Test” (BASD) is a test battery designed to assess central auditory processing. It consists of both non-speech psychoacoustic methods and speech-based subtests. The BASD was tested on 30 adults aged 18 yr 9 mo to 26 yr 9 mo, normed on 293 children aged 4 yr to 5 yr 11 mo and validated in 158 children aged 4 yr to 4.5 yr. The test showed significant differences between children with abnormal language development or auditory processing and perception disorders and the control group.

1. Einleitung

1.1. Motivation

„Die Sprache ist die Gesittung selbst ... Das Wort, selbst das widersprechendste, ist so verbindend ... Aber die Wortlosigkeit vereinsamt“

Thomas Mann, Der Zauberberg

Im Alltag ist das miteinander Sprechen derart selbstverständlich, dass kaum jemand bewusst darüber nachdenkt wie anspruchsvoll das Erlernen und Verstehen von Sprache ist und welche Fähigkeiten dafür benötigt werden. Bei diesem Prozess kommt dem Gehör eine Schlüsselrolle zu. Im jungen Alter beeinträchtigen Hörschäden die Entwicklung und führen im weiteren Verlauf zu Schulversagen (Wiesner et al., 2019). Bei älteren Menschen führen sie zu Vereinsamung und Rückzug aus dem sozialen Leben.

In Deutschland kommen etwa 2 bis 3 von 1000 Neugeborenen mit einer behandlungsbedürftigen Hörminderung zur Welt (Neumann et al., 2006; Ptok, 2011). Untersuchungen zeigen, dass bei schwerhörigen Kindern das Lernvermögen und die Sprachfähigkeiten mit dem Alter bei Behandlungsbeginn in kausalem Zusammenhang stehen. Je früher die Versorgung erfolgt, desto besser sind die sprachlichen Fähigkeiten (Nelson et al., 2008; Neumann et al., 2006; Yoshinaga-Itano et al., 1998). Folglich hat die frühe Erkennung und Versorgung einer Schwerhörigkeit eine entscheidende Bedeutung für die Hör- und Sprachentwicklung eines Kindes (Gravel, 2001; Yoshinaga-Itano, 2001).

Es herrscht Konsens, dass kindliche Hörstörungen frühest möglich erkannt, diagnostiziert und wirkungsvoll behandelt werden müssen, um das Risiko von Störungen in der Hör-, Sprach- und Allgemeinentwicklung zu minimieren (Hoth et al., 2015). Dank des universellen Neugeborenen-Hörscreenings, das 2009 in Deutschland eingeführt wurde, ist die frühe Erkennung einer Schwerhörigkeit auch in Deutschland flächendeckend möglich geworden. Dabei werden Messungen an beiden Ohren mittels objektiver Messmethoden vorgenommen mit dem Ziel Hörverluste ab 35 dB HL zu erkennen (Bundesministerium für Gesundheit, 2008). Vor Einführung des Screenings wurden insbesondere einseitige Hörschäden besonders spät entdeckt und behandelt (Wiesner et al., 2019).

Zur Behandlung stehen, je nach Indikation, vielfältige Therapiemöglichkeiten bereit wie Hörverbessernde Operationen, Hörgeräteversorgung und Hörimplantate (Cochlea- oder Mittelohrimplantate). Diese Behandlungsschritte müssen stetig von Überprüfungs-, Rehabilitations- und Fördermaßnahmen begleitet werden, um einen optimalen Versorgungserfolg zu erzielen. (DGPP, 2013; Hoth et al., 2015)

Da bei Neugeborenen Hörprüfungen, die eine Kooperation erfordern, nicht möglich sind, kommen zunächst objektive Messmethoden zum Einsatz. Beim Neugeborenen-Hörscreening sind das in erster Linie otoakustische Emissionen (OAE) und frühe auditorisch evozierte Potenziale (FAEP). Erstere testen das Gehör bis zu den äußeren Haarzellen des Innenohrs, letztere liefern zusätzlich Information über die Funktion des Hörnervs. Mit zunehmendem Alter können auch subjektive Testverfahren wie Ton- und Sprachaudiometrie durchgeführt werden. Sofern die Versuchsperson in der Lage und willens ist zu kooperieren, lassen sich damit schnell und zuverlässig Information zur gesamten Hörleistung gewinnen.

Ziel der vorliegenden Arbeit ist die Erstellung, Normierung und Validierung neuer Testverfahren zur Erweiterung des Inventars der pädaudiologischen Diagnostik. Es gibt zwar zahlreiche Tests für Erwachsene, aber nur wenige, teils veraltete, für Kinder, insbesondere Kleinkinder. Die neuen Tests werden auf einer mobilen Plattform implementiert und verfügbar gemacht.

Es handelt sich um drei psychoakustische Testverfahren:

- eines zur Bestimmung der Reintonschwelle „Multiple-choice Auditory Graphical Interactive Check“ (MAGIC)
- einen sprachaudiometrischen Test „Mainzer Audiometric Test for Children“ (MATCH)
- eine Testbatterie zur Prüfung der zentralen Schallverarbeitung, „Bochum Auditory Speech Discrimination Test“ (BASD)

Alle verfolgen den Ansatz eines bildgestützten Selbsttests, bei dem das Kind den Test selbstständig mithilfe von Bildern auf einem Touchscreen steuern und durchführen kann. Beim Testdesign stand die Eignung für den Einsatz bei jungen Kindern im Mittelpunkt.

1.2. Überblick über die vorliegende Arbeit

Im hier anschließenden Kapitel 2 werden die Grundlagen kurz zusammengefasst, die zum Verständnis der Problemstellung und Lösungsansätze hilfreich sind. Das umfasst zum einen den anatomischen Überblick über das Hörsystem und zum anderen die Funktion der Schallverarbeitung. Darüber hinaus werden die subjektiven audiologischen Testverfahren vorgestellt sowie die Verfahren zur Kalibrierung der Messsysteme und der Signale. Abschließend werden die angewendeten Methoden und Hilfsmittel beschrieben.

Das darauf folgende Kapitel 3 enthält eine Beschreibung der verwendeten und weiterentwickelten Hardware. Ferner erläutert es die Basisfunktionen der Firmware, die als Grundlage für die Implementierung der neuen Tests dient.

Die neu entwickelten Untersuchungsmethoden werden anschließend in Kapitel 4 im Detail beschrieben. Es handelt sich hierbei um einen тонаudiometrischen Test (MAGIC), einen Sprachtest (MATCH) und einen Test (BASD) für den Einsatz im Screening bei Verdacht auf

Sprachauffälligkeiten, die einer pädagogischen Sprachförderung bedürfen, auf Sprachentwicklungsstörungen sowie auf Auditive Verarbeitungs- und Wahrnehmungsstörungen. Sämtliche Tests werden als Selbsttest per Touchscreen durchgeführt, was das Fachpersonal entlastet und die Motivation der untersuchten Kinder steigert.

Kapitel 5 ist den Ergebnissen der klinischen Studien an mehreren hundert Kindern, bei denen die hier erstellten psychoakustischen Testverfahren eingesetzt wurden, gewidmet.

Daran schließt sich die Diskussion der Ergebnisse in Kapitel 6 an, die die Resultate im Kontext des Stands der Technik und Literatur bespricht.

Die Arbeit endet mit einem Fazit und einem Ausblick in Kapitel 7. Hier werden die wichtigsten Ergebnisse zusammengefasst und die Möglichkeiten für zukünftige Verbesserung und Erweiterung skizziert.

2. Grundlagen

2.1. Überblick über das Hörsystem

2.1.1. Schall

Schall kann als Bewegung kleiner Teilchen in Medien aller Aggregatzustände betrachtet werden, die um ihre Ruhelage schwingen. Diese Schwingungen breiten sich als longitudinale Druckwellen im Raum aus. Ihre Amplitude wird als Schalldruck mit dem Formelzeichen p bezeichnet und in der Maßeinheit Pascal (Pa) dargestellt.

2.1.2. Anatomie des Ohres

Das Gehör hat die Funktion der Aufnahme, Weiterleitung, Transformation und Verarbeitung akustischer Schallsignale bis zur bewussten Wahrnehmung. Es gliedert sich in das periphere Hörorgan mit Außen-, Mittel- und Innenohr (Abb. 2.1) sowie die zentrale Hörbahn. (Abb. 2.3) (Hoth et al., 2015)

Das Außenohr besteht aus der Ohrmuschel und dem äußeren Gehörgang. Es wird vom Trommelfell abgeschlossen, das gleichzeitig die Grenze zum Mittelohr und der Paukenhöhle bildet. Die Eustachische Röhre verbindet den Nasenrachenraum mit der Paukenhöhle. Ihre Hauptfunktion ist der Druckausgleich zwischen dem Druck der Außenluft und dem Druck im Mittelohr, um eine optimale Schwingungsfähigkeit des Trommelfells sicherzustellen (Lenarz & Boeninghaus, 2012). Im Mittelohr verbindet die Gehörknöchelchenkette, bestehend aus Hammer, Amboss und Steigbügel, das Trommelfell mit dem ovalen Fenster der Cochlea

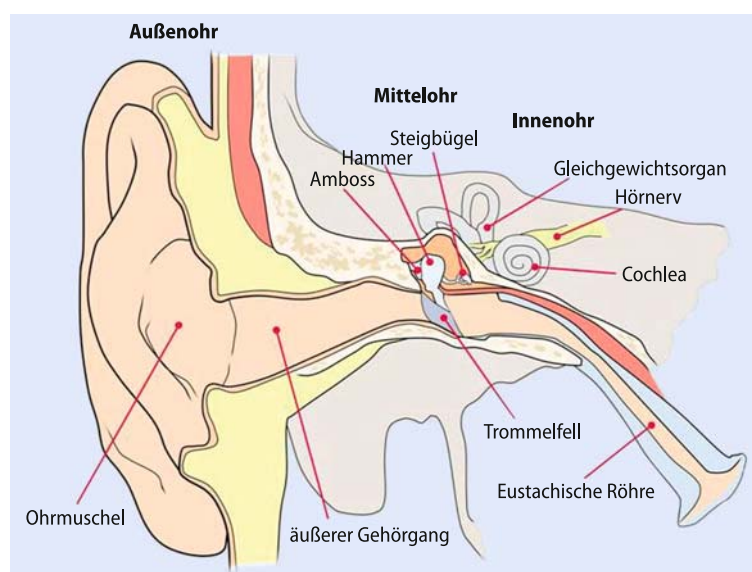


Abb. 2.1 Übersicht über das periphere Hörorgan mit Außen-, Mittel- und Innenohr (Hoth et al., 2015, S. 7, Abb. 2.1)

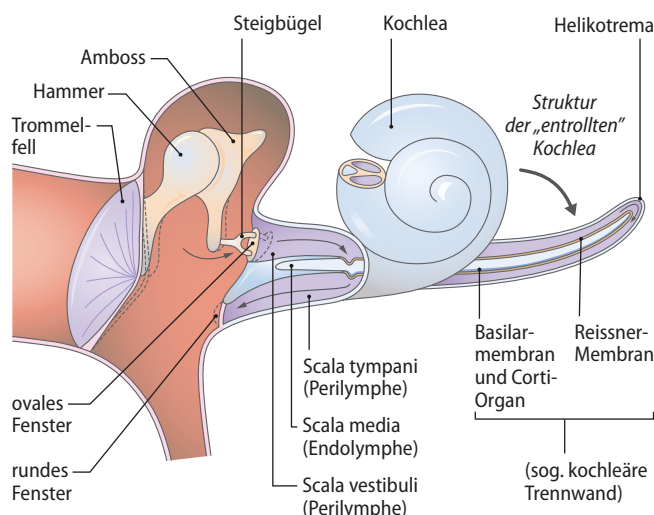


Abb. 2.2 Schema von Mittelohr und Cochlea. Die Cochlea ist auch ausgerollt dargestellt, um die Skalen besser zu visualisieren (Moser & Zenner, 2019a, S. 688, Abb. 52.4)

(Schnecke) im Innenohr. Diese Art der Ankoppelung bewirkt eine Impedanzanpassung, so dass 60 % der Schallenergie in das Innenohr eintreten können. (Reiß et al., 2021)

Im Innenohr liegen das Gleichgewichtsorgan und die Cochlea. Diese beinhaltet 3 übereinanderliegende flüssigkeitsgefüllte Gangsysteme. Am ovalen Fenster beginnt die Scala vestibuli, in der Mitte befindet sich die Scala media und mit dem runden Fenster bildet die Scala tympani einen Abschluss zur Paukenhöhle (Abb. 2.2). An der Schneckenspitze gehen die mit Perilymphe gefüllten Scala vestibuli und Scala tympani ineinander über. Nach unten hin ist die mit Endolymphe gefüllte Scala media von der Scala tympani durch die Basilarmembran getrennt. Auf dieser, in der Scala media, befindet sich das Corti'sche Organ mit den äußeren und inneren Haarzellen, das von der Tektorialmembran abgedeckt ist (Hoth et al., 2015). Mit ihrer Hilfe erfolgt die Umwandlung von mechanischer Schallenergie in elektrische Impulse, die vom Hörnerv weitergeleitet werden. Hier beginnt die zentrale Hörbahn, die beidseitig angelegt ist und gekreuzt und ungekreuzt verläuft (Abb. 2.3). Sie besteht aus hierarchisch geordneten Stationen zwischen Hörnerv und primärem auditorischen Kortex, die afferent (aufsteigend, in Richtung Hörrinde) und efferent (absteigend, in Richtung Cochlea) durch Synapsen miteinander verbunden sind (Moser & Zenner, 2019b; Reiß et al., 2021). Die absteigende Hörbahn ist im Wesentlichen parallel zur aufsteigenden, allerdings reziprok vom Kortex zu den Haar-sinneszellen. Ihre Funktion ist die Modulation niederer durch höhere Hörzentren. (Reiß et al., 2021)

Die aufsteigende Hörbahn besteht aus zwei Systemen: dem tonotopen und dem „polysensorischen“ System. Im tonotopen System gibt es eine Zunahme der Nervenfasern durch zahlreiche Aufzweigungen und häufig beidseitige Verbindungen durch zusätzliche Kreuzung der Fasern, besonders im oberen Olivenkomplex, Colliculus inferior und auditorischen Kortex. Es besteht aus 3 Hauptrelaisstationen zwischen Hörnerv und primärem auditorischen Kortex. Die erste Umschaltung der aufsteigenden Hörbahn erfolgt im Hirnstamm vom Nucleus coch-

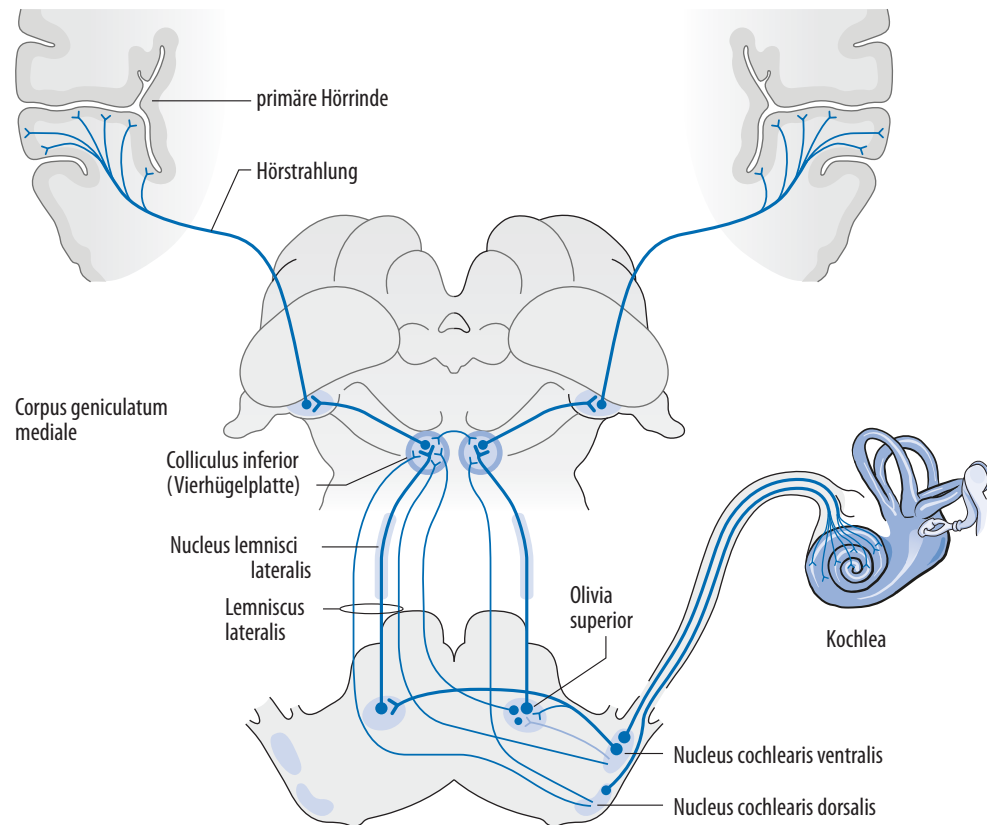


Abb. 2.3 Schematische Darstellung der Hörbahn (Reiß et al., 2021, S. 258, Abb. 19.5)

learis auf den oberen Olivenkomplex der ipsilateralen und kontralateralen Seite sowie Lemniscus lateralis der Gegenseite. Die zweite Umschaltung findet im Colliculus inferior, dem größten auditorischen Kern, statt. Im Corpus geniculatum mediale vollzieht sich die dritte Umschaltung über die Hörstrahlung zur primären Hörrinde. (Reiß et al., 2021)

Das „polysensorische“ System projiziert über den dorsalen Anteil des Corpus geniculatum mediale (z. B. zum sekundären auditorischen Kortex) unter Umgehung des primären auditorischen Kortex und verfügt über eine kürzere Verbindung subkortikal (unbewusst) zur Amygdala. Dies ist besonders für die emotionale Komponente bedeutsam. Das „polysensorische“ System empfängt auch Informationen anderer sensorischer Systeme, wie dem visuellen und somatosensorischen System. (Reiß et al., 2021)

2.1.3. Schallverarbeitung

Die Schallverarbeitung beginnt beim Auftreffen des Schallsignals an der Ohrmuschel. Ihre Form beeinflusst den Frequenzgang, was beim Richtungshören bedeutsam ist. Das Außenohr erzeugt als Trichter zusätzlich eine Verstärkung, indem über die Ohrmuschel Schall gesammelt und durch den Gehörgang zum Trommelfell geleitet wird. Das Mittelohr bewirkt durch das Flächenverhältnis von Trommelfell zu ovalem Fenster und dem Hebelverhältnis der Gehörknöchelchen eine Verstärkung um den Faktor 18 (Moser & Zenner, 2019a) bis 22 (Reiß et al., 2021) sowie eine Impedanzanpassung. Ohne das Mittelohr kämen nur 2% der Schall-

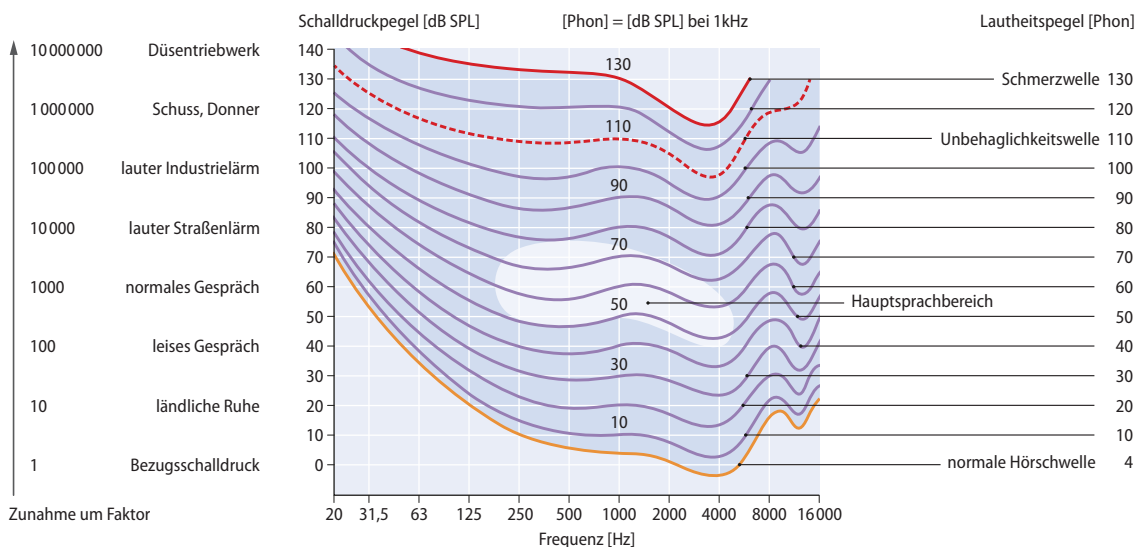


Abb. 2.4 Isophone, Hörfläche und Hauptsprachbereich (hell). Isophone sind Kurven gleicher Lautstärkepegel in Phon. (Moser & Zenner, 2019a, S. 687, Abb. 52.3)

energie im Innenohr an. Vom ovalen Fenster aus breitet sich das Schallsignal als Wanderwelle aus, mit einem Amplitudenmaximum an dem zur Reizfrequenz korrespondierenden Ort auf der Basilarmembran. Bei dieser Frequenz-Orts-Transformation werden die hohen Frequenzen an der Basis und die tiefen Frequenzen an der Spitze der Cochlea abgebildet. Dieses Konzept der tonotopen Organisation setzt sich bei den Nervenfasern der aufsteigenden Hörbahn und im auditorischen Kortex fort. (Moser & Zenner, 2019a)

Die äußeren Haarzellen des Hörorgans haben zwei wichtige Funktionen: Sie erhöhen die Frequenzselektivität und fungieren mittels aktiver schallinduzierter Längenausdehnung als nichtlineare cochleäre Verstärker von leisen Schallsignalen um ca. 50 dB (Hoth & Steffens, 2016). Ohne sie wäre der Dynamikbereich des Hörens (s. Abb. 2.4) deutlich geringer (Reiß et al., 2021). Die in Abb. 2.4 gezeigte Hörfläche des Menschen hat einen Frequenzumfang von etwa 10 Oktaven (ca. 20 Hz bis 20 kHz) und einen Dynamikumfang von 7 Dekaden (20 μ Pa bis 200 Pa, Faktor 1 bis 10 000 000). Da die Zahlenwerte für den Schalldruck unübersichtlich und unhandlich werden, wird er üblicherweise als logarithmisch skaliertes Schalldruckpegel L angegeben.

$$L = 20 \cdot \log_{10} \frac{p}{p_0} \text{ dB SPL}$$

Aus dem Schalldruck p wird der Schalldruckpegel L (sound pressure level, SPL) in dB relativ zu einem Bezugsschalldruck $p_0 = 20 \mu Pa$ nach obiger Formel berechnet. In Abb. 2.4 sind die Isophone (Kurven gleicher empfundener Lautheit) dargestellt. Sie verlaufen im Wesentlichen parallel zur orange eingezeichneten Hörschwelle. Die Empfindlichkeit des Hörsystems nimmt zu hohen und tiefen Tönen hin ab. Ihr Maximum liegt bei ungefähr 4 kHz, bedingt durch die Resonanz des Gehörgangs. Der Hauptsprachbereich umfasst die Frequenzen und Pegel der menschlichen Sprache. Die Schmerzgrenze ist bei 130 Phon erreicht. Bei längerer

Lärmeinwirkung reichen auch schon kleinere Schalldruckpegel aus, um die Hörsinneszellen irreversibel zu zerstören.

Via Innenohr gelangt die akustische Information zur Weiterverarbeitung in die zentrale Hörbahn. Durch die besondere Verschaltung in den einzelnen Stufen der zentralen Verarbeitung wird die Fähigkeit einer differenzierten Frequenzunterscheidung erreicht. Denn für die Erkennung von Sprache muss das Hörsystem in der Lage sein, die Zeitstruktur und die Intensitätsunterschiede zu analysieren. Grundlagen des Richtungshörens sind Pegel-, Frequenz- und Zeitdifferenzen, die infolge des Schallschattens des Kopfes zustande kommen. (Lenarz & Boenninghaus, 2012)

Die zentrale Hörbahn ist von besonderer Bedeutung für das binaurale Hören, das Informationen und Signale beider Ohren und deren Integration und Interpretation beinhaltet. Für das räumliche Hören oder Sprachverstehen im Störschall sind binaurale Hörleistungen wesentliche Voraussetzungen. (Hoth et al., 2015)

2.2. Audiologische Testverfahren

Audiologische Testverfahren dienen der Untersuchung der Funktionsfähigkeit des Gehörs. Es wird grundsätzlich zwischen subjektiven und objektiven Testverfahren unterschieden. Bei subjektiven Verfahren ist die aktive Mitarbeit der Versuchsperson zwingend erforderlich, wohingegen es bei objektiven Verfahren genügt, wenn die Versuchsperson entspannt und ruhig sitzt oder liegt. Im Rahmen dieser Arbeit wurden ausschließlich subjektive Messmethoden eingesetzt, die sich für die Anwendung bei Kindern eignen.

Vor einer Untersuchung, bei der eine subjektive Testmethode zum Einsatz kommt, muss die Testperson mit der Aufgabe vertraut gemacht werden. Dabei sind die Fähigkeit und der Wille zur Kooperation sowie die Konzentration entscheidend für die Zuverlässigkeit der Ergebnisse. Von großer Bedeutung ist auch die präzise Platzierung der Schallwandler. Variationen der Kopfhörerposition relativ zur Ohrmuschel nach oben/unten oder vorwärts/rückwärts und Undichtigkeiten zwischen Ohrmuschel und Kopfhörer können insbesondere bei höheren Frequenzen zu fehlerhaft erhöhten Hörschwellen führen. Bei der Messung von Ruhehörschwellen oder Sprachverständlichkeitsschwellen in Ruhe können Umgebungsgeräusche das Ergebnis ebenfalls verfälschen.

Schwerhörigkeit liegt vor, wenn das Hörvermögen auf einer oder mehreren Stufen der Schallverarbeitung (Abschnitt 2.1.3) schlechter im Vergleich zu einem Normkollektiv ist. Das Normkollektiv können beispielsweise junge Erwachsene im Alter zwischen 18 und 25 Jahren sein, die keine Vorerkrankungen oder sonstige Schädigungen des Hörorgans aufweisen. Eine Auflistung von Anforderungen an otologisch normale Testpersonen für Hörversuche ist in ISO 389-9 (2009) festgelegt.

2.2.1. Tonaudiometrie

Tonaudiometrische Messungen gehören zu den ältesten und am weitesten verbreiteten Methoden zur Überprüfung der Hörfunktion. Wie der Name schon nahelegt, werden hierbei Töne, im Allgemeinen Sinustöne oder frequenz- und/oder amplitudenmodulierte Sinustöne verwendet, üblicherweise im Bereich zwischen 125 Hz und 8 kHz, oder bis 16 kHz bei der Hochfrequenz-Audiometrie. Die Schalldarbietung kann über Einsteckhörer, Kopfhörer, Knochenhörer oder im Freifeld erfolgen.

Die Geräte (DIN 60645-1, 2018) und der Messablauf (ISO 8253-1, 2011) sind genormt. Das führt bei Durchführung von erfahrenem Personal mit kooperativen Testpersonen zu zuverlässigen Messergebnissen. Daher wird die Tonaudiometrie beim Vergleich mit anderen Messmethoden oft als Gold-Standard angesehen. Reinton-Hörschwellen bilden die Basis in der klinischen Diagnostik und werden als Grundlage für eine Hörgeräte-Anpassung verwendet. Durch den Vergleich von Knochenleitungsschwellen und Luftleitungsschwellen kann der Einfluss des Mittelohres auf eine Hörstörung ermittelt werden.

Bei jungen Kindern hängt das Ergebnis einer Hörprüfung nicht nur von der tatsächlichen Funktionsfähigkeit des Hörorgans ab, sondern auch von ihrem Entwicklungsstand sowie von der Beobachtungsgabe und Erfahrung des untersuchenden Personals. Die Auswahl der Untersuchungsmethode sollte sich am Entwicklungsalter des Kindes orientieren, das meist nicht mit dem Lebensalter korreliert. Mangels Verfügbarkeit dieser Information wird in der Praxis hilfsweise das Lebensalter als erster Anhaltspunkt bei der Wahl verwendet. Bis zu einem Alter von ungefähr 2 Jahren finden die Messungen im freien Schallfeld statt. Das bedeutet, dass keine seitengesonderte Überprüfung des Hörvermögens stattfindet, sondern das bessere Ohr das Ergebnis dominiert. Innerhalb der ersten Lebensmonate umfasst das Inventar der Messmethoden im Wesentlichen die Reflexaudiometrie, die Verhaltensaudiometrie bzw. Ablenkaudiometrie und die Visual Reinforcement Audiometry (VRA) bzw. Konditionierte Reaktionsaudiometrie. (Hoffmann & Schäfer, 2020)

Bei der Reflexaudiometrie wird das Baby aufmerksam beobachtet, ob es unwillkürliche Reflexe zeigt, die im Zusammenhang mit einer Schalldarbietung stehen. Im Alter zwischen ca. 3 und 6 Monaten wird die Methode der Verhaltensaudiometrie angewendet, die darauf abzielt, reproduzierbare Verhaltensänderungen (wie z.B. Kopfdrehen), auf Schallreize zu erzeugen (Hoffmann & Schäfer, 2020). Ähnlich hierzu ist die Konditionierte Reaktionsaudiometrie, bei der eine Konditionierungsphase vor der eigentlichen Messphase stattfindet. Dabei werden Schall, und ein visueller Reiz parallel gegeben um eine definierte Reaktion, beispielsweise das Drehen des Kopfes zur Schallquelle, zu erreichen (Mrowinski et al., 2017).

Ab dem 2. Lebensjahr ist die tonaudiometrische Untersuchung als Spielaudiometrie durchführbar. Ab diesem Alter tolerieren Kinder üblicherweise auch Kopf- oder Einsteckhörer, was eine seitengesonderte Ermittlung der Hörschwelle möglich macht. Die Spielaudiometrie startet ebenso wie die Reaktionsaudiometrie mit einer Konditionierungsphase, in der dem Kind

die Spielhandlung und Aufgabe vermittelt wird, die oftmals das Stecken von Klötzchen in eine Schablone ist (Hoffmann & Schäfer, 2020). Allerdings ist die Messung, bei der das Kind ununterbrochen aufmerksam sein muss, anstrengend und die Spielhandlung wenig abwechslungsreich. Das kann nachteilige Effekte auf die Genauigkeit haben.

Wünschenswert ist daher ein neuer Test mit einer abwechslungsreicheren Spielhandlung und bei dem kurzzeitige Unterbrechungen der Aufmerksamkeit das Ergebnis nicht besonders beeinträchtigen.

2.2.2. Sprachaudiometrie

Sprache ist mit das wichtigste Kommunikationsmittel im zwischenmenschlichen Austausch. Sprachaudiometrische Verfahren sind aus diesem Grund ein wichtiges Hilfsmittel zur Untersuchung von durch Hörstörungen verursachten Einschränkungen im alltäglichen Leben. In der Diagnostik wird, sobald vom Entwicklungsalter her möglich, die Sprachaudiometrie eingesetzt. Sprachtests untersuchen das Diskriminationsverhalten für komplexe Signale mit semantischer Information (Billich, 1981).

Bei den Sprachtests werden die Wörter entweder vom Kind nachgesprochen, oder das Kind zeigt auf die dem Wort oder Satz zugeordnete Bildkarte. Bei PC-gesteuerten Audiometern werden die Bilder auf dem Bildschirm gezeigt. Bildkarten ermöglichen die Durchführung der Hörtests unabhängig von der Artikulationsfähigkeit des Kindes.

Beim Einsatz sprachaudiometrischer Tests bei Kleinkindern und sehr jungen Kindern sind einige Besonderheiten zu beachten, die sich auf das Testdesign auswirken. An erster Stelle ist das limitierte Vokabular der Zielgruppe zu nennen. Auch sind offene Tests, bei denen die Aufgabe in der korrekten Wiederholung der gehörten Wörter besteht, zu komplex und sie sind darüber hinaus von der Sprachproduktion des Kindes und dem Hörvermögen der Testleitung abhängig. Zudem wird auch aus der klinischen Praxis berichtet, dass Kinder irritiert sein können, wenn für sie nicht sichtbar ist, wer spricht. Grundsätzlich hat sich eine weibliche Sprecherstimme als vorteilhaft erwiesen.

Gegenwärtig werden bei Kindern in Deutschland folgende Sprachtests am häufigsten eingesetzt: Der Göttinger Kindersprachverständnistest (Chilla et al., 1976), der Heidelberger Konsonant-Vokal-Konsonant Test (CVC) (Billich, 1981), der Oldenburger Kinder-Reimtest (OLKI) (Brand et al., 1999; Kliem & Kollmeier, 1995), der Oldenburger Kinder-Satztest (Neumann et al., 2012; Wagener & Kollmeier, 2005; Weißgerber et al., 2012), der AAST-Sprachtest (Coxinix, 2005, 2006), der Würzburger Kindersprachtest (Baumann, 2006; Völter et al., 2005), und der Mainzer Kindersprachtest (Biesalski et al., 1974). Lediglich die Teile I der beiden letztgenannten Tests sind gemäß den Testanleitungen für Kinder ab 2 Jahren geeignet.

Der Mainzer Kindersprachtest besteht aus drei Tests. Das Wortmaterial für Kinder unter 4 Jahren enthält 10 Wörter, die weiter unten aufgelistet werden. Das Wortmaterial für Kinder

zwischen 4 und 5 Jahren besteht aus 25 Wörtern, das Wortmaterial für Kinder zwischen 6 und 8 Jahren beinhaltet 50 verschiedene Wörter. Der Göttinger Kindersprachverständnistest besteht aus zwei Wörterlisten (20 Wörter für Kinder zwischen 3 und 4 Jahren, 100 Wörter für Kinder zwischen 5 und 6 Jahren). Der Heidelberger Konsonant-Vokal-Konsonant Test verwendet Wörter, die nach phonologischen Kriterien zusammengestellt wurden. Dabei werden Lautbildungsort (labial, alveolar, postalveolar und palatal) und Lautbildungsart der Konsonanten (plosiv, nasal, lateral und frikativ) variiert. Das Wortmaterial des Oldenburger Kinder-Reimtest besteht aus Dreiergruppen von Reimwörtern mit insgesamt 78 Begriffen. Hierbei hat das Kind die Aufgabe, ein Bild aus einem Set von drei Bildern auszuwählen, das dem gehörten Begriff entspricht. Die Bilder sind farbig und von Kindern gemalt. Der Test ist für Kinder der 2. bis 6. Schulklasse geeignet. Der Würzburger Kindersprachtest besteht aus Listen mit einsilbigen und zweisilbigen Wörtern für ein- bis zweijährige Kinder sowie für Vorschul- und Schulkinder. Der von Coninx entwickelte AAST-Sprachtest (Coninx, 2005, 2006) besteht aus 6 Wörtern. Dieser ist jedoch erst für Kinder ab einem Alter von 3–4 Jahren geeignet. Darüber hinaus wurde die Sprachverständlichkeitsschwelle an Erwachsenen auf HL („hearing level“) normiert.

Der Mainzer Kindersprachtest I, „Mainzer I“, (Biesalski et al., 1974) wird in der Praxis am häufigsten für Kinder ab 2 Jahren verwendet. Er besteht insgesamt aus 10 Wörtern und enthält 5 Einsilber (Ball, Bär, Bahn, Uhr, Ei), 3 Zweisilber (Mama, Puppe, Auto) und 2 Lautmalereien (Wauwau, Miau) (DIN 45621-3, 1985). Zur Unterstützung der Durchführung finden Bildtafeln Verwendung. Die wesentlichen Kritikpunkte am „Mainzer I“ sind das sehr kleine und veraltete Vokabular, welches den heutigen Wortschatz von 2-Jährigen nicht mehr genügend abbildet. Hinzu kommt, dass das Bildmaterial veraltet ist und gerade sehr junge Kinder, bei Zuhilfenahme der Bildtafeln, die Begriffe „Ball“ und „Ei“ häufig verwechseln. Schwerhörnde Kinder müssen sowohl während der Diagnostik als auch bei der anschließenden Hörgeräteversorgung sehr häufig Sprachtests durchlaufen. Das kann bei dem alten „Mainzer I“ zu Lern- bzw. Deckelungseffekten führen. Vergleichbare Kritikpunkte gelten auch für das Wortmaterial und die Bildvorlagen des „Würzburger Kindersprachtests“.

Wünschenswert ist daher ein neuer Test, der durch ein größeres und bildlich besser dargestelltes Vokabular beliebig oft wiederholbar ist.

2.3. Psychoakustik

In der Psychophysik werden Zusammenhänge zwischen einem objektiv messbaren Reiz und der dadurch ausgelösten Empfindung untersucht. Als Teilgebiet davon befasst sich die Psychoakustik mit Schallreizen im hörbaren Bereich (s. Abb. 2.4). Während die physikalischen Größen eines Reizes exakt mit Formeln und Einheiten beschrieben werden können, lassen sich Empfindungen oft nur in Adjektiven ausdrücken. Diese Beschreibungen quantitativ zu erfassen ist eine der schwierigsten Aufgaben auf diesem Gebiet, da spezielle Messmethoden

erforderlich sind. Die Messergebnisse der Empfindungen hängen vom gesamten Hörsystem ab, beginnend am Außenohr und in der Großhirnrinde endend. Die psychoakustischen Testverfahren werden sowohl für rein wissenschaftliche Zwecke, als auch für Hörprüfungen bei den subjektiven Hörtests eingesetzt. (Hoth & Steffens, 2016; Zwicker, 1982)

2.3.1. Methoden und Verfahren der Psychoakustik

Im Folgenden werden Methoden beschrieben, die in der Psychoakustik häufig Verwendung finden. Die Methoden unterscheiden sich darin, für welche Aufgaben sie gedacht sind und in welcher Dauer sie zu einem Ergebnis konvergieren. Direkte Verfahren wie die Anpassungsmethode, die Methode des Trackings und die Größenschätzung sind sehr zeitsparend. Ihre Ergebnisse sind aber abhängig von der Kooperation der Versuchspersonen. Ja-Nein-Verfahren, Intervall-Zwangswahlverfahren und adaptive Verfahren sind robuster aber auch zeitaufwändiger. (Zwicker & Fastl, 1999)

Bei der **Anpassungsmethode** wird der Versuchsperson die vollständige Kontrolle über den Stimulus übertragen, mit dem Ziel, beispielsweise die Lautstärke eines Sinustons in der Art einzustellen, dass der Ton gerade eben wahrnehmbar ist (Hörschwelle), oder die Tonhöhe so zu verändern, bis sie derjenigen eines Referenztons entspricht. (Zwicker & Fastl, 1999)

Auch bei der **Methode des Trackings** liegt die Kontrolle über den Stimulus bei der Versuchsperson. Allerdings liegt nur die Änderungsrichtung des Reizes in ihrer Hand, sodass, am Beispiel der Hörschwellenmessung, die Fälle „Ton hörbar“ und „Ton nicht hörbar“ aufeinanderfolgen. Das Ergebnis ist der Mittelwert der Umkehrpunkte. (Zwicker & Fastl, 1999)

Bei der **Größenschätzung** besteht die Aufgabe darin, einer wahrgenommenen Größe des Stimulus eine Zahl zuzuordnen. Hierbei ist es hilfreich, eine Referenz als Anker (z.B. 100) vorzugeben. So kann der Lautheit eines Schalls ein Zahlenwert 200 zugeordnet werden, wenn er als doppelt so laut empfunden wird. (Zwicker & Fastl, 1999)

Beim **Ja-Nein Verfahren** handelt es sich um ein „Ein Intervall Zwei Alternativen Zwangswahl Verfahren“. Die Aufgabe besteht darin anzugeben, ob ein Signal vorhanden ist oder nicht („Ja“ oder „Nein“). (Zwicker & Fastl, 1999)

Zwei-Intervall-Zwangswahlverfahren: Hierbei werden zwei Intervalle dargeboten und die Aufgabenstellung lautet, anzugeben in welchem der beiden Intervalle das Signal auftritt. Es können auch mehr als zwei Intervalle verwendet und gefragt werden, welches sich von den anderen unterscheidet. Die Versuchsperson erhält nach jedem Versuch üblicherweise ein Feedback, ob die Angabe korrekt war oder nicht. (Zwicker & Fastl, 1999)

Bei **Adaptiven Verfahren**, auch als „Up-Down“-Verfahren bezeichnet, hängt der präsentierte Stimulus eines Versuchs von der Antwort aus den vorangehenden Versuchen ab. Gewöhnlich wird die gesuchte Messgröße (z.B. Schalldruckpegel, Stimulusdauer oder Frequenzun-

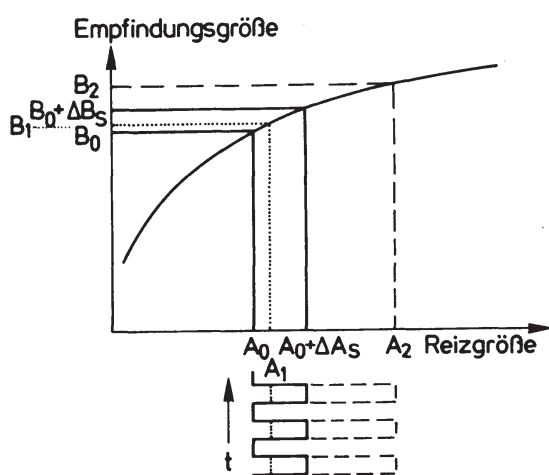
terschied) bei korrekten Antworten verringert bis falsche Antworten kommen und anschließend wieder erhöht. Beim Erreichen von Umkehrpunkten wird die Schrittweite reduziert. Die Schwelle wird, wie bei der Methode des Trackings, aus den Umkehrpunkten berechnet. (Zwicker & Fastl, 1999)

Der **Vergleich von Stimuluspaaren** kommt zur Anwendung, wenn mehr als eine Reizdimension gleichzeitig untersucht wird. So könnte sich das Paar A-B in der Lautstärke und das Paar C-D in der Tonhöhe unterscheiden mit der Aufgabenstellung, ob der wahrgenommene Unterschied des Paares A-B größer ist als derjenige des Paares C-D. (Zwicker & Fastl, 1999)

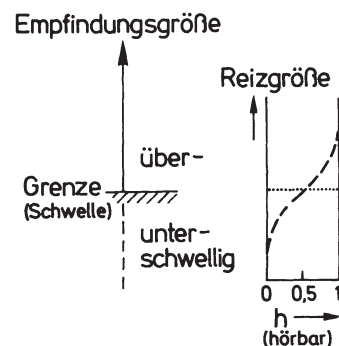
2.3.2. Unterschiedsschwellen

Die Empfindungsfunktion (Abb. 2.5a) zeigt die Empfindungsgröße in Abhängigkeit von der Reizgröße. Ihre Darstellung ist meist eine kontinuierliche Kurve, obwohl die wahrgenommene Empfindung dem Reiz nicht streng monoton folgt. Eine unterschwellige Reizerhöhung führt zu einer unveränderten Wahrnehmung. Die Änderung von A_0 nach A_1 in Abb. 2.5a führt zur Empfindung B_1 . Diese ist jedoch so nah bei B_0 , dass sie nicht wahrgenommen wird. Erst wenn die Reizänderung größer als ΔA_S ist, z.B. A_2 , ändert sich die Wahrnehmung. Die Schwelle (ΔA_S) ist dabei derjenige Wert, der in der Hälfte der Fälle, d.h. mit 50 % Wahrscheinlichkeit, zu einer Wahrnehmung führt (Abb. 2.5b). (Zwicker, 1982)

Bei der Betrachtung von eben merklichen Unterschieden (just noticeable differences, JND) sind 2 Arten von Unterschieden zu beachten: Applizieren einer kontinuierlichen Änderung als Modulation des Signals (Modulationsschwellen) und sequentielle Darbietung von akustischen Reizen mit einer Pause dazwischen (Unterschiedsschwellen). Audiologischen Testverfahren liegt das Konzept der Unterschiedsschwellen zugrunde. (Hoth & Steffens, 2016)



(a) Empfindungsfunktion (S. 4, Abb. 1.1)



(b) Bestimmung des Grenzwertes (S. 8, Abb. 1.2)

Abb. 2.5 (a) Empfindungsfunktion: Zusammenhang zwischen Reizgröße A und Empfindungsgröße B. Die gerade wahrnehmbare Reizänderung ΔA_S , die Reizstufe, führt zur Empfindungsstufe ΔB_S (b) Bestimmung des Grenzwertes: Reizgröße, bei welcher in 50% der Fälle die gefragte Empfindungsgröße hörbar ist (Zwicker, 1982)

2.4. Kalibrierung

Es ist eine wichtige Voraussetzung für eine zuverlässige Diagnostik und für den Vergleich von Studien verschiedener Arbeitsgruppen, dass ein Messergebnis – wie z.B. die Ruhehörschwelle für Sinustöne – nur vom Hörvermögen der Versuchsperson abhängt. Unabhängig von der verwendeten Messapparatur sollen bei vergleichbaren Messbedingungen im Wesentlichen die gleichen Ergebnisse erzielt werden.

Zu diesem Zweck kooperieren nationale Normungsinstitute mit dem Ziel international gültige Normen auszuarbeiten, die entscheidende Gerätespezifikationen festlegen. Für Reinton- und Sprachaudiometer ist dies die DIN 60645-1 (2018), in der unter anderem Anforderungen an die Frequenz- und Pegelgenauigkeit sowie den Frequenzgang aufgeführt sind. Grundlegende Messverfahren werden in ISO 8253-1 (2011) für die Tonaudiometrie und ISO 8253-3 (2022) für Sprachaudiometrie genannt. Letztgenannte beinhaltet darüber hinaus Vorgaben zur Erstellung, Normierung und Validierung von sprachaudiometrischen Tests.

Zur Messung der Ausgabepegel von Schallwandlern kommen ebenfalls genormte Messmittel und Kuppler zum Einsatz, wie z.B. das künstliche Ohr (DIN 60318-1, 2010) für Messungen an Kopfhörern. Da jeder Wandler einen eigenen Frequenzgang hat, wurde als Referenz bei tonaudiometrischen Messungen die Hörschwelle hörgesunder Erwachsener zugrunde gelegt. ISO 389-9 (2009) legt Kriterien für Studien zur Ermittlung von Bezugswerten fest, unter anderem das Alter der Teilnehmenden und ihre Mindestanzahl, sowie Kriterien für otologisch normale Versuchspersonen.

Das Ergebnis ist die Bezugshörschwelle, welche bei Audiometern als 0 dB HL auf der Pegelskala gekennzeichnet ist. Für verschiedene Kopfhörer sind diese Werte in der Norm ISO 389-1 (2018) festgehalten. An der Hörschwelle ist der am Kuppler gemessene Schalldruckpegel eines Kopfhörers frequenzabhängig und meistens verschieden von 0 dB SPL.

Die unterschiedlichen Frequenzgänge der verschiedenen Schallwandlertypen beeinflussen die Sprachverständlichkeit. Zur Sicherstellung von reproduzierbaren und vergleichbaren Ergebnissen bei Sprachverständlichkeitstests wurde das Konzept der Freifeldentzerrung entwickelt (Brinkmann & Richter, 1989; Brinkmann & Richter, 1997). Dabei fungiert die Wahrnehmung im freien Schallfeld über Lautsprecher als Referenz. In psychoakustischen Vergleichsmessungen wird mittels Schmalbandrauschen ermittelt, welcher Schalldruckpegel, z.B. eines Kopfhörers bei einer bestimmten Frequenz, die gleiche Lautheitswahrnehmung erzeugt wie die Freifeldlautsprecher (Brinkmann & Richter, 1997).

Diese Korrekturen werden als Equalizer zwischen Abspielgerät und Kopfhörer eingefügt, um per Kopfhörer die gleiche Wahrnehmung zu generieren (wie im freien Schallfeld) und, um damit auch dieselbe Sprachverständlichkeit zu erzielen (Fastl & Zwicker, 1983).

2.5. Methoden

Die statistische Auswertung der Studienergebnisse erfolgte mittels der Software MATLAB R2012a (Fa. The MathWorks Inc., Natick/MA, USA). Deskriptiv wurden die Messergebnisse mittels Streudiagrammen, linearer Regression, Histogrammen und Boxplots visualisiert. In einem Boxplot sind Minimum und Maximum (unterer und oberer Whisker), Ausreißer (s.u.), Median (rote Linie), 95 % Konfidenzintervall des Medians (Einkerbung) und unteres und oberes Quartil (untere und oberer Grenze der Box) gezeigt. Die Länge der Whisker beträgt maximal 1,5 Interquartilsabstände, was ungefähr $2,7 \sigma$ (Standardabweichung) bzw. 99,3 % der Werte bei normalverteilten Daten entspricht. Ausreißer sind definiert als Ergebnisse, die weiter als zwei Interquartilsabstände von Median entfernt liegen. Mediane unterscheiden sich statistisch signifikant (5 %-Niveau), wenn die 95 %-Konfidenzintervalle überlappungsfrei sind.

Zur Beschreibung von Gruppenergebnissen werden Median, Interquartilsabstand, arithmetisches Mittel und Standardabweichung herangezogen. Der Korrelationskoeffizient nach Pearson (Steland, 2010) wurde zur Bestimmung der Test-Retest-Reliabilität verwendet. Die Rangkorrelation nach Spearman wird in den Fällen berechnet, in denen die Voraussetzungen für den Korrelationskoeffizient nach Pearson nicht erfüllt sind. Zur Prüfung, ob sich die Mittelwerte zweier unabhängiger Stichproben unterscheiden wurde der t-Test (Steland, 2010) eingesetzt.

Für die Ermittlung der Standardabweichung der Diskriminationsfunktionen (s. Tabelle 5.5) kam das Bootstrapping-Verfahren zum Einsatz. Dieses Verfahren basiert auf der Annahme, dass eine repräsentative Stichprobe vorliegt. Aus dieser werden neue Stichproben durch Ziehen mit Zurücklegen erzeugt, die zur Berechnung von Mittelwert und Standardabweichung dienen (Efron, 1979).

Zur Analyse der Phonemstatistik wurde das Munich Automatic Segmentation (MAUS) System (Schiel, 1999) verwendet, das Schallsignale automatisiert verarbeitet und als Ergebnis eine in SAMPA (Abb. 2.6) transkribierte Textdatei zurückliefert. Das gleiche Werkzeug kam auch bei der Ermittlung der Phonemverteilung deutscher Spontansprache zum Einsatz (Schiel, 2010), mit welcher das neu erstellte Sprachmaterial verglichen wird. MAUS steht als Webservice (Kisler et al., 2017) im Rahmen der BAS Web Services (Institut für Phonetik und Sprachverarbeitung, 2023) zur Verfügung. In der vorliegenden Arbeit kam *WebMAUS Basic* in der Version 3.13 zum Einsatz.

Beispielwort	IPA	SAMPA	Beispielwort	IPA	SAMPA bzw. Erweiterung
Ried	i:	i:	bei	b	b
Politik	i	i	Peter	p	p
Ritt	ɪ	ɪ	du	d	d
Beet	e:	e:	Tag	t	t
Metall	e	e	Gang	g	g
Bett	ɛ	ɛ	Kamm	k	k
Käse	ɛ:	ɛ:	ich	ʔ	ʔ*
Kahn	a:	a:	Mann	m	m
kann	a	a	Nah	n	n
bog	o:	o:	eng	ŋ	n
Politik	o	o	Vase	v	v
Bock	ɔ	ɔ	fast	f	f
Mus	u:	u:	sah	z	z
Musik	u	u	fast	s	s
muß	ʊ	ʊ	Tasche	ʃ	ʃ
Hüte	y:	y:	Loge	ʒ	ʒ
Kymogramm	ʏ	ʏ	dich	ç	ç*
Hütte	ʏ	ʏ	Dach	x	x*
Höhle	ø:	2:	Hast	χ	χ*
Ökonom	ø	2	ja	j	j
Hölle	œ	9	Last	l	l
Lesen	ə	ɐ	Rauch	r	r
Leser	ɐ	6		R	R*
frei	aɪ	aɪ		ʀ	ʀR*
Frau	aʊ	aʊ		ɹ	ɹr*
neu	ɔɪ	oɪ		ɻ	ɻr*

Abb. 2.6 Übersicht SAMPA-Zeichen. Quelle: (Machelett, 1996) Umsetzung von IPA (Internationales Phonetisches Alphabet, wird im Rahmen dieser Arbeit nicht weiter referenziert) nach SAMPA (Speech Assessment Methods Phonetic Alphabet), welches in Abb. 4.8 verwendet wird. Abweichungen von bzw. Ergänzungen zu der SAMPA-Notation wurden von Machelett (1994) vorgenommen und sind mit * gekennzeichnet: [ʔ], [χ], [ç], [R], [RR] und [rɹ]

3. Messapparatur

Im Rahmen verschiedener von der Deutschen Forschungsgemeinschaft und der Bayerischen Forschungsstiftung geförderter Forschungsprojekte wurde in mehreren Iterationen ein Laborsystem (*Otobox*) entwickelt, das nahezu alle Messmethoden der audiologischen Diagnostik abdeckt. Dieses umfassende, vielseitige System sowie die einzelnen Messverfahren sind in mehreren, sich ergänzenden Vorgängerarbeiten ausführlich beschrieben (Kandzia, 2005, 2011; Müller, 2002; Oswald, 2005; Rosner, 2004, 2011; Schirkonyer, 2006, 2007). Es besteht aus einer Hardware-Einheit ohne Display mit der Möglichkeit, verschiedene Wandler und weiteres Zubehör anzuschließen. Diese führt eine Datenvorverarbeitung durch und wird von einem PC gesteuert, der auch für die Visualisierung der Ergebnisse und Speicherung zuständig ist. Der Datenaustausch zwischen PC und Messhardware erfolgt über Bluetooth. Für die Stromversorgung ist ein externes Netzteil erforderlich. (Kandzia, 2011; Rosner, 2011).

Die Erfahrungen, die bei der Entwicklung und Anwendung des Laborsystems gesammelt wurden, waren hilfreich bei der Neuentwicklung eines Messsystems zusammen mit der Firma PATH MEDICAL GmbH (Germering). Das neue Gerät sollte abweichend vom Laborsystem einen Bildschirm mit Touchscreen haben, über einen Akku mit Strom versorgt werden und unabhängig von einem PC betrieben werden können.

3.1. Hardware

Das Ergebnis der Neuentwicklung ist die *Senti-Plattform*, die alle notwendigen Komponenten zur Durchführung der gängigen subjektiven audiometrischen Tests bereitstellt und als Medizinprodukt der Klasse IIa gemäß MDD (Medical Device Directive, EU-Medizinprodukte-Richtlinie 93/42/EWG) zertifiziert wurde. Sie kann über Firmwarelizenzcodes flexibel konfiguriert und an konkrete Bedürfnisse angepasst werden.

Ein *Senti (Typ Desktop)* mit zugehörigem Kopfhörer mit Peltor-Kapseln zur Dämpfung externer Störgeräusche (Sennheiser HDA 200) sind als Produktfotos in Abb. 3.1 dargestellt. Die Stromversorgung erfolgt über einen Li-Ion-Akku, wodurch das Gerät mobil einsetzbar ist. Das farbige Touchdisplay hat eine Diagonale von ca. 13 cm und ist damit groß genug, um auch von Kleinkindern, deren Feinmotorik noch nicht ausgereift ist, bedient zu werden.

Das Hardwareschema in Abb. 3.2 zeigt die verwendeten Hardwarekomponenten und deren Verschaltung. Herzstück des Systems ist ein digitaler Signalprozessor (DSP), der die Signal- und Datenverarbeitung sowie die Ansteuerung der anderen Systemkomponenten übernimmt. Firmware, Gerätekonfiguration und Messdaten speichert das Gerät in einem internen Flash Speicher. Für große Datenmengen, die Bilder und Sounddateien benötigen, steht zusätzlich eine SD-Card mit einer Kapazität von bis zu 16 GB zur Verfügung. Als Arbeitsspeicher



Abb. 3.1 Produktfoto des neuen Messsystems: Senti (Type Desktop) mit dem Kopfhörer Sennheiser HDA 200.
Quelle: PATH MEDICAL und Sennheiser

für den Datenaustausch zwischen den Komponenten und als Zwischenspeicher für Bilder ist SDRAM vorhanden. Das Display wird direkt vom Prozessor mit bis zu 60 Bildern pro Sekunde angesteuert. Es hat ein Seitenverhältnis von 3:4 und wird im Hochformat mit einer Auflösung von 240 x 320 Pixel und 15 Bit Farbtiefe betrieben. Über der Anzeigefläche befindet sich eine resistive, druckempfindliche Ebene. Bei Druck auf die Oberfläche berechnet der Touch-Controller die Koordinaten der Berührungsposition und übermittelt sie an den DSP.

Für subjektive Messungen steht ein Antworttaster zur Verfügung, mit dem beispielsweise angegeben werden kann, ob ein Ton gehört wurde. Per USB besteht die Möglichkeit, Messergebnisse an einen PC zu übertragen oder bei Bedarf Informationen des Gerätedisplays auf dem Computerbildschirm darzustellen und Eingaben entgegenzunehmen. Für die Digital-Analog-Wandlung steht ein CODEC mit insgesamt 8 Ausgangs- und 4 Eingangskanälen zur Verfügung. Das analoge Signal wird zur Tonausgabe über die zugehörigen Verstärker an Kopfhörer, Einsteckhörer und Knochenleitungshörer weitergeleitet. Alternativ können aktive Freifeldlautsprecher an den Kopfhörerausgang angeschlossen werden. Zusätzlich gibt es ein Mikrophon für die Hinsprecheinrichtung zur Kommunikation vom Fachpersonal zur Versuchsperson.

3.2. Firmwarestruktur

Das Messsystem ist mit einem einzigen Prozessor ausgestattet, der alle anfallenden Aufgaben übernimmt. Bei der unidirektionalen Übertragung vom DSP zum Display und der bidirektionalen Kommunikation zwischen DSP und CODEC fallen die größten Datenmengen an. Wie bereits von Kandzia (2011) beschrieben, war es eine wesentliche Anforderung, die Firmwarekomponenten so zu entwerfen, dass sie möglichst einfach und vielfältig in anderen Projekten des Kooperationspartners PATH MEDICAL wiederverwendet werden können. Aus diesem Grund wurde unter anderem auf den Einsatz eines (Realzeit-)Betriebssystems

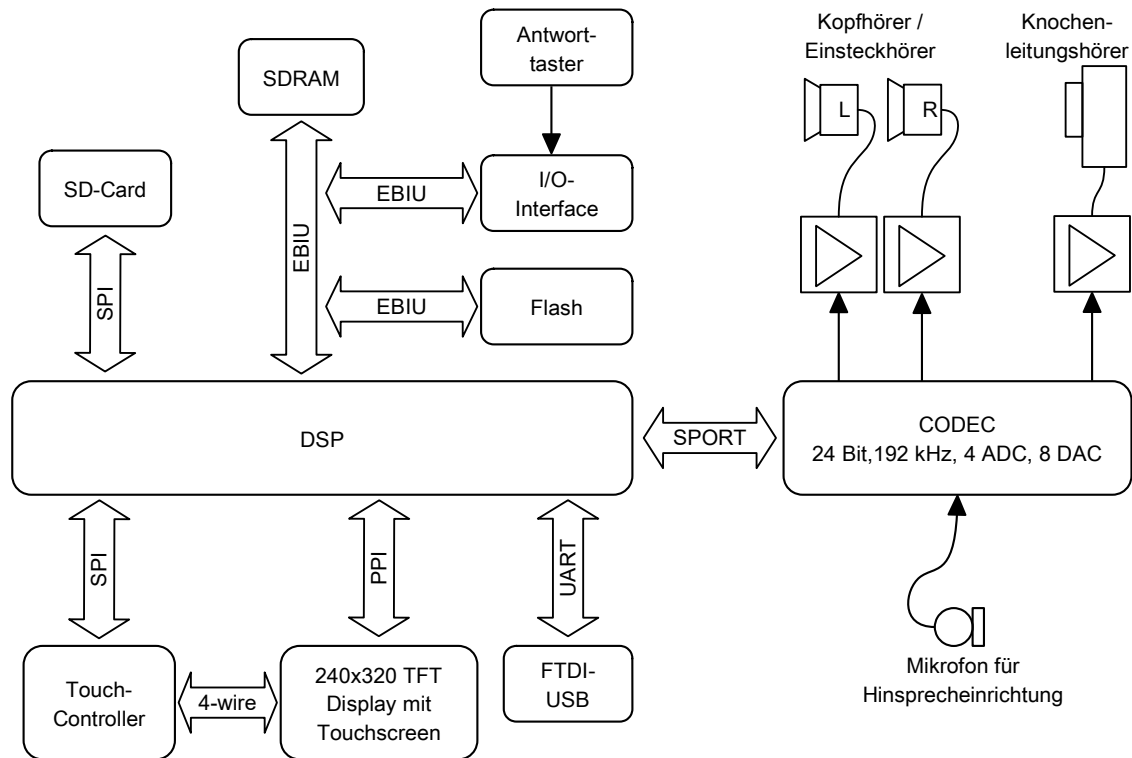


Abb. 3.2 Hardwareschema des Tischgeräts zur Durchführung subjektiver Tests. (in Anlehnung an Kandzia, 2011, S. 56, Abb. 3.11)

verzichtet. Daraus folgt, dass nur eine Prioritätsebene neben den Interrupts verfügbar ist. Der gesamte Anwendungscode läuft daher im *supervisor mode* mit vollem Zugriff auf alle Prozessorressourcen. (Kandzia, 2011)

Im Gegensatz zu den von Kandzia (2011) beschriebenen 3 Ebenen der Zeitanforderung mit den zugehörigen Tasks T_0 , T_1 und T_2 für objektive Messmodule sind subjektive Messmodule aufgrund ihrer geringeren Anforderungen an die Rechenleistung einfacher zu realisieren, sodass die Ebene mit T_1 nahezu entfällt. T_1 muss hierbei keine aufwendigen Berechnungen oder Datenverarbeitungen durchführen, sondern wartet lediglich auf Eingaben per Touchscreen und initiiert eine Reaktion darauf. Die Ebenen und Tasks T_0 und T_2 sind weitgehend identisch: T_0 ist für die sampleweise Verarbeitung der Daten zuständig, d.h. für die synchrone Tonausgabe und -aufnahme mit einer Abtastrate von 48 kHz. Sie enthält außerdem die Signalfilterung, die für die Freifeldentzerrung bei Sprachtests erforderlich ist. T_2 aktualisiert den Bildschirminhalt mit einer für die Anwendung adäquaten Wiederholrate von mehreren Bildern pro Sekunde.

Der modulare Aufbau der Firmware ist in Abb. 3.3 dargestellt. Innerhalb einer Hauptschleife findet beim Systemstart zunächst die Initialisierung der Hardware statt, gefolgt vom Start des ersten Moduls, das z.B. eine Übersicht über alle verfügbaren Messmodule anzeigt. Sobald ein Messmodul ausgewählt wird, beginnt die modulspezifische Initialisierung, die unter anderem einen Zeiger auf die verwendete *user_isr*-Funktion setzt, die als CODEC-ISR für die

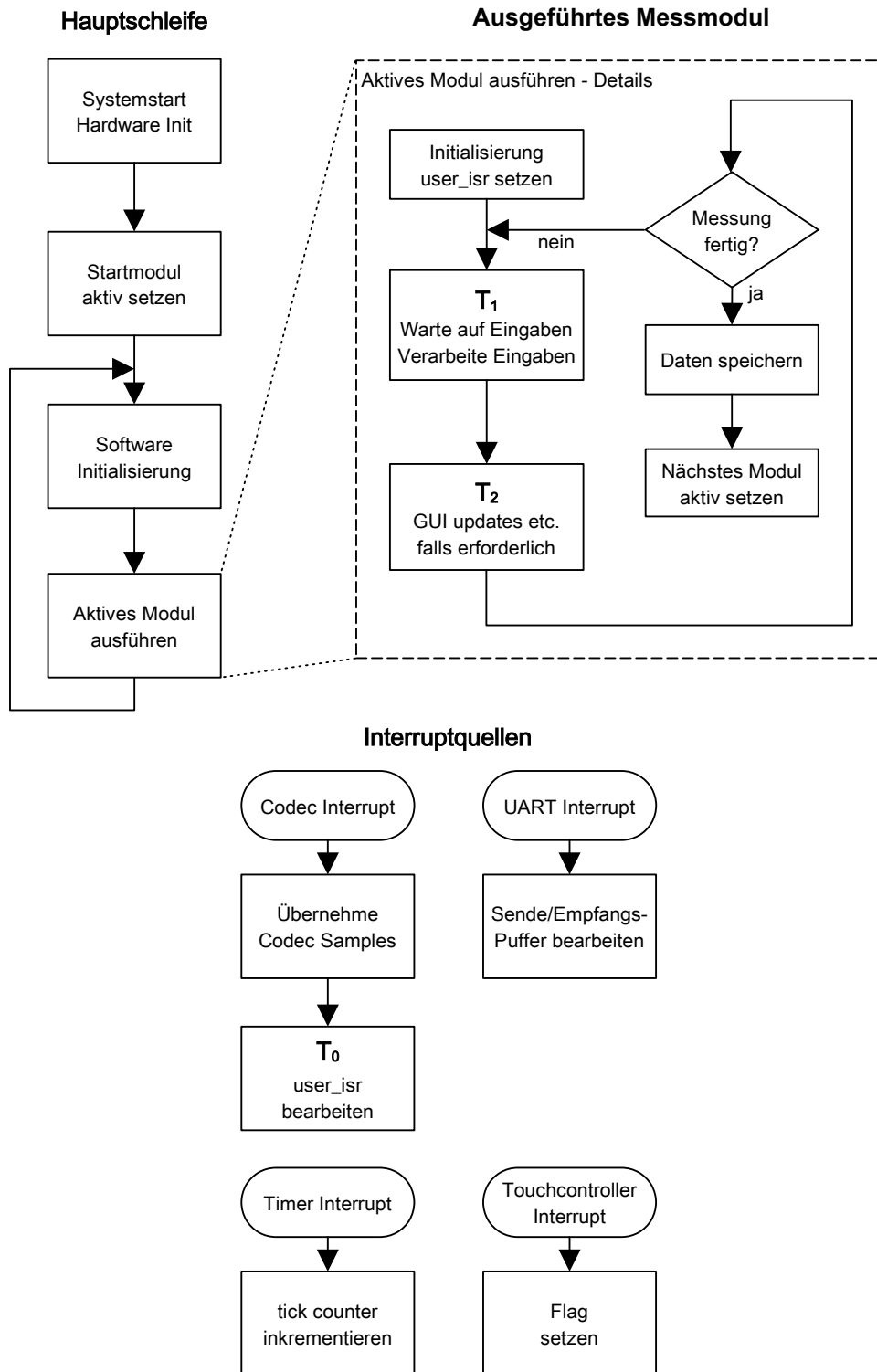


Abb. 3.3 Firmwarestruktur des Messgerätes. Sequentieller Programmablauf mit Priorisierung zeitkritischer Aufgaben durch Interrupts. Während einer Messung wird im Modul die Schleife um T_1 und T_2 durchlaufen, wobei in T_2 nur Rechenzeit anfällt, wenn sich der Bildschirminhalt verändert. Von den dargestellten Interrupt-Quellen sind in dieser Phase im Wesentlichen der Codec- und Timer-Interrupt relevant. Die durch den Timer-Interrupt, Touch-Ereignis und UART-Kommunikation erzeugte Last ist im Vergleich zur Codec-ISR vernachlässigbar. (modifiziert nach Kandzia, 2011, S. 60, Abb. 3.13, mit freundlicher Genehmigung des Autors)

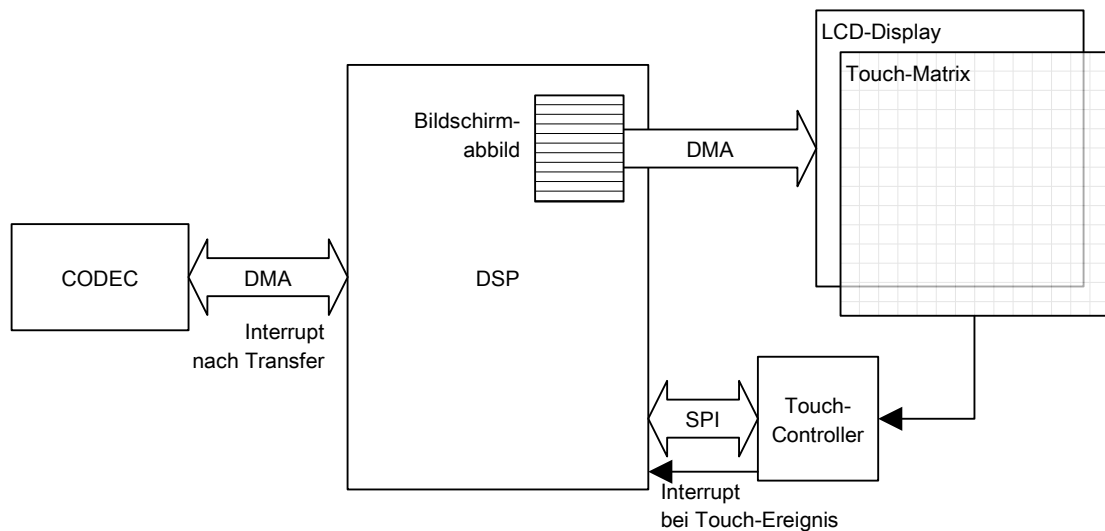


Abb. 3.4 Schema der Kommunikation zwischen DSP und Peripherie. Der Datenaustausch mit dem CODEC findet über DMA-Transfers statt, jeweils nach Übertragung eines Blockes (8 Kanäle à 32 Bit) erfolgt ein Interrupt. Das im Speicher des DSPs gehaltene Bildschirmabbild wird mittels DMA zum Display übertragen. Bei einem Touch-Ereignis löst der Touch-Controller einen Interrupt aus. Die Daten des Ereignisses werden im Controller gespeichert und können auch zu einem späteren Zeitpunkt über SPI ausgelesen werden. (modifiziert nach Kandzia, 2011, S. 57, Abb. 3.12, mit freundlicher Genehmigung des Autors)

sampleweise Datenverarbeitung verantwortlich ist (T_0). Das aktive Messmodul führt wiederum selbst eine Schleife aus, die im Wesentlichen aus T_1 und T_2 besteht. (Kandzia, 2011)

Abb. 3.4 veranschaulicht den Datenaustausch zwischen DSP und Peripherie. Die zeitkritischen und großen Datentransfers zwischen DSP und CODEC und vom DSP zum Display erfolgen direkt zwischen den Komponenten mittels DMA-Transfer. Sobald der CODEC einen Transfer beendet hat, wird ein Interrupt ausgelöst und die Abarbeitung der entsprechenden Interrupt Service Routine (ISR) gestartet (T_0). Berührungen des Touchscreens lösen ebenfalls einen Interrupt aus. Die Daten des Ereignisses stehen dann im Touch-Controller zum Auslesen durch den DSP bereit. (Kandzia, 2011)

4. Testverfahren

In diesem Kapitel werden drei neue Verfahren vorgestellt, die speziell auf die Besonderheiten der Hördiagnostik bei Kindern ab 2 Jahren zugeschnitten sind. Es handelt sich hierbei um den „MAGIC“ genannten tonaudiometrischen Test, den Sprachtest „MATCH“ und den „BASD“-Test für den Einsatz im Screening bei Verdacht auf Sprachauffälligkeiten, die einer pädagogischen Sprachförderung bedürfen, auf Sprachentwicklungsstörungen sowie auf Auditive Verarbeitungs- und Wahrnehmungsstörungen (AVWS). Sämtliche Tests wurden als Selbsttest zur Durchführung und Steuerung per Touchscreen entworfen und implementiert. Dadurch wird das Fachpersonal entlastet und die Motivation der untersuchten Kinder gesteigert. Dieses Kapitel beschreibt das Testdesign, die Konstruktion und die Implementierung der Testverfahren. Dabei flossen auch die Resultate erster Erprobungen in die Weiterentwicklung mit ein. Im nachfolgenden Kapitel 5 werden die klinischen Studiendaten vorgestellt, die unter Verwendung der jeweils abschließenden Versionen der Tests erhoben wurden.

4.1. Bildgestützte Tonschwellenaudiometrie (MAGIC)

4.1.1. Einleitung

Mehrere Faktoren beeinflussen das Ergebnis bei einer Untersuchung des Hörorgans bei Kleinkindern. Diese sind insbesondere der Entwicklungsstand und die Aufmerksamkeit des Kindes sowie die Beobachtungsgabe und Erfahrung des Fachpersonals. Die Spielaudiometrie ist ein etabliertes Verfahren zur Bestimmung der Reinton-Hörschwelle bei Kindern. Zur Durchführung ist neben einer Audiometrieanlage auch Spielmaterial erforderlich. Die Aufgabe besteht darin, mit Gegenständen auf ein akustisches Signal hin eine Spielhandlung auszuführen. Für das Klinikpersonal ist die Durchführung herausfordernd, weil das Kind genau beobachtet werden muss und gleichzeitig das Audiometer so zu bedienen ist, dass keinerlei (auch keine unbewusste) Beeinflussung des Kindes stattfindet. Der im Folgenden vorgestellte Test entlastet das Personal, indem die Steuerung automatisiert erfolgt. Durch die Verwendung eines Touchscreens und von Bildern ist kein zusätzliches Spielmaterial mehr erforderlich.

Die einzelnen Arbeitsschritte zur Veröffentlichung des MAGIC waren:

1. Entwicklung des tonaudiometrischen Tests
 - a. Entwicklung eines kindgerechten Konzepts und Verfahrens
 - b. Erstellung des erforderlichen Bildmaterials
 - c. Implementierung des Tests auf einem geeigneten Messsystem
2. Validierung des tonaudiometrischen Tests
 - a. Rekrutierung der Versuchspersonen
 - b. Durchführung des Tests
 - c. Statistische Auswertung und Ergebnisdokumentation

4.1.2. Testdesign

Entwicklung eines kindgerechten Konzepts und Verfahrens

Das in dieser Arbeit vorgestellte Test-Verfahren für Kinder greift auf Ideen zurück, die von Heller und Opp-Enzinger (2006) vorgestellt wurden und stellt eine Weiterentwicklung dieser dar. Die wesentliche Neuerung an der von Heller und Opp-Enzinger präsentierten Methode – im Vergleich zur Praxis bei Hörtests – ist, dass das Kind die Stimulusdarbietung selbst starten und wiederholen darf. Dadurch sinkt der mögliche Einfluss der Vigilanz und die Antworten („Ton gehört“ oder „Ton nicht gehört“) sind sicher. Dies erlaubt es, die Pegelwiederholungen zum Finden der Schwelle, wie sie in ISO 8253-1 (2011) beschrieben sind, wegzulassen, was die Messung beschleunigt. Der Algorithmus zur Pegelsteuerung ist eine Kombination aus der Methode des Trackings (Ja/Nein Antwort bewirkt Sinken/Steigen des Reizpegels) und adaptiven Verfahren (die Schrittweite sinkt mit Annäherung an die Schwelle).

Für einen reibungslosen Ablauf ist die Konditionierung am Anfang des Hörtests wichtig, um die Aufgabe zu vermitteln. Sie besteht darin, verschiedene Tiere zu untersuchen, ob sie Töne von sich geben, die manchmal sehr leise und nur gerade so wahrnehmbar sind. Dabei kommt auch ein sogenannter „stummer“ oder „mute“ Pegel zum Einsatz, bei dem keinerlei Tonausgabe erfolgt, um zu demonstrieren, dass es normal ist, manchmal nichts zu hören. Er dient außerdem zur Kontrolle, ob auf unhörbare Darbietungen die richtige Antwort gegeben wird. Als Stimuli werden frequenzmodulierte Töne verwendet (Heller et al., 1995).

Das Verfahren von Heller und Opp-Enzinger (2006) wurde in einem prototypischen Demonstrator zur Durchführung einer Machbarkeitsstudie implementiert. Dies erfolgte in Zusammenarbeit mit der HNO-Klinik für Kommunikationsstörungen in Mainz. Das System bestand aus einer Vielzahl von Komponenten. Die gesamte Steuerung des Ablaufs und die Eingabe der Antworten der Kinder erfolgte manuell durch das Fachpersonal.

Die Hauptziele der hier vorgestellten Weiterentwicklung waren, (a) den Test auf einem Messsystem verfügbar zu machen, das eine Zulassung als Medizinprodukt hat, und (b) die Ablaufsteuerung zu verbessern und zu automatisieren. Des Weiteren sollte der Frequenzbereich von 500 Hz bis 4 kHz auf 250 Hz bis 8 kHz erhöht werden. Die Erfahrungen aus der Erprobung in Mainz sollten darüber hinaus Berücksichtigung finden.

Der überarbeitete und erweiterte Test erhielt den Namen MAGIC. Er erfordert nach der Konditionierung keine weitere aktive Mitarbeit der Testleitung, wenn das Kind die Aufgabe vollständig beherrscht. Da zur Durchführung des Tests ausschließlich das in Kapitel 3 beschriebene Messsystem notwendig ist, kann die Sitzanordnung im Untersuchungsraum freier gestaltet werden als die bei der Spielaudiometrie übliche Positionierung. Statt gegenüber des Kindes kann das Untersuchungspersonal auch daneben oder im 90°-Winkel Platz nehmen und bei Bedarf eingreifen oder das Kind unterstützen. Die Eingabe, ob das gerade untersuchte Tier hörbar ist, erfolgt direkt am Messgerät. Auf diese Weise wird der Medienbruch der alten

Testvariante, nämlich von Tonauslösen per Touchscreen zu mündlichem Mitteilen des Ergebnisses, vermieden. Dies erfordert zusätzliche Bilder und stellt einen wesentlichen Schritt in der Konditionierung dar. Die Festlegung der Bilder und die Wortwahl, wenn die Aufgabe dem Kind erklärt wird, muss so gestaltet sein, dass es keine Vorbehalte gibt. So haben sich in Vorversuchen die Begriffe „glücklich“ und „traurig“ oder „gesund“ und „krank“ als nur begrenzt geeignet herausgestellt, um die Eingabe „gehört“ und „nicht gehört“ zu repräsentieren. Manche Kinder wollen nicht, dass Tiere traurig sind, während anderen von ihren Eltern beigebracht wurde, aufgrund der potenziellen Ansteckungsgefahr niemals kranke Tiere anzufassen. Daher wurde die Variante gewählt, dass die Tiere singen, wenn der Ton hörbar ist, und heiser sind sowie einen Schal erhalten, falls kein Ton hörbar ist (s. auch Abb. 4.3). Als Reiz kommen ausschließlich Sinus- und frequenzmodulierte Töne zum Einsatz, die der Norm für Tonaudiometer (DIN 60645-1, 2018) entsprechen. Die Kalibrierung erfolgt gemäß den in Abschnitt 2.4 beschriebenen Regeln und Vorgehensweisen.

Der Algorithmus zur Bestimmung der Hörschwelle wurde mit dem Ziel entwickelt, die Messdauer durch Minimierung der Anzahl der Prüfpegel so kurz wie möglich zu halten. Dies wird zum einen durch eine Halbierung der Schrittweite von 20 dB zu Beginn der Messung auf 5 dB nach 3 Reizdarbietungen erreicht und zum anderen durch die Möglichkeit, eine Präsentation bei einem bestimmten Pegel zu wiederholen, bis ein sicheres Urteil gefällt werden kann, ob das gerade untersuchte Tier einen Ton abgibt oder nicht. Während der Messung wird zur Überprüfung der Zuverlässigkeit nach jeder 3. „gehört“ Eingabe ein „stummer“ Pegel präsentiert. Bei einem Ohr- oder Frequenzwechsel wird der „stumme“ Pegel um einen Schritt nach hinten verschoben, damit das Kind mit der neuen Teilaufgabe vertraut gemacht wird. Für den Fall, dass der 1. Prüfpegel nicht gehört wird, erhöht sich der Pegel in 20 dB Schritten bis der Ton gehört oder der eingestellte Maximalpegel erreicht wird.

Abb. 4.1 zeigt die Messung beider Ohren bei der ersten Testfrequenz, einschließlich der Konditionierung. Die Konditionierung startet mit einem vergleichsweise leisen, aber deutlich hörbaren Ton, damit sich das Signal, auf das es ankommt, besser einprägt. Hierbei soll das Kind den Ton mehrfach selbst auslösen und die Gerätebedienung kennenlernen, insbesondere, dass das Drücken auf den Bauch des Tieres zur Tonausgabe führt. Der 2. Ton soll leiser, aber noch immer überschwellig hörbar sein, um das Kind für die Aufgabe des genauen Hinhörens zu sensibilisieren und zu trainieren. Am Ende der Konditionierung wird mit dem „stummen“ Pegel die mögliche Abwesenheit des Tones demonstriert. Auch beim 2. und 3. Schritt ist es wichtig, mehrfach auszuprobieren und genau zuzuhören, und zwar vor der jeweiligen Eingabe auf dem Gerät.

Während der Konditionierung wird die Aufgabe in einer kindgerechten Rahmengeschichte erklärt. Sie könnte in der Art lauten, dass Tiere verschiedener Tierarten, die alle unterschiedliche Töne erzeugen, zu untersuchen sind. Allerdings sind manche Tiere heute heiser und die Aufgabe ist herauszufinden, welche das sind und einen Schal erhalten. Bei besonders jungen Kindern, die sehr lange auf den Touchscreen drücken, ist es hilfreich, ihnen eine kur-

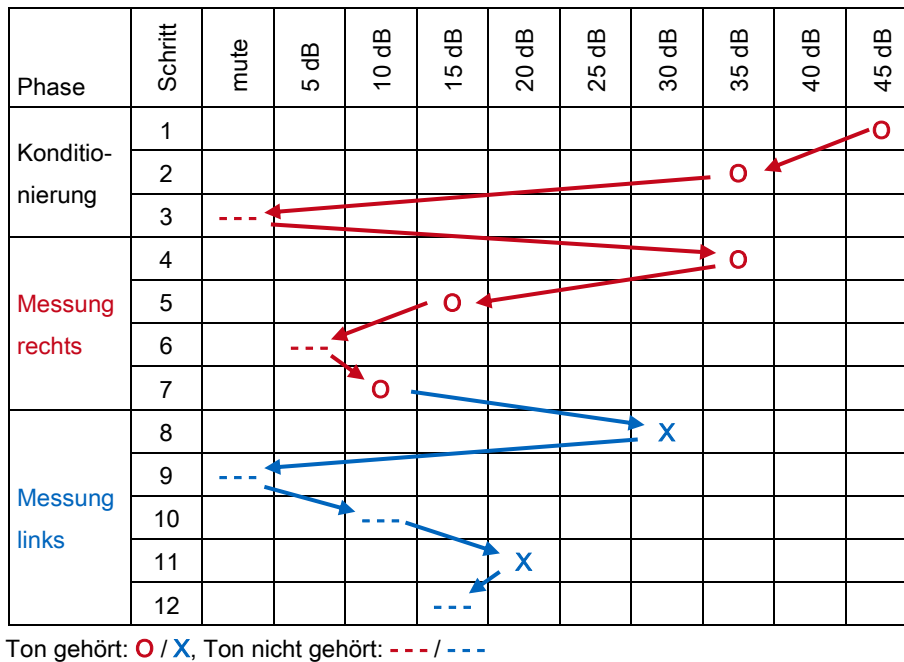


Abb. 4.1 MAGIC adaptiver Algorithmus. Der Test beginnt mit einer Konditionierungsphase, in der dem Kind die Art des Tons sowie das Konzept von leisem und keinem Ton vermittelt werden. Anschließend startet die Messung des rechten Ohres (rot dargestellt). Sobald die Hörschwelle des rechten Ohrs bestimmt wurde, wird das linke Ohr gemessen (blau). Das Ziel ist es durch Pendeln um die Hörschwelle und Verkleinern der Schrittweite mit möglichst wenigen (ca. 4–5) Pegeln die Schwelle (hier: Rechts 10 dB HL; Links 20 dB HL) zu finden.

ze Geschichte zu erzählen, z.B. dass man das Tier nicht so stark drücken darf, weil es weh tun könnte. Je nach Situation kann die Eingabe von „gehört“ oder „nicht gehört“ auch durch die Untersuchenden erfolgen, und nur die Tonpräsentation vom Kind selbst, um im richtigen Moment aufmerksam zuzuhören.

Nach der Konditionierung, die im Beispiel in Abb. 4.1 auf dem rechten Ohr stattfindet, startet die Messung am gleichen Ohr mit dem niedrigeren der beiden Pegel der Konditionierung (35 dB). Da dieser hörbar ist, wird der Pegel um 20 dB auf 15 dB abgesenkt. Der 15-dB-Ton ist ebenfalls wahrnehmbar, sodass nach Halbierung der Schrittweite ein 5-dB-Ton folgt der nicht gehört wird. Nach erneuter Halbierung der Schrittweite folgt ein 10-dB-Ton, der gehört wird. Die Schwelle ist hier als leisester gehörter Pegel definiert, sodass sie für das rechte Ohr bei 10 dB HL liegt.

Falls beide Ohren gemessen werden, folgt die Messung des linken Ohres, bevor die Frequenz geändert wird. Der Startpegel ist immer 20 dB über der letzten Hörschwelle oder der in den Einstellungen gewählte Startpegel, wenn dieser niedriger ist. Dadurch soll der erste Ton wahrnehmbar sein, unter Sicherstellung, dass keine zu lauten Pegel im Fall von großen Seitendifferenzen erzeugt werden. Der nächste Pegel im 9. Schritt ist ein „stummer“ Pegel zur Überprüfung der Zuverlässigkeit. Anschließend folgen die weiteren Pegel mit Schrittweiten von 20, 10 und 5 dB. Die Richtung der Änderung hängt davon ab, ob „gehört“ oder „nicht gehört“ eingegeben wurde. In Abb. 4.1 liegt die Schwelle für das linke Ohr bei 20 dB HL.

Bei einer falschen Eingabe auf einen „stummen“ Pegel wird der Darbietungspegel erhöht, und eine erneute Überprüfung mit einem „stummen“ Pegel erfolgt bereits nach zwei „gehört“ Eingaben.

Erstellung von Bildmaterial

Das Bildmaterial ist bei diesem Test ein wichtiges Element zur Steigerung der Motivation und Aufmerksamkeit. Die Gestaltung der Bilder für Kinder ab 2 Jahren muss nach speziellen Kriterien erfolgen (Wechtenbruch, 1996). Gegenstände sollen vollständig abgebildet werden, von vorne oder von der Seite aus betrachtet. Auf perspektivische Effekte, Schatten sowie Farbgebung soll verzichtet werden (Wechtenbruch, 1996). Ausgehend von diesen Vorgaben sollte ein moderner, einfacher und kindgerechter Stil zur Anwendung kommen, der nur die für die Erkennung wesentlichen Details enthält. Hierfür wurde schon von Beginn des Planungsprozesses an eng mit der Firma DAAL VISION (Prien a. Chiemsee) zusammengearbeitet. Es wurden 8 Tiere ausgesucht, die Kindern ab 2 Jahren aus Bilderbüchern bekannt sind. Die Tierbilder sind in Abb. 4.2 dargestellt. Jedes Tier repräsentiert, soweit möglich, eine Testfrequenz, die zu den vom Tier normalerweise erzeugten Lauten passt:

250 Hz – Kuh; 500 Hz – Bär; 1 kHz – Elefant; 2 kHz – Katze;
3 kHz – Schaf; 4 kHz – Maus; 6 kHz – Vogel; 8 kHz – Delfin.

Alle Tiere wurden, wie für den Testalgorithmus erforderlich, in 4 Varianten erstellt. Abb. 4.3 stellt sie am Beispiel des Bären dar. Es handelt sich um folgende Zustände:

- Neutral: Die Tonpräsentation kann durch Drücken auf den Bauch ausgelöst werden.
- Aktiv: Visualisierung der laufenden Tonausgabe.
- Heiser: Eingabe, wenn kein Ton wahrnehmbar ist.
- Singend: Eingabe von „Ton gehört“

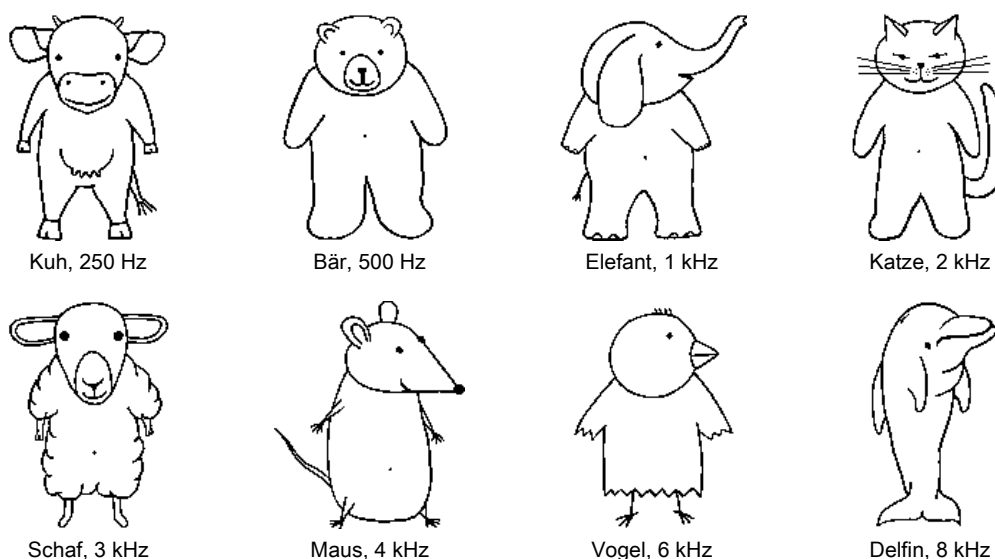


Abb. 4.2 MAGIC Bildkarten: Darstellung aller 8 Tierbilder mit der zugehörigen Stimulusfrequenz.

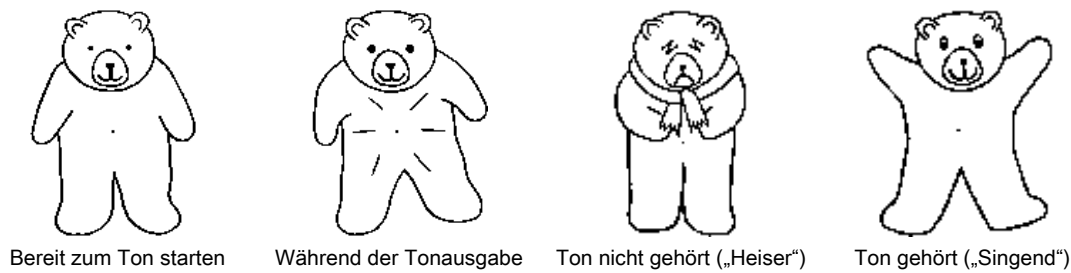


Abb. 4.3 Zustände der Tiere am Beispiel des Bären: Normal (Tonausgabe kann durch Drücken auf den Bären gestartet werden), Tonausgabe aktiv, Ton nicht gehört („heiser“) und Ton gehört („singend“)

4.1.3. Implementierung

Die Implementierung des MAGIC erfolgte in der Programmiersprache C auf der in Kapitel 3 beschriebenen Hardwareplattform unter Zuhilfenahme einer auf *Eclipse* (Eclipse Foundation, Brüssel, Belgien) basierenden Softwareentwicklungsumgebung für eingebettete Systeme. Bildschirminhalte, die zur Interaktion von Versuchspersonen mit dem Gerät gedacht sind, sollten so gestaltet werden, dass sie nur diejenigen Elemente enthalten, die für den Messablauf erforderlich sind. Alle Bedienelemente sollten möglichst groß sein, um damit die Bedienung des Tests zu vereinfachen. Lediglich eine Kopfzeile mit Informationen zur Überwachung des Tests sollte zusätzlich sichtbar sein.

In Abb. 4.4 ist dargestellt, was eine Testperson auf dem Gerätebildschirm sieht. Bis auf eine schmale Kopfzeile mit Information zum aktuellen Messstatus wird die gesamte verfügbare Bildschirmfläche für die Interaktion mit der Versuchsperson genutzt. Da es keine zusätzlichen Knöpfe gibt, ist eine Fehlbedingung unwahrscheinlich und das Gerät sicher während eines Selbsttests bedienbar. Abb. 4.4a zeigt die Auswahl des Tieres (der Messfrequenz) an.

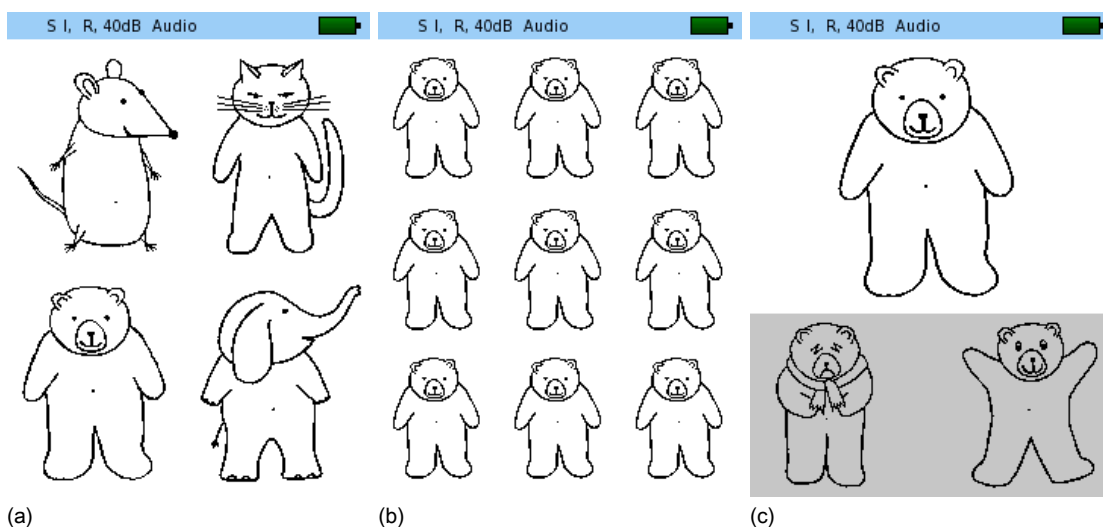


Abb. 4.4 MAGIC Messverlauf: Die Prozedur aus (a)–(c) wird wiederholt bis alle Testfrequenzen gemessen wurden. (a) Auswahl der Frequenz. (b) Regal mit Tieren aus dem der Reihe nach eines zur Prüfung, ob es einen Ton erzeugt, entnommen wird. (c) Prüfung des Tieres und Eingabe.

Da in dieser Darstellung insgesamt 4 Bilder Platz haben, wurden die 8 zur Verfügung stehenden Frequenzen in je 4 Haupt- und Nebenfrequenzen eingeteilt. In Übereinstimmung mit den Empfehlungen der Weltgesundheitsorganisation (World Health Organization, 2021) sind 500 Hz, 1 kHz, 2 kHz und 4 kHz die Hauptfrequenzen und werden immer auf der ersten Seite präsentiert (Abb. 4.4a). Dementsprechend sind 250 Hz, 3 kHz, 6 kHz und 8 kHz Nebenfrequenzen. Diese Töne werden erst nach den Hauptfrequenzen dargeboten (Abb. 4.5a).

Die Tiere wurden so angeordnet, dass jeweils die größeren und schwereren unten im Bild erscheinen (Abb. 4.4a und 4.5a). In Abb. 4.4b ist das Regal der zu untersuchenden Tiere zu sehen. Es beinhaltet am Anfang jeweils 9 Exemplare der ausgewählten Tierart. Die im Verlauf der Untersuchung aus dem Regal entnommenen Tiere verschwinden, sodass das sich leerende Regal den Messfortschritt anzeigt. Das Auslösen der Tondarbietung und die Eingabe erfolgen auf dem in Abb. 4.4c dargestellten Bildschirm. Die Eingabe ist zunächst gesperrt, was durch einen dunkelgrauen Hintergrund angezeigt wird (hier nicht dargestellt). Sie wird kurze Zeit nach einem Druck auf das Tier und der dadurch ausgelösten Stimuluspräsentation freigegeben. Zur Anzeige, dass die Eingabe möglich ist, wird die Hintergrundfarbe des Eingabebereichs heller. In der blauen Kopfzeile sind, von links nach rechts, abgesehen von der Akkuanzeige, folgende Informationen enthalten (Abb. 4.5c): Stimulus (**S**inus-/**F**requenzmodulierter Ton), Phase (**M**essung oder **I**nstruktion/**K**onditionierung), aktuelles Testohr (**R**echts/**L**inks), gegenwärtiger Darbietungspegel, Modus (**A**udiometrie oder **S**creening), vorheriger Darbietungspegel und letzte Eingabe (Ohrsymbol für „gehört“ und durchgestrichenes Ohrsymbol für „nicht gehört“).

Während der Messung stehen für das Fachpersonal die in Abb. 4.5 gezeigten Optionen zur Verfügung. Durch einen Druck auf den Einschalttaster wird eine Fußzeile eingeblendet

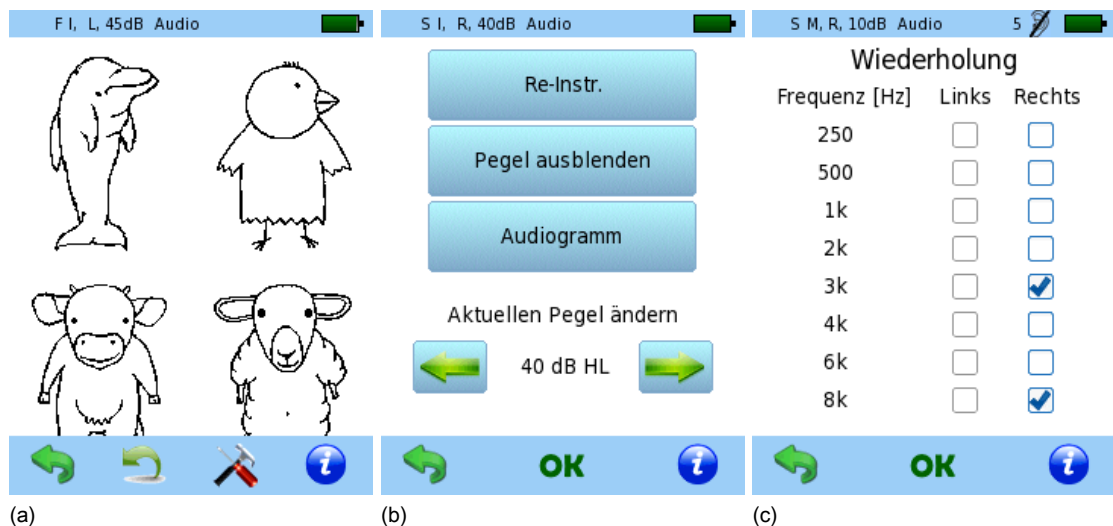


Abb. 4.5 MAGIC Einstelloptionen während des Tests: (a) Die Fußzeile zum Eingriff durch Untersuchende erscheint beim Druck auf den Einschalttaster. (b) Möglichkeiten den Testablauf zu verändern: Konditionierung starten, aktuellen Pegel ausblenden, bisheriges Audiogramm anzeigen und Darbietungspegel ändern. (c) Auswahl von Frequenzen, die nochmals geprüft werden sollen.

(Abb. 4.5a). Über die dort vorhandenen Knöpfe kann die Messung vorzeitig verlassen werden, die letzte Eingabe rückgängig gemacht werden und eine Seite mit Testeinstellungen aufgerufen werden (Abb. 4.5b). In den Testeinstellungen besteht die Möglichkeit den Darbietungspegel manuell zu verändern, die Konditionierung erneut zu starten und die Information zum aktuellen Ton in der Kopfzeile auszublenden. Des Weiteren kann das Audiogramm, das bisher ermittelt wurde, angezeigt werden. Nach einem Messdurchlauf können einzelne Frequenzen wiederholt werden (Abb. 4.5c). Testfrequenzen, bei denen eine „gehört“ Eingabe als Antwort auf einen „stummen“ Pegel gemacht wurde, sind bereits vorausgewählt.

Vor einer Messung können die Einstellungen und Optionen vom Untersuchungspersonal eingesehen und geändert werden (Abb. 4.6). Auf der 1. Seite (Abb. 4.6a) werden der Messmodus, die Testfrequenzen und der Stimulus festgelegt. Das Testmodul kann zur adaptiven Messung der Hörschwelle an den ausgewählten Frequenzen verwendet werden (*Audio*) oder ein Screening bei ausgewählten Pegeln durchführen (Abb. 4.6c). Die Stimulation kann mittels Sinus- (*Reinton*) und frequenzmodulierten (*Wobbel*) Tönen erfolgen. Zusätzlich zur Darbietung eines 1,2s langen Tons beim Druck auf das Tier ist eine Tonfolge aus einem kurzen (ca. 0,4 s) und dem 1,2s langen Ton (*2 Töne*) verfügbar, um die Aufmerksamkeit stärker auf die Tonpräsentation zu lenken. Immer wenn Personen gemessen werden, die lesen können, ist es empfehlenswert den aktuellen Darbietungspegel aus der Kopfzeile mit der Option *Pegel ausblenden* zu entfernen (Abb. 4.6a). Wenn als Messmodus *Audio* gewählt wird, können auf der 2. Seite der Einstellungen (Abb. 4.6b) die Parameter zum adaptiven Messalgorithmus verändert werden. Der Anfangspegel sollte zu einer leisen, aber dennoch gut hörbaren Reizwahrnehmung führen. Die oberen und unteren Grenzen sollten zum erwarteten Hörverlust und zur Versuchsperson passen. Beispielsweise kann der Minimalpegel bei Kindern auf 5 dB gestellt werden, um Zeit zu sparen, da eine Schwelle unter 5 dB keinen diagnostischen Mehrwert hat, aber die Messung verlängert. Sollte die Wahl eines Tieres aus dem Regal



Abb. 4.6 MAGIC Testeinstellungen: (a) Einstellung von Messmodus, Testfrequenzen und Stimulus. (b) Pegeleinstellungen für den adaptiven Modus zur Messung des Audiogramms. (c) Auswahl der Screeningpegel.

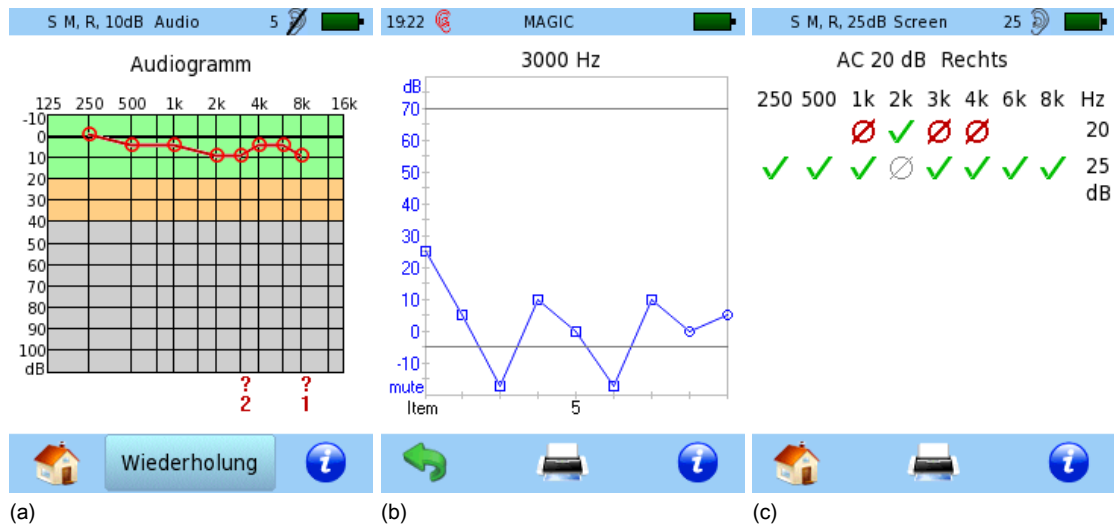


Abb. 4.7 MAGIC Ergebnisdarstellung: (a) Audiogramm und Angabe, ob „stumme“ Pegel als „gehört“ angegeben wurden (? und Zahlen). (b) Messverlauf bei einer Frequenz. Quadrate bedeuten „gehört“, Kreise stehen für „nicht gehört“. (c) Ergebnis eines Screeningtests. Grüne Haken bedeuten „gehört“.

(Abb. 4.4b) die Versuchsperson überfordern oder viel Zeit in Anspruch nehmen, kann dieser Schritt durch Setzen der Option *Tier-Regal überspringen* ausgelassen werden. Stattdessen wandert das Tier nach der „gehört“/„nicht gehört“ Eingabe automatisch aus dem Bildschirm und von der anderen Seite erscheint ein neues.

In Abb. 4.7 sind Ergebnisse von Messungen dargestellt, die mit den Einstellungen gemäß Abb. 4.6 ermittelt wurden. Abb. 4.7a zeigt ein Audiogramm mit 8 Testfrequenzen des rechten Ohrs direkt nach Beenden der 8. Testfrequenz. In der Kopfzeile sind die Daten der letzten Eingabe sichtbar (5 dB, „nicht gehört“). Unter dem Audiogramm stehen bei 2 Frequenzen (3 kHz und 8 kHz) rote Fragezeichen und darunter Zahlen. Das bedeutet, dass bei diesen Frequenzen „stumme“ Pegel 2 x bzw. 1 x als „gehört“ angegeben wurden. Über den *Wiederholen*-Knopf wird die Seite zum Einstellen der Frequenzen geladen, die wiederholt werden sollen (Abb. 4.5c). Dadurch kann unmittelbar eine Überprüfung der fraglichen Ergebnisse erfolgen. Zu jeder Testfrequenz wird der vollständige Messverlauf aus Pegel und „gehört“ bzw. „nicht gehört“ gespeichert. Die Darstellung dieser Daten (Abb. 4.7b) wird durch Tippen auf die jeweilige Frequenz aufgerufen. Hier sind zusätzlich die eingestellten Minimum- und Maximum-Pegel (-5 dB bzw. 70 dB) durch graue horizontale Linien dargestellt. Quadrate stehen für eine „gehört“- und Kreise für „nicht gehört“-Eingabe. Mute ist der „stumme“ Pegel, bei dem keinerlei Tonausgabe erfolgt. Bei einem Screeningtest wird lediglich eine „gehört“- bzw. „nicht gehört“-Eingabe gespeichert (Abb. 4.7c). Die Messung beginnt beim niedrigsten gewählten Pegel und wird nur mit einem höheren Pegel fortgesetzt, falls der Ton nicht gehört wird.

Vorläufige Ergebnisse aus Abschnitt 4.1 wurden in Auszügen vorab auf Tagungen präsentiert (DGA Jahrestagung 2010, DGPP Jahrestagung 2010, EFAS 2011, HNO 88. Jahresversammlung 2017)

4.2. Bildgestützte Sprachaudiometrie (MATCH)

4.2.1. Einleitung

Der Mainzer Kindersprachtest I besteht aus zehn ein- und zweisilbigen Items (siehe auch Abschnitt 2.2.2). Die Zusammensetzung aus Lautmalereien und Wörtern mit überwiegend hochfrequenten Signalanteilen entspricht jedoch nicht dem Wortschatz von Zweijährigen. Eine Erweiterung des Testmaterials ist daher notwendig. Grundsätzlich sollte das vorhandene Wortmaterial des Mainzer Kindersprachtests I aufgrund der fehlenden phonetischen und phonematischen Ausgewogenheit und der unspezifischen Auswahl der Testitems durch anderes Wortmaterial ersetzt werden.

Als Testitems kommen nur solche Wörter infrage, die von der jeweiligen Altersgruppe sicher erkannt werden und die sich gut bildlich darstellen, um mit Bildkarten arbeiten zu können. Bei zweijährigen Kindern steht daher nur ein eingeschränkter Wortschatz für die Testung zur Verfügung. Ein Ziel der vorliegenden Arbeit war es, einen neuen kindgerechten Sprachtest als Alternative zum Mainzer Kindersprachtest I zu entwickeln und klinisch zu validieren.

Das Wortmaterial eines Kindersprachtests sollte auf einer standardisierten Erhebung des rezeptiven Wortschatzes des jeweiligen Lebensalters beruhen. Die Altersangemessenheit von Tests ist eine wesentliche Voraussetzung für die Testgüte. Tests sollten valide, reliabel und standardisiert sein. Eine Standardisierung von Sprachtests ist nur möglich, wenn die Sprache von Ton- oder Datenträgern abgespielt wird. Der Nachteil dieser Methode besteht darin, dass das Tonband für das Kind weniger attraktiv ist als das unmittelbar gesprochene Wort, was eine geringere Aufmerksamkeit des Kindes zur Folge hat. Das bedeutet, dass bei Tests mit Schallereignissen von Ton- oder Datenträgern audio-visuelle Stimuli eingesetzt werden müssen, um den Spannungsbogen über die gesamte Testdauer aufrechtzuerhalten.

Die einzelnen Arbeitsschritte zur Veröffentlichung des MATCH waren:

1. Entwicklung des Sprachtests
 - a. Auswahl des Wortmaterials auf der Grundlage aktueller Erhebungen
 - b. Erstellung von digitalen Bildtafeln für das Wortmaterial
 - c. Anfertigung von Tonaufnahmen im Tonstudio
 - d. Entwicklung eines kindgerechten adaptiven Verfahrens
 - e. Implementierung des Tests auf einem geeigneten Messsystem
2. Validierung des Sprachtests
 - a. Rekrutierung der Versuchspersonen
 - b. Durchführung des Sprachtests
 - c. Statistische Auswertung und Ergebnisdokumentation

4.2.2. Erstellung von Sprach- und Bildmaterial

Identifikation und Auswahl des Wortmaterials

Wie eingangs erwähnt, sollte das Wortmaterial erneuert und erweitert werden. Dabei stand vor allem die aktiv gesprochene Sprache und deren Wortschatz im Fokus und nicht das Vokabular aus Kinder- und Bilderbüchern, um eventuelle Einschränkungen im Alltag möglichst realitätsnah zu erfassen.

Kurz vor Beginn der Arbeiten am neuen Sprachtest widmete sich die Entwicklungspsychologie dem Thema Wortschatz mit dem Ziel, einen Elternfragebogen als diagnostisches Screeningverfahren zur Früherkennung von Kindern mit Sprachentwicklungsverzögerungen zu entwickeln. Die Konstruktion des Fragebogens erfolgte in mehreren Schritten. Die Normierungsstichprobe umfasste 685 einsprachig deutsch aufwachsende Kinder im Alter von 21–24 Monaten mit 119–221 Kindern pro Lebensmonat. Ausgangspunkt war eine Liste von 712 Wörtern, die auf Grundwortschatztabelle basierte. Zusätzlich bestand für die Eltern die Möglichkeit, fehlende Wörter, die ihre Kinder sprechen, zu ergänzen. Weiterhin fand eine freie Erhebung statt, in der die Eltern aufgefordert wurden, alle Wörter zu nennen, die ihre Kinder bereits sprechen. Für die nächste Iteration mit 406 Wörtern wurden aus diesem Gesamtwortschatz diejenigen Wörter entfernt, deren Häufigkeit unter 25 % lag, die regionale Varianten aufweisen oder die von der Jahreszeit oder von spezifischen Erfahrungen abhängig sind. In einem weiteren Filterschritt wurden unter anderem Wörter entfernt, die je nach Geschlecht eine deutlich unterschiedliche Häufigkeit aufwiesen. (Suchodoletz & Sachse, 2008a, 2012)

Aus der endgültigen Liste von 57 Wörtern wurden von Phoniater*innen und Logopäd*innen geeignete Testwörter identifiziert, die sich zum einen gut bildlich darstellen lassen und zum anderen mit hoher Wahrscheinlichkeit zum kindlichen Wortschatz gehören. Weitere Kriterien bei der Auswahl waren, dass nur geschlechtsneutrale Substantive infrage kommen und dass das Testmaterial möglichst phonetisch und phonematisch ausgewogen ist. Als Ergebnis wurden folgende Wörter ausgewählt: Affe, Apfel, Auge, Auto, Bär, Ball, Baum, Bett, Boot, Butter, Eier, Ente, Gurke, Haare, Hase, Hose, Licht, Milch, Mund, Nase, Ohr, Opa, Pferd, Tür, Uhr und Wasser. Damit stehen nun im neuen Test 26 statt bisher 10 Testitems zur Verfügung.

Obwohl der Wortschatz von Zweijährigen deutlich kleiner ist als der von Erwachsenen und 26 Wörter eine sehr kleine Stichprobe mit begrenzter Aussagekraft darstellen, wurde dennoch die Phonemverteilung analysiert, um eine Einschätzung der Übereinstimmung von der Phonemverteilung im Testmaterial mit der Phonemverteilung in der deutschen Alltagssprache zu erhalten. Zusätzlich zum neuen Test wurden der alte Mainzer I und der OLKI zu Vergleichszwecken in die Auswertung einbezogen. Die Ermittlung der in den Aufnahmen enthaltenen Phoneme erfolgte analog zu MATCH mittels *WebMAUS Basic* (s. Abschnitt 2.5). Da beim OLKI, im Gegensatz zum MATCH und Mainzer I bei einem Test nicht das gesamte zur Verfügung stehende Testmaterial zum Einsatz kommt, wurde eine typische Untersuchung mit 3 Testlisten postuliert und dafür die Listen 1–3 für die hier dargestellten Vergleiche ausgewählt.

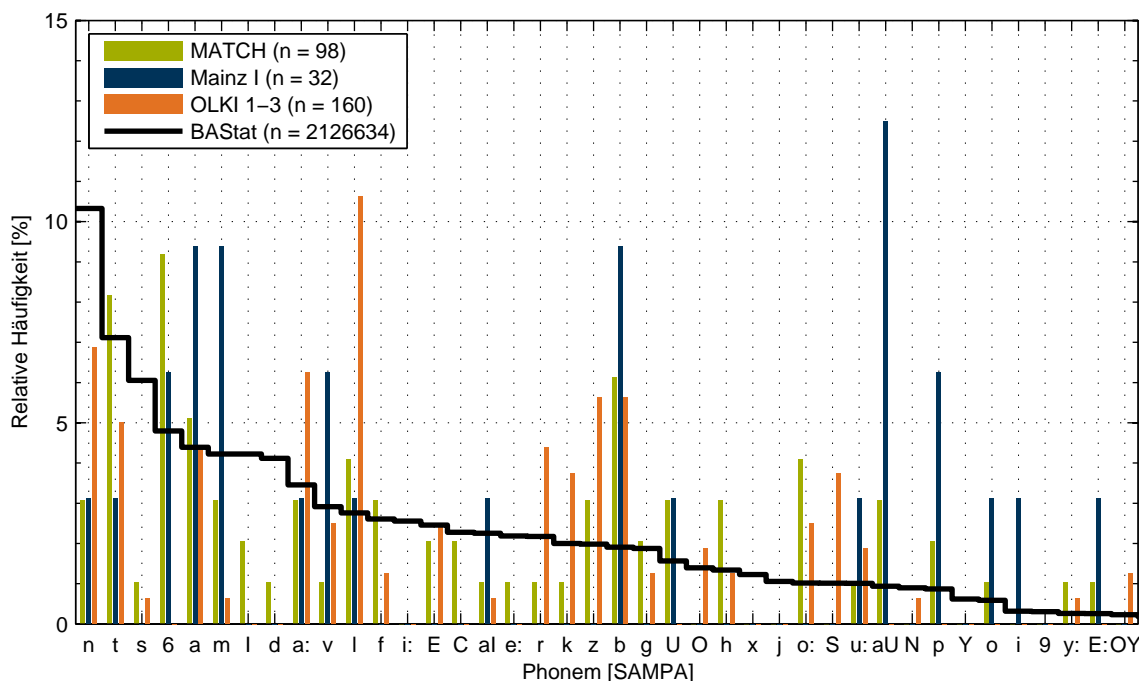


Abb. 4.8 Relative Häufigkeit der 40 häufigsten Phoneme der deutschen Spontansprache (BASat), des MATCH, des Mainzer Kindersprachtests I (Mainzer I) und des OLKI (Testlisten 1–3)

Die spontan gesprochene Sprache in zwischenmenschlichen Unterhaltungen kann sich deutlich von der Schriftsprache aus Literatur und Presse unterscheiden. Bei der Konstruktion eines Sprachtests wurde jedoch in den meisten Fällen das Testmaterial mit der Schriftsprache verglichen (Baumann, 2006; Chilla et al., 1976; Hahlbrock, 1953; Kliem & Kollmeier, 1995). Da aktuelle Daten zur Spontansprache des Deutschen vorliegen (Schiel, 2010), werden diese als Referenz für den Vergleich der verschiedenen Sprachtests herangezogen. Die unter der Bezeichnung *BASat* verfügbaren Daten liegen in einem Alphabet mit insgesamt 52 Zeichen vor (German SAM-PA, Speech Assessment Methods Phonetic Alphabet). Da die *BASat*-Stichprobe nicht alle 52 Phoneme enthält und die relative Häufigkeit einiger Phoneme sehr nahe bei 0 liegt, sind in Abb. 4.8 nur die 42 wichtigsten Phoneme für den Vergleich dargestellt.

Abb. 4.8 zeigt die Phonemverteilung spontaner Alltagssprache von Erwachsenen der *BASat*-Stichprobe (Schiel, 2010), des Mainzer Kindersprachtests I, des MATCH und des OLKI exemplarisch für einen typischen Testdurchlauf mit 3 Testlisten (hier: Listen 1–3). Zur Quantifizierung der Übereinstimmung mit der deutschen Spontansprache wurde die Summe der absoluten Differenzen zwischen den Phonemhäufigkeiten der *BASat*-Stichprobe und den Items des jeweiligen Tests herangezogen. Für MATCH beträgt die Summe der absoluten Differenzen 61,3%, für den Mainzer I 108,4% und für die Testlisten 1–3 des OLKI 77,0%, wobei eine kleinere Differenz eine höhere Übereinstimmung bedeutet. Der MATCH weist im Vergleich zu den beiden anderen Tests die geringste Abweichung zur Spontansprache und auch die wenigsten Lücken in der Verteilung auf.

Aufzeichnung und Verarbeitung des Testmaterials

Die Sprachaufnahmen sind der wichtigste Teil eines Sprachverständlichkeitstests. In Abschnitt 2.2.2 wurde bereits erwähnt, dass eine weibliche Stimme bei Sprachtests für Kleinkinder von Vorteil ist. Weitere Kriterien bei der Suche nach einer Sprecherin waren, dass sie eine eher jugendliche, angenehme Stimme mit fehlerfreier Artikulation ohne Dialektfärbung haben sollte. Für diese zentrale Aufgabe wurde eine Kooperation mit BAS Services Schiel (BASSS) eingegangen, einer hoch spezialisierten Einrichtung auf dem Gebiet der Sprachforschung und Sprachaufnahmen mit eigenem Tonstudio. BASSS ist eine Ausgründung aus dem Bayerischen Archiv für Sprachsignale (BAS) am Institut für Phonetik und Sprachverarbeitung der Ludwig-Maximilians-Universität München.

Das grundsätzliche Vorgehen bei den Aufnahmen war wie folgt:

Im Vorfeld: Vorbereitung und Eingabe der Wortlisten, damit bei der Aufnahme auf einem Bildschirm das jeweils aktuelle Item angezeigt wird. Dies verhindert Verrutschen in der Liste und Störgeräusch durch Papierrascheln.

Vor Ort: Vor Beginn der Aufnahme erfolgt ein Training mit der Sprecherin u.a. bzgl. Abstand zum Mikrofon und Prosodie. Anschließend wird die Mikrofonverstärkung angepasst und das Signal auf ca. -10dB Full Scale eingepegelt. Die Aufnahme erfolgt Wort für Wort mit Display-Prompting und auditiver und visueller Kontrolle während der Aufnahmesitzung. Fehlerhafte Aufnahmen werden dabei sofort wiederholt.

Im Anschluss: Schneiden ohne Pause vor oder nach dem Stimulus und jeweils 1 msec sinusförmiges fade-in und fade-out. Übergabe der fertigen Aufnahmen.

Die Aufnahmen erfolgten in einem reflexionsarmen Raum mit einer Nachhallzeit von weniger als 0,5 s. Zur Vermeidung von Pegelschwankungen bei minimalen Abstandsänderungen und Blasgeräuschen (z.b. beim Sprechen von /b/ oder /p/) kam ein extrem rauscharmes Großmembran-Kondensatormikrofon (Georg Neumann GmbH, Berlin) mit Blasschutz zum Einsatz. Der Abstand zwischen Mikrofon und Mund betrug 30 cm. Auf diese Art und Weise konnte sichergestellt werden, dass der Signal-Rausch-Abstand der finalen Aufnahmen mindestens 40 dB beträgt. Die Aufzeichnung erfolgte digital mit einer Abtastrate von 48 kHz und einer Signalquantisierung von 16 Bit. Die Sprecherin wurde gebeten, die Wörter so zu artikulieren und zu betonen, als ob sie in der Mitte eines Satzes stünden. Insgesamt wurden 4 Varianten aufgenommen: 3 Sprechgeschwindigkeiten (langsam, normal und schnell) und eine eindringlich, mit Nachdruck gesprochene Variante in normaler Geschwindigkeit. In Absprache mit den Kooperationsbeteiligten, insbesondere aus den Fachbereichen Logopädie, Phoniatrie und Pädaudiologie, wurde die Version mit normaler Sprechgeschwindigkeit und neutraler Betonung ausgewählt.

Im nächsten Verarbeitungsschritt wurde der Pegel jedes Wortes nach dem in ISO 8253-3 (2022) empfohlenen Verfahren gemessen (Parameter: Frequenzbewertung C, Zeitbewertung „slow“) und ggf. angepasst, sodass alle Wörter gleich laut waren. Aus dem auf diese Weise gewonnenen Testmaterial wurde das Langzeitspektrum nach dem von Byrne et al. (1994) be-

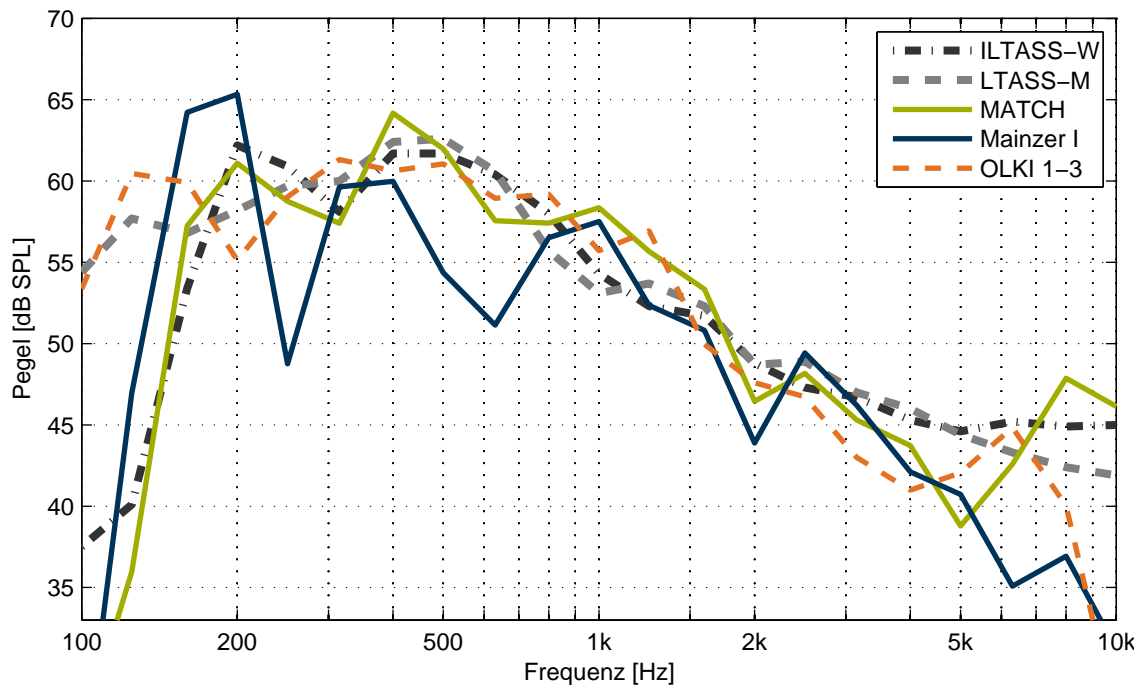


Abb. 4.9 Langzeit-Sprachspektren für Sprecherinnen (ILTASS-W), für Sprecher (ILTASS-M), den MATCH, den Mainzer Kindersprachtest I (Mainzer I) und den OLKI Testlisten 1–3 (OLKI 1–3), normiert auf einen Gesamtpegel von 70 dB SPL

schriebenen Verfahren berechnet. Das eigene Material und bereits vorhandene Tests wurden mit den *International (Universal) Long-Term Average Speech Spectra (ILTASS)* von Männern (ILTASS-M) oder Frauen (ILTASS-W) (Byrne et al., 1994) verglichen, je nachdem, ob das jeweilige Testmaterial von einem Mann oder einer Frau gesprochen wurde. Entsprechend Byrne et al. (1994) wurden alle Signale auf 70 dB SPL normiert. Abb. 4.9 zeigt die Langzeit-spektren für Frauen (ILTASS-W), Männer (ILTASS-M), MATCH, Mainzer I und OLKI Testlisten 1–3 (OLKI 1–3). Insgesamt zeigen die neueren Tests (OLKI, MATCH) eine wesentlich bessere Übereinstimmung mit dem jeweiligen ILTASS als die alte Aufnahme des Mainzer I. Diese zeigt mehr Schwankungen mit Abweichungen von teilweise mehr als 8 dB, insbesondere bei niedrigen (< 1 kHz) und hohen (> 5 kHz) Frequenzen. Demgegenüber weisen die neueren Aufnahmen bei nahezu allen Frequenzen zwischen 200 Hz und 10 kHz Abweichungen von höchstens 3 dB auf, mit Ausnahme von 1 kHz: + 4 dB und 5 kHz: - 5,5 dB beim MATCH und + 3,5 dB bei 800 Hz, - 4 dB bei 3 kHz, - 5 dB bei 4 kHz und - 15 dB bei 10 kHz beim OLKI.

Im Anschluss an die technischen Messungen und Verarbeitungsschritte wurden erste Messungen mit erwachsenen Testpersonen durchgeführt. Da zu erwarten war, dass sich die Verständlichkeit der Wörter bei Kindern von derjenigen bei Erwachsenen unterscheidet, war das Ziel der ersten subjektiven Messungen lediglich zu überprüfen, ob es deutliche Ausreißer in der Sprachverständlichkeit zwischen den Testitems gab und daraus ggf. notwendige Pegelkorrekturen zu ermitteln. Das Testkollektiv bestand aus 6 Erwachsenen im Alter zwischen 21 und 35 Jahren (Mittelwert 29,6 Jahre). Insgesamt wurden 144 Messungen mit allen 26 Testitems bei verschiedenen Darbietungspegeln durchgeführt, um die Verständlichkeit jedes

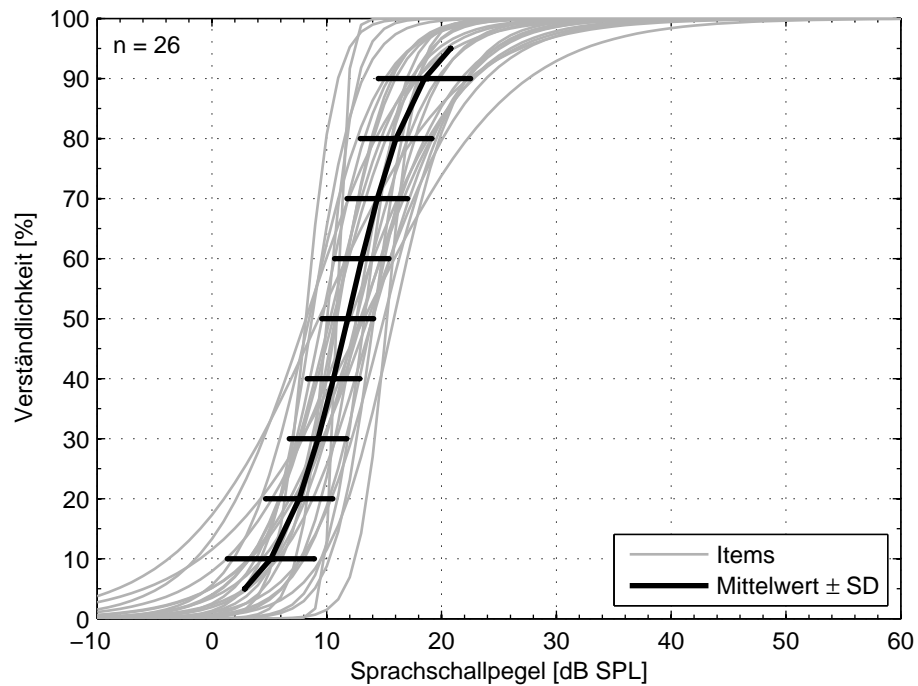


Abb. 4.10 Individuelle (dünne graue Linien) und mittlere (schwarz) Ergebnisse für die psychometrische Funktion der Sprachverständlichkeit als Funktion des Darbietungspegels bei normal hörenden Erwachsenen. Die Standardabweichung (SD) wird zwischen 10 und 90 % Verständlichkeit angezeigt (schwarze horizontale Linien)

einzelnen Wortes in Abhängigkeit des Pegels zu ermitteln. Die Ergebnisse sind in Abb. 4.10 dargestellt. Die Verständlichkeit unterscheidet sich zwischen den Testitems. Ausreißer waren jedoch keine vorhanden. Die mittlere SVS (50 % Verständlichkeit) im Sinne des Mittelwerts und der Standardabweichung (SD) der wortspezifischen Sprachverständlichkeitsfunktionen liegt bei $11,6 \pm 2,3$ dB SPL. Aus diesen Daten wurden wortspezifische Pegelkorrekturen zur Reduktion der Streuung abgeleitet. Wörter, die überdurchschnittlich gut verständlich waren – d.h. bei geringerem Pegel als der Durchschnitt richtig identifiziert wurden (eine Linksverschiebung der Verständlichkeitsfunktion auf der Pegelachse) – erhielten eine Absenkung des Pegels. Analog dazu wurden Wörter, die schlechter verständlich waren, im Pegel angehoben. Dabei wurde darauf geachtet, dass der Mittelwert aller Korrekturen möglichst 0 ist, um somit die durchschnittliche Verständlichkeit des Tests beizubehalten. Die einzelnen Korrekturwerte sind in Tabelle 4.1 angegeben. Ihr Mittelwert beträgt 0,01 dB. Positive Werte bedeuten eine Erhöhung, negative eine Absenkung des Pegels.

Tab. 4.1 Pegelkorrekturen auf Grundlage der Messungen

Wort	Affe	Apfel	Auge	Auto	Bär	Ball	Baum	Bett	Boot
dB	-1.4	-1.2	-3.0	-2.6	-3.0	-0.2	1.1	0.5	4.0
Wort	Butter	Eier	Ente	Gurke	Haare	Hase	Hose	Licht	Milch
dB	0.1	-3.0	-3.0	-0.2	-0.4	-3.0	2.5	2.2	2.2
Wort	Mund	Nase	Ohr	Opa	Pferd	Tür	Uhr	Wasser	
dB	1.1	2.2	3.0	-0.9	3.0	0.6	1.8	-2.1	



Abb. 4.11 MATCH Bildkarten: Darstellung aller 26 Wörter und der Zauberbox zum Start der Darbietung.

Anfertigung von Bildmaterial

Das Bildmaterial spielt für die Durchführung des Tests und die Motivation der Kinder eine wesentliche Rolle. Aus diesem Grund wurde bei der Entwicklung und Erstellung des Bildmaterials besonders sorgfältig vorgegangen. Der Stil sollte für alle Tests (MATCH, MAGIC (Abschnitt 4.1), und BASD (Abschnitt 4.3)) einheitlich sein. Es sollte ein moderner, einfacher und kindgerechter Stil zur Anwendung kommen, der nur die für die Erkennung wesentlichen Details enthält. Hierfür wurde schon von Beginn des Planungsprozesses an eng mit der Firma DAAL VISION (Prien a. Chiemsee) zusammengearbeitet. Bei der Entwicklung des Bildmaterials wurde iterativ vorgegangen und aus mehreren Vorschlägen die endgültigen Bilder ausgewählt. Als Beispiel ist der Prozess für das Bild des Ohres in Abb. 4.12 gezeigt. Aus verschiedenen Varianten (OhrF1, OhrSW4, etc.) entstand schließlich die finale Darstellung mit der Farbgebung und den abgerundeten Ecken. Ziel bei der Erstellung der Bilder war es, dass sowohl Erwachsene als auch Kinder der Zielgruppe im Alter von 2 bis 8 Jahren das abgebildete Wort sofort erkennen und benennen können. Nach dieser Validierung wurden die Bilder digital als Pixelgrafiken in verschiedenen Größen geliefert. Das Ergebnis dieser Teilarbeit inklusive der Zauberkiste zum Auslösen der Sprachdarbietung zeigt Abb. 4.11.

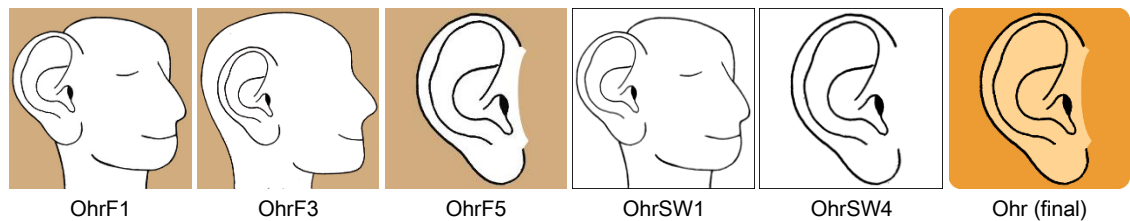


Abb. 4.12 MATCH Ohr-Abbildung: Beispiel für das iterative Vorgehen

4.2.3. Design und Implementierung

Testdesign

Der MATCH wurde mit der Anforderung entworfen, die auditive Funktion (Sprachverstehen) möglichst isoliert von anderen Faktoren der Entwicklung des Kindes zu untersuchen. Zu diesem Zweck wurde der „picture-pointing task“ als Grundlage gewählt, bei dem die Versuchsperson auf dem Touchscreen ein Bild auswählt, das das zuvor gehörte Wort repräsentiert. Dadurch wird die Sprachwahrnehmung des Kindes unabhängig von seiner Fähigkeit, korrekt zu sprechen, und unabhängig vom Hörvermögen des Untersuchenden gemessen. Die Verwendung eines Touchscreens ermöglicht eine intuitive Interaktion, die insbesondere für Kinder leicht zugänglich ist. Dieser Ansatz trägt dazu bei, eine zuverlässige Diagnose und eine angemessene Intervention für Kinder mit Hörstörungen zu gewährleisten. Die Aufgabe besteht darin, die Präsentation eines Wortes zu starten, indem man auf ein Feld mit dem Bild einer „Zauberkiste“ drückt und dann das präsentierte Element aus einer Auswahl von 4 Bildern auf dem Bildschirm auswählt. Da die Motivation bei jüngeren Kindern schnell nachlässt, wenn sie Wörter nicht hören und trotzdem durch Raten ein Wort auswählen müssen, besteht die Möglichkeit über einen 5. Knopf mit einem Fragezeichen („?“-Knopf) anzugeben, dass das Element nicht gehört oder nicht verstanden wurde. Es handelt sich folglich um ein

4-Alternativen-Verfahren ohne Zwangswahl. Das Design ohne Zwangswahl vermeidet Frustration in Fällen, in denen ein Wort nicht gehört oder nicht verstanden wurde. Es ist auch ausdrücklich vorgesehen, dass die Darbietung des Testitems bei Bedarf wiederholt werden kann, um eine sichere, von der Vigilanz unabhängige Eingabe zu gewährleisten. Die Wahl von 4 Alternativen wurde getroffen, um den Test interessanter zu gestalten, indem immer neue Bilder präsentiert werden. Außerdem kann die gleichzeitige Präsentation von zu vielen Bildern die Kinder überfordern und den Test unnötig in die Länge ziehen, wenn das Wort zwar verstanden, das Bild aber erst gesucht werden muss.

Abb. 4.13 zeigt die Bedienoberfläche, die auf dem Klinik-PC zur Kontrolle des Testverlaufs angezeigt wird. In der linken Bildhälfte befindet sich die große Zauberkiste und rechts ist eine beispielhafte Auswahl von 4 Bildern für die Eingabe dargestellt. In diesem Beispiel war *Baum* das Prüfitem. *Tür*, *Ente* und *Boot* wurden vom Gerät als Alternativen gewählt.

Analog zum Mainzer I wird nur das jeweilige Testitem präsentiert, ohne einen einleitenden Trägersatz wie „Zeig mir . . .“. Ein Trägersatz dient in der Regel dazu, die Aufmerksamkeit auf das folgende Testwort zu lenken. Dies ist vorteilhaft bei Tests, bei denen die Versuchsperson den Zeitpunkt der Darbietung nicht kennt. Bei dem hier vorgestellten selbstgesteuerten Test ist eine solche Lenkung der Aufmerksamkeit nicht notwendig, da die Testperson weiß, wann die Präsentation beginnt. Dadurch verkürzt sich die Gesamtdauer der Testdurchführung. Die Gestaltung des MATCH als spielerischer, interaktiver Selbsttest führt zudem zu einer erhöhten Motivation und längeren Aufmerksamkeitsspanne.

Vor jeder Testung muss gewährleistet sein, dass das Kind alle einbezogenen Wörter zumindest im rezeptiven Wortschatz zur Verfügung hat und die auf den Bildern dargestellten Gegenstände erkennt. Außerdem muss die Handlungsanweisung verstanden und umgesetzt werden können. Sie könnte lauten: „Drücke auf die Zauberkiste und höre genau zu was diesmal drinnen ist. Wähle nun was du gehört hast.“ Grundsätzlich ist ein überschwelliger, binauraler Test mit festem Pegel gut geeignet, um das Kind mit dem Tonmaterial und der Aufgabe vertraut zu machen.

Bei jedem Testdurchlauf werden immer alle Items präsentiert, sodass jederzeit der aktuelle Fortschritt der Messung und die noch zu erledigende Aufgabe ersichtlich sind. Jedes Bild erscheint insgesamt bis zu 4-mal: Einmal als Testitem und zusätzlich bis zu 3-mal als eine der 3 Alternativen. Die Reihenfolge der Wörter während eines Tests, die angebotenen Alternativen und die Positionen auf dem Bildschirm werden randomisiert und sind somit bei jedem Test unterschiedlich. Dies verhindert die Beeinflussung der Messergebnisse durch das Auswendiglernen von Testlisten und macht den Test trotz des relativ kleinen Vokabulars beliebig oft wiederholbar. Es wurden zwei Testmodi vorgesehen, die beide in der Praxis häufig zur Anwendung kommen:

1. Ermittlung der Sprachverständlichkeit mit einem festen Darbietungspegel
2. Adaptive Bestimmung der Sprachverständlichkeitsschwelle (SVS)

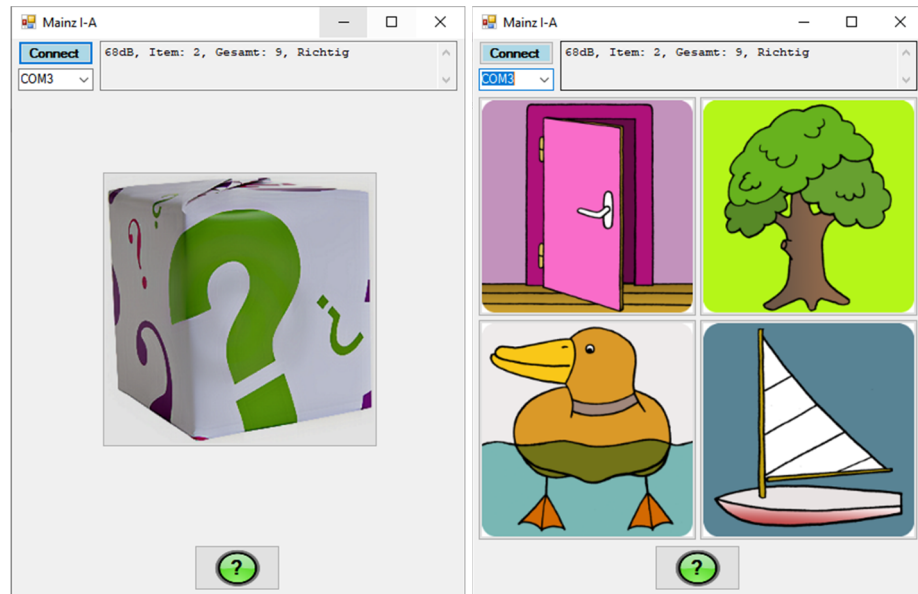


Abb. 4.13 MATCH PC Darstellung. Erste Version zur Darstellungen des Messverlaufs auf einem separaten Bildschirm des PCs für Untersuchende

Bei der adaptiven Messung der SVS wird der Pegel bei richtiger Eingabe um 4 dB abgesenkt, während er bei falscher Eingabe um 10 dB erhöht wird. Zu Beginn des Tests wird der Pegel in Schritten von 10 dB ausgehend vom eingestellten Startpegel bis zur ersten Umkehrung (Falscheingabe oder „?“) abgesenkt. Die SVS ist der Mittelwert der Prüfpegel und demjenigen Pegel, der sich aus der Antwort auf das letzte Item ergibt, unter Ausschluss der ersten 5. Bei Anwendung dieses gewichteten adaptiven Up-Down-Verfahrens entspricht der Schwellenwert einer Sprachverständlichkeit von etwa 71,4 % (Kaernbach, 2001). Zur Einschätzung der Zuverlässigkeit wird neben dem Mittelwert auch die Standardabweichung berechnet. Die Verwendung einer Sprachverständlichkeitsschwelle von 71,4 % hat den Vorteil, dass während des Tests eine größere Anzahl an Testelementen wahrgenommen werden kann. Dies ist insbesondere für jüngere Kinder von Vorteil, um ihre Motivation aufrechtzuerhalten.

Die Sprachdiskriminationsfunktion (psychometrische Funktion) wird als logistische Funktion mit den Parametern m (Steigung), L (Darbietungspegel) und SVS (50 %-Sprachverständlichkeitsschwelle) nach Gleichung (4.1) modelliert. Diese Gleichung ist die Basis der Darstellungen in Abb. 4.10, 5.8, 5.10, sowie für die Ergebnisanzeige auf dem Gerät, die im nächsten Unterabschnitt beschrieben wird.

$$f(L) = \frac{1}{1 + e^{4m(SVS-L)}} \quad (4.1)$$

- L Darbietungspegel [dB]
- m Steigung im Wendepunkt (50 % Sprachverständlichkeit) [% / dB]
- SVS 50 % Sprachverständlichkeitsschwelle [dB]

Implementierung des Test

Die Implementierung erfolgte auf der in Kapitel 3 beschriebenen Hardwareplattform in C unter Verwendung einer auf *Eclipse* (Eclipse Foundation, Brüssel, Belgien) basierenden Softwareentwicklungsumgebung für eingebettete Systeme. Bei eingebetteten Systemen ist ein sparsamer Umgang mit Speicher und Rechenleistung erforderlich, insbesondere wenn es sich um batteriebetriebene Systeme handelt. Dies hat Auswirkungen auf die Implementierung des Tests. So sind z.B. alle Bildschirmdarstellungen derart zu gestalten, dass sie bei einer Auflösung von 240 x 320 Punkten gut erkennbar und lesbar sind. Dabei ist immer ein Kompromiss zwischen Informationsfülle und Bedienbarkeit zu finden, denn einerseits soll die Bedienoberfläche nur aus wenigen Seiten bestehen, um das Gerät schnell und effektiv bedienen zu können, andererseits sollen die Schaltflächen möglichst groß sein, um eine flüssige Bedienung mit dem Finger zu ermöglichen. Beim Speichern von Testergebnissen ist sorgfältig abzuwägen zwischen einer möglichst großen Anzahl an Tests, die gespeichert werden können, und vielen Details jedes einzelnen Tests zum Messverlauf und Messergebnis. Die Prozessorlast wird beispielsweise davon beeinflusst, mit welcher Frequenz der Bildschirm aktualisiert und neue Inhalte zeigt oder wie die Filter zur Freifeldentzerrung der Wandler ausgelegt sind.

In Abb. 4.14 sind Screenshots dargestellt, die den Messverlauf schematisch repräsentieren. Die Tonpräsentation wird mit der Zauberkiste gestartet (Abb. 4.14a). Der Bildschirminhalt ändert sich entsprechend Abb. 4.14b und zeigt die 4 Bilder sowie den „?“-Knopf. Die Eingabe ist bis kurz vor Ende des Stimulus zur Vermeidung von versehentlichen vorschnellen Eingaben gesperrt, was durch die zunächst graue Hintergrundfarbe des „?“-Knopfs visualisiert wird. Sobald die Eingabe erlaubt ist, wird der Hintergrund grün. Nachdem alle Testitems abgearbeitet sind, erscheint eine Animation als Belohnung (Abb. 4.14c). Der Status der Messung kann jederzeit aus der Kopfzeile abgelesen werden. Von links, mit dem aktuellen Pegel beginnend,



Abb. 4.14 MATCH Testverlauf: (a) Zauberkiste und (b) Eingabe des verstandenen Testitems oder „?“ für nicht gehört/verstanden. (c) Nach der Messung erscheint eine Animation als Belohnung



Abb. 4.15 MATCH Testeinstellungen. (a) Messmodus und Pegelinstellungen, (b) Auswahl der Sprache, Option zum Ausblenden der Statusinformationen in der Kopfzeile, (c) Auswahl der Wörter, die getestet werden

zeigt sie außerdem, welches Wort gerade zu hören ist, das wievielt Wort es ist und ob die vorangegangene Eingabe richtig oder falsch war, symbolisiert durch ein grünes oder rotes Quadrat am Bildschirm oben rechts. Die Nummerierung für die Anzeige des aktuell geprüften Wortes ist in Abb. 4.14b angegeben. Mittels Antworttaster stehen für Untersuchende zwei zusätzliche Funktionen bereit: Wiederholung der Testitem-Darbietung und Rücknahme einer versehentlichen oder vom Kind als falsch gemeldeten Eingabe.

Vor Beginn einer Messung können verschiedene Optionen und Parameter gewählt werden. Abb. 4.15 zeigt die entsprechenden Screenshots. Messmodus (adaptive Pegelsteuerung oder ein fester Pegel) und zugehörige Pegel werden auf der 1. Seite eingestellt (Abb. 4.15a). Die Implementierung des Tests wurde derart durchgeführt, dass der Test in verschiedenen Sprachen mit den gleichen Bildern und Messalgorithmen verwendbar ist (Abb. 4.15b). Ebenso wurde Rauschen vorgesehen für die Ermittlung der Sprachverständlichkeit im Störgeräusch. Für den Fall, dass die Versuchsperson die Information der Kopfzeile entschlüsseln kann, besteht die Möglichkeit sie auszublenden. Sollte es einzelne Begriffe geben, die im Vokabular des Kindes fehlen und ihre Vermittlung während der Eingewöhnungsphase nicht erfolgreich sein, können einzelne Wörter des Testmaterials weggelassen werden. Dies kann ebenfalls in den Einstellungen vorgenommen werden (Abb. 4.15c).

Während der Messung wird eine Reihe von Daten erhoben, die zur Berechnung des Ergebnisses und zur Einschätzung der Zuverlässigkeit dienen. Im Einzelnen sind das: die Prüfitems in der Reihenfolge des Tests, die als Alternativen angezeigten Bilder, das ausgewählte Bild, der zu jedem Item dazugehörige Darbietungspegel, die Reaktionszeiten vom Ende des Stimulus bis zur Eingabe, die Testdauer, der Testmodus (adaptiv oder mit festem Pegel), die Testsprache, die Anzahl der geprüften und der richtig verstandenen Wörter, die Schwelle mit

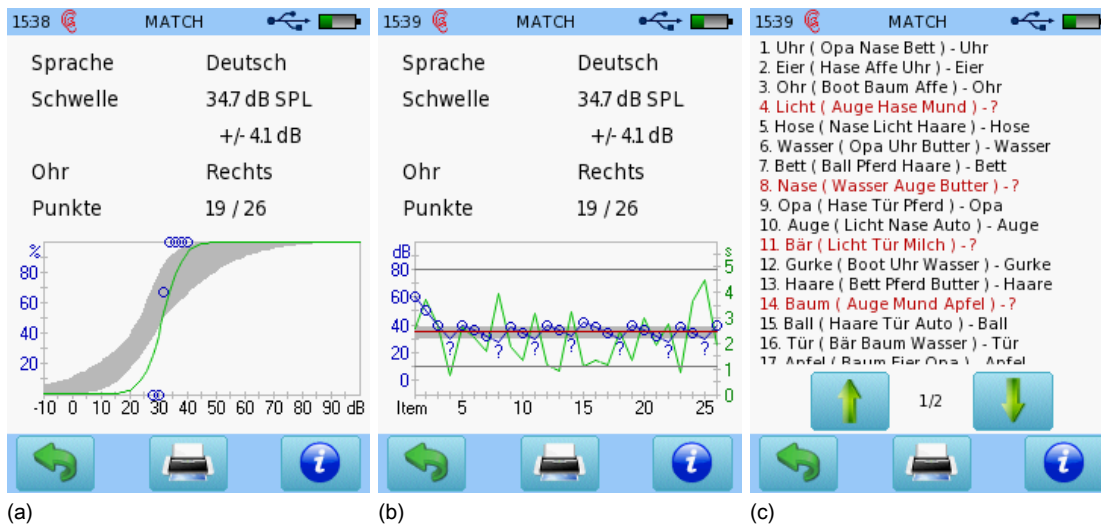


Abb. 4.16 MATCH Ergebnisdarstellung adaptive Messung. (a) Schwelle und Testohr, grafische Darstellung des Normbereichs und des aktuellen Testergebnisses als Streudiagramm mit geschätzter Diskriminationsfunktion, (b) Darstellung des Testverlaufs und Reaktionszeiten, (c) Listenansicht des ganzen Tests

der dazugehörigen Standardabweichung und die als Minimum und Maximum ausgewählten Pegel bei adaptiven Messungen.

Abb. 4.16 und 4.17 zeigen die Darstellung der gespeicherten Daten am Beispiel eines adaptiven Tests und einer Messung mit festem Pegel. In der oberen Hälfte in Abb. 4.16a und (b) sind die wichtigsten Daten zusammengefasst: Schwelle und Standardabweichung, Sprache, Ohr, die Anzahl der richtig verstandenen Wörter und die Gesamtzahl der Testitems. Darüber hinaus werden alle geprüften Wörter in der Reihenfolge des Tests, die Alternativen und die Eingabe des Kindes angezeigt (Abb. 4.16c). Bei Messungen mit festem Darbietungspegel wird zusätzlich die erzielte Sprachverständlichkeit berechnet (Abb. 4.17a). Die Grafik in Abb. 4.16a zeigt den Normbereich (graue Fläche) für die Sprachverständlichkeit in Abhängigkeit des Darbietungspegels für 3–7 Jährige, der in der Studie aus Abschnitt 5.2 ermittelt wurde. Blau werden die Werte des aktuellen Tests als Streudiagramm dargestellt, aus denen die grün dargestellte psychometrische Funktion geschätzt wird.

Die Darstellung in Abb. 4.16b visualisiert mittels der grünen Kurve, deren y-Achse sich rechts befindet, die Reaktionszeit in Sekunden zu jedem einzelnen Testitem. Die Reaktionszeit kann Hinweise zur Höranstrengung geben, da einfache Entscheidungen (Wort deutlich gehört und verstanden) schneller getroffen und eingegeben werden als schwierige, nahe der Verständlichkeitsschwelle, die durch den Mittelwert (rote horizontale Linie) und die Standardabweichung (grauer Bereich um die rote Linie) repräsentiert wird. Die beiden grauen horizontalen Linien (in Abb. 4.16b bei 10 dB und 80 dB) stellen die eingestellten Grenzen des adaptiven Algorithmus dar. Sollte die blaue Kurve zu oft an die Grenzen stoßen, kann das auf Einstellungen hindeuten, die nicht zum Hörverlust des untersuchten Kindes passen, oder auf ein generelles Verständnis- oder Kooperationsproblem. Die blaue Kurve, deren y-Achse sich links befindet, zeigt den Darbietungspegel für jedes Testitem und die Bewertung der Eingabe

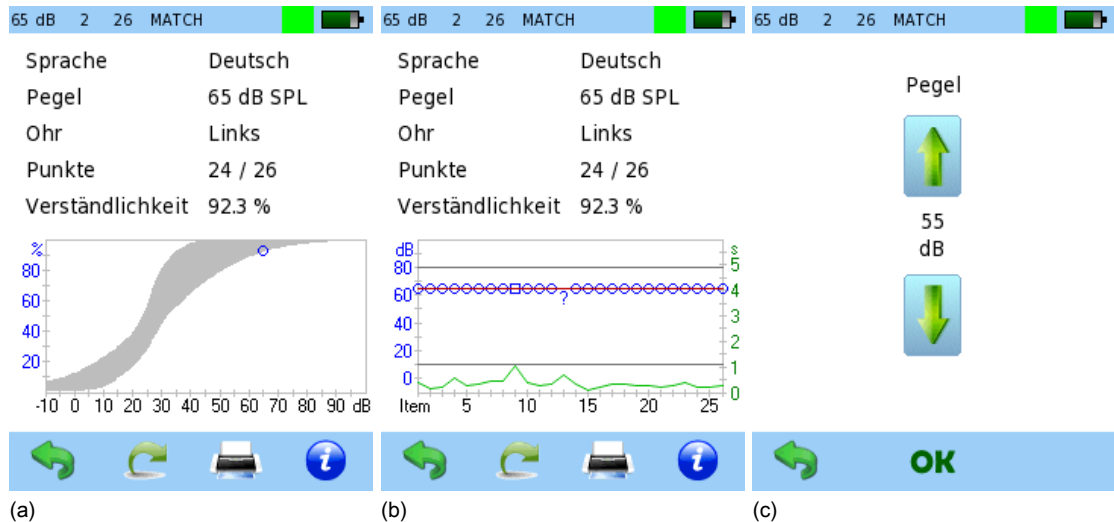


Abb. 4.17 MATCH Ergebnisdarstellung fester Messpegel. (a) Schwelle und Testohr, grafische Darstellung des Normbereichs und der erzielten Verständlichkeit, (b) Darstellung des Testverlaufs und Reaktionszeiten, (c) Auswahl des Prüfpegels für die nächste Messung am aktuellen Testohr.

als richtig (Kreis), falsch (Quadrat) oder „weiß nicht“ (?). Da es im Falle von Messungen mit festem Prüfpegel nur einen einzigen Pegel gibt, reduziert sich das Streudiagramm der Grafik in Abb. 4.17a zu einem einzigen blauen Kreis. In Abb. 4.17b ist der Mittelwert gleich dem Prüfpegel und die Standardabweichung 0.

Die in Abb. 4.17 gezeigte Variante der Ergebnisdarstellung ist diejenige, die direkt nach der Testdurchführung erscheint. Die Eingaben zum letzten Testitem sind noch in der Kopfzeile sichtbar und in der Fußzeile befindet sich an der 2. Position von links ein Knopf, der direkt zu einem Retest des gleichen Ohres führt und damit die Bedienung des Gerätes vereinfacht. Bei Messungen mit festem Pegel besteht zudem die Möglichkeit, vor der erneuten Testung den Prüfpegel zu verändern (Abb. 4.17c). Im Gerät ist die in der Praxis gängige Pegelfolge 65–55–85 dB hinterlegt, wobei als Voraussetzung zur Reduktion des Pegels von 65 dB auf 55 dB eine Sprachverständlichkeit von mindestens 50 % bei 65 dB erreicht werden muss.

Auszüge aus Abschnitt 4.2 wurden vorab auf der 17. Jahrestagung der Deutschen Gesellschaft für Audiologie präsentiert und veröffentlicht als: Schirkonyer et al. (2020): Der neue Mainzer Sprachtest für Kinder 3–7 Jahre (MATCH). HNO, 68(7), 526–533.

4.3. Test zur zentralen Verarbeitung (BASD)

4.3.1. Einleitung

Die zentral auditorische Verarbeitung ist sehr vielfältig und komplex, sodass ihre Funktion nicht durch einen einzelnen Test zuverlässig bestimmbar ist. Insbesondere wenn Aussagen zur Sprachentwicklung oder zur auditiven Verarbeitung und Wahrnehmung gemacht werden sollen, kommen in der Regel Testbatterien zum Einsatz, deren Untertests jeweils eine Teilfunktion untersuchen (DGPP, 2019). Die empfohlene Testbatterie zur Diagnose von AVWS umfasst Elemente der zentralen Verarbeitung, die das Sprachverstehen und damit die Sprachentwicklung beeinflussen: die Fähigkeit Tonhöhen, Lautstärken und Signaldauern zu unterscheiden (DGPP, 2019). Darüber hinaus beinhaltet sie unter anderem dichotische Tests mit unterschiedlichen Signalen auf beiden Ohren, binaurale Integrationstests, Phonemdiskrimination und -identifikation sowie Tests zum phonologischen Kurzzeitgedächtnis (DGPP, 2019). Die Durchführung aller diagnostischen Testverfahren der empfohlenen Testbatterie ist sehr zeitaufwendig. Daher ist es wünschenswert, in einem ersten Schritt mithilfe eines Screeningtests nur diejenigen Verdachtsfälle zu identifizieren, bei denen eine weitere Abklärung erforderlich ist.

Die einzelnen Arbeitsschritte zur Veröffentlichung des BASD waren:

1. Entwicklung des zentral auditorischen Tests
 - a. Zusammenstellung der Testbatterie
 - b. Entwicklung eines kindgerechten Konzepts und Verfahrens
 - c. Erarbeitung eines Algorithmus zur Schwellenbestimmung, der sowohl bei Kindern als auch bei Erwachsenen anwendbar ist
 - d. Implementierung des Tests auf einem geeigneten Messsystem
2. Validierung des zentral auditorischen Tests
 - a. Rekrutierung von Kindern und Erwachsenen als Versuchspersonen
 - b. Durchführung des Tests
 - c. Statistische Auswertung und Ergebnisdokumentation

4.3.2. Testdesign

Da ein einzelner Test lediglich einen begrenzten Ausschnitt der zentralen Verarbeitung abdeckt, wurde der BASD als Testbatterie konzipiert, bestehend aus 6 Subtests, von denen jeder 4 Arten der Signalpräsentation umfasst. Damit soll eine möglichst gute Übereinstimmung mit den einleitend genannten Empfehlungen erreicht werden, wobei gleichzeitig die Möglichkeit einer einfachen und schnellen Durchführung im Vordergrund steht. In Abb. 4.18 ist die Signalpräsentation dargestellt, die grundsätzlich aus 3 Tönen mit jeweils einer Pause dazwischen besteht. Dabei sind immer 2 Töne identisch und 1 Ton unterscheidet sich in der untersuchten Eigenschaft. Die Positionen der Referenztöne und des Testsignals wird vor jeder Darbietung eines Tripletts randomisiert. Durch die Verwendung eines Intervalls mit 3

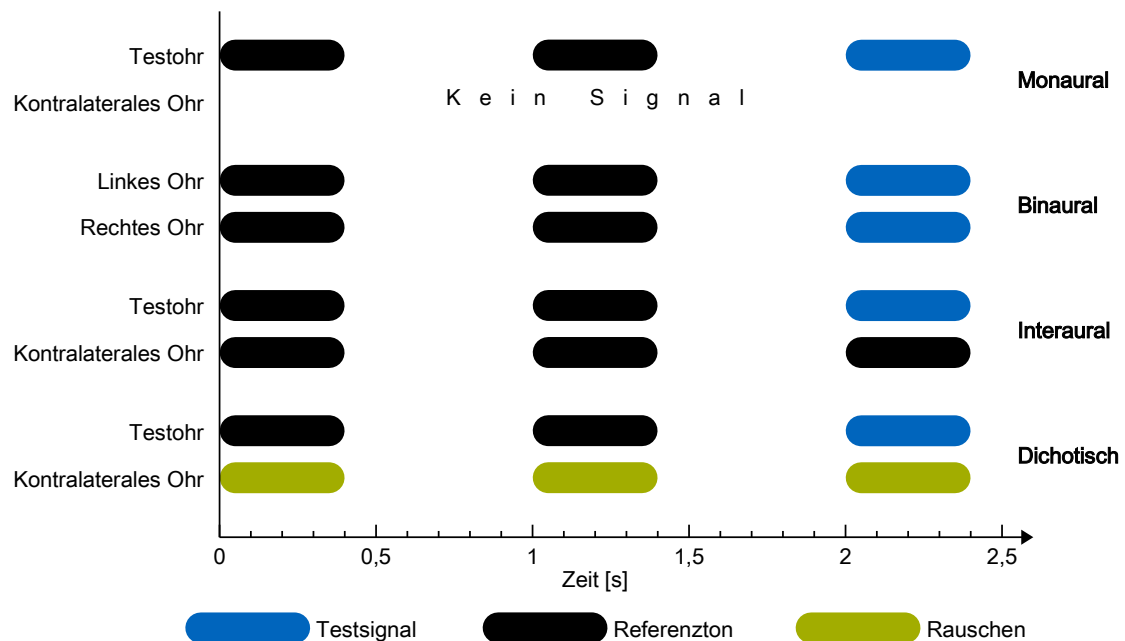


Abb. 4.18 BASD Signalpräsentation: Das Testohr erhält immer eine Sequenz aus Testsignal und 2 Referenztönen. Die „Ein“/„Aus“-Zeiten betragen 400 ms bzw. 600 ms. Das kontralaterale Ohr erhält: nichts (monaural), das Testohr-Signal (binaural), 3x den Referenzton (interaural), 3x Rauschen (dichotisch)

Signalen wird zusätzlich das Arbeitsgedächtnis geprüft, da sich die Versuchspersonen alle Signale und deren Reihenfolge merken müssen, um sie miteinander zu vergleichen. Damit kein Knacksen durch das Ein- und Ausschalten entsteht, werden die Signale über eine Dauer von 60 ms cosinusförmig ein- und ausgeblendet. Die „Ein“/„Aus“-Zeiten gemäß der Definition für gepulste Töne der Audiometer-Norm (DIN 60645-1, 2018), betragen 400 ms bzw. 600 ms.

Die 4 Möglichkeiten der Signalpräsentation aus Abb. 4.18 sind wie folgt:

Monaural Bei der monauralen Präsentation wird das Stimulstriplett aus 2 Referenztönen und dem Testsignal ausschließlich auf dem Testohr dargeboten. Das kontralaterale Ohr erhält kein Signal.

Binaural Bei der binauralen Präsentation wird dasselbe Stimulstriplett auf beiden Ohren dargeboten. Sofern kein asymmetrischer Hörverlust vorliegt, wird das Signal in der Mitte des Kopfes und bei gleichem Darbietungspegel lauter als monaural wahrgenommen. Bei asymmetrischem Hörverlust scheint das Signal auf der Seite mit dem besseren Hörvermögen zu sein.

Interaural Bei der interauralen Präsentation wird auf dem Testohr eine Signalfolge wie im monauralen Fall abgespielt, während auf dem anderen Ohr 3-mal der Referenzton erklingt. Aufgrund der zentralen Fusion der Signale von beiden Ohren entsteht ein Raumeindruck, als ob das Signal im Raum wandern würde.

Dichotisch Bei der dichotischen Präsentation wird auf dem Testohr eine Signalfolge wie im monauralen Fall abgespielt, während auf dem anderen Ohr 3-mal Breitbandrauschen dargeboten wird.

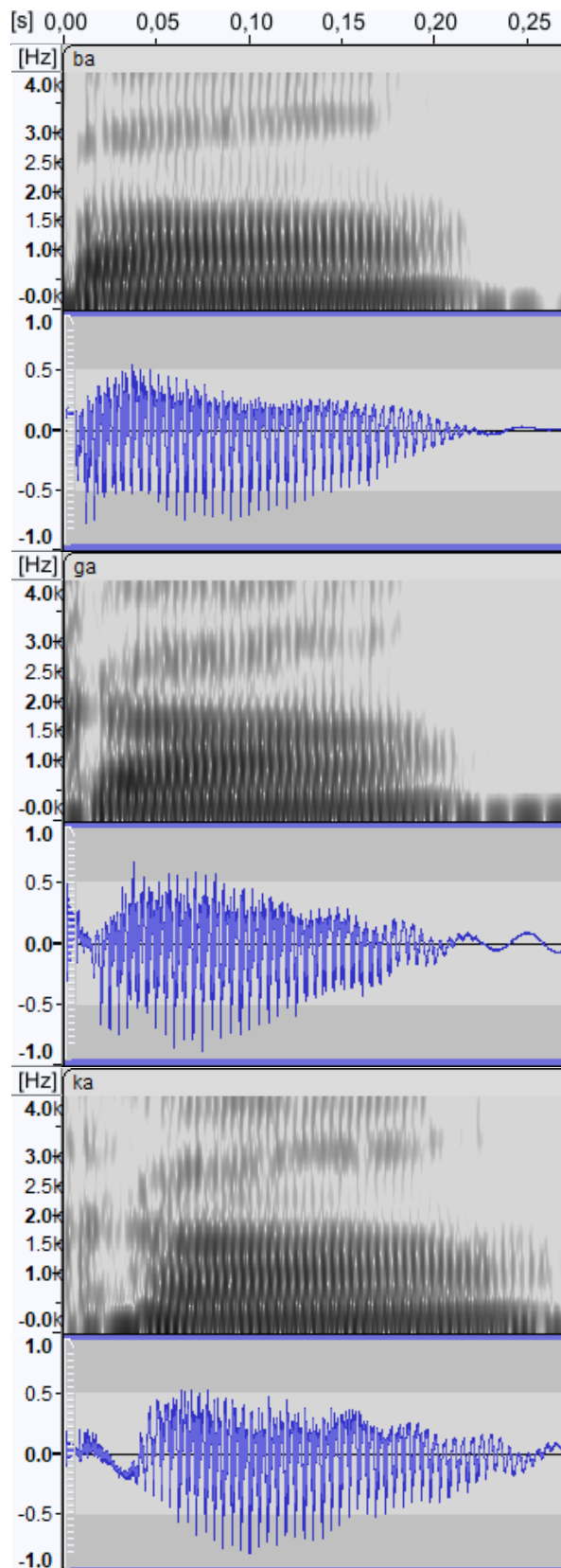


Abb. 4.19 Spektrogramm (grau) und Zeitsignal (blau) der Sprachstimuli. Von oben nach unten: /ba/, /ga/ und /ka/. /ba/ und /ga/ unterscheiden sich im Artikulationsort (bilabial und velar). Bei /ba/ sind die Formanten F1 bis F3 fallend zum Plosiv hin, während F2 und F3 bei /ga/ zusammenlaufen. /ga/ und /ka/ unterscheiden sich in der Stimmhaftigkeit. Insbesondere die längere Verschlussphase am Anfang ist gut sichtbar.

Von den 6 Untertests des BASD sind 4 sprachfrei. Bei diesen Untertests wird eine Messung der Unterschiedsschwellen (just noticeable differences, JND) für Frequenz, Pegel, Tondauer und Amplitudenmodulationsfrequenz durchgeführt. Die Aufgabe besteht darin, ein Testsignal von einem stets gleichbleibenden Referenzton zu unterscheiden. Die Änderungen erfolgen dabei immer in Richtung einer Vergrößerung – höhere Tonfrequenz, höherer Schalldruckpegel, höhere Modulationsfrequenz – mit Ausnahme bei der Messung der JND für die Tondauer. Hierbei wird ein Signal verwendet, das stets kürzer als der Referenzton ist. Die Frequenz des Referenztons ist einstellbar und beträgt wahlweise 500 Hz oder 1 kHz.

In 2 Subtests werden die 3 Silben /ba/, /ga/ und /ka/ als Stimuli verwendet. Mit ihrer Hilfe wird die Fähigkeit zur Diskrimination und Identifikation von Phonemen überprüft. Dabei kommt die Aufnahme einer Frauenstimme mit einer Grundfrequenz von ca. 200 Hz zum Einsatz. Die Stimulusdauer beträgt jeweils ca. 275 ms. Die Aufgabe für die Versuchspersonen besteht darin, Plosiv-/a/-Paare zu vergleichen, die sich durch die Formanttransitionen (/ba/ und /ga/) oder durch die Stimmhaftigkeit (/ga/ und /ka/) unterscheiden. Das Ergebnis dieser Subtests ist keine Unterschiedsschwelle, sondern die Anzahl, wie oft die Aufgabe richtig gelöst wurde. Die Bildung eines Plosivs geschieht in 3 Phasen: Verschlussbildung (1), Verschlussphase (2) und Verschlusslösung (3) (Machelett, 1996). In der 1. Phase bildet das Artikulationsorgan (Lippen, Zungenspitze, Zungenrücken) einen vollständigen Verschluss mit der Artikulationsstelle (labial, alveolar, velar) (Machelett, 1996). Während dieser Zustand eine Zeit lang aufrechterhalten wird, staut sich der Ausatemluftstrom hinter dem Verschluss (Phase 2) bis der Verschluss in der 3. Phase unter explosionsartiger Geräuschentwicklung aktiv gelöst wird (Machelett, 1996). In Abb. 4.19 sind das Spektrogramm (grau) und das Zeitsignal (blau) für jeden der 3 Sprachstimuli dargestellt. In allen Spektrogrammen sind die ersten 3 Formanten des Vokals /a/ – nach Peterson und Barney (1952) für Frauenstimmen 850 Hz (F1), 1220 Hz (F2) und 2810 Hz (F3) – gut erkennbar. Das /b/ ist ein bilabialer stimmhafter Plosiv. Kennzeichen dieser Laute ist, dass F1 bis F3 zum Plosiv hin fallen (negative Transitionen) (Machelett, 1996). Beim /g/ handelt es sich um einen stimmhaften velaren Plosiv. Kennzeichen velarer Plosive ist das Zusammenlaufen von F2 und F3, wenn ein Vokal folgt (außer bei /o/ und /u/) (Machelett, 1996). Das /k/ ist ein stimmloser velarer Plosiv. Die Verschlussphase ist bei stimmlosen deutlich länger als bei stimmhaften Plosiven.

Bei der Entwicklung des Konzepts und des Verfahrens wurde darauf geachtet, dass der BASD als Selbsttest für Kinder ab 4 Jahren zuverlässig durchführbar ist. Die Entwicklung erfolgte in mehreren Schritten. Es wurde zunächst mit einem einfachen Ja/Nein-Verfahren begonnen, bei dem ein Referenzton und das Testsignal dargeboten wurden mit der Fragestellung, ob beide Signale gleich sind. Zur Überprüfung der Zuverlässigkeit der Antworten wurden Intervalle eingefügt, in denen anstelle des Testsignals ein 2. Mal der Referenzton dargeboten wurde. Die hohe Ratewahrscheinlichkeit sowie die Schwierigkeit das Konzept gleich oder verschieden zu vermitteln und mit bestimmten Schaltflächen einer grafischen Oberfläche zu verknüpfen, führten dazu, dass dieser Weg nicht weiterverfolgt wurde. Die Verwendung eines 3-Intervall-Zwangswahlverfahrens senkt die Ratewahrscheinlichkeit auf $\frac{1}{3}$. Die Aufgabe be-

steht darin, das Testsignal zu finden, das sich von den beiden Referenztönen unterscheidet. Vorversuche haben gezeigt, dass 4-Jährige nahe der Schwelle und darunter Schwierigkeiten haben, sich für einen Ton zu entscheiden und schnell die Mitarbeit einstellen, wenn sie aufgefordert werden, in so einem Fall einfach zu raten. Um diesem Umstand Rechnung zu tragen, wurde eine weitere Änderung vorgenommen: Anstatt ein Testsignal zu benennen, das sich von den anderen unterscheidet, sollen nach Art des Memoryspiels 2 identische Signale (Referenzsignale) angegeben werden. Es zeigte sich, dass es den Kindern in der Nähe der Schwelle wesentlich leichter fällt 2 gleiche Items zu wählen, wenn alle 3 subjektiv gleich klingen. Zur Steigerung der Motivation und Aufmerksamkeit wurden die Bilder des MATCH (Abschnitt 4.2.2) wiederverwendet. Während der Stimulusdarbietung sind nur die Rückseiten der Karten sichtbar. Zur Darstellung der Zuordnung eines Stimulus zu einer Eingabeschaltfläche wird die zum momentan dargebotenen Stimulus gehörende Karte farblich hervorgehoben. Die Reizwiedergabe wird von der Versuchsperson selbst gestartet und kann zur Absicherung der Entscheidung wiederholt werden.

Der Algorithmus zur Schwellenbestimmung muss mehrere Anforderungen erfüllen: Es sollen nur wenige Versuche notwendig sein. Es soll möglich sein, mit einer großen, deutlich überschwelligen Startdifferenz zu beginnen (z.B. 200 Hz beim Frequenzdiskriminationstest) und anschließend eine kleine Schwelle zu messen (z.B. unter 1 Hz beim Frequenzdiskriminationstest). Raten soll das Ergebnis möglichst wenig beeinflussen. Wie in Abschnitt 2.3.1 beschrieben, sind direkte Verfahren zeitsparend, sodass als Basis die Methode des Trackings ausgewählt wurde. Bei diesem Verfahren ist die verwendete Schrittweite die entscheidende Stellgröße. Um deren Einfluss genauer zu untersuchen, wurden Simulationen mit unterschiedlichen festen und variablen Schrittweiten durchgeführt. Eine Simulation erzeugte jeweils 10000 Werte für jede Schwelle. Zur Vereinfachung wurde die psychometrische Funktion ($f(x)$, s. auch Abb. 2.5) durch 3 Geraden nach Gleichung (4.2) approximiert.

$$f(x) = \begin{cases} \frac{1}{3} & \text{für } x \leq \text{unteres Limit} \\ m \cdot x + c & \text{für } \text{unteres Limit} < x < \text{oberes Limit} \\ 1 & \text{für } x \geq \text{oberes Limit} \end{cases} \quad (4.2)$$

$$\text{unteres Limit} = 0,9967 \cdot \text{wahre Schwelle}$$

$$\text{oberes Limit} = 1,01 \cdot \text{wahre Schwelle}$$

$$m = \frac{1}{0,02 \cdot \text{wahre Schwelle}}$$

$$c = -49,5$$

Dabei nimmt die psychometrische Funktion an der Stelle *wahre Schwelle* den Wert 0,5 an. Zwischen der Ratewahrscheinlichkeit von $\frac{1}{3}$ und der stets richtigen Eingabe ($f(x) = 1$) verläuft die psychometrische Funktion im verwendeten Modell sehr steil mit einer Steigung von 50 % pro Prozent Änderung der Reizgröße. Als Startwert für die Simulation wurde der Wert 200 (ohne spezifische Einheit) gewählt. Der zu messende Schwellwert lag zwischen 1 und

100, um den Fall abzudecken, dass bei realen Messungen deutlich überschwellig begonnen werden kann. Die maximale Versuchszahl zur Ermittlung einer Schwelle wurde auf 36 begrenzt. Zur Berechnung der Schwelle wurden 4 Umkehrpunkte (je 2 Minima und Maxima) gefordert. Die Methode des Trackings wurde geändert, sodass ab dem 4. Versuch 2 richtige Eingaben notwendig sind, um den Messwert zu reduzieren. Dadurch wird der Einfluss von Raten auf das Messergebnis verringert.

Die Ergebnisse der Simulation sind in Abb. 4.20 dargestellt. Dabei stellt die grau gestrichelt gezeichnete Winkelhalbierende das ideale Ergebnis dar, für den Fall, dass die gemessene Schwelle identisch mit der wahren Schwelle ist. Bei der Verwendung von festen Schrittweiten (blau: 25, orange: 10) bleibt der Wert der gemessenen Schwelle für mehrere Werte der wahren Schwelle gleich und es können große Sprünge in der gemessenen Schwelle bei kleinen Änderungen der wahren Schwelle auftreten. Außerdem ist die Schwelle mit einer Schrittweite von 10 unter den gegebenen Rahmenbedingungen nicht im gesamten Bereich messbar. Unter 40 konnte bei jeweils 10000 simulierten Messungen keine einzige Schwelle ermittelt werden. Bei der Verwendung einer variablen Schrittweite, die im Verlauf der Messung verkleinert

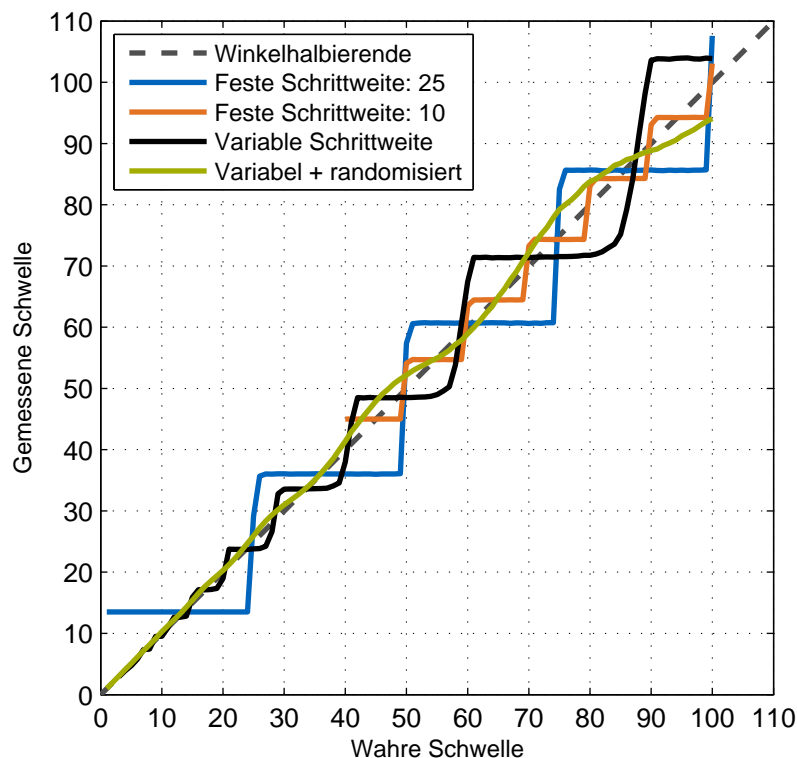


Abb. 4.20 Simulation der BASD Schwellenbestimmung. Bei festen Schrittweiten ergibt sich ein Stufenmuster (blaue und orange Kurve), sodass unterschiedliche Werte der wahren Schwelle auf denselben Wert als gemessene Schwelle abgebildet werden. Außerdem treten bei kleinen Änderungen der wahren Schwelle große Sprünge im gemessenen Schwellwert auf. Bei kleinen Schrittweiten (orange Kurve) verringert sich die Abweichung, wobei jedoch nur ein begrenzter Bereich messbar ist. Bei variabler Schrittweite (schwarze Kurve) sind fast alle Schwellen messbar, aber bei größeren Werten entsteht ebenfalls ein Stufenmuster. Wird der variablen Schrittweite bei jedem Versuch ein zufälliger Betrag hinzugefügt (grüne Linie), liegen die Ergebnisse nahe der Winkelhalbierenden (grau gestrichelte Linie), die das ideale Ergebnis darstellt.

wird (schwarze Linie), zeigen sich ähnliche Stufenmuster wie mit festen Schrittweiten. Hierbei sind allerdings Schwellen bis hinab zu 3 messbar. Wird die Schrittweite zusätzlich um einen zufälligen Wert verändert (grüne Kurve), gibt es keine Sprünge oder konstante Werte wie bei den 3 anderen Varianten. Alle Schwellen sind messbar und die gemessene Schwelle ist in guter Übereinstimmung mit der wahren Schwelle. Die Schrittweite S_{n+1} für den nächsten Versuch ($n+1$) wird dabei wie folgt berechnet, wobei die numerischen Werte in Gleichung (4.3) mithilfe der beschriebenen Simulation empirisch ermittelt wurden:

$$\begin{aligned}
 S_{n+1} &= \frac{S_n}{\text{mod}} \quad (\text{richtige Eingabe bei Versuch } n) \\
 S_{n+1} &= S_n \cdot \text{mod} \quad (\text{falsche Eingabe bei Versuch } n)
 \end{aligned}
 \tag{4.3}$$

$$\begin{aligned}
 \text{mod} &= -0.0144 \cdot n + 1.5144 + \text{rand} \\
 0 &\leq \text{rand} \leq 0,4 \quad (\text{Zufallszahl})
 \end{aligned}$$

4.3.3. Implementierung

Die BASD-Implementierung erfolgte unter Zuhilfenahme einer auf Eclipse (Eclipse Foundation, Brüssel, Belgien) basierenden Softwareentwicklungsumgebung für eingebettete Systeme in der Programmiersprache C auf der in Kapitel 3 beschriebenen Hardwareplattform. Die Bedienoberfläche sollte einfach, weitgehend selbsterklärend und für alle 6 Subtests gleich sein. Eine Kopfzeile mit Informationen zur Überwachung des Tests sollte immer sichtbar sein.

In Abb. 4.21 ist der Messverlauf dargestellt. Bei jedem Versuch wird zunächst die Signalsequenz mittels eines großen roten „Play“-Buttons gestartet (Abb. 4.21a). Während der Reizdarbietung leuchtet die jeweils zum Ton gehörende Karte gelb auf (Abb. 4.21b). Nachdem die Präsentation aller 3 Signale beendet ist, wird die Eingabe über die unten dargestellten 3 Karten freigeschaltet, von denen bisher nur die Rückseite sichtbar ist (Abb. 4.21b). Der Knopf zum Starten der Reizdarbietung ist grau dargestellt, um die Aufmerksamkeit auf die Karten zu lenken. Er ist jedoch weiterhin verwendbar, falls eine Wiederholung für eine sichere Entscheidung notwendig ist. Die Aufgabe besteht darin, ähnlich wie beim „Memory“ spielen, 2 gleiche Signale zu finden. Beim Antippen der Karten werden Bilder angezeigt, sodass es eine unmittelbare Rückmeldung gibt, ob die Wahl richtig oder falsch war. Für den Fall, dass die Wahl falsch war, wird die 3. Karte vom Gerät aufgedeckt zur Verdeutlichung, welche beiden Signale identisch waren (Abb. 4.21c). Anschließend kehrt das Gerät zur ursprünglichen Darstellung zurück (Abb. 4.21a) und ein neuer Versuch kann gestartet werden. Dabei ist es wichtig, dass die Tondarbietung immer von der Versuchsperson selbst ausgelöst wird, damit die Aufmerksamkeit auf das Zuhören gelenkt wird. Die 3 Schritte in Abb. 4.21 werden so lange wiederholt, bis die Schwelle feststeht, oder das eingestellte Maximum an Versuchen erreicht wird. Zur weiteren Steigerung der Motivation und zur Visualisierung des Messfortschritts wird auf der linken Seite ein Kartenvorrat durch einen im Messverlauf kleiner werdenden Balken angezeigt. Auf der rechten Seite wird ein Balken bei jeder richtigen Antwort größer, um auf

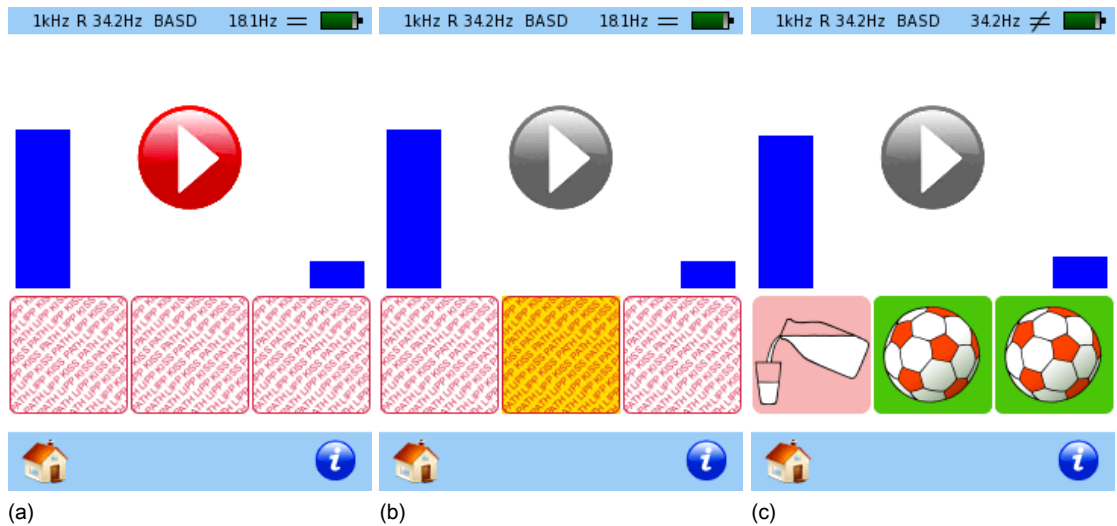


Abb. 4.21 BASD Messverlauf: (a)–(c) wird wiederholt bis die Schwelle feststeht, oder das eingestellte Maximum an Versuchen erreicht wird. (a) Starten der Tonsequenz (b) Tondarbietung aktiv. Die zum aktuellen Ton gehörende Karte leuchtet auf (c) Eingabe. Sollte die Wahl falsch sein, wird die 3. Karte aufgedeckt als unmittelbare Rückmeldung an die Versuchsperson

diese Weise den Kartenstapel der gewonnenen Karten zu veranschaulichen. In der Kopfzeile werden die Frequenz des Referenztons, das Testrohr und die Differenz zwischen Referenz- und Testsignal für den aktuellen Versuch angezeigt. Bei den sprachhaltigen Subtests werden statt der Differenz die verwendeten Silben angezeigt. Auf der rechten Seite ist die Differenz des vorangehenden Versuchs angegeben und als \neq und $=$ kodiert, ob die Identifikation des Testsignals richtig (\neq) oder falsch ($=$) war.

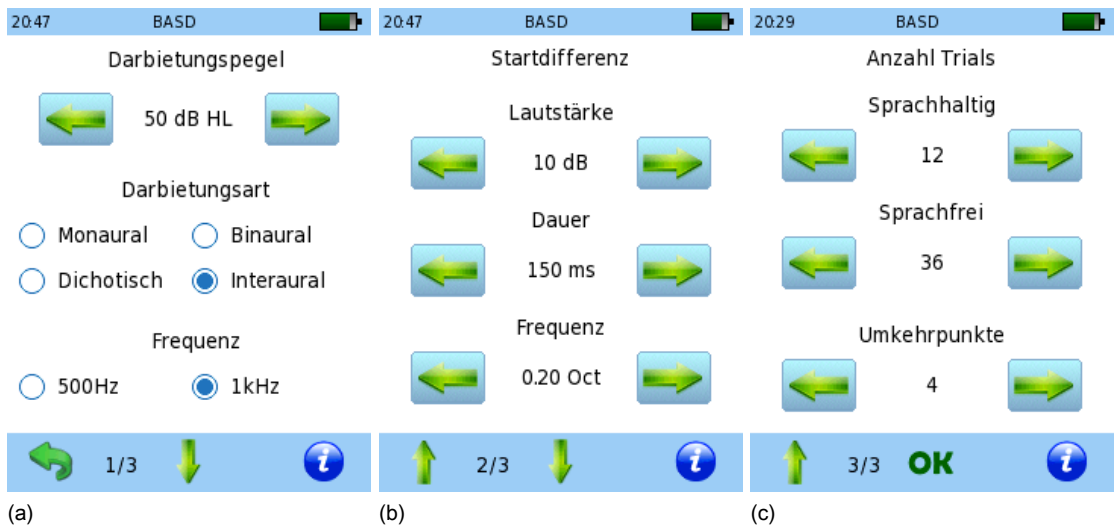


Abb. 4.22 BASD Testeinstellungen: (a) Globale Einstellungen für alle Subtests: Darbietungspegel und -art, Frequenz des Referenztons (b) Startdifferenzen für die sprachfreien Tests (c) Einstellungen für die Schwellenbestimmung: Anzahl der Versuche bzw. Trials getrennt für sprachfreie und sprachhaltige Untertests, Anzahl der Umkehrpunkte

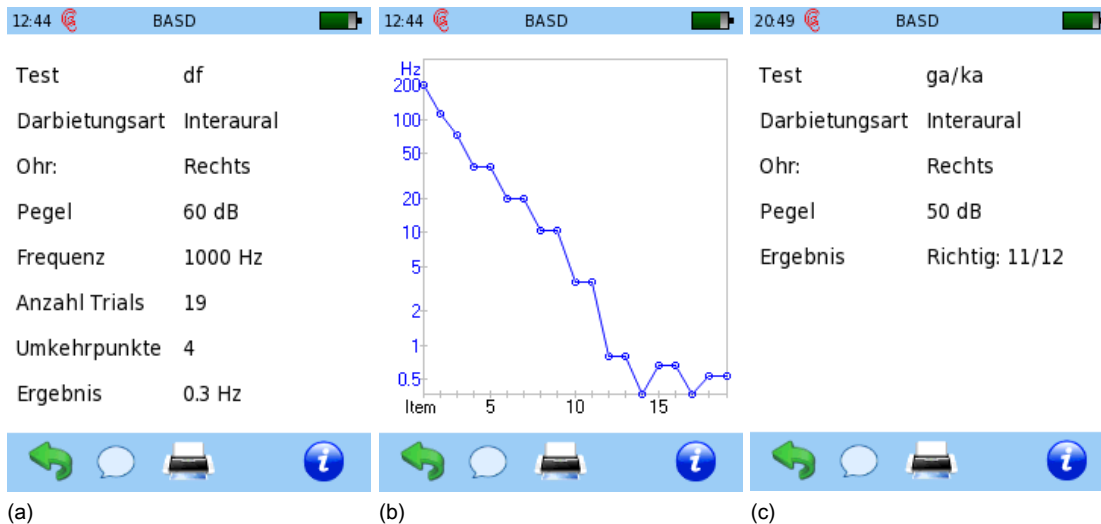


Abb. 4.23 BASD Ergebnisdarstellung: (a) tabellarische Zusammenfassung der Werte eines sprachfreien Tests (b) Messverlauf mit der Differenz für die Versuche bzw. Items eines sprachfreien Tests (c) tabellarische Zusammenfassung des Ergebnisses eines sprachhaltigen Tests

Abb. 4.22 zeigt die Testeinstellungen, die jeweils vor dem Start einer Messung angepasst werden können. Sie sind in 3 Bereiche gegliedert: Globale Einstellungen, die für alle 6 Subtests gelten (Abb. 4.22a), Startdifferenzen für die sprachfreien Tests (Abb. 4.22b) und Einstellungen, die die Schwellenbestimmung beeinflussen (Abb. 4.22c). Als Darbietungspegel kommt in der Regel ein Wert zum Einsatz, der ca. 30–40 dB über der Hörschwelle der gewählten Referenztonfrequenz liegt. Für Kinder ist es zur leichteren Erklärung und zum besseren Verständnis der Aufgabe hilfreich, die Startdifferenzen verhältnismäßig groß einzustellen. Mit der Erhöhung der geforderten Anzahl an Umkehrpunkten wird das Ergebnis genauer. Allerdings verlängert sich die Messzeit.

Abb. 4.23 zeigt die Darstellung der Ergebnisse. Für alle Subtests gibt es eine tabellarische Zusammenfassung, die das Ergebnis selbst und die Messparameter enthält (Abb. 4.23a und c). Zusätzlich wird bei allen sprachfreien Untertests eine grafische Repräsentation des Messverlaufs gezeigt (Abb. 4.23b). Hierbei ist auch die Funktionsweise des Schwellenbestimmungsalgorithmus gut sichtbar: Ausgehend von einer Startdifferenz von 200 Hz wird der Unterschied zunächst bei jedem Versuch bzw. Item in großen Schritten reduziert. Ab dem 4. Versuch sind immer 2 aufeinanderfolgende richtige Eingaben notwendig, um die Differenz weiter abzusenkten. Nach 4 Umkehrpunkten an den Stellen 14, 15–16, 17 und 18–19 kann die Schwelle bereits nach nur 19 Versuchen berechnet werden. Sie beträgt im abgebildeten Beispiel 0,3 Hz (Abb. 4.23b).

Auszüge aus Abschnitt 4.3 wurden vorab auf Tagungen präsentiert (DGPP Jahrestagung 2012, EUHA 2014) und veröffentlicht als: Pürner, Schirkonyer und Janssen (2022): Changes in the peripheral and central auditory performance in the elderly – A cross-sectional study. Journal of Neuroscience Research, 100(9), 1791-1811.

5. Klinische Studien

Vor dem Einsatz eines neuen Testverfahrens in der klinischen Routine ist ein Nachweis der Effektivität und Zuverlässigkeit erforderlich. Die in Kapitel 4 entwickelten Verfahren wurden auf der in Kapitel 3 vorgestellten Mess-Plattform implementiert. Die Tests wurden in Zusammenarbeit mit der Abteilung für Kommunikationsstörungen (HNO Klinik Mainz), dem Schwerpunkt für Phoniatrie und Pädaudiologie (HNO Klinik Frankfurt), der Phoniatrie und Pädaudiologie (HNO Klinik Bochum), dem Cyprus Audiology Center (Nikosia, Zypern) und der Experimentellen Audiologie (Klinikum rechts der Isar München) normiert und validiert. Die Ergebnisse der einzelnen Studien zu MAGIC, MATCH und BASD werden in diesem Kapitel vorgestellt.

5.1. MAGIC

5.1.1. Studiendesign

Die Studie wurde multizentrisch mit Messungen in Mainz, Nikosia und München durchgeführt. Für sämtliche MAGIC Messungen wurde einheitlich das in Kapitel 3 beschriebene Messsystem verwendet. Bei der zum Vergleich gemessenen Reintonaudiometrie (pure tone audiometry, PTA), die als Spielaudiometrie ausgeführt wurde, kam entweder das Audiometer-Modul des Senti Gerätes oder ein in der jeweiligen Klinik vorhandenes Audiometer zum Einsatz. Die Auswahl der Startmessung, MAGIC oder PTA, erfolgte randomisiert. Es wurden immer beide Ohren nacheinander gemessen und nach einer kurzen Pause die jeweils andere Untersuchungsmethode durchgeführt. Die Messungen fanden in Audiometrieräumen statt, sodass der Einfluss von Störgeräuschen auf die Ergebnisse weitgehend ausgeschlossen werden konnte. Bei Test-Retest-Messungen, betrug der zeitliche Abstand zwischen der ersten und zweiten Messung im Allgemeinen 24 ± 1 h. Alle beschriebenen Untersuchungen wurden von der jeweils zuständigen Ethik-Kommission genehmigt und sind im Einklang mit nationalem Recht sowie gemäß der Deklaration von Helsinki von 1975 (in der aktuellen, überarbeiteten Fassung) durchgeführt worden.

Insgesamt 156 Kinder im Alter zwischen 3;6 und 10;3 Jahren (Mittelwert: $6;0 \pm 1;3$ Jahre) nahmen an der Studie teil. Die Messungen mit beiden Messmethoden, MAGIC und PTA, erfolgten im Frequenzbereich von 250 Hz bis 8 kHz. Dabei wurden, sofern möglich, beide Ohren gemessen. Das war bei 154 Kindern der Fall. Der mittels PTA gemessene Hörverlust lag zwischen -5 und 95 dB HL. Die MAGIC Messungen erfolgten im Bereich von 5 bis 70 dB HL. In der Auswertung wurden nur vollständige Wertepaare bestehend aus PTA und MAGIC berücksichtigt. Bei 2 Kindern konnte aufgrund nachlassender Aufmerksamkeit und daraufhin fehlender Kooperation nur jeweils 1 Ohr gemessen werden, sodass Daten von 310 Ohren vorliegen. Test-Retest-Messungen wurden an 49 Kindern (98 Ohren) aus dem Testkollektiv durchgeführt. Als zusätzlicher Parameter wurde die Durchführungszeit des MAGIC automa-

tisch aufgezeichnet. Allerdings konnte die Messdauer von 7 Kindern nicht erfasst werden, da der hierfür verwendete Uhrenbaustein des Gerätes eine Fehlfunktion aufwies, die sich durch unplausible Datumsangaben wie z.B. „2040-00-15“ bemerkbar machte. Der Uhrenbaustein hat jedoch keinen Einfluss auf die Messung oder das Ergebnis. Anhand der handschriftlichen Aufzeichnungen konnte das Messdatum rekonstruiert werden, nicht jedoch die Prüfdauer.

Die MAGIC-Messung wurde in der Regel durch das Kind selbst gesteuert. Dies umfasste die Darbietung des Tons sowie die Eingabe, ob der Ton gehört wurde oder nicht. Das Fachpersonal überwachte den Messablauf und unterstützte die Kinder, wenn erforderlich, bei der Eingabe. Es hatte auch jederzeit die Möglichkeit zum Eingreifen in den Messablauf, wenn dies vonnöten war. Im Rahmen der Studie wurde jedoch weitgehend darauf verzichtet, um den Anteil der Ausreißer zu schätzen, der bei unbeaufsichtigter Ausführung auftreten könnte. Als Kriterium zur Untersuchung von Ausreißern wurde die Differenz zwischen zwei Messungen derselben Frequenz am selben Ohr des Kindes verwendet. Gemäß Steland (2010) wurden Werte als Ausreißer betrachtet, die weiter als 1,5 Interquartilsabstände vom 1. oder 3. Quartil entfernt lagen.

Nach den Spielaudiometrie und MAGIC Messungen wurden die ersten 66 Versuchspersonen befragt, welches Testverfahren mehr Spaß gemacht hatte. Zusätzlich wurden die ersten 27 dieser 66 Kinder gefragt, welche Methode sie als anstrengender empfanden.

5.1.2. Ergebnisse

Für die Messungen bei dieser Studie hatte das Untersuchungspersonal die Vorgabe, möglichst wenig einzugreifen sowie die Eingaben der Kinder nur in Ausnahmefällen zu korrigieren. Dabei konnte es auch zu Resultaten kommen, die offensichtlich unplausibel waren. Daher wurde eine Analyse der Ausreißer durchgeführt. In Abb. 5.1 sind exemplarisch „Ausreißer-Audiogramme“ dargestellt, die mittels MAGIC gemessen wurden. So traten teilweise steile Hörverluste bei tieferen Frequenzen auf, während das Hörvermögen im mittleren und höheren Frequenzbereich sehr gut war (Abb. 5.1a). Oder es zeigten sich Zickzackmuster, die wahrscheinlich nicht der wahren Hörschwelle entsprachen (Abb. 5.1b).

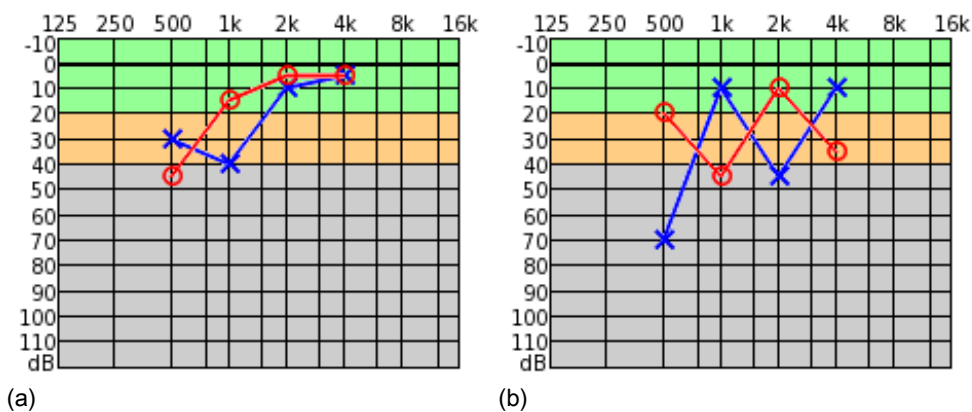


Abb. 5.1 Beispiele für Audiogramme mit Ausreißern

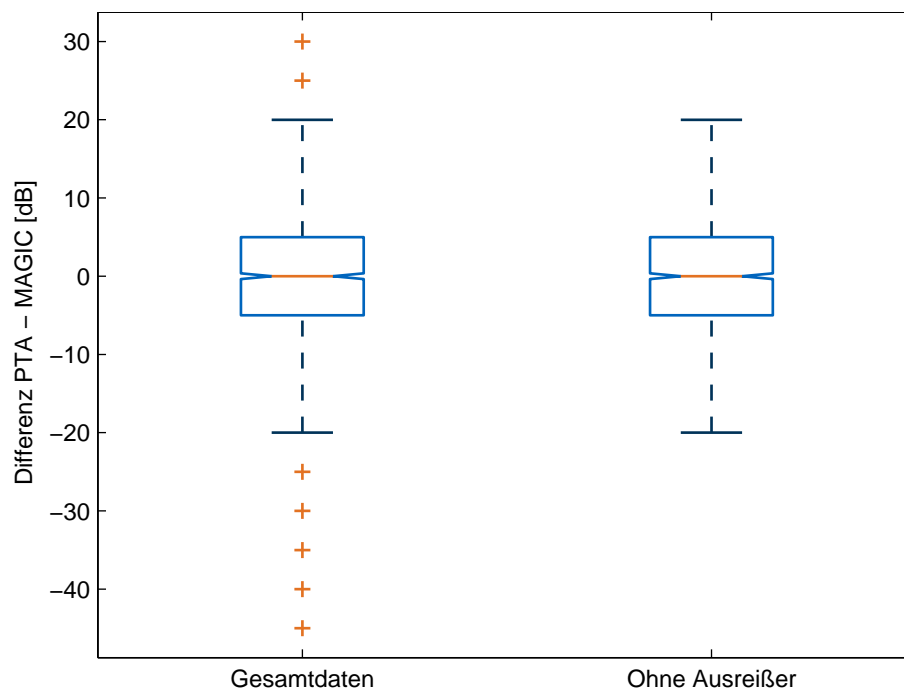


Abb. 5.2 Boxplot der Differenzen. Links sind die Gesamtdaten bestehend aus 1829 Werten dargestellt. Rechts wurden die 55 Ausreißer weggelassen. Median und Interquartilsabstand sind jeweils 0 dB bzw. 10 dB.

Tab. 5.1 Ausreißer in Abhängigkeit der Frequenz

250 Hz	500 Hz	1 kHz	2 kHz	3 kHz	4 kHz	6 kHz	8 kHz	Alle
6,9%	6,5%	2,5%	1,7%	0,0%	1,8%	3,9%	1,8%	3,0%

Insgesamt ergab sich eine Anzahl von 1829 Wertepaaren aus jeweils einer PTA und einer MAGIC Hörschwelle bei derselben Frequenz am selben Ohr. In Abb. 5.2 ist die Differenz der beiden Schwellen in einem Box-Whisker-Plot dargestellt, wobei links alle Daten abgebildet sind, und rechts die Ausreißer gemäß obiger Definition weggelassen wurden. Der Median ist 0 dB und der Interquartilsabstand beträgt 10 dB. Die 55 Ausreißer befinden sich im Bereich von -45 bis -25 dB und 25 bis 30 dB. Ihre Anzahl entspricht 3 % der Werte. Tabelle 5.1 listet den Anteil der Ausreißer für jede Frequenz einzeln auf. Dabei sind die tiefen Frequenzen, 250 Hz und 500 Hz mit 6,9 % und 6,5 % besonders betroffen. Für 8 kHz sowie zwischen 1 kHz und 4 kHz liegt der Anteil an Ausreißern zwischen 0 % und 2,5 %. Bei 6 kHz beträgt er 3,9 %. Die weitere Datenauswertung erfolgt anhand der 1774 Wertepaare ohne Ausreißer.

Die Validierung des MAGIC erfolgte durch Korrelationsanalysen, bei denen einerseits die Ergebnisse des MAGIC mit den Ergebnissen der PTA verglichen wurden, und andererseits ein Test-Retest-Vergleich durchgeführt wurde. Darüber hinaus wurde der Einfluss des Alters auf den Grad der Übereinstimmung von MAGIC und PTA sowie Test und Retest mittels linearer Regression untersucht. In Abb. 5.3 wird der Vergleich zwischen MAGIC und PTA gezeigt. Zur besseren Darstellung der Anzahl von Datenpunkten an einer bestimmten Stelle (z.B. MAGIC 5 dB / PTA 10 dB) werden die Symbole an diesen Koordinaten in der Grafik dicker gezeich-

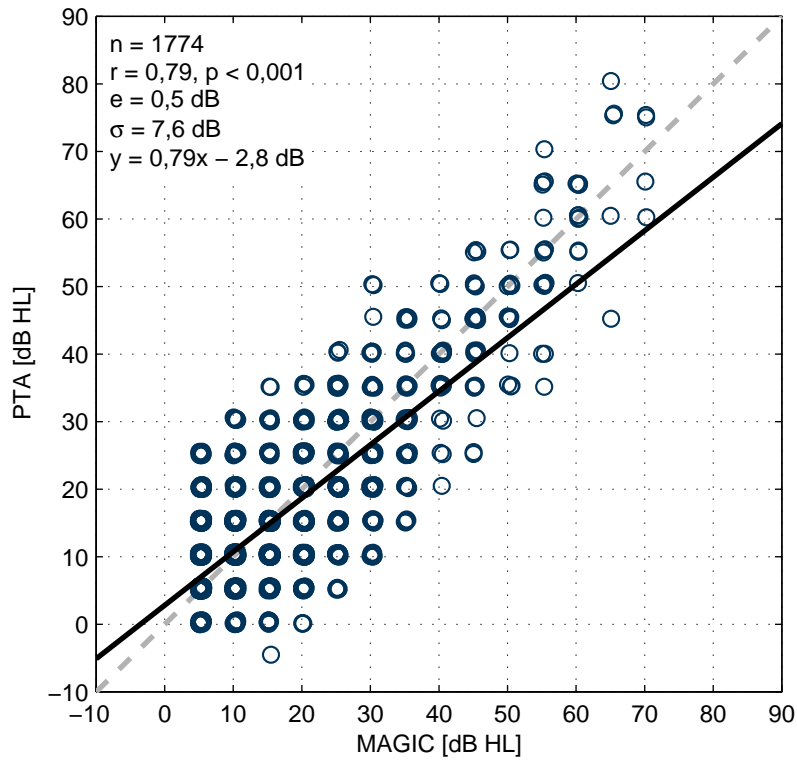


Abb. 5.3 Korrelation zwischen MAGIC und PTA. Der Mittelwert und die SD der Unterschiede zwischen den Tests liegen bei 0,5 dB und 7,6 dB. Die Korrelation für die 1774 Hörschwellen beträgt 0,79. Die lineare Regressionsgleichung (durchgezogene Linie) lautet $y = 0,79x - 2,8 \text{ dB}$

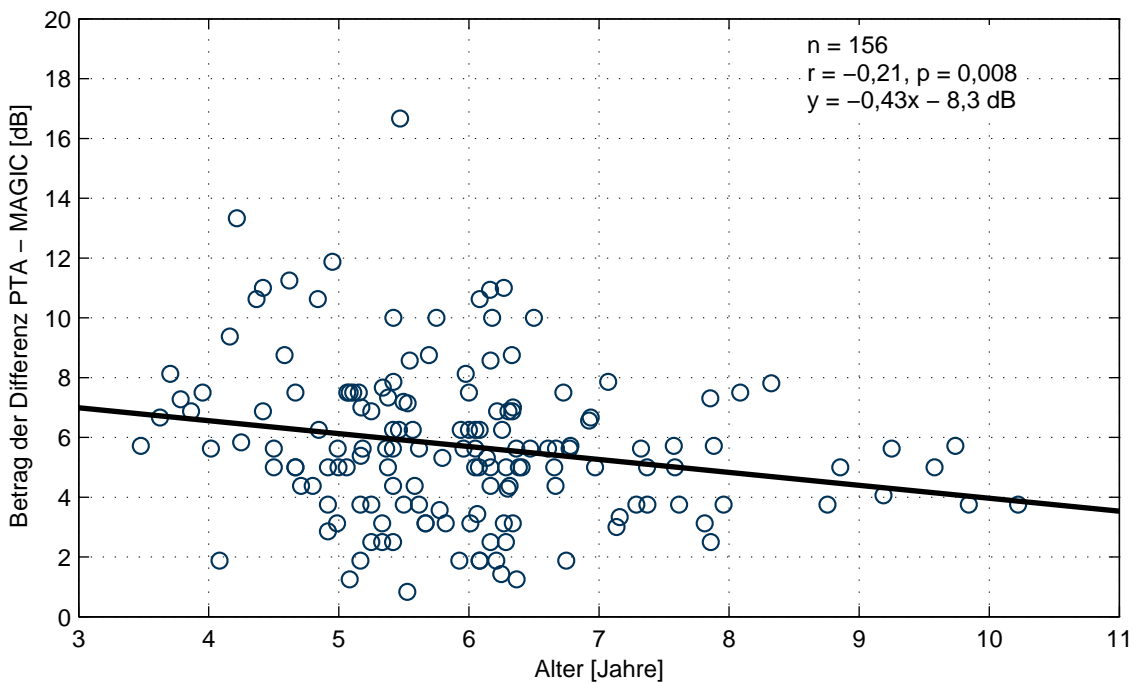


Abb. 5.4 Betrag der mittleren Differenz zwischen MAGIC und PTA in Abhängigkeit des Alters. Das Lebensalter und der Betrag der Differenz zwischen den beiden Messmethoden sind nur schwach negativ korreliert ($r = -0,21$; $p = 0,008$). Mit zunehmendem Alter nimmt der Betrag der Differenz um ungefähr 0,4 dB pro Jahr ab. Die lineare Regressionsgleichung (durchgezogene Linie) lautet $y = -0,43x - 8,3 \text{ dB}$

net. 71,1 % der Daten liegen innerhalb des Bereichs von -5 dB bis 5 dB Differenz zwischen MAGIC und PTA und 89,5 % befinden sich zwischen -10 dB und 10 dB. Es besteht eine starke positive Korrelation zwischen MAGIC und PTA, was sich am Korrelationskoeffizienten über alle 1774 Einzelmessungen von 0,79 ($p < 0,001$) widerspiegelt. Der Mittelwert der Differenzen (e) beträgt 0,5 dB bei einer SD von 7,6 dB. Die Regressionsgeraden-Gleichung lautet $y = 0,79x - 2,8$ dB.

Das Streudiagramm in Abb. 5.4 veranschaulicht die Abhängigkeit des durchschnittlichen Betrags der Differenz zwischen MAGIC und PTA vom Alter für alle 156 Kinder des Testkollektivs. Dabei wurde für jedes Kind ein Wert nach der Formel $\frac{1}{n} \sum_{i=1}^n |L(i)_{PTA} - L(i)_{MAGIC}|$ berechnet, mit $L(i)_{MAGIC}$ mittels MAGIC gemessene i -te Hörschwelle und $L(i)_{PTA}$ mittels PTA gemessene Hörschwelle am selben Ohr und derselben Frequenz sowie der Anzahl n an paarweise vorliegenden Hörschwellen. Das Lebensalter und die Differenz zwischen den beiden Messmethoden sind schwach negativ korreliert ($r = -0,21$; $p = 0,008$). Mit zunehmendem Alter nimmt der Betrag um ungefähr 0,4 dB pro Jahr ab. Die Regressionsgeraden-Gleichung lautet $y = -0,43x - 8,3$ dB. Der Mittelwert \pm SD der Differenzbeträge über alle Kinder beträgt $5,7 \pm 2,6$ dB. Die Werte reichen von 0,8 bis 16,7 dB. Ihr Median liegt bei 5,6 dB.

Im Streudiagramm in Abb. 5.5 sind die Daten der 49 Kinder des Test-Retest-Vergleichs visualisiert. Zur besseren Darstellung sind in der Grafik die Symbole an Stellen mit vielen Daten-

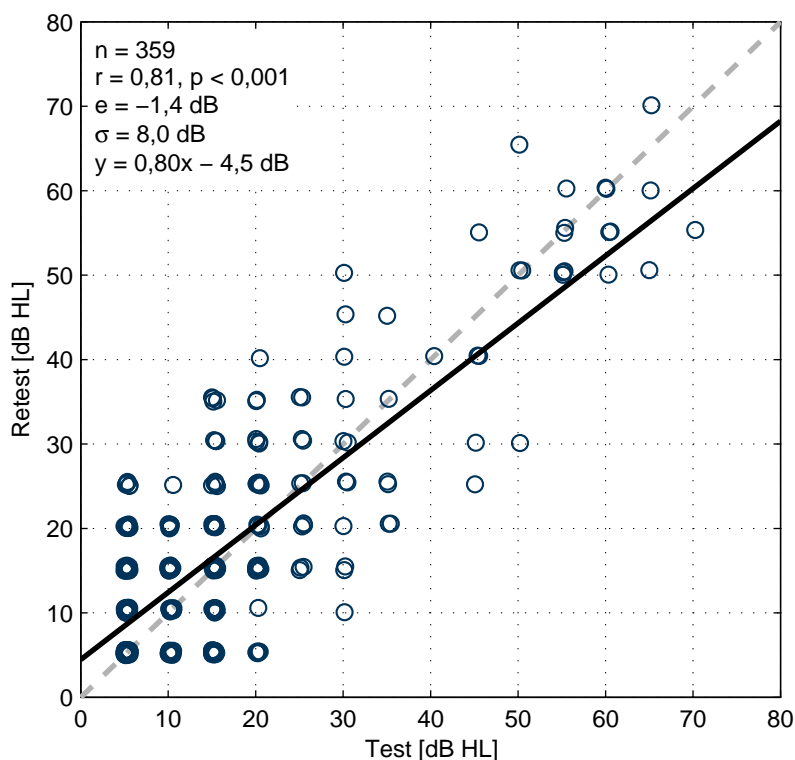


Abb. 5.5 Reliabilität der MAGIC Test-Retest-Messungen: Streudiagramm der ersten versus zweiten Messung. Der Mittelwert und die SD der Unterschiede zwischen den Tests betragen -1,4 dB und 8,0 dB. Die Korrelation für die 359 Schwellen beträgt 0,81. Die lineare Regressionsgleichung (durchgezogene Linie) lautet $y = 0,80x - 4,5$ dB.

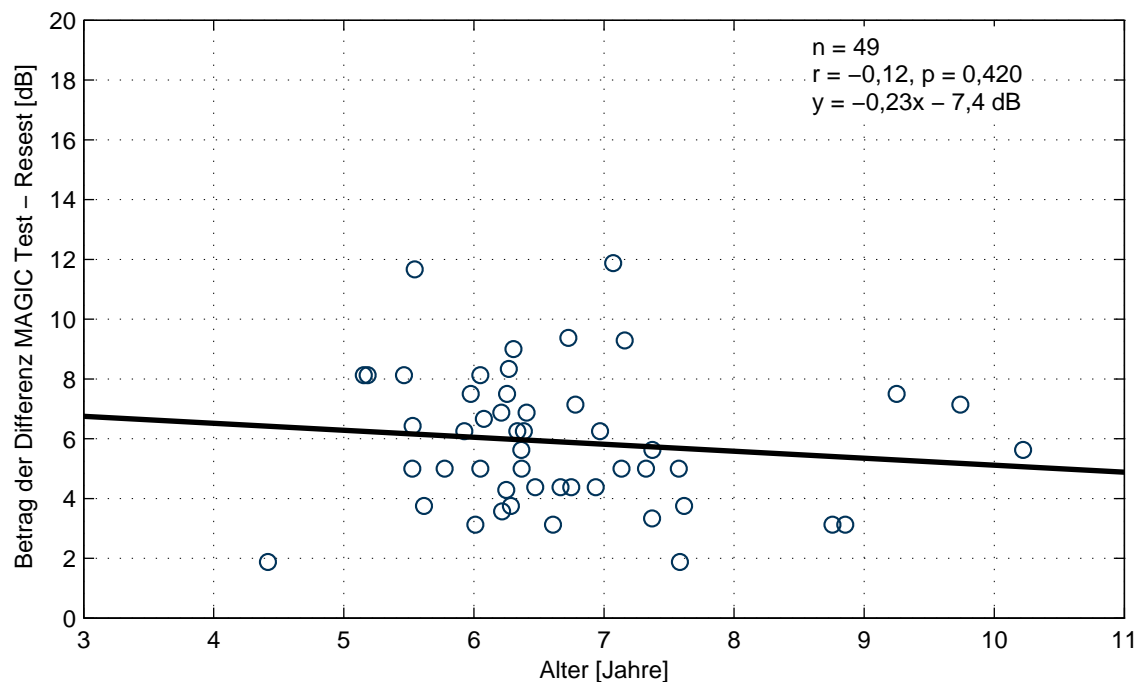


Abb. 5.6 Durchschnittlicher Betrag der Differenz zwischen Test und Retest in Abhängigkeit des Alters. Das Lebensalter und der durchschnittliche Betrag der Differenz zwischen den beiden Testdurchläufen sind in der vorliegenden Stichprobe nicht korreliert.

punkten dicker gezeichnet. 68,2% der Daten liegen innerhalb des Bereichs von -5 dB bis 5 dB Differenz zwischen Test und Retest und 86,1% befinden sich zwischen -10 dB und 10 dB. Es besteht eine starke positive Korrelation zwischen Test und Retest. Der Korrelationskoeffizient über alle 359 Schwellen liegt bei 0,81 ($p < 0,001$). Test und Retest unterscheiden sich im Mittel um 1,4 dB bei einer SD von 8,0 dB. Die Beziehung von Test- und Retest-Ergebnissen wird durch die Regressionsgleichung $y = 0,80x - 4,5$ dB beschrieben.

In Abb. 5.6 ist der durchschnittliche Betrag der Differenz zwischen Test und Retest in Abhängigkeit des Alters als Streudiagramm dargestellt. Diese Messungen wurden mit 49 Kindern durchgeführt, von denen die meisten zwischen 5 und 8 Jahre alt waren. Die Berechnung des Differenzbetrags zwischen Test und Retest erfolgte analog zu oben gemäß der Formel $\frac{1}{n} \sum_{i=1}^n |L(i)_{\text{Test}} - L(i)_{\text{Retest}}|$. Es besteht kein linearer Zusammenhang: Das Lebensalter und der durchschnittliche Betrag der Differenz zwischen den beiden Testdurchläufen sind also nicht korreliert ($r = -0,12$; $p = 0,42$). Über alle 49 Kinder, die an den Retest-Messungen teilnahmen, beträgt der Mittelwert \pm SD der Differenzbeträge $5,9 \pm 2,3$ dB. Die Werte reichen von 1,9 bis 11,9 dB. Ihr Median liegt bei 5,6 dB.

Das Streudiagramm in Abb. 5.7 zeigt die durchschnittliche Messdauer zur Bestimmung einer monauralen Schwelle für eine Testfrequenz in Abhängigkeit des Alters für 149 Kinder des Testkollektivs. Es besteht eine moderate negative Korrelation ($r = -0,38$; $p < 0,001$) zwischen Lebensalter und Messdauer. Die Regressionsgeraden-Gleichung lautet $y = -4,67x - 63,2$ s. Entsprechend diesem linearen Modell nimmt die durchschnittliche Testdauer pro Frequenz

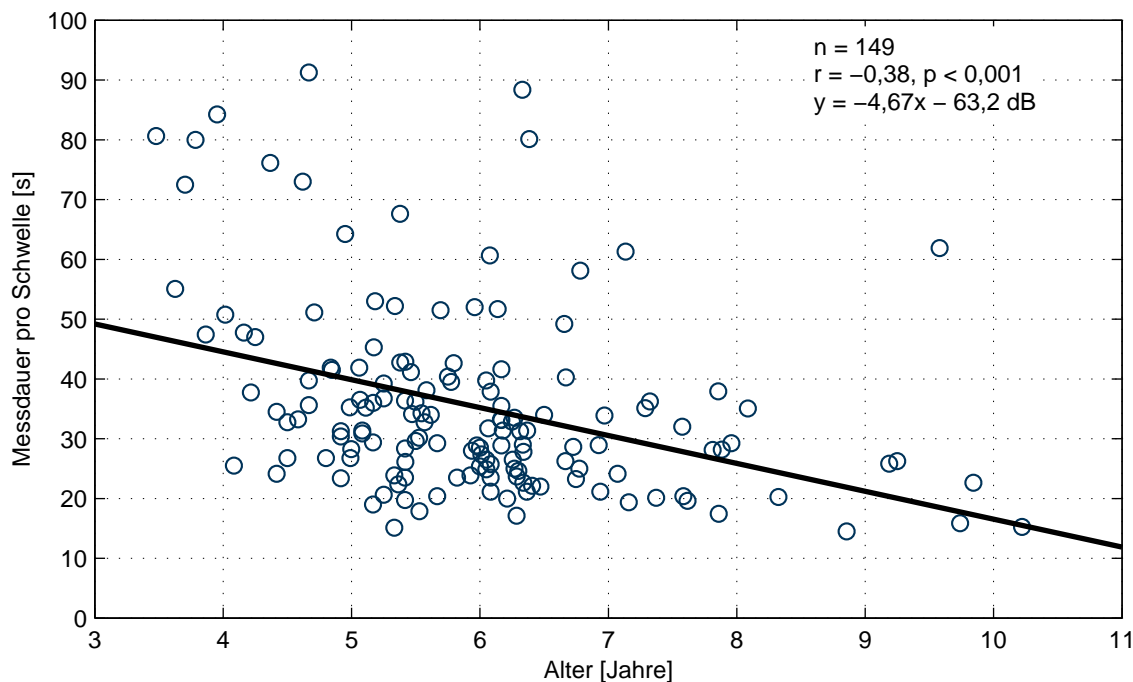


Abb. 5.7 Messdauer pro Schwelle in Abhängigkeit des Alters. Es besteht eine moderate negative Korrelation ($r = -0,38$; $p < 0,001$) zwischen Lebensalter und Messdauer. Die lineare Regressionsgleichung (durchgezogene Linie) lautet $y = -4,67x - 63,2$ s. Mit zunehmendem Alter nimmt die durchschnittliche Testdauer pro Frequenz um ungefähr 5 s pro Jahr ab.

mit zunehmendem Alter um ungefähr 5 s pro Jahr ab. Der Mittelwert \pm SD der durchschnittlichen Messdauer beträgt $36,1 \pm 15,4$ s. Die Werte reichen von 14,5 bis 91,3 s. Ihr Median liegt bei 32,8 s.

Die Ergebnisse der Kinderbefragung, welches Messverfahren mehr Spaß gemacht hat und welches anstrengender war, sind in Tabelle 5.2 angegeben. Einer deutlichen Mehrheit der Kinder (51 von 66) hat das neue bildgestützte Audiometrieverfahren (MAGIC) mehr Spaß gemacht. 6 von 66 Kindern gaben an, dass sie die konventionelle Spielaudiometrie (PTA) bevorzugten. Auf „beide“ und „keine Angabe“ entfielen jeweils 5 bzw. 4 Stimmen. Bei der Frage, welches Verfahren anstrengender war, haben 4 bzw. 9 von 27 für „beide“ und „keine Angabe“ gestimmt, sodass diese Frage nicht mehr gestellt wurde. Im Vergleich zwischen den Testverfahren empfanden 5 Kinder MAGIC als anstrengender und 9 Kinder PTA.

Tab. 5.2 Kinderbefragung welcher Test mehr Spaß gemacht hat und welcher anstrengender war.

Parameter	MAGIC	PTA	Beide	k.A.	n
Spaß	51	6	5	4	66
Anstrengung	5	9	4	9	27

Vorläufige Ergebnisse aus Abschnitt 5.1 wurden vorab in Auszügen auf Tagungen präsentiert (DGA 2010, DGPP 2010, EFAS 2011, HNO 88. Jahresversammlung 2017).

5.2. MATCH

5.2.1. Studiendesign

Messungen

Für die Messungen wurde das in Kapitel 3 beschriebene Messsystem verwendet, das Sprachsignale freifeldentzerrt an den angeschlossenen Kopfhörer ausgibt. Der Kopfhörer hat eine hohe passive Schalldämpfung (> 45 dB bei 4 kHz), was beim Einsatz außerhalb von Audiometrikabinen von Vorteil ist. Die Messungen an normalhörenden Kindern erfolgten größtenteils in Kindergärten im Großraum Mainz. Dabei wurden Ruheräume genutzt und darauf geachtet, dass der Umgebungslärm 40 dB(A) nicht überschreitet. Die Untersuchungen an schwerhörenden Kindern wurden in einer schallisolierten Audiometrikabine der Abteilung für Kommunikationsstörungen (HNO Klinik Mainz) durchgeführt. Dort fanden auch die Messungen statt, die zur Beurteilung der Test-Retest-Reliabilität der Sprachverständlichkeitsschwelle (SVS) und zur Korrelation mit den Ergebnissen der Tonaudiometrie dienten. Der zeitliche Abstand zwischen der 1. und 2. Messung betrug im Allgemeinen 24 ± 1 h. Für den Vergleich der tonaudiometrischen Daten mit den sprachaudiometrischen Ergebnissen wurde der mittlere Hörverlust (mHV PTA) gemäß der Rec. 02/1 des International Bureau for Audiophonology (1996) zur Klassifizierung von Hörschäden bei den Frequenzen 0,5, 1, 2 und 4 kHz verwendet.

Vor dem Beginn der Tests wurde das Kind mit den Abbildungen vertraut gemacht zur Sicherstellung, dass alle Begriffe Teil des rezeptiven Vokabulars sind. Unmittelbar danach erfolgte eine Messung mit adaptiver Pegelsteuerung zur Bestimmung der SVS pro normalhörendem Ohr. Alle beschriebenen Untersuchungen wurden mit Zustimmung der zuständigen Ethik-Kommission der Landesärztekammer Rheinland-Pfalz genehmigt und sind im Einklang mit nationalem Recht sowie gemäß der Deklaration von Helsinki von 1975 (in der aktuellen, überarbeiteten Fassung) durchgeführt worden.

Versuchspersonen

Die Studie zur Standardisierung und Validierung umfasste insgesamt 179 Kinder im Altersbereich von 2 Jahren und 11 Monaten bis 6 Jahren und 9 Monaten. Die Teilnehmenden wurden zum einen in Kindergärten (überwiegend Normalhörende), zum anderen in der pädaudiologischen Sprechstunde der Klinik (überwiegend Schwerhörende) rekrutiert. Angaben zum Geschlecht wurden nicht erhoben. Für die Standardisierung des Sprachtests konnten aus dem Gesamtkollektiv 234 Ohren mit normalem Hörvermögen herangezogen werden. Um als normalhörend klassifiziert zu werden, musste entweder ein Screening mit 30 dB HL bei den Frequenzen 500 Hz, 1 kHz, 2 kHz, 4 kHz und 6 kHz erfolgreich abgeschlossen werden, oder es mussten sämtliche Hörschwellen in diesem Frequenzbereich ≤ 20 dB HL sein. Die Auswertung der Hörschwellen erfolgte pro Ohr. Bei 51 Kindern erfüllte keines der beiden Ohren die beschriebenen Kriterien für Normalhörigkeit. Mindestens 1 Ohr mit normalem Hörvermögen war bei 128 Kindern vorhanden. Von 22 Teilnehmenden wurde lediglich 1 Ohr in die Normierung einbezogen, da die Aufmerksamkeitsspanne für die Messung beider Ohren zu

Tab. 5.3 Testkollektiv: Verteilung nach Alter und Hörstatus. Die Einteilung in normal- und schwerhörend erfolgte ohrweise, sodass ein Kind bei einseitiger Schwerhörigkeit mit je 1 Ohr in beiden Gruppen auftauchen kann.

Altersgruppe	< 4,25 Jahre	4,25–5,5 Jahre	> 5,5 Jahre	Gesamt
Normalhörend: Kinder	42	50	36	128
Normalhörend: Ohren	75	94	65	234
Schwerhörend: Kinder	15	13	29	57
Schwerhörend: Ohren	20	18	49	87
Gesamt: Kinder	52	63	64	179
Gesamt: Ohren	95	112	114	321

kurz oder das Kriterium für Normalhörigkeit nicht erfüllt war. Die normalhörenden Kinder wurden ihrem Alter entsprechend in 3 Gruppen eingeteilt: Kinder bis zu einem Alter von 4 Jahren und 3 Monaten (< 4,25 Jahre; $n = 42$), Kinder zwischen 4 Jahren und 3 Monaten und 5 Jahren und 6 Monaten (4,25–5,5 Jahre; $n = 50$) und Kinder ab 5 Jahren und 6 Monaten (> 5,5 Jahre; $n = 36$). Die Festlegung der Altersgrenzen soll die Entwicklung der Kinder widerspiegeln: Kinder bis 4,25 Jahre sind in der Regel weniger als ein Jahr im Kindergarten, während Kinder zwischen 4,25 und 5,5 Jahren bereits länger als ein Jahr den Kindergarten besuchen. Die Altersgruppe > 5,5 Jahre repräsentiert diejenigen Kinder, die kurz vor der Einschulung stehen. Tabelle 5.3 listet das gesamte Testkollektiv auf sowie die Verteilung nach Altersgruppen und Hörstatus. Insgesamt konnte an 37 Ohren kein Resultat ermittelt werden, weil die Konzentration nachließ, die Zeitplanung im Kindergarten es nicht mehr zuließ oder die Kooperation des Kindes aus verschiedensten Gründen nicht mehr gegeben war. Bei 6 Kindern erfüllte nur je 1 Ohr die Kriterien für Normalhörigkeit, sodass diese jeweils als normalhörende und schwerhörende Kinder in der Tabelle 5.3 aufgeführt sind.

Die Test-Retest-Reliabilität der SVS wurde anhand der Ergebnisse von $n = 85$ Ohren ermittelt, von denen 74 gemäß obiger Definition ein normales Hörvermögen und 11 eine Hörschädigung aufwiesen. Für die Messung des Zusammenhangs zwischen Reintonaudiometrie und SVS standen die Daten von $n = 91$ Ohren zur Verfügung. Eine Hörschädigung wiesen 68 auf und 23 waren normalhörend.

Auswertung der Daten

Die psychometrischen Funktionen wurden auf der Grundlage der Sprachverständlichkeit bei Schalldruckpegeln zwischen 15 und 50 dB SPL in Schritten von 5 dB durch Schätzung der Parameter aus Gleichung (4.1) ermittelt. Für die altersspezifischen Sprachdiskriminationsfunktionen wurde ein Parametersatz je Altersgruppe berechnet. Zur Analyse der Verständlichkeit des Testmaterials wurde ein Parametersatz, unter Verwendung sämtlicher Messungen, für jedes einzelne Wort bestimmt. Mittelwert und Standardabweichung der altersspezifischen Sprachverständlichkeit wurden mittels Bootstrapping mit je 1000 Stichproben pro Altersgruppe berechnet.

5.2.2. Ergebnisse

Sprachverständlichkeit des Testmaterials

Die Gewährleistung einer möglichst gleichen Verständlichkeit aller Testitems ist eine wichtige Voraussetzung für die Konsistenz und Validität eines sprachaudiometrischen Verfahrens. Aus diesem Grund wurde die Verständlichkeit des Wortmaterials auf 2 verschiedene Arten untersucht. Zum einen erfolgte die Schätzung einer Diskriminationsfunktion für jedes einzelne Wort unter Zuhilfenahme aller Messungen an Normalhörenden. Zum anderen wurde auf Basis der Messungen aller Kinder analysiert, ob es Wörter gab, die bei Schalldruckpegeln deutlich unterhalb der durchschnittlichen Schwelle verstanden wurden („leichte Wörter“) oder deutlich oberhalb des Schwellwertes nicht verstanden wurden („schwere Wörter“).

Abb. 5.8 zeigt die Sprachverständlichkeit der 26 Wörter des Testmaterials in Abhängigkeit des Darbietungspegels bei Kindern ($n=128$) nach der Pegelanpassung gemäß Tabelle 4.1 als graue Linien. In schwarz sind der Mittelwert zwischen 5 und 95 % und die SD im Bereich von 10 bis 90 % Verständlichkeit in 10 %-Schritten dargestellt. In diagnostisch relevanten Bereichen zwischen 30 und 70 % liegt die SD zwischen 1,1 dB bei 50 % und 1,7 dB bei 30 % Verständlichkeit. Sie beträgt 1,3 dB bei 70 %.

Mit dem in Gleichung (4.1) dargestellten Model und den ermittelten Werten (Tabelle 5.5) wurde individuell für jede einzelne Messung untersucht, ob es hierbei für das Kind Wörter gab, die besonders leicht oder besonders schwer zu verstehen waren. Für diese Auswertung wurde

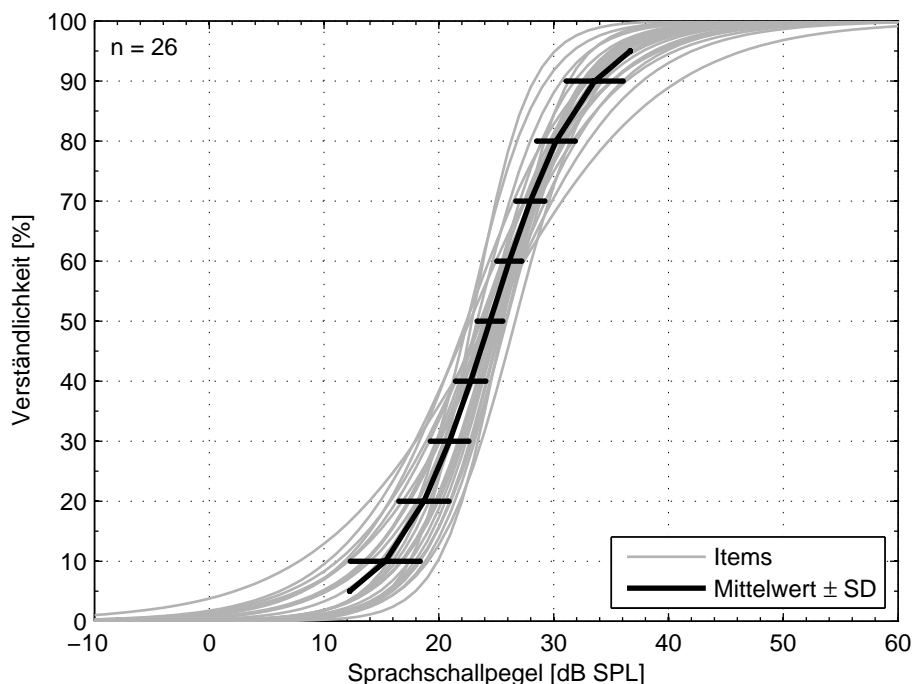


Abb. 5.8 Individuelle (dünne graue Linien) und mittlere (schwarz) Ergebnisse für die psychometrische Funktion der Sprachverständlichkeit als Funktion des Darbietungspegels bei normalhörenden Kindern. Die Standardabweichung (SD) wird zwischen 10 und 90 % Verständlichkeit angezeigt (schwarze horizontale Linien)

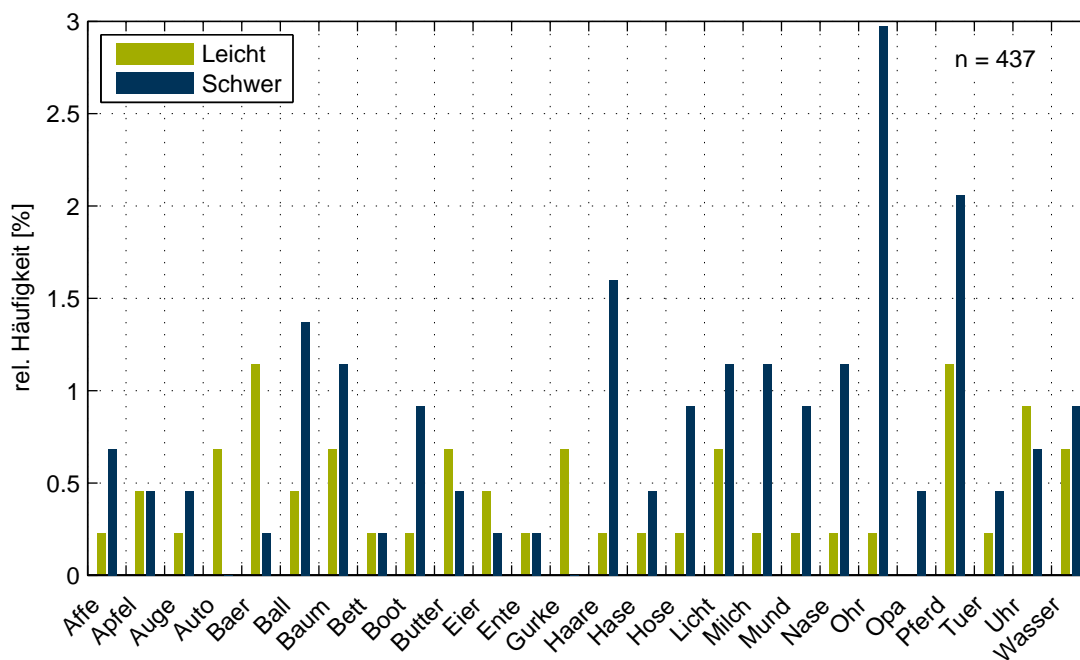


Abb. 5.9 Leichte und schwere Worte des Testmaterials. Relative Häufigkeit der Messungen, bei denen das betreffende Wort als „Schwer“ (im 85 %-Punkt der Diskriminationskurve nicht verstanden) und als „Leicht“ (im 30 %-Punkt der Diskriminationskurve verstanden) eingestuft wurde.

die Annahme getroffen, dass die individuelle psychometrische Funktion mit den Ergebnissen aus Abb. 5.10 übereinstimmt und bei einem Testergebnis, das vom Mittel abweicht, nur eine Verschiebung entlang der Pegelachse (x-Achse) stattfindet. „Schwere Items“ wurden demnach definiert als Wörter, die bei einem Sprachschallpegel, der 85 % Sprachverständlichkeit unter Verwendung der individuell verschobenen psychometrischen Funktion entspricht, nicht verstanden wurden. Analog dazu wurden „leichte Items“ festgelegt als Wörter, die bereits bei einem Pegel, der 30 % Verständlichkeit entspricht, richtig verstanden wurden. Die Ergebnisse sind in Abb. 5.9 und Tabelle 5.4 veranschaulicht. Sie geben den prozentualen Anteil

Tab. 5.4 Leicht- und schwerverständliche Wörter des Testmaterials. Tabellarische Auflistung von Abb. 5.9

Wort	Affe	Apfel	Auge	Auto	Bär	Ball	Baum	Bett	Boot
leicht [%]	0,2	0,5	0,2	0,7	1,1	0,5	0,7	0,2	0,2
schwer [%]	0,7	0,5	0,5	0,0	0,2	1,4	1,1	0,2	0,9
Wort	Butter	Eier	Ente	Gurke	Haare	Hase	Hose	Licht	Milch
leicht [%]	0,7	0,5	0,2	0,7	0,2	0,2	0,2	0,7	0,2
schwer [%]	0,5	0,2	0,2	0,0	1,6	0,5	0,9	1,1	1,1
Wort	Mund	Nase	Ohr	Opa	Pferd	Tür	Uhr	Wasser	
leicht [%]	0,2	0,2	0,2	0,0	1,1	0,2	0,9	0,7	
schwer [%]	0,9	1,1	3,0	0,5	2,1	0,5	0,7	0,9	

der Messungen an, in denen ein bestimmtes Wort als leicht oder schwer vorkommt. Es sind alle 3 theoretisch möglichen Kombinationen vertreten: Wörter, die schwerer verständlich waren (z.B. Ohr, 3 %, Haare, 1,6 %), Wörter, die besser verstanden wurden (z.B. Auto, 0,7 %, Gurke, 0,7 %, Bär, 1,1 %) und Wörter, die für manche Kinder leicht und für andere schwer verständlich waren (z.B. Apfel je 0,5 %, Baum, 0,7 % / 1,1 %, Pferd 1,1 % / 2,1 %).

Normdaten

Normdaten sind bei der Anwendung des Tests in der klinischen Diagnostik von essenzieller Bedeutung, weil sie es ermöglichen, individuelle Testergebnisse in einen Kontext zu setzen. Hierdurch kann beurteilt werden, ob das individuelle Hörvermögen im Vergleich zur Normgruppe innerhalb akzeptabler Grenzen liegt oder ob es Abweichungen gibt, die auf Hörbeeinträchtigungen hinweisen. Die für den diagnostischen Einsatz des MATCH ermittelten Normwerte und Kennzahlen sind in Tabelle 5.5 zusammengefasst und in Abb. 5.10 mit den jeweiligen Ergebnissen für die 3 Altersgruppen grafisch dargestellt. Das Ergebnis der adaptiven Messung entspricht aufgrund seiner Pegelschritte einer SVS von 71,4 %.

Zur besseren Vergleichbarkeit mit der Literatur wurden zusätzlich die Werte der 50 %-Schwelle für jede Altersgruppe berechnet und angegeben. Sie betragen für Kinder unter 4,25 Jahren 29,3 dB, für Kinder zwischen 4,25 und 5,5 Jahren 26,7 dB und für Kinder über 5,5 Jahren 24,6 dB. Die Steigung der Diskriminationsfunktion beträgt an diesem Punkt 2,6 % / dB (< 4,25 Jahre), 3,2 % / dB (4,25–5,5 Jahre) und 4,3 % / dB (> 5,5 Jahre).

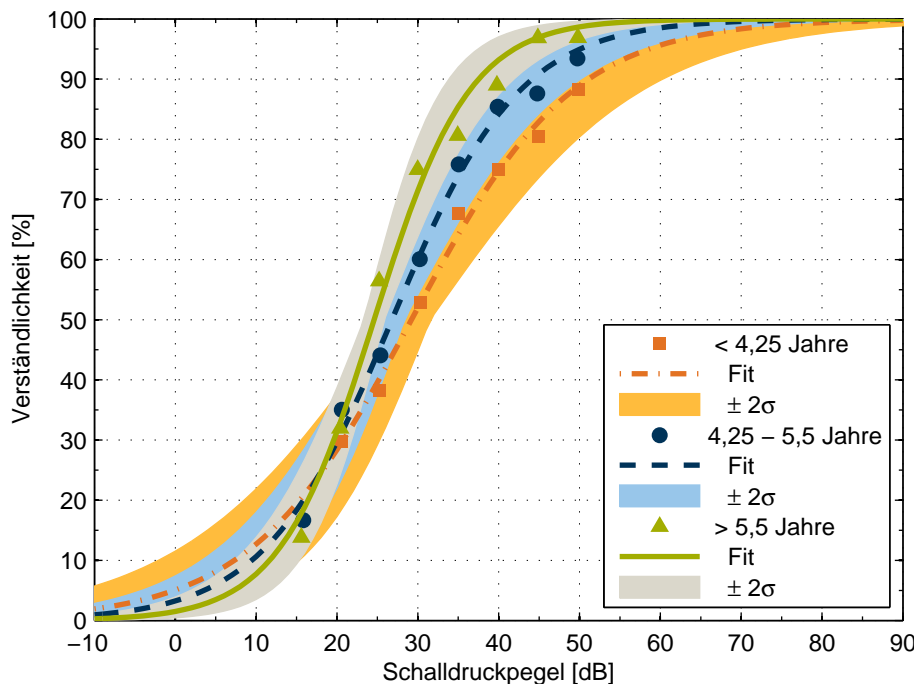


Abb. 5.10 Psychometrische Funktion der Sprachverständlichkeit und Schätzung der Diskriminationsfunktionen unter Verwendung des logistischen Modells für die 3 Altersgruppen. Die Schätzungen von ± 2 SD werden als schattierte Bereiche dargestellt

Tab. 5.5 Kennzahlen der angepassten Diskriminationsfunktionen für normalhörende Kinder aller Altersgruppen

Altersgruppe	< 4,25 Jahre	4,25–5,5 Jahre	> 5,5 Jahre
Mittelwert 50 %-Schwelle \pm SD [dB]	29,3 \pm 1,3	26,7 \pm 0,9	24,6 \pm 0,6
Mittelwert Steigung \pm SD bei 50 % Sprachverständlichkeit [%/ dB]	2,6 \pm 0,4	3,2 \pm 0,3	4,3 \pm 0,5
Mittelwert 71,4 %-Schwelle \pm SD [dB]	38,3 \pm 0,8	34,0 \pm 0,7	30,0 \pm 0,6
Mittelwert Steigung \pm SD bei 71,4 % Sprachverständlichkeit [%/ dB]	2,1 \pm 0,3	2,6 \pm 0,3	3,5 \pm 0,4

Bei einer Sprachverständlichkeit von 71,4 % wurden SVS-Werte von 38,3 dB, 34,0 dB und 30,0 dB für die 3 Altersgruppen ermittelt. An dieser Stelle hat die Diskriminationsfunktion der jeweiligen Altersgruppe eine Steigung von 2,1 %/dB, 2,6 %/dB und 3,5 %/dB. In Abb. 5.10 stellen die Symbole Dreieck (\blacktriangle), Kreis (\bullet) und Quadrat (\blacksquare) die psychometrische Funktion der Sprachverständlichkeit für die verschiedenen Altersgruppen dar. Die daraus unter Verwendung des Modells in Gleichung (4.1) geschätzten Diskriminationsfunktionen sind als Linien eingezeichnet. Die schattieren Bereiche umfassen den Bereich von ± 2 SD, gemäß der Schätzung mittels Bootstrapping. Ihre Zahlenwerte sind in Tabelle 5.5 angegeben. Mit zunehmendem Alter nehmen die SVS und SD ab und die Steigung zu. Dies entspricht einer besseren Sprachverständlichkeit der älteren Kinder im Vergleich zu den jüngeren.

Tabelle 5.6 fasst die Ergebnisse der 71,4 %-Schwellen zusammen, die mithilfe des adaptiven Messalgorithmus ermittelt wurden. Sie stimmen gut mit dem Modell der logistischen Diskriminationsfunktionen überein. Der Median und der Interquartilsabstand nehmen mit zunehmendem Alter ab. Ebenso verringert sich die Spannweite zwischen Minimum und Maximum. In Abb. 5.11 sind die Resultate in einem Box-Whisker-Plot nach Altersgruppen getrennt dargestellt. Die 71,4 % SVS unterscheidet sich zwischen 2 Gruppen statistisch signifikant (5 %-Niveau), wenn die 95 % Konfidenzintervalle überlappungsfrei sind. Dies ist für alle Gruppen der Fall.

Tab. 5.6 Zusammenfassung der gemessenen 71,4 %-Schwellenwerte für normalhörende Kinder

Altersgruppe	< 4,25 Jahre	4,25–5,5 Jahre	> 5,5 Jahre
Median 71,4 %-Schwelle (95 %-Konfidenzintervall) [dB]	39,2 (37,9–40,5)	35,5 (34,5–36,5)	30,9 (29,8–32,0)
Unteres – oberes Quartil (Interquartilsabstand) bei 71,4 % Sprachverständlichkeit [dB]	35,2–42,4 (7,2)	32,0–38,3 (6,3)	28,3–34,2 (5,9)
Minimum – Maximum der 71,4 % Sprachverständlichkeitsschwelle [dB]	25,3–52,0	25,1–47,1	22,7–41,9
Ausreißer: Anzahl (Werte [dB])	2 (53,4; 53,9)	6 (22,0; 48,0 (2x); 48,1; 49,6; 52,1)	3 (43,2; 43,8; 43,9)

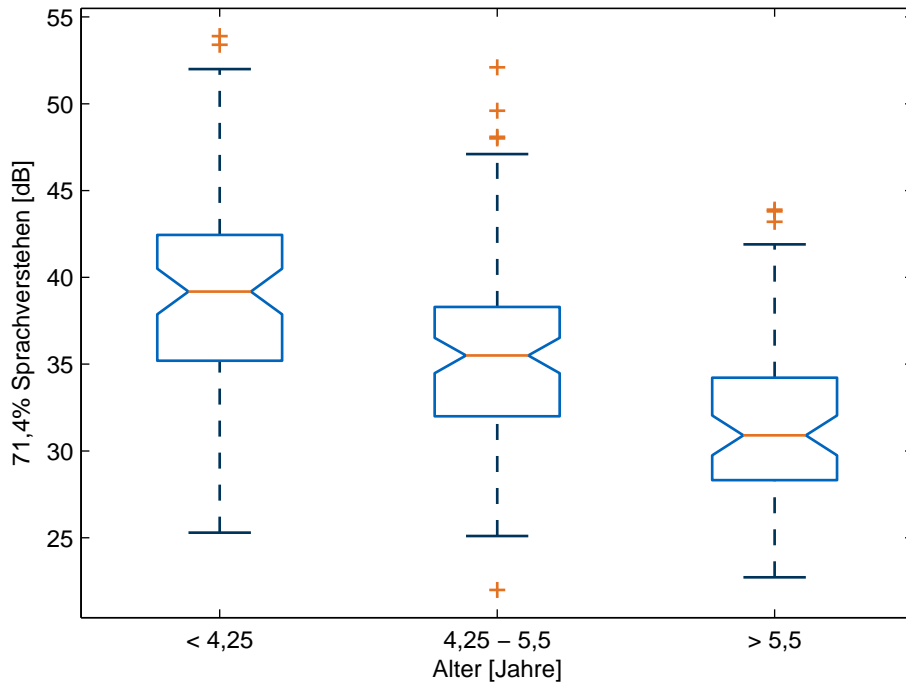


Abb. 5.11 Boxplot der 71,4 % Sprachverständnisschwelle für normalhörende Kinder der 3 Altersgruppen.

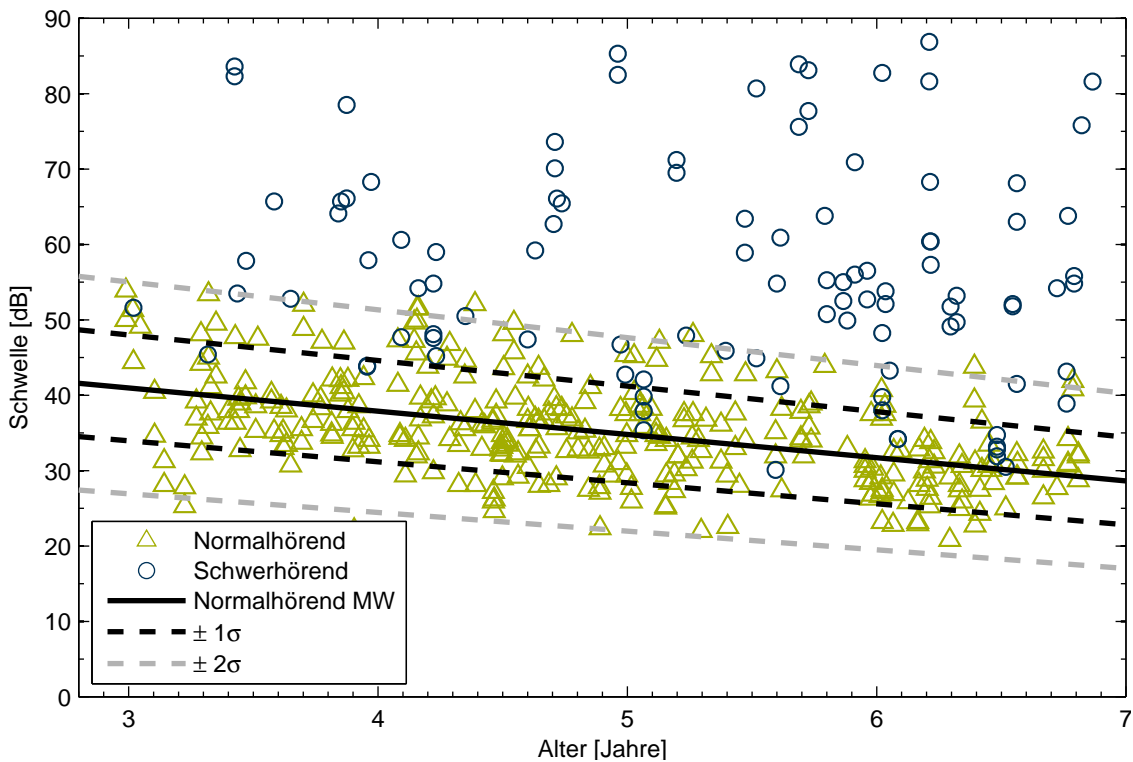


Abb. 5.12 Streudiagramm aller Schwellen (71.4% Sprachverständnis, normal- und schwerhörender Kinder) in Abhängigkeit des Alters. Mittlere Schwelle als lineare Regression der Normalhörenden mit 1 bzw. 2 Standardabweichungen (σ) um den Mittelwert. Nahezu alle Normalhörenden und kaum Schwerhörende befinden sich im 2σ Bereich.

Abb. 5.12 zeigt ein Streudiagramm, das die 71,4% SVS in Abhängigkeit vom Alter darstellt. Die Daten werden durch grüne Dreiecke (\triangle) für normalhörende Kinder und blaue Kreise (\circ) für schwerhörende Kinder repräsentiert. Die durchgezogene schwarze Linie ist die Regressionsgerade der Normalhörenden mit jeweils 2 gestrichelten Linien für ± 1 SD (schwarz) und ± 2 SD (grau). Erwartungsgemäß liegt die überwiegende Mehrheit der Normalhörenden-Ergebnisse im ± 2 SD Bereich. Es ist zu beobachten, dass es Ergebnisse von als schwerhörend eingestuften Kindern gibt, die ebenfalls im Normalbereich liegen. Eine genaue Betrachtung dieser Ergebnisse ergab, dass in diesen Fällen eine geringgradige Schwerhörigkeit vorlag mit einem mittleren Hörverlust bei 0,5, 1, 2 und 4 kHz von höchstens 25 dB. Die Regressionsgerade zeigt einen eindeutigen Trend: Mit zunehmendem Alter der Kinder nimmt die SVS um ca. 10 dB im Alter von ca. 3 bis 7 Jahren ab. Auch die SD verringert sich um ca. 2 dB mit zunehmendem Alter.

Validierung

Die Validierung ist ein notwendiger Schritt, um sicherzustellen, dass das Testverfahren tatsächlich das misst, was es messen soll, und dass die erzielten Ergebnisse reproduzierbar und vertrauenswürdig sind. Eine hohe Korrelation und eine geringe Streuung deuten auf eine hohe Reliabilität des Verfahrens hin. Die Validierung des MATCH erfolgte durch Korre-

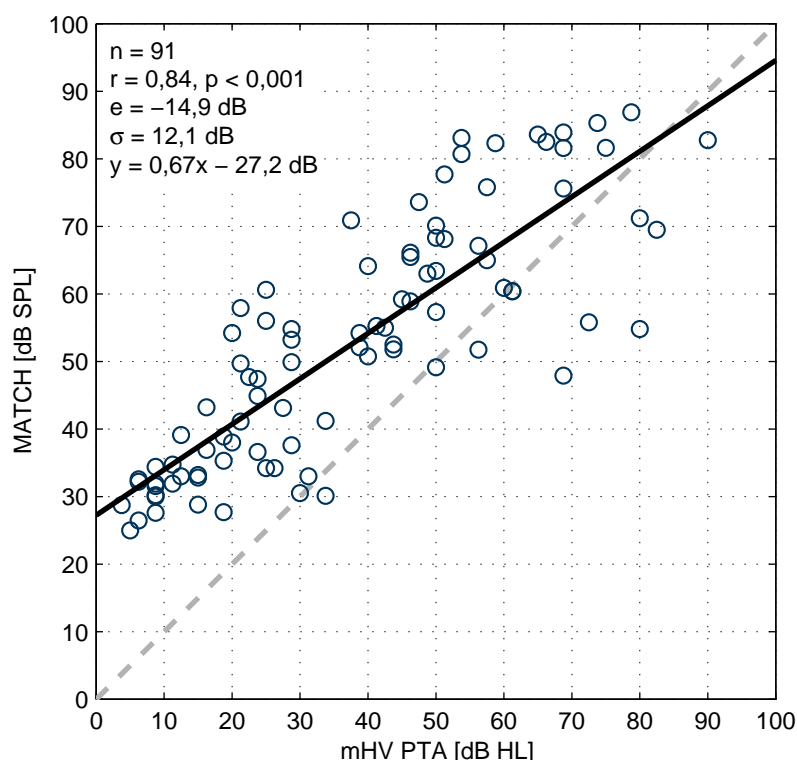


Abb. 5.13 Korrelation zwischen dem mittleren Hörverlust (mHV PTA, 0,5-1-2-4 kHz) und der MATCH-Sprachverständlichkeitsschwelle. Der Mittelwert und die SD der Unterschiede zwischen den Tests liegen bei -14,9 dB und 12,1 dB. Die Korrelation für die 91 Messungen beträgt 0,84. Die meisten Datenpunkte befinden sich oberhalb der Winkelhalbierenden (gestrichelte Linie). Die lineare Regressionsgleichung (durchgezogene Linie) lautet $y = 0,67x - 27,2 \text{ dB}$

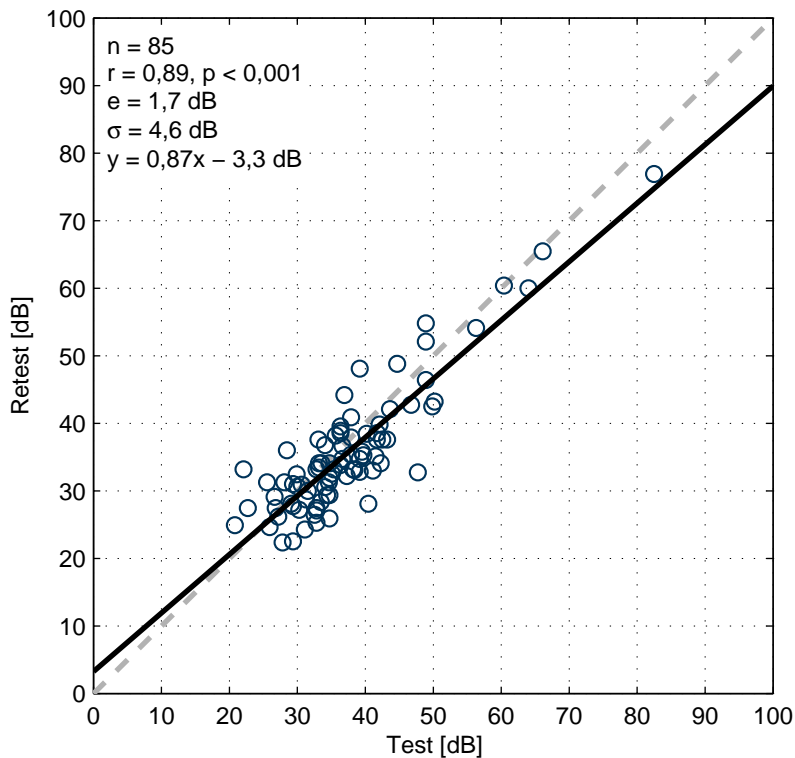


Abb. 5.14 Reliabilität der MATCH Test-Retest-Messungen: Streudiagramm der ersten vs. zweiten Messung. Der Mittelwert und die SD der Unterschiede zwischen den Tests betragen 1,7 dB und 4,6 dB. Die Korrelation für die 85 Messungen beträgt 0,89. Die meisten Datenpunkte befinden sich in der Nähe der Winkelhalbierenden (gestrichelte Linie). Die lineare Regressionsgleichung (durchgezogene Linie) lautet $y = 0,87x - 3,3$ dB.

lationsanalysen, bei denen einerseits die Ergebnisse des MATCH mit den Ergebnissen der Tonaudiometrie verglichen wurden, und andererseits ein Test-Retest-Vergleich durchgeführt wurde.

In Abb. 5.13 wird der Vergleich zwischen der SVS des MATCH und dem mittleren Hörverlust bei den Frequenzen 500 Hz, 1 kHz, 2 kHz und 4 kHz dargestellt. Die Lautheit beim MATCH wird als Schalldruckpegel (dB SPL) angegeben, während sie in der Audiometrie als Hörpegel (dB HL) gemessen wird. Dadurch liegen die Datenpunkte in der Regel oberhalb der Winkelhalbierenden (gestrichelte Linie). Der Mittelwert der Differenzen (e) beträgt -14,9 dB bei einer SD von 12,1 dB. Der Pearson-Korrelationskoeffizient zwischen PTA und MATCH beträgt 0,84 ($p < 0,001$). Die Regressionsgeraden-Gleichung lautet $y = 0,67x - 27,2$ dB.

Im Streudiagramm in Abb. 5.14 sind die Daten des Test-Retest-Vergleichs visualisiert. Hierbei zeigt sich eine mittlere Differenz (e) von 1,7 dB zwischen den Test- und Retest-Werten, während die Standardabweichung (σ) der Differenzen 4,6 dB beträgt. Der Pearson-Korrelationskoeffizient (r) zwischen den beiden Tests ist 0,89 ($p < 0,001$). Die meisten Datenpunkte liegen nahe der Winkelhalbierenden (gestrichelte Linie). Die Regressionsgleichung $y = 0,87x - 3,3$ dB beschreibt die Beziehung zwischen den Test- und Retest-Ergebnissen. Die Steigung ist < 1 das bedeutet, dass ein kleiner Lerneffekt zwischen 1. und 2. Test vorliegt und damit die Schwelle des 2. Tests etwas niedriger ausfällt.

Testzeiten

Eine möglichst kurze Testdauer ist bei der Untersuchung von Kindern besonders wichtig, weil deren Aufmerksamkeitsspanne wesentlich kürzer als diejenige von Erwachsenen ist und die Zuverlässigkeit der Testergebnisse mit zunehmender Testdauer bei nachlassender Konzentration abnehmen kann. Beim MATCH beträgt die mittlere Messdauer zwischen 3 und 4 Minuten pro Ohr, je nach Altersgruppe, wobei die Unterschiede zwischen den Gruppen statistisch signifikant sind. Eine Übersicht über die Messzeiten stellen die Abb. 5.15 und 5.16 und Tabelle 5.7 dar. Abb. 5.15 zeigt einen Box-Whisker-Plot für alle 3 Altersgruppen. Die Zahlenwerte in Form von Median (mit 95 %-Konfidenzintervall), unterem und oberem Quartil (einschließlich Interquartilsabstand), Minimum und Maximum sowie allen Ausreißern sind in Tabelle 5.7 aufgelistet. In allen Gruppen konnten Messzeiten von deutlich unter 2 Minuten erzielt werden. Die größte Testdauer von 8,5 Minuten trat in der Gruppe der jüngsten Kinder

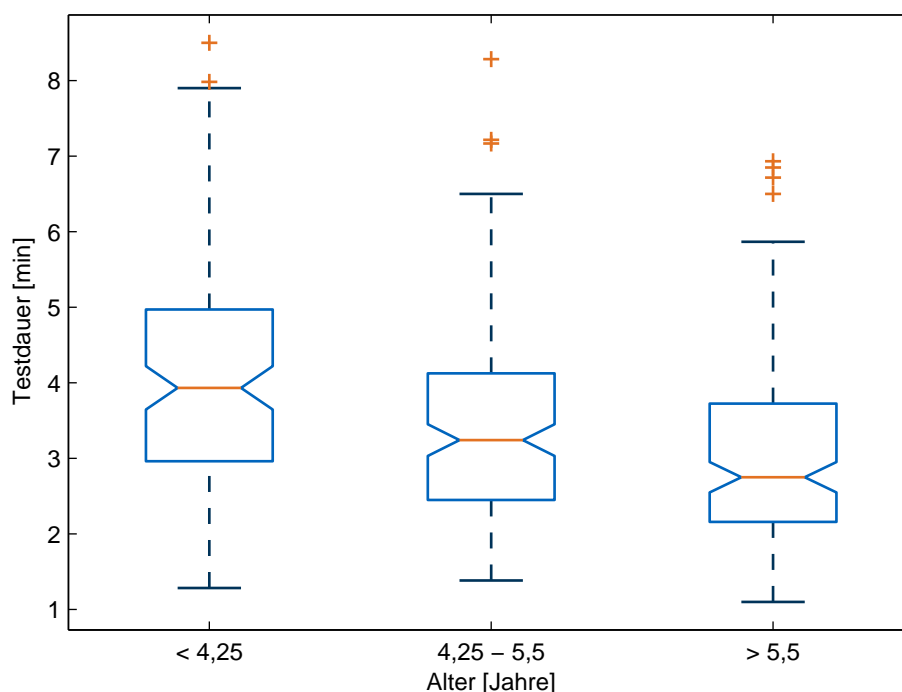


Abb. 5.15 Boxplot der Testdauer aller Kinder für ein Ohr. Jüngere Kinder benötigen tendenziell etwas länger, die Unterschiede sind statistisch signifikant. In allen Altersgruppen waren Testzeiten von weniger als 2 Minuten möglich.

Tab. 5.7 MATCH Testdauer pro Ohr aller Kinder in Abhängigkeit des Alters in Minuten

Altersgruppe	< 4,25 Jahre	4,25–5,5 Jahre	> 5,5 Jahre
Median	3,9	3,2	2,8
(95 %-Konfidenzintervall)	(3,6–4,2)	(3,03–3,5)	(2,5–2,95)
Unteres – oberes Quartil (Interquartilsabstand)	3,0–5,0 (2,0)	2,5–4,1 (1,6)	2,2–3,7 (1,6)
Minimum – Maximum	1,3–7,9	1,4–6,5	1,1–5,9
Ausreißer: Anzahl (Werte)	3 (8,0 (2x); 8,5)	3 (7,2 (2x); 8,3)	4 (6,5; 6,7; 6,8; 6,9)

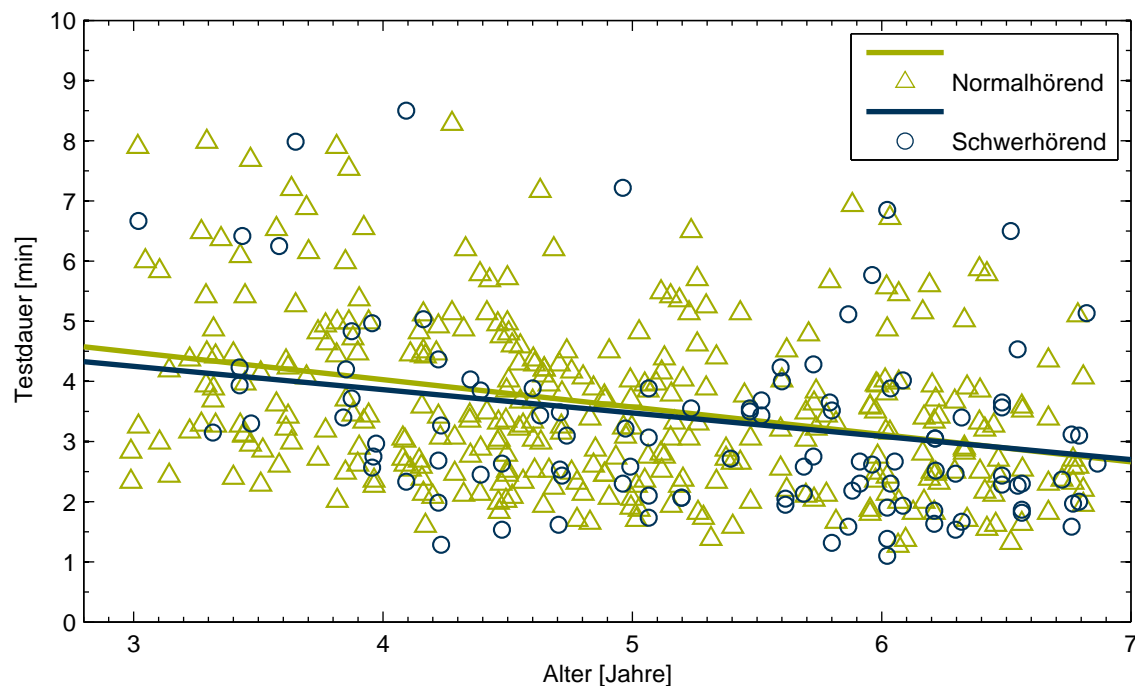


Abb. 5.16 Streudiagramm der Testdauern in Abhängigkeit des Alters. Bei normal- und schwerhörnden Kindern verringert sich die Testdauer mit zunehmendem Alter.

auf. Das Streudiagramm in Abb. 5.16 veranschaulicht die Verteilung sämtlicher Testdaten. Ergebnisse der Normalhörenden bzw. Schwerhörnden sind in grün bzw. blau dargestellt. Die Regressionen über die Normalhörenden- sowie über die Schwerhörnden-Daten zeigen eine Abnahme der Testdauer mit zunehmendem Alter um ca. 1 Minute. Es besteht kaum ein Unterschied zwischen Normalhörenden und Schwerhörnden.

Reaktionszeiten

Die Reaktionszeit ist im Rahmen dieser Arbeit die Zeitspanne zwischen dem Ende der Präsentation des akustischen Signals und der darauf folgenden Eingabe per Touchscreen des Messsystems. Bevor eine Eingabe erfolgen kann, muss das Wort gehört, verstanden und das zugehörige Bild auf dem Bildschirm identifiziert und gefunden werden. In den Abb. 5.17 und 5.18 ist ein Streudiagramm bzw. eine Heatmap zu sehen, welche die Reaktionszeiten für jede Stimulus-Darbietung aus allen durchgeführten Messungen veranschaulichen. In beiden Darstellungen ist die Zeit in Abhängigkeit der Differenz des aktuellen Präsentations-Pegels zur individuellen SVS abgebildet. Bedingt durch das verwendete adaptive Verfahren befindet sich die überwiegende Mehrheit der Datenpunkte im Bereich zwischen ± 15 dB. Die Heatmap in Abb. 5.18 zeigt die Dichteverteilung der Datenpunkte. An der Schwelle (Differenz = 0) befindet sich ein Maximum der Reaktionszeit. Die Reaktionszeit ist folglich niedrig bei einfachen Entscheidungen im Falle von deutlich überschwellig gehörten Wörtern, oder wenn gar nichts gehört wurde. In der Nähe der Schwelle gibt es laut diesen Daten eine verlängerte Reaktionszeit, was auf eine erhöhte Höranstrengung hindeuten kann.

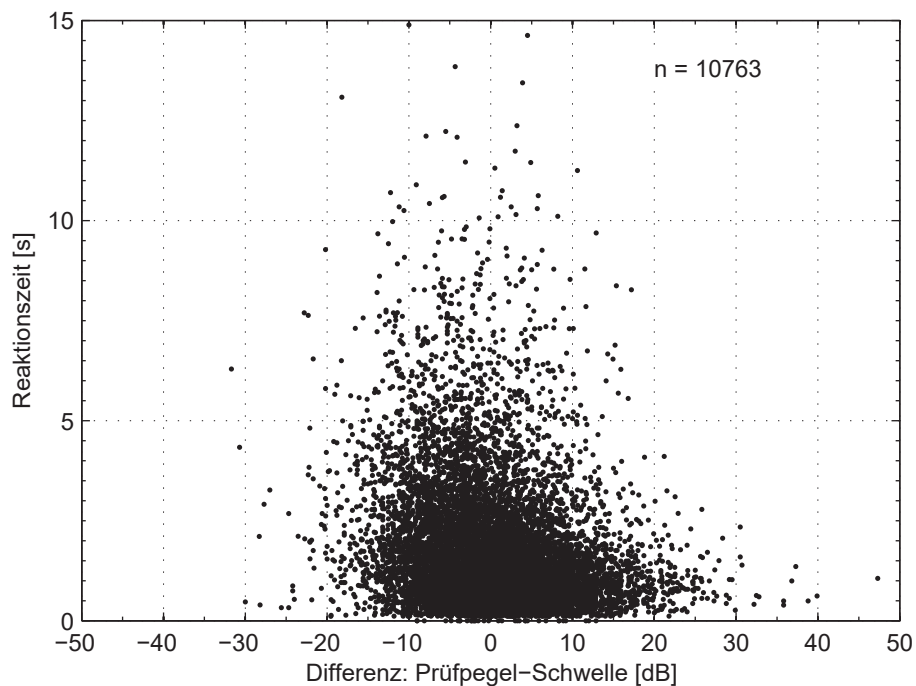


Abb. 5.17 Reaktionszeiten als Indiz für Höranstrengung. Darstellung als Streudiagramm

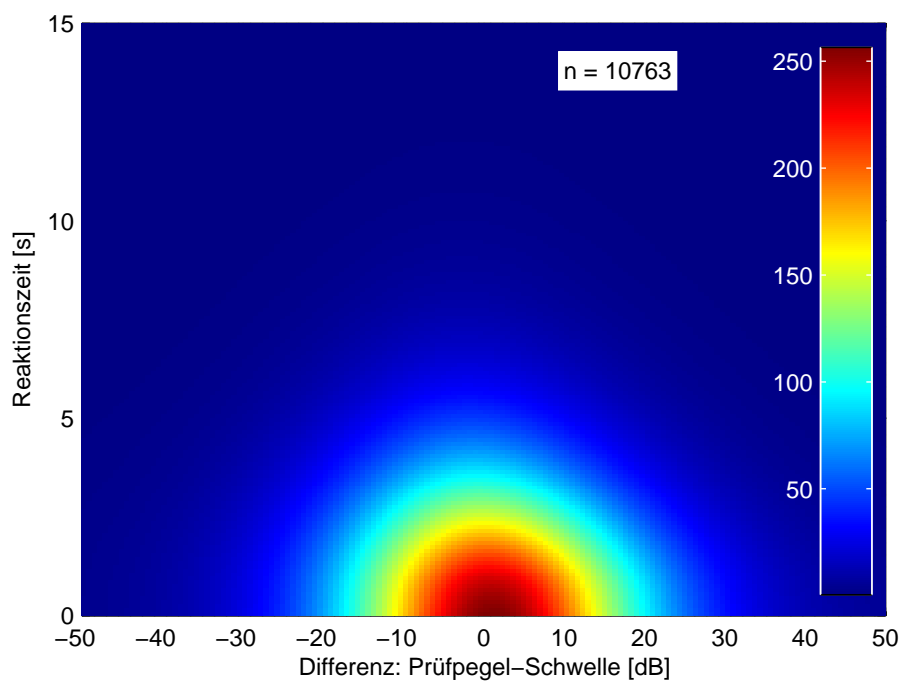


Abb. 5.18 Reaktionszeiten als Indiz für Höranstrengung. Darstellung als Heatmap

Auszüge aus Abschnitt 5.2 und der zugehörigen Diskussion wurden vorab auf der 17. Jahrestagung der Deutschen Gesellschaft für Audiologie präsentiert und veröffentlicht als: Schirkonyer et al. (2020): Der neue Mainzer Sprachtest für Kinder 3–7 Jahre (MATCH). *HNO*, 68(7), 526–533.

5.3. BASD

5.3.1. Studiendesign

In diesem Abschnitt werden die Ergebnisse dreier Studien vorgestellt, bei denen jeweils die in Kapitel 3 beschriebene Messplattform zum Einsatz kam:

1. Erprobung des BASD an 30 Erwachsenen und Erhebung von Normdaten für normalhörende Erwachsene
2. Normierung des BASD an 293 Kindern im Alter zwischen 4;0 und 5;11 Jahren
3. Validierung des BASD an 158 Kindern im Alter zwischen 4;0 bis 4;5 Jahre, die in 3 Gruppen eingeteilt wurden: (a) sprachlich regelrecht entwickelte Kinder, (b) sprachförderbedürftige Kinder und (c) sprachentwicklungsgestörte Kinder. Des Weiteren fand eine Beurteilung dieser Kinder hinsichtlich eines Verdachts auf Auditive Verarbeitungs- und Wahrnehmungsstörungen (AVWS) statt.

Alle nachfolgend beschriebenen Untersuchungen wurden von der jeweils zuständigen Ethik-Kommission genehmigt und sind im Einklang mit nationalem Recht sowie gemäß der Deklaration von Helsinki von 1975 (in der aktuellen, überarbeiteten Fassung) durchgeführt worden.

Erwachsenenstudie

Im Rahmen einer Untersuchung der Änderungen der peripheren und zentralen Schallverarbeitungsleistungen mit dem Alter (Pürner, Schirkonyer & Janssen, 2022) wurde der BASD-Schwellenalgorithmus im Labor der Experimentellen Audiologie des Klinikums rechts der Isar (München) an 30 jungen Erwachsenen im Alter zwischen 18;9 und 26;9 Jahren (Mittelwert 24;2, SD: 1;8) erprobt. Dabei wurden auch exemplarische Normwerte erhoben für die Subtests Frequenz-, Hörpegel- und Tondauer-Diskriminierung bei Verwendung der Signalpräsentationsarten binaural, dichotisch und interaural. Alle Messungen fanden in einer schallisolierten Audiometrikabine statt. Die Versuchspersonen waren normalhörend mit Hörschwellen von maximal 20 dB HL. Zur Übung und Gewöhnung an die Aufgabenstellung wurde immer mit der binauralen Messung begonnen. Die dichotischen und interauralen Messungen wurden auf jedem Ohr durchgeführt mit randomisierter Reihenfolge hinsichtlich Darbietungsart und Ohr. Aufgrund der binauralen zentralen Verschaltung sind die Ergebnisse von rechtem und linkem Ohr nicht statistisch unabhängig, sodass aus den beiden Werten für das rechte und linke Ohr ein einziger Wert durch Mittelwertbildung für die Darbietungsarten binaural und dichotisch berechnet wurde. Die Referenztonfrequenz betrug 1 kHz. Als Darbietungspegel wurde 40 dB SL (Sensation Level) gewählt d.h. er lag 40 dB oberhalb der 1 kHz-Hörschwelle der Versuchsperson. Die Startdifferenzen zwischen Referenzton und Testsignal wurden für Erwachsene deutlich überschwellig eingestellt: Frequenz-Differenz $\Delta f = 0,05$ Oktaven; Hörpegel-Differenz $\Delta L = 10$ dB; Tondauer-Differenz $\Delta t = 100$ ms. Ein Test wurde nach Erreichen von 6 Umkehrpunkten oder spätestens nach 42 Versuchen, sofern bis dahin keine Schwelle ermittelbar war, beendet.

Kinderstudien

Die BASD-Kinder-Messungen wurden im Rahmen eines BMBF Projektes zur Etablierung von Screening-Verfahren zur flächendeckenden Erfassung des Sprachstandes 4;0 bis 4;5 jähriger Kinder durchgeführt (Neumann, 2012). An den Messungen zur Normierung des BASD nahmen 293 Kinder im Alter zwischen 4;0 und 5;11 Jahren teil, die monolingual deutschsprachig aufgewachsen sind. Alle waren normalhörend mit Hörschwellen von maximal 20 dB HL. Die Kinder wurden dem Alter nach in 4 Gruppen eingeteilt: 4;0 bis 4;5 Jahre ($n = 143$), 4;6 bis 4;11 Jahre ($n = 77$), 5;0 bis 5;5 Jahre ($n = 50$), 5;6 bis 5;11 Jahre ($n = 23$). Die Messungen fanden in Kindertagesstätten des Rhein-Main-Gebietes statt. Dabei wurden Ruheräume genutzt und darauf geachtet, dass der Umgebungslärm 40 dB(A) nicht überschreitet.

Die Messungen zur Validierung des BASD wurden in Audiometrieräumen der Abteilung für Phoniatrie und Pädaudiologie der Ruhr-Universität Bochum durchgeführt. Die gesamte Stichprobe umfasst 189 Kinder im Alter von 4;0 bis 4;5 Jahren. An 158 Kindern wurden Messungen mit der BASD-Testbatterie durchgeführt. Alle waren normalhörend mit Hörschwellen von maximal 20 dB HL. Als Gold-Standard zur Beurteilung hinsichtlich eines Verdachts auf AVWS und für die Zuordnung zu einer der 3 Kategorien – (a) sprachlich regelrecht entwickelt, (b) sprachförderbedürftig, (c) sprachentwicklungsgestört – diente das Urteil zweier ausgewiesener Fachleute auf dem Gebiet. Zunächst fällten sie unabhängig voneinander für jedes Kind ein Urteil auf Basis von Ergebnissen einer umfassenden diagnostischen Testbatterie gemäß der AVWS-Leitlinie (DGPP, 2019) und unter Zuhilfenahme von Sprachaufnahmen des Kindes. Bei Nichtübereinstimmung wurde die Klassifizierung als Konsens im Gespräch gefunden. Durch eine Korrelationsanalyse wurde der Zusammenhang zwischen dieser Klassifizierung und den Ergebnissen der Untertests des BASD untersucht. Unterschiede in den Mittelwerten der BASD-Ergebnisse zwischen den Gruppen (a) bis (c) und AVWS-Verdacht versus kein AVWS-Verdacht wurden mittels t-Test analysiert.

Bei allen BASD Messungen betrug die Referenztonfrequenz 1 kHz. Als Startdifferenzen zwischen Referenzton und Testsignal wurden deutlich überschwellige Werte gewählt: Frequenz-Differenz $\Delta f = 1$ Oktave; Hörpegel-Differenz $\Delta L = 20$ dB; Tondauer-Differenz $\Delta t = 300$ ms. Als Darbietungspegel wurde 45 dB SL für den Hörpegel-Differenz Subtest und 55 dB SL für alle anderen durchgeführten Untertests eingestellt. Damit lag er 45 dB bzw. 55 dB oberhalb der 1 kHz-Hörschwelle der Kinder und war deutlich hörbar, jedoch nicht unangenehm laut. Um sicherzustellen, dass keine lauten Töne die Kinder verschrecken, wurde der Darbietungspegel vom Gerät auf 80 dB HL begrenzt. Aus diesem Grund und wegen der hohen Startdifferenz beim ΔL -Subtest betrug der Darbietungspegel hierbei 45 dB SL. Alle Messungen erfolgten binaural (s. Abb. 4.18). Die Eingewöhnung und das Üben der Testausführung wurden anhand des jeweils ersten Versuchs mit den genannten großen überschwelligen Startdifferenzen vorgenommen. Sofern Kriterien für eine nicht erfolgreiche Durchführung festgelegt wurden, sind sie und der Anteil der davon betroffenen Ergebnisse aufgeführt (siehe z.B. Tabellen 5.11 und 5.13). Darüber hinaus konnte bei manchen Kindern für einzelne Subtests kein Resultat ermittelt werden, weil die Konzentration nachließ, die Zeitplanung im Kindergarten es nicht

mehr zuließ, die Kooperation des Kindes aus verschiedensten Gründen nicht mehr gegeben war, oder das Kind nicht in der Lage war, die Aufgabe durchzuführen.

5.3.2. Ergebnisse

Erwachsenenstudie

In diesem Abschnitt sind die Ergebnisse der BASD Messungen an Erwachsenen für die Unterschiedsschwellen von Frequenz-, Hörpegel- und Tondauer-Differenz bei Verwendung der drei Präsentationsarten binaural, dichotisch und interaural dargestellt. Die Werte für die binaurale Darbietung sind der Vollständigkeit halber aufgeführt, da dies stets die erste Messung zur Übung und Eingewöhnung war. Die Ergebnisse für die dichotische und interaurale Darbietung können als Normwerte für junge Erwachsene betrachtet werden. Zur übersichtlicheren Darstellung sind die Einzelergebnisse eines Untertests jeweils grafisch als Box-Whisker-Plot und als tabellarische Zusammenfassung wiedergegeben (z.B. Abb. 5.19 und Tabelle 5.8).

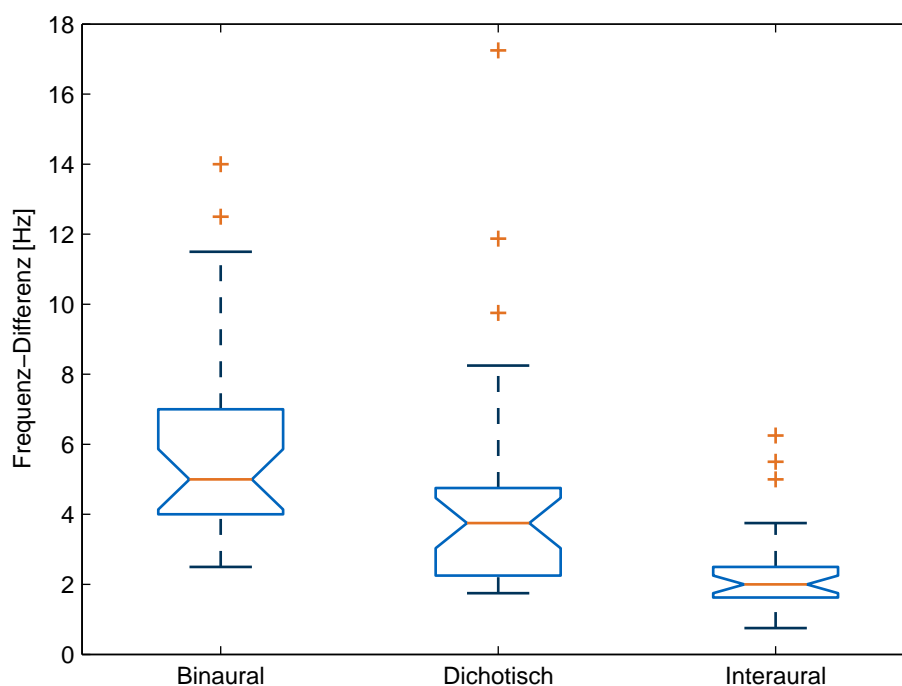


Abb. 5.19 Frequenz-Differenz Unterschiedsschwellen

Tab. 5.8 Frequenz-Differenz Unterschiedsschwellen

Präsentationsart	Binaural	Dichotisch	Interaural
Median	5,0	3,8	2,0
(95%-Konfidenzintervall) [Hz]	(4,1–5,9)	(3,0–4,5)	(1,7–2,3)
Unteres – oberes Quartil (Interquartilsabstand) [Hz]	4,0–7,0 (3,0)	2,3–4,8 (2,5)	1,6–2,5 (0,9)
Minimum – Maximum [Hz]	2,5–11,5	1,8–8,3	0,8–3,8
Ausreißer: Anzahl (Werte [Hz])	2 (12,5; 14,0)	3 (9,8; 11,9; 17,3)	3 (5,0; 5,5; 6,3)

In Abb. 5.19 und Tabelle 5.8 sind die Unterschiedsschwellen des Frequenz-Differenz-Untertests dargestellt. Die niedrigsten Schwellen werden bei den interauralen Messungen erzielt, wobei ein 1002 Hz Testsignal vom 1000 Hz Referenzton unterscheidbar ist. Bei dieser Präsentationsart sind auch der Wertebereich (0,8–3,8 Hz) und die Streuung am kleinsten (Interquartilsabstand 0,9 Hz). Die größte Unterschiedsschwelle (Median 5,0 Hz), der größte Wertebereich (2,5–11,5 Hz) und die größte Streuung (Interquartilsabstand 2,5 Hz) werden bei der binauralen Messung beobachtet. Die Ergebnisse der dichotischen Präsentation liegen dazwischen mit einem Median von 3,8 Hz, einem Wertebereich zwischen 1,8–8,3 Hz und einem Interquartilsabstand von 2,5 Hz. Pro Präsentationsart sind 2–3 Ausreißer vorhanden, wobei es sich außer in einem Fall (binaural 12,5 Hz und interaural 6,3 Hz gehören zur selben Versuchsperson) um unterschiedliche Versuchspersonen handelt. Die 95 %-Konfidenzintervalle der binauralen und dichotischen Messungen überlappen sich. Somit unterscheiden sich die Mediane nicht statistisch signifikant. Der Median der interauralen Messung ist statistisch signifikant von den beiden anderen verschieden.

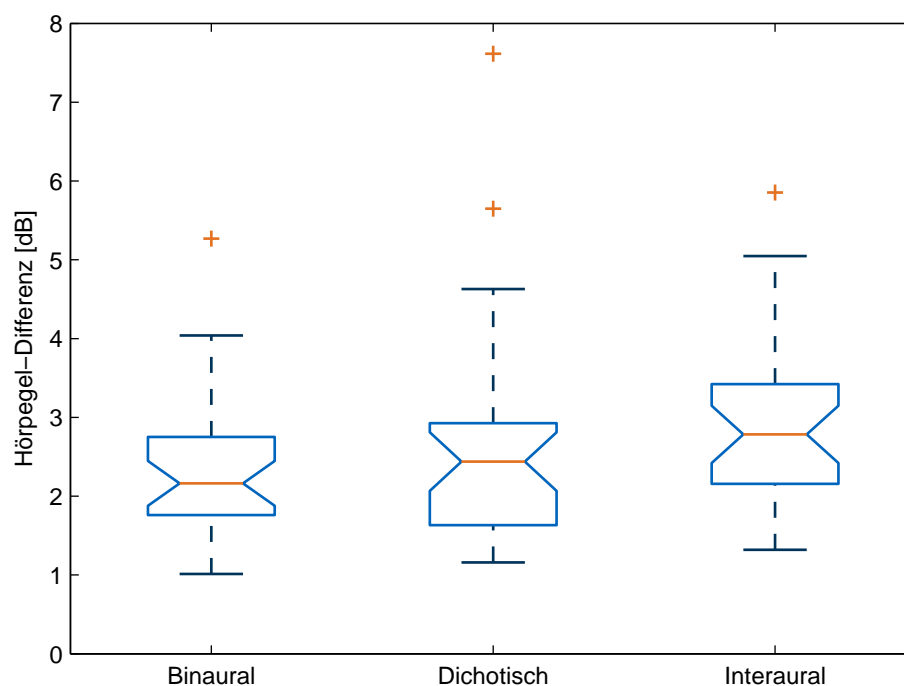


Abb. 5.20 Hörpegel-Differenz Unterschiedsschwellen

Tab. 5.9 Hörpegel-Differenz Unterschiedsschwellen

Präsentationsart	Binaural	Dichotisch	Interaural
Median	2,2	2,4	2,8
(95 %-Konfidenzintervall) [dB]	(1,9–2,45)	(2,1–2,8)	(2,42–3,1)
Unteres – oberes Quartil (Interquartilsabstand) [dB]	1,8–2,8 (1,0)	1,6–2,9 (1,3)	2,1–3,4 (1,3)
Minimum – Maximum [dB]	1,0–4,0	1,2–4,6	1,3–5,0
Ausreißer: Anzahl (Werte [dB])	1 (5,3)	2 (5,6; 7,6)	1 (5,9)

In Abb. 5.20 und Tabelle 5.9 sind die Unterschiedsschwellen des Hörpegel-Differenz-Untertests dargestellt. Hierbei werden die niedrigsten Schwellen bei den binauralen Messungen erzielt, wobei ein Pegelunterschied von 2,2 dB zwischen Testsignal und Referenzton unterscheidbar ist. Bei dieser Präsentationsart sind auch der Wertebereich (1,0–4,0 dB) und die Streuung am kleinsten (Interquartilsabstand 1,0 dB). Die größte Unterschiedsschwelle (Median 2,8 dB) und der größte Wertebereich (1,3–5,0 dB) treten bei den interauralen Messungen auf. Die Ergebnisse der dichotischen Präsentation liegen dazwischen mit einem Median von 2,4 dB und einem Wertebereich zwischen 1,2–4,6 dB. Der Interquartilsabstand beträgt sowohl bei den dichotischen als auch den interauralen Messungen 1,3 dB. Pro Präsentationsart sind 1–2 Ausreißer vorhanden, wobei es sich um unterschiedliche Versuchspersonen handelt. Die 95 %-Konfidenzintervalle aller Messungen überlappen sich. Somit unterscheiden sich die Mediane nicht statistisch signifikant.

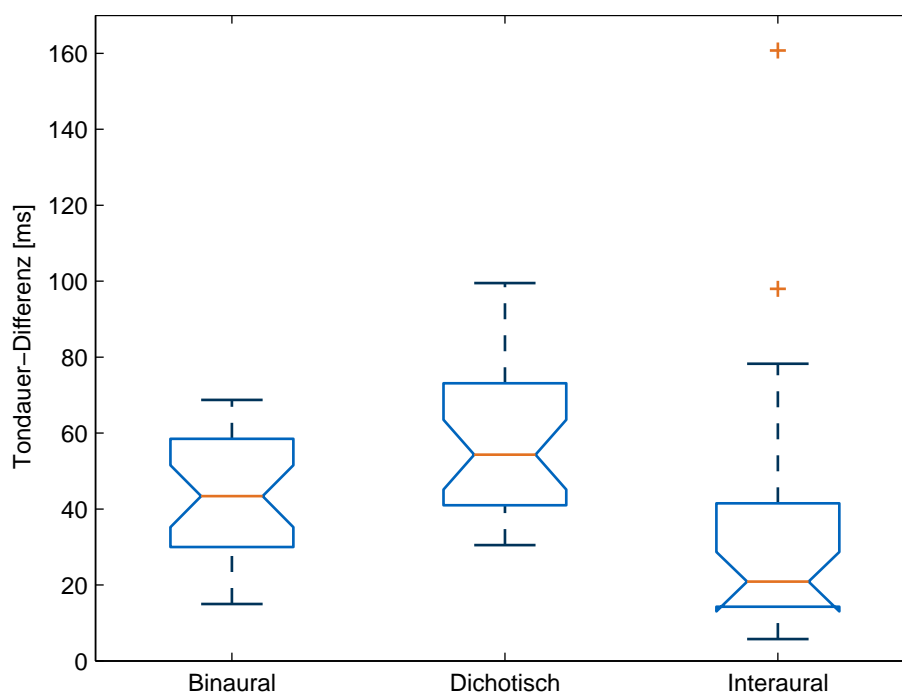


Abb. 5.21 Tondauer-Differenz Unterschiedsschwellen

Tab. 5.10 Tondauer-Differenz Unterschiedsschwellen

Präsentationsart	Binaural	Dichotisch	Interaural
Median	43,4	54,3	20,9
(95 %-Konfidenzintervall) [ms]	(35,2–51,5)	(45,1–63,5)	(13,1–28,7)
Unteres – oberes Quartil (Interquartilsabstand) [ms]	30,0–58,5 (28,5)	41,0–73,1 (32,1)	14,2–41,5 (27,3)
Minimum – Maximum [ms]	15,0–68,8	30,5–99,5	5,8–78,3
Ausreißer: Anzahl (Werte [ms])	0	0	2 (98,0; 160,8)

In Abb. 5.21 und Tabelle 5.10 sind die Unterschiedsschwellen des Tondauer-Differenz-Untertests dargestellt. Die niedrigsten Schwellen werden bei den interauralen Messungen erzielt (Median 20,9 ms). Der Wertebereich und der Interquartilsabstand betragen 5,8–78,3 ms bzw. 27,3 ms. Die größte Unterschiedsschwelle (Median 54,3 ms) und die größte Streuung (Interquartilsabstand 32,1 ms) sind bei den dichotischen Messungen zu finden. Der Wertebereich erstreckt sich zwischen 30,5–99,5 ms. Die Ergebnisse der binauralen Präsentation sind: Median 43,4 ms, Wertebereich 15,0–68,8 ms, Interquartilsabstand 28,5 ms. Ausreißer treten lediglich bei der interauralen Messung auf. Die 95 %-Konfidenzintervalle der binauralen und dichotischen Messungen überlappen sich. Somit unterscheiden sich die Mediane nicht statistisch signifikant. Der Median der interauralen Messung ist statistisch signifikant von den beiden anderen verschieden.

Kinderstudien

In diesem Abschnitt sind zunächst die Ergebnisse der Messungen zur Normierung des BASD an 4;0–5;11 Jahre alten Kindern für die Unterschiedsschwellen von Frequenz-, Hörpegel- und Tondauer-Differenz und die Anzahl der richtigen Unterscheidungen von /ba/ und /ga/ bzw. /ga/ und /ka/ bei Verwendung der Präsentationsart binaural aufgeführt. Die Ergebnisse der Messungen zur Normierung des Tests liegen in zwei verschiedenen Darstellungsformen vor: (1) Median und Interquartilsabstand und (2) Werte für das 80. bis 95. Perzentil in 5er-Schritten. Anschließend folgen die Ergebnisse der Messungen an 4;0 bis 4;5-jährigen Kindern zur Validierung und die Ergebnisse der statistischen Analyse zur Untersuchung des Zusammenhangs zwischen der Klassifizierung auf Basis einer umfangreichen Testbatterie und den Untertests des BASD. Die Daten in Tabellen 5.11 und 5.12 und die Resultate der Klassifikation

Tab. 5.11 Altersabhängige Mediane von Diskriminationsschwellen für Frequenz-, Hörpegel- und Tondauer-Differenz sowie Anzahl der richtigen Eingaben bei der Diskrimination von /ba/ und /ga/ bzw. /ga/ und /ka/. Interquartilsabstand in runden Klammern (), Anzahl Kinder mit gültigen Messungen in eckigen Klammern []. Werte aus Neumann (2012, S. 44, Tab. 7).

Alter [Jahre]	4;0–4;5	4;6–4;11	5;0–5;5	5;6–5;11
Hörpegel [dB]	8,7 (7,6) [143]	8,6 (6,3) [77]	7,6 (8,4) [50]	6,0 (9,0) [23]
Frequenz [Hz]	67,8 (167,1) [103]	69,0 (165,4) [59]	62,3 (137,4) [41]	40,0 (41,4) [21]
% für $\Delta > 500$ Hz	35	23	18	9
Tondauer [ms]	146,0 (123,0) [143]	146,7 (151,6) [77]	113,4 (130,4) [50]	65,3 (104,0) [23]
/ba/ und /ga/ [n/12]	6 (4) [125]	8 (5) [74]	9 (4) [50]	11 (3) [23]
/ga/ und /ka/ [n/12]	6 (4) [122]	5 (3) [73]	6 (4) [50]	10 (4) [23]

und der statistischen Analyse stammen aus dem Abschlussbericht des BMBF-Projektes von Neumann (2012).

Tabelle 5.11 zeigt für alle 4 genannten Altersgruppen und 5 Untertests jeweils den Median, in runden Klammern den Interquartilsabstand und in eckigen Klammern die Anzahl der gültigen Messungen, die für die Berechnung verwendet wurden. Damit eine Messung der Diskriminationsschwelle beim Subtest Frequenz-Differenz als gültig betrachtet wurde, durfte die Schwelle nicht größer als 500 Hz sein. Bei einem Referenzton von 1 kHz entspricht dieser Tonhöhenabstand in der Musik dem Intervall einer Quinte (Frequenzverhältnis 3:2). Der prozentuale Anteil der Messungen, die bei Anwendung dieses Kriteriums entfernt wurden, ist in Tabelle 5.11 ebenfalls angegeben. Mit zunehmendem Alter verbessert sich die Diskriminationsleistung der Kinder, sodass die Unterschiedsschwellen sinken und die Anzahl richtiger Unterscheidungen der Silben in den sprachhaltigen Subtests steigt. Auch nimmt der Anteil an nicht erfolgreich absolvierten Frequenz-Differenz-Tests mit zunehmendem Alter ab.

In Tabelle 5.12 sind die in Halbjahresabständen für 5 Untertests erhobenen Normwerte für das jeweils 80., 85., 90. und 95. Perzentil wiedergegeben. Für die Bestimmung der Werte im Untertest Frequenz-Differenz wurden Messungen mit Resultaten für $\Delta f > 500$ Hz verworfen. In dieser Darstellung der Normwerte ist ebenfalls eine Verbesserung der Leistung sichtbar, mit Ausnahme der 4;6 bis 4;11 Jahre alten Kinder bei den Untertests Frequenz-

Tab. 5.12 Normwerte von Diskriminationsschwellen für Frequenz-, Hörpegel- und Tondauer-Differenz sowie Anzahl der richtigen Eingaben bei der Diskrimination von /ba/ und /ga/ bzw. /ga/ und /ka/ für das jeweils 80., 85., 90. und 95. Perzentil für Kinder im Alter von 4;0–5;11 Jahre. Bei /ba/ und /ga/ bzw. /ga/ und /ka/ wurde auf ganze Zahlen gerundet. Werte aus Neumann (2012, S. 46, Tab. 8).

Alter [Jahre]	4;0–4;5				4;6–4;11			
	80.	85.	90.	95.	80.	85.	90.	95.
Hörpegel [dB]	12,8	13,0	13,4	14,4	11,8	12,8	13,4	13,9
Frequenz [Hz]	198	231	296	354	218	230	344	404
Tondauer [ms]	225	236	249	260	240	249	254	288
/ba/ und /ga/ [n/12]	3	3	2	1	4	4	3	2
/ga/ und /ka/ [n/12]	3	3	2	1	3	3	2	1

Alter [Jahre]	5;0–5;5				5;6–5;11			
	80.	85.	90.	95.	80.	85.	90.	95.
Hörpegel [dB]	11,7	12,2	12,7	13,0	11,7	12,2	12,7	13,0
Frequenz [Hz]	171	198	223	255	57	60	70	118
Tondauer [ms]	140	146	158	162	139	145	157	164
/ba/ und /ga/ [n/12]	5	4	5	3	7	6	6	5
/ga/ und /ka/ [n/12]	3	3	2	1	5	5	4	3

und Tondauer-Differenz. Bei diesen Untertests schneiden die Kinder dieser Altersgruppe am schlechtesten ab.

In der Stichprobe für die Validierung des Tests kamen die Fachleute auf Grundlage der Referenzmessungen zu folgendem Ergebnis: Von 189 Kindern wurden 120 (63,5 %) als sprachlich regelrecht entwickelt eingestuft, bei 37 Kindern (19,6 %) wurde ein Sprachförderbedürfnis festgestellt und 32 (16,9%) Kinder wiesen einen Sprachtherapiebedarf wegen einer Sprachentwicklungsstörung auf. Zur Beurteilung hinsichtlich eines Verdachtes auf AVWS lagen von 188 der Kinder Ergebnisse von Referenzmessungen vor. Auf dieser Basis wurde von den Fachleuten bei 13 Kindern ein Verdacht auf AVWS festgestellt. Von diesen 13 Kindern wurden lediglich 2 als sprachlich regelrecht entwickelt eingestuft. Bei 4 wurde ein Sprachförderbedürfnis und bei 7 eine Sprachentwicklungsstörung festgestellt. Die Rangkorrelation nach Spearman zwischen AVWS und Sprachentwicklung beträgt $\rho = 0,3$ ($p < 0,001$). Es besteht folglich ein Zusammenhang zwischen den beiden Diagnosen.

Mittels Korrelationsanalyse wurde kein statistisch signifikanter linearer Zusammenhang zwischen Ergebnissen der AVWS-Klassifizierung auf Basis der Referenz-Testbatterie und den einzelnen Subtests des BASD gefunden. Gemäß t-Test unterscheiden sich allerdings die Ergebnisse der Diskrimination von /ga/ und /ka/ von Kindern mit AVWS-Verdacht ($n = 8$; durchschnittlich 4 Richtige von 12) von den Ergebnissen der Kinder ohne AVWS-Verdacht ($n = 134$; durchschnittlich 5,74 Richtige von 12) statistisch signifikant ($p = 0,028$).

Zwischen der Klassifizierung der Sprachentwicklung und den BASD-Subtests bestehen gemäß Rangkorrelation nach Spearman statistisch signifikante Korrelationen mit den Untertests Frequenz-Differenz ($\rho = 0,24$; $p = 0,014$), Diskrimination von /ba/ und /ga/ ($\rho = 0,27$; $p = 0,001$) und Diskrimination von /ga/ und /ka/ ($\rho = 0,27$; $p = 0,001$). Gemäß t-Test unterscheiden sich auch die Ergebnisse der genannten Subtests von als unauffällig eingestuften Kindern, wie nachfolgend beschrieben, statistisch signifikant von den Ergebnissen der als auffällig eingestuften Kinder. Für die Unterschiedsschwellen der Frequenz-Differenz ergaben sich folgende Ergebnisse: 77 Kinder unauffällig, 27 auffällig, $\Delta f = 111$ Hz verglichen mit $\Delta f = 174$ Hz im Mittel bei auffälligen Kindern, $p = 0,023$. Für die Diskrimination von /ba/ und /ga/ ergaben sich folgende Ergebnisse: 88 Kinder unauffällig, 56 auffällig, im Mittel 7,06 Richtige von 12 verglichen mit 5,52 Richtigen von 12 bei auffälligen Kindern, $p = 0,002$. Für die Diskrimination von /ga/ und /ka/ ergaben sich folgende Ergebnisse: 87 Kinder unauffällig, 55 auffällig, im Mittel 6,26 Richtige von 12 verglichen mit 5,20 Richtigen von 12 bei auffälligen Kindern, $p = 0,001$.

Für eigene Auswertungen im Rahmen dieser Arbeit lagen die Daten der BASD-Testbatterie von 75 Kindern im Alter zwischen 4;0 und 4;5 Jahren, die als sprachlich regelrecht entwickelt eingestuft wurden, vor. In einem Fall waren nur die Ergebnisse der Untertests mit sprachhaltigen Stimuli vorhanden. Von 16 Kindern fehlten die Ergebnisse der Untertests mit sprachhaltigen Stimuli. Bei der Auswertung der sprachfreien Untertests wurden zur Festlegung von

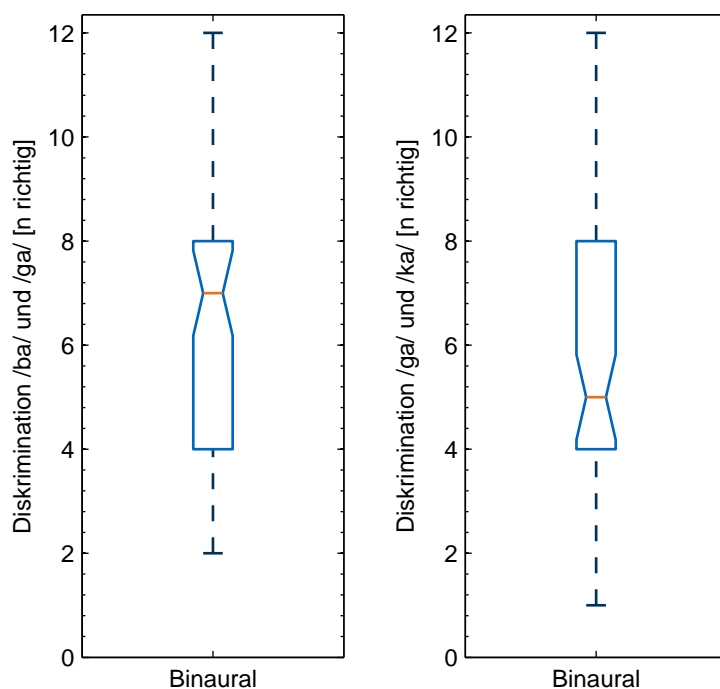


Abb. 5.22 Kinderergebnisse für sprachhaltige Tests

Ausreißern bzw. zur Identifikation von Kindern, die die Aufgabe wahrscheinlich nicht vollständig verstanden hatten und daher möglicherweise schlechtere Ergebnisse lieferten als es ihrer wirklichen Leistung entsprechen würde, die nachfolgend beschriebenen Kriterien angewendet: (1) für den Untertest der Frequenz-Differenz wurde das musikalische Intervall von einer großen Terz (Frequenzverhältnis 5:4) als Grenze festgelegt. Das entspricht bei der Referenztonfrequenz von 1 kHz einem Frequenzunterschied von 250 Hz; (2) bei der Tondauer-Differenz wurden Werte, die größer als 249 ms waren, als Ausreißer betrachtet. In diesen Fällen gab es mehrfach falsche Eingaben bei der Unterscheidung eines 100 ms langen Tons von einem 400 ms langen Ton, wobei beide als identisch statt unterschiedlich angegeben wurden. Eine Differenz von 300 ms war das Maximum, das beim Test auftreten konnte. Das Gleiche gilt analog (3) für die Hörpegel-Differenz. Hierbei wurden Werte, die größer als 14 dB waren, als Ausreißer eingestuft, da im Verlauf der Messung die maximale Differenz vom Gerät auf

Tab. 5.13 Kinderergebnisse für sprachhaltige Tests

Diskrimination Anzahl richtige Eingaben	/ba/ und /ga/	/ga/ und /ka/
Median	7	5
(95 %-Konfidenzintervall)	(6,2–7,8)	(4,2–5,8)
Unteres – oberes Quartil (Interquartilsabstand)	4–8 (4)	4–8 (4)
Minimum – Maximum	2–12	1–12
Gültige Messungen	59	59

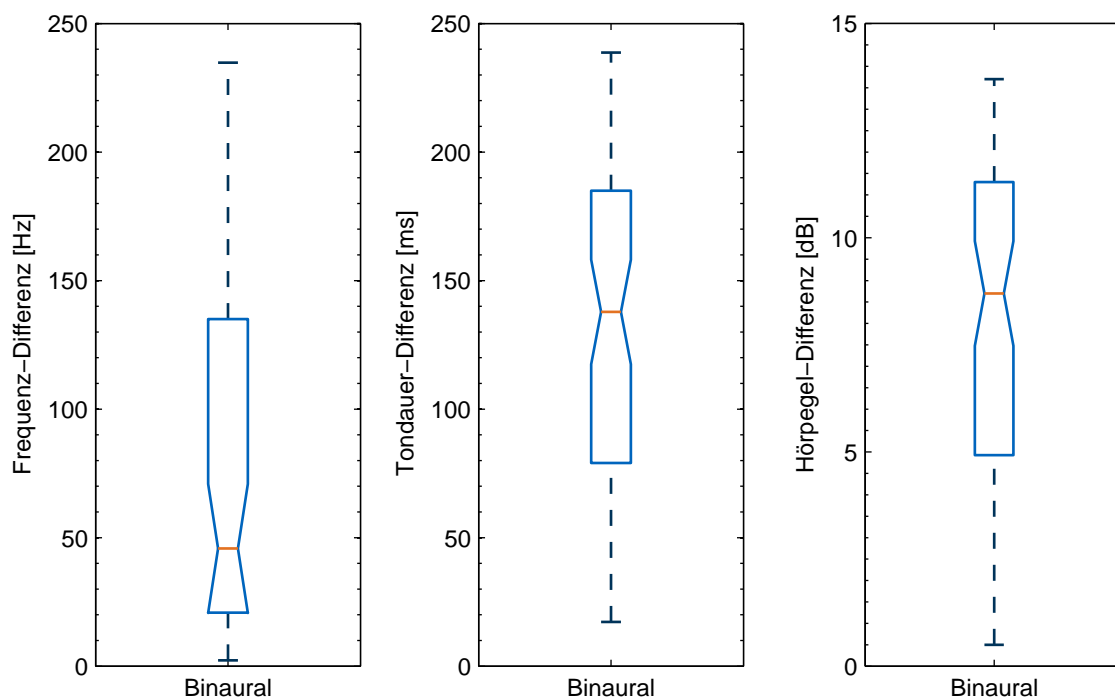


Abb. 5.23 Kinderergebnisse für sprachfreie Tests

15 dB begrenzt wurde. Der Anteil an Ausreißern ist als Zahlenwert und prozentualer Anteil in Tabelle 5.14 angegeben.

In Abb. 5.22 und Tabelle 5.13 sind die Ergebnisse der sprachhaltigen Untertests dargestellt. Die Mediane 7 für /ba/ und /ga/ bzw. 5 für /ga/ und /ka/ liegen im Bereich der Normwerte aus Tabelle 5.11, weichen allerdings vom dort genannten Wert 6 ab. Mangels Angabe eines Konfidenzintervalls für die Normdaten ist nicht zweifelsfrei prüfbar, ob die Abweichung statistisch signifikant ist. Aufgrund des verhältnismäßig großen Interquartilsabstands ist es jedoch wahrscheinlich, dass der Unterschied der Mediane zwischen Normdaten und dieser Stichprobe nicht statistisch signifikant ist. Der Interquartilsabstand ist 4 und somit in beiden

Tab. 5.14 Kinderergebnisse für sprachfreie Tests

Differenz	Frequenz (f) [Hz]	Tondauer (t) [ms]	Hörpegel (L) [dB]
Median	45,8	137,8	8,7
(95%-Konfidenzintervall)	(20,7–70,9)	(117,5–158,1)	(7,5–9,9)
Unteres – oberes Quartil (Interquartilsabstand)	20,8–135,0 (114,2)	79,1–185,0 (105,9)	4,9–11,3 (6,4)
Minimum – Maximum	2,3–234,8	17,2–238,7	0,5–13,7
Ausreißer Kriterium	$\Delta f > 250$	$\Delta t > 249$	$\Delta L > 14$
Ausreißer Anteil	22 (30,1 %)	7 (9,5 %)	4 (5,4 %)
Gültige Messungen	51	67	67

Stichproben gleich. Das Minimum ist innerhalb der in Tabelle 5.12 angegebenen Werte für das 95. Perzentil. Es zeigt sich insgesamt eine Übereinstimmung mit den Normwerten.

In Abb. 5.23 und Tabelle 5.14 sind die Ergebnisse der sprachfreien Untertests dargestellt. Die Mediane liegen im Bereich der Normwerte aus Tabelle 5.11, wobei die Ergebnisse für Unterschiedsschwellen der Frequenz- und Tondauer-Differenz etwas niedriger ausfallen. Die Normwerte liegen jedoch im Bereich des Konfidenzintervalls dieser Stichprobe. Auch die Streuung der Ergebnisse gemessen als Interquartilsabstand ist etwas niedriger als in der Normstichprobe. Der Anteil an Ausreißern ist in dieser Stichprobe gering und selbst mit dem etwas strengeren Kriterium im Frequenz-Differenz Untertest im Vergleich zur Normstichprobe mit ungefähr 30 % gegenüber 35 % kleiner. Das Maximum ist innerhalb der in Tabelle 5.12 angegebenen Werten für das 95. Perzentil teilweise sogar innerhalb des 90. Perzentils. Es zeigt sich insgesamt eine gute Übereinstimmung mit den Normwerten.

*Auszüge aus Abschnitt 5.3 wurden vorab auf Tagungen präsentiert (DGPP Jahrestagung 2012, EUHA 2014) und veröffentlicht als: Pürner, Schirkonyer und Janssen (2022): Changes in the peripheral and central auditory performance in the elderly – A cross-sectional study. *Journal of Neuroscience Research*, 100(9), 1791–1811. sowie Pürner (2019): Änderungen der peripheren und zentralen Schallverarbeitungsleistungen mit dem Alter, *Dissertation, TU München*. und Neumann und Euler (2013): Kann ein Sprachstandsscreening zwischen dem Bedarf für Sprachförderung und für Sprachtherapie trennen? in A. Redder & S. Weinert (Hrsg.), *Sprachförderung und Sprachdiagnostik : Interdisziplinäre Perspektiven*, (S. 174–198), Waxmann.*

6. Diskussion

Im Rahmen dieser Dissertation wurden drei neue psychoakustische Testverfahren zur Untersuchung des Hörorgans von Kleinkindern entwickelt, implementiert und in Studien erprobt. Dabei handelt es sich um ein Verfahren zur Bestimmung der Reintonschwelle „Multiple-choice Auditory Graphical Interactive Check“ (MAGIC), einen sprachaudiometrischen Test „Mainzer Audiometric Test for Children“ (MATCH) und eine Testbatterie zur Prüfung der zentralen Schallverarbeitung, „Bochum Auditory Speech Discrimination Test“ (BASD). Die Verfahren sind unabhängig voneinander anwendbar und werden in diesem Kapitel in jeweils eigenen Abschnitten behandelt. Als Messsystem diente die in Kapitel 3 vorgestellte Senti/Sentiero Plattform.

6.1. MAGIC

In der vorliegenden Arbeit wurde ein neuer bildgestützter tonaudiometrischer Selbsttest für den Einsatz bei Kindern ab 2 Jahren entwickelt und in einer klinischen Studie validiert. Die Tonaudiometrie spielt in der Diagnostik und Therapie von Hörstörungen eine wichtige Rolle. So erfolgt die Beurteilung, ob eine therapiebedürftige Hörstörung vorliegt, anhand von Reintonschwellen (World Health Organization, 2021). Auch für die Anpassung von Hörgeräten bilden die Ergebnisse der Tonaudiometrie die Basis. Ebenso wird der Erfolg einer Hörgeräteversorgung im ersten Schritt mittels Tonaudiometrie überprüft und nachgewiesen. Daher ist es von großem Interesse, die tonaudiometrische Messung mit möglichst geringer Abweichung von der tatsächlichen Hörleistung und einer geringen Streuung zwischen Test- und Retest-Daten durchzuführen. Für die Abschätzung der erreichbaren Messgenauigkeit bei Kindern können die bei Erwachsenen erzielten Ergebnisse als Obergrenze angesehen werden, da bei Erwachsenen in der Regel Unsicherheiten aufgrund mangelnder Kooperation ausgeschlossen werden können und die Aufgabenstellung verstanden und korrekt umgesetzt wird. Auf der Grundlage dieses Wissens können Kriterien dafür festgelegt werden, welche Veränderungen der Hörschwelle als zufällig und welche als klinisch relevant anzusehen sind (Schlauch & Edward, 2007; Studebaker, 1967).

Frühere Studien an Erwachsenen haben gezeigt, dass die Standardabweichung (SD) bei Test-Retest-Messungen in der Regel etwa 5 dB beträgt (Dobie, 1983; Hickling, 1966; Jerlvall & Arlinger, 1986; Katz et al., 2014; McCormick, 1993; Robinson, 1991). Eine Metaanalyse von Mahomed et al. (2013) über 29 Studien zum Vergleich der Test-Retest-Übereinstimmung zwischen manuellen und automatisierten Tonaudiometrie-Messungen kam zu dem Ergebnis, dass es im Mittel keinen signifikanten Unterschied gibt. Test und Retest unterschieden sich im Mittel über alle Studien um etwa 1 dB bei einer Standardabweichung von 6,1 dB für das manuelle Verfahren und 6,9 dB für das automatisierte Verfahren, wobei dieser Unterschied statistisch nicht signifikant war (Mahomed et al., 2013). Vor allem die Testfrequenz,

der Sitz des Schallwandlers und der zeitliche Abstand zwischen Test und Retest beeinflussen die Streuung (Hickling, 1966; Katz et al., 2014; Mahomed et al., 2013; Robinson, 1991; Schlauch & Edward, 2007; Schmuziger et al., 2004). In Anhang A der Norm über audiometrische Prüfverfahren für die Reintonschwellenaudiometrie (ISO 8253-1, 2011) wird anhand eines Modells die Messunsicherheit für die Ermittlung des Hörschwellenpegels einer Testperson mittels Luftleitungs-Audiometrie abgeschätzt. Sie entspricht in diesem Kontext der SD. Die Berechnung in ISO 8253-1 (2011) erfolgt unter der Annahme, dass keine Verdeckung notwendig ist und keine weiteren Faktoren wie z.B. Hintergrundgeräusche das Ergebnis beeinflussen. Der Wert der Messunsicherheit beträgt 4,9 dB für Hörschwellen bei Frequenzen bis 4 kHz und 6,7 dB bei höheren Frequenzen.

Die Ergebnisse der eigenen Messungen zum Vergleich von MAGIC- und Spielaudiometrie-Messungen ergaben eine mittlere Differenz von 0,5 dB und eine SD von 7,6 dB. In den Test-Retest-Messungen unterschieden sich die beiden Messungen um 1,5 dB. Die SD war 8,0 dB. Alle Zahlen beziehen sich jeweils auf sämtliche gemessenen Hörschwellen im Frequenzbereich von 250 Hz bis 8 kHz. Diese Werte sind zwar etwas höher als die Werte für Erwachsene, liegen aber in dem Bereich, der in der Literatur für Kinder angegeben wird. Stuart et al. (1991) untersuchten normalhörende ($HV \leq 20$ dB HL) Kinder im Alter von 6 bis 13 Jahren und junge Erwachsene bei den Frequenzen 0,25, 0,5, 1, 2, 4 und 8 kHz. Sie verwendeten jeweils einen supraauralen Kopfhörer und Einsteckhörer mit zwei verschiedenen Ohrstöpseltypen (Schaumstoff- und Gummistöpsel). Die SD lag bei den Kindermessungen zwischen 2,6 dB und 8,9 dB und war abhängig von der Frequenz und dem Wandler bzw. Ohrstöpsel. Ein Gesamtwert über alle Frequenzen fehlt in der Veröffentlichung. Bei Erwachsenen war die Spannbreite erwartungsgemäß kleiner und reichte von 2,6 dB bis 7,1 dB. Höhere Werte traten gehäuft bei höheren Frequenzen auf, insbesondere bei 8 kHz (Stuart et al., 1991).

Anstelle oder zusätzlich zu der SD wird in Publikationen der prozentuale Anteil der Messungen angegeben, die um 5 dB oder 10 dB voneinander abweichen. Eikelboom et al. (2013) berichteten als Ergebnis ihrer Studie an 44 Erwachsenen, in der ein automatisiertes Verfahren mit manuell durchgeführter Audiometrie verglichen wurde, dass 86,1 % der Differenzen zwischen den beiden Methoden innerhalb eines Bereichs von ± 10 dB lagen. An einer Studie von Lowell et al. (1956) nahmen 21 Kinder im Alter von 3;6–6;11 Jahren teil. Die longitudinale Studie erstreckte sich über 4 Jahre und umfasste die Testfrequenzen 0,25, 0,5, 1, und 2 kHz. Beim Vergleich der jeweils ersten und letzten Messung wichen 66 % der Differenzen um höchstens 5 dB ab und 89 % lagen im Bereich von ± 10 dB (Lowell et al., 1956). Orchik und Rintelmann (1978) führten Test-Retest-Messungen an 40 normalhörenden Kindern ($HV \leq 20$ dB HL) durch, die in 4 Altersgruppen eingeteilt wurden: 3;6, 4;6, 5;6 und 6;6 Jahre. Die Wiederholungsmessungen fanden innerhalb von einer Woche nach der Erstmessung statt. Mit den beiden jüngsten Altersgruppen wurden Messungen bei den Frequenzen 0,5, 1 und 2 kHz durchgeführt. Die älteren Kinder wurden zusätzlich bei 0,25 und 4 kHz gemessen. In dieser Studie wurde über eine sehr gute Übereinstimmung zwischen Test und Retest berichtet: 93 % der Differenzen lagen innerhalb von ± 5 dB (Orchik & Rintelmann, 1978). Smith

et al. (1988) untersuchten bei 33 Kindern im Alter von 4;1 bis 10;1 Jahren die Unterschiede zwischen audiometrischen Messungen in einer Hausarztpraxis und in einer HNO-Klinik. Die Messungen wurden an beiden Orten am selben Tag bei den Frequenzen 0,5, 1 und 4 kHz durchgeführt. Dabei wurde 1 Kind (3 %) wegen großen Abweichungen zwischen den Messungen von der Datenanalyse als Ausreißer ausgeschlossen. Pro Kind wurde jeweils ein Mittelwert über alle Frequenzen in der Hausarztpraxis und in der HNO-Klinik gebildet. In 62,5 % der Fälle unterschieden sich die Ergebnisse der Hausarztpraxis von denen der HNO-Klinik um nicht mehr als ± 5 dB und in 92,2 % um nicht mehr als ± 10 dB (Smith et al., 1988). Die Ergebnisse der im Rahmen dieser Arbeit durchgeführten MAGIC-Messungen stimmen mit den Ergebnissen aus der Literatur überein. Beim Vergleich zwischen MAGIC und Spielaudiometrie lagen 71,1 % der Differenzen im Bereich ± 5 dB und 89,5 % im Bereich ± 10 dB. Für Test und Retest sind die Werte: 68,2 % der Differenzen im Bereich ± 5 dB und 86,1 % im Bereich ± 10 dB.

Die Kenntnis der erreichbaren Messgenauigkeit ermöglicht die Festlegung von Kriterien für klinisch relevante Veränderungen. Entsprechende Überlegungen reichen weit zurück. Studebaker (1967) schlug einen Ansatz vor, der die Einzelergebnisse und damit jede Testfrequenz isoliert betrachtet. Er nahm an, dass die Differenzen zwischen zwei Messungen normalverteilt mit einer SD von 5 dB sind. Schlauch und Edward (2007) stellten ein multinomiales Modell vor, das die Ergebnisse mehrerer Frequenzen gleichzeitig berücksichtigt, um eine Aussage darüber zu treffen, ob Veränderungen in einem Audiogramm eher zufällig sind oder ob tatsächlich eine Verschlechterung des Hörvermögens eingetreten ist. Darüber hinaus ist es bei der Diagnostik des Hörvermögens üblich, anstelle der Ergebnisse bei einzelnen Testfrequenzen den Mittelwert z. B. über 0,5, 1, 2 und 4 kHz zu verwenden (World Health Organization, 2021) oder die Ergebnisse benachbarter Frequenzen daraufhin zu untersuchen, ob eine Verschlechterung bei allen Frequenzen auftritt (Schlauch & Edward, 2007). Zur Absicherung des Befundes bei Veränderungen über 10 dB wird auch ein sofortiger Retest empfohlen (Katz et al., 2014; Schlauch & Edward, 2007). Dementsprechend sind Audiogrammverläufe, wie in Abb. 5.1 dargestellt, bei einer diagnostischen Untersuchung durch erfahrenes Personal in den Kliniken äußerst selten anzutreffen. Im Rahmen der Studie für die vorliegende Dissertation wurde auf die Untersuchung von Ausreißern mithilfe von Nachmessungen bewusst verzichtet. Deren Anteil lag bei 3 % und zeigte bei tiefen Frequenzen und bei 6 kHz höhere Werte. Auch in der Literatur wird eine erhöhte SD bei diesen Frequenzen bei Kopfhörmessungen erwähnt (Hickling, 1966; Jerlvall & Arlinger, 1986; Schlauch & Edward, 2007).

Der Messalgorithmus des MAGIC fand die Schwelle in wenigen Schritten und die „stummen“ Pegel stellten sicher, dass die Messung nicht einfach durch ständiges Drücken auf den „gehört“-Knopf beendet wurde. Das Konzept der „stummen“ Pegel oder „no sound“ bzw. „control trials“ findet sich auch in den Empfehlungen für die Untersuchung von Kindern unter 2 Jahren wieder (Katz et al., 2014; McCormick, 1993). Auch wenn es sich um einen Selbsttest handelt, ist die Anwesenheit und Unterstützung durch erfahrenes Fachpersonal nach wie vor unerlässlich. Sei es zu Beginn zur Konditionierung, dem mit Abstand wichtigsten Teil der Messung,

oder zur Überwachung und Hilfestellung während der Messung. Der im Gerät implementierte Messablauf entlastet jedoch das Personal, sodass keine Ressourcen durch die Bedienung des Audiometers gebunden sind und die volle Aufmerksamkeit dem Kind gewidmet werden kann.

Die Ergebnisse von Kindern bei subjektiven Messungen sind reproduzierbar und konsistent (Olsho et al., 1988; Orchik & Rintelmann, 1978). In der vorliegenden Arbeit konnte für die untersuchte Altersgruppe kein statistisch signifikanter Zusammenhang zwischen Alter und Test-Retest-Übereinstimmung nachgewiesen werden. Beim Vergleich von MAGIC und Spielaudiometrie besteht jedoch eine schwache negative Korrelation. Die Übereinstimmung der Resultate beider Messmethoden verbessert sich mit zunehmendem Alter.

Das Alter hat jedoch einen Einfluss auf die Messzeit. Zwischen der Messzeit und dem Alter besteht eine moderate negative Korrelation, sodass die Messzeit mit zunehmendem Alter sinkt. Je älter die Kinder sind, desto besser verstehen sie die Aufgabe und ältere Kinder führen sie auch schneller aus.

Das neu entwickelte Messsystem (Kapitel 3) erwies sich als zuverlässig, einfach zu bedienen, flexibel einsetzbar und aufgrund des geringen Gewichts und der kompakten Größe leicht zu transportieren. Der Testablauf über den Touchscreen und die Bilder zur Motivation funktionierten gut, wie die Befragung der Kinder zeigte. Insgesamt war die Akzeptanz der bildgestützten Audiometrie sowohl beim Fachpersonal als auch bei den Kindern sehr hoch. Damit steht neben den bestehenden Methoden eine neue Möglichkeit zur Messung der Reintonschwellen bei Kindern zur Verfügung.

6.2. MATCH

Auf der Grundlage aktueller Forschungsergebnisse zum Wortschatz und zur Sprachentwicklung von Zweijährigen (Suchodoletz & Sachse, 2008b, 2012) wurde ein neuer deutscher Sprachtest für Kinder (MATCH) entwickelt. Das Langzeitsprachspektrum der Sprachaufnahmen zeigte eine bessere Übereinstimmung mit dem „ILTASS W“ als der derzeit häufig verwendete Mainzer I. Auch die Phonemverteilung des MATCH lag näher an der Alltagssprache als die des alten Mainzer I (s. Abb. 4.8).

Aufgrund der im Allgemeinen viel geringeren Ausdauer von Kindern im Vergleich zu Erwachsenen sollten Studien mit Kindern anders durchgeführt werden als mit Erwachsenen. Die Aufmerksamkeitsspanne hängt von vielen Faktoren ab, wie z.B. der Entwicklung des Kindes (Hoffmann & Schäfer, 2020; Martin, 1997; Ptok & Nickisch, 2018; Steffens, 2017), und wird in der Literatur von 10–15 Minuten (Hnath-Chisolm et al., 1998) bis zu 30 Minuten für 2-Jährige angegeben (Sachse & Buschmann, 2016). Jedoch wird an anderer Stelle von Kindern im Alter von 5 bis 9 Jahren berichtet, dass einige von ihnen nach einer Untersuchungsdau-

er von etwa 30 Minuten gegen Ende des Tests Ermüdungserscheinungen und Konzentrationsschwächen zeigten (Steigenberger, 2009). Die Verwendung eines zu anspruchsvollen Designs oder sehr strenge Einschlusskriterien für die Gruppe der Normalhörenden können bei Kindern zu Ergebnissen führen, die für eine bestimmte Altersgruppe nicht repräsentativ sind. Aus diesem Grund wurde in den klinischen Studien der vorliegenden Arbeit ein Paradigma verfolgt, bei dem so viele Versuchspersonen wie möglich einbezogen wurden, indem der Testaufwand reduziert wurde. Dieser Ansatz könnte zu einer größeren Varianz geführt haben, repräsentiert aber die Zielgruppe besser.

Der MATCH unterscheidet sich in mehreren Punkten von bekannten Tests und basiert zum Teil auf anderen Annahmen als die Norm für sprachaudiometrische Verfahren (ISO 8253-3, 2022). Statt vordefinierte Testlisten zu verwenden, werden die Testelemente und Alternativen völlig zufällig präsentiert. Der Test eignet sich daher gut für wiederholte Untersuchungen, da es keine Listen gibt, die sich eine Versuchsperson merken und damit das Ergebnis beeinflussen könnte. Im Rahmen der verwendeten Pegelschritte ist die Verständlichkeit aller Wörter gleich.

Sehr junge Kinder, die sich noch im Spracherwerb befinden, haben häufig Probleme mit der Aussprache. Da der MATCH als „picture-pointing task“ über einen Touchscreen durchgeführt wird, zeigt das Kind auf das entsprechende Testelement. Dadurch wird sichergestellt, dass die Sprachverständlichkeit unabhängig von der Aussprache des Kindes und dem Hörvermögen des Untersuchenden ermittelt wird. Die in der Literatur empfohlene Präsentation von jeweils 4 Bildern gleichzeitig, anstelle einer größeren Anzahl (Elliott et al., 1979; Foster & Haggard, 1987), trägt dazu bei, den Schwierigkeitsgrad angemessen für die Zielgruppe zu gestalten und ermöglicht eine zügige Testdurchführung.

Kinder im Alter von 1;9 bis 2;11 Jahren zeigten trotz Beherrschung des erforderlichen Wortschatzes häufig das Verhalten, nicht auf das Element zu zeigen oder es zu wiederholen, wenn sie nicht sehen können, wer spricht. Dies trat bei dem neuen Testverfahren MATCH auf und ist aus der klinischen Praxis auch vom Mainzer I bekannt. Daher konnten für die klinische Studie nur Daten von Kindern ab 2;11 Jahren erhoben werden. Der randomisierte Testablauf erfordert eine Implementierung des MATCH als Softwareprogramm. Er ist nämlich nicht mittels vorgefertigter Testlisten per CD durchführbar. Die hier ermittelten Bezugswerte für den Test (s. Abschnitt 5.2.2) jeweils für die 3 Altersgruppen (< 4,25 Jahre; 4,25–5,5 Jahre; > 5,5 Jahre) und für den 50 %-Punkt der Diskriminationsfunktion sind: Sprachverständlichkeitsschwelle (SVS) \pm Standardabweichung (SD) [dB] $29,3 \pm 1,3$; $26,7 \pm 0,9$; $24,6 \pm 0,6$; Steigung [%/dB] $2,6 \pm 0,4$; $3,2 \pm 0,3$; $4,3 \pm 0,5$. Sie liegen im gleichen Bereich wie die von anderen bekannten deutschen Sprachtests für Kinder, die in den folgenden Absätzen zusammengefasst werden.

Für den alten Mainzer I lag die SVS bei 21 dB, es wurde aber keine SD angegeben (Biesalski et al., 1974). Beim Oldenburger Kinder-Reimtest wurde eine Steigung der Diskriminations-

funktion von ca. 5,5%/dB angegeben. Die SVS lagen bei ca. 19–22 dB für Kinder der 1. bis 4. Klasse (Wagener et al., 2006). Der Würzburger Kindersprachtest Teil 2, der für Kinder im Alter von 3 bis 5 Jahren entwickelt wurde, weist eine Ähnlichkeit in der Gesamtzahl von Testitems und den getesteten Altersgruppen auf. Die SVS und die Steigung der psychometrischen Funktion betragen 34,2 dB bzw. 4,4%/dB und waren damit etwas höher als beim MATCH (Baumann, 2006; Völter et al., 2005). Leider fehlt die Angabe der SD.

Der Göttinger Kindersprachverständnistest Teil I ist für Kinder im Alter von 3 bis 4 Jahren vorgesehen (Chilla et al., 1976). Nach Chilla et al. basierte der Test auf 20 Wörtern aus dem Grundwortschatz von Kleinkindern im Jahr 1976. Es wurden immer vier Bilder gezeigt, wie beim MATCH, und das Kind zeigte auf das Testitem. Im Gegensatz zum MATCH gab es Bilder ohne Tondateien, die nie geprüft wurden. Die mittlere SVS \pm betrug 35 ± 7 dB. Coninx (2005, 2006) stellte einen Sprachtest (AAST) vor, der dem Paradigma folgte, ihn so einfach wie möglich zu gestalten. Er besteht aus nur sechs Wörtern und wurde für Kinder im Alter von 3 bis 4 Jahren entwickelt. Da die Schwelle bei Erwachsenen auf HL normiert wurde, ist ein Vergleich mit dem vorliegenden Test nicht möglich. Die Standardabweichung der SVS bei dreieinhalbjährigen Kindern betrug 8 dB (Coninx, 2005). Für Erwachsene wird in Deutschland häufig der Freiburger Sprachtest (Hahlbrock, 1953) eingesetzt. Seine SVS liegt bei 29,3 dB und die mittlere Steigung zwischen 40 % und 60 % Verständlichkeit beträgt 3,6%/dB (DIN 45626-1, 1995).

Die Zuverlässigkeit des MATCH wurde mittels Test-Retest-Messungen bestimmt, wobei eine hohe Reliabilität festgestellt wurde. Es bestand eine hohe Korrelation von $r = 0,89$, die hochsignifikant war ($p < 0,001$). Der Vergleich mit der Tonaudiometrie ergab Ergebnisse im erwarteten Bereich. Die hochsignifikante hohe Korrelation ($r = 0,84$; $p < 0,001$) zwischen der SVS des MATCH und der Tonaudiometrie entspricht den in der Literatur beschriebenen Resultaten. Aufgrund fehlender Studiendaten von Kindern wurden für die Beurteilung Vergleichsdaten von Erwachsenen herangezogen. Die Korrelation zwischen Sprachaudiometrie und Tonaudiometrie ist bei Erwachsenen im Allgemeinen hoch und liegt, je nach Studie, zwischen 0,8 bis 0,9 (Martin, 1997) oder nahe bei 1 (American Speech-Language-Hearing Association, 1988). Im Vergleich zur Tonaudiometrie ist die Sprachaudiometrie ein aussagekräftigerer Indikator für hörbedingte Einschränkungen im Alltag. Sie spielt eine entscheidende Rolle in der Differenzialdiagnose von Hörschäden und bei der Hörgeräteanpassung.

Angaben zur Testdauer werden in der Literatur selten gemacht und fehlen auch für nahezu alle hier beschriebenen Verfahren wie den Mainzer I, den Heidelberger CVC, und den OLKI (Biesalski et al., 1974; Billich, 1981; Kliem & Kollmeier, 1995). Auch in Übersichtsartikeln sind keine Testzeiten genannt, jedoch besteht Einigkeit darüber, dass die verfügbare Zeit besonders durch die Aufmerksamkeitsspanne und Kooperationsbereitschaft der Kinder begrenzt ist (Ptok, 2014; Ptok & Nickisch, 2018; Steffens, 2017). Wie lange eine Untersuchung dauert, ist nicht für jeden Einzelfall exakt planbar, jedoch unterliegen die Kliniken einer gewissen Terminplanung, sodass die Kenntnis der voraussichtlichen Untersuchungsdauer hilfreich ist.

Die erzielten Messzeiten von maximal ca. 8 Minuten (8,5 Minuten inkl. Ausreißer) liegen im Bereich der in der Literatur beschriebenen Aufmerksamkeitsspanne der Zielgruppe. Die erwartete Messdauer liegt für die überwiegende Mehrheit des Testkollektivs bei höchstens 5 Minuten pro Ohr. Zum Vergleich finden sich Angaben zu Tests mit Erwachsenen, wobei im Fall einer 30er Liste des Oldenburger Satztests (Wagener et al., 1999) die durchschnittliche Testzeit etwa $6,2 \pm 1,2$ Minuten betrug (Heyn, 2019).

Zur Messung der Höranstrengung bei Erwachsenen wird in der Regel ein „Parallel-Aufgaben-Paradigma“ („Dual-Task-Paradigma“) verwendet. Bei diesem Verfahren führen die Versuchspersonen zwei Aufgaben gleichzeitig aus, wobei die primäre Aufgabe eine auditorische Aktivität wie die Worterkennung in Ruhe oder im Störgeräusch ist. Die sekundäre Aufgabe kann verschiedene Formen annehmen, wie etwa Gedächtnisaufgaben oder Reaktionszeittests. Das Dual-Task-Paradigma basiert auf der Annahme, dass das kognitive System nur über begrenzte Ressourcen verfügt. Folglich können bei geringer Belastung beide Aufgaben fehlerfrei durchgeführt werden, während es bei hoher Belastung zu Fehlern bei der sekundären Aufgabe kommt. (Downs, 1982; Downs & Crum, 1978; Feuerstein, 1992; Fraser et al., 2010; Gosselin & Gagné, 2010; Hick & Tharpe, 2002; Rabbitt, 1966; Rakerd et al., 1996). In der klinischen Routine sind solche komplexen und anstrengenden Untersuchungen selbst bei Erwachsenen nur selten durchführbar. Dennoch können die während der MATCH-Testdurchführung gespeicherten Reaktionszeiten als gute Näherung für die Höranstrengung dienen. Zum einen zeigt sich, dass Versuchspersonen desto sicherer auf einen Reiz reagieren, je höher die Signalintensität ist (Ptok, 2002). Zum anderen haben Pals et al. (2015) nachgewiesen, dass bereits die Reaktionszeit der primären Aufgabe die Höranstrengung ähnlich gut misst wie ein Dual-Task-Paradigma. Somit steht mit den Reaktionszeiten ein zusätzliches Merkmal zur Verfügung, anhand dessen die Zuverlässigkeit des Messergebnisses im Nachhinein beurteilt werden kann, wenn eine Erhöhung der Reaktionszeit mit der Annäherung an die Schwelle korreliert.

6.3. BASD

Im Rahmen dieser Arbeit wurde eine Testbatterie für die Messung der Unterschiedsschwellen von Frequenz-, Hörpegel-, Tondauer- und Amplitudenmodulationsfrequenz-Differenz entwickelt und normiert. Außerdem verfügt die Testbatterie über sprachhaltige Untertests, bei denen jeweils zwei Konsonanten richtig zu erkennen und zu unterscheiden sind. Störungen bei diesen grundlegenden Funktionen können ein Hinweis auf eine nicht regelgerechte Sprachentwicklung oder AVWS sein (Leurpendeur et al., 2006; Nickisch & Kiese-Himmel, 2009; Nickisch et al., 2019; Ptok et al., 2019).

Die Wahrnehmung von Differenzen in der Lautstärke hängt vom Darbietungspegel (Zwicker & Fastl, 1999, S.180), der Stimulus- und Pausenlänge (Zwicker & Fastl, 1999, S.181, S.265ff) sowie dem Testsignal (Zwicker & Fastl, 1999, S. 176, S.181f) ab. Die Wahrnehmung von

Differenzen der Tonhöhe wird ebenfalls vom Darbietungspegel, der Stimulus- und Pausenlänge sowie dem Testsignal beeinflusst (Zwicker & Fastl, 1999, S.186f). Außerdem hat der Messalgorithmus einen Einfluss auf das erzielte Ergebnis (s. Abschnitt 4.3.2 und Abb. 4.20). Bisher gibt es keine Normen in denen verbindliche Vorgaben für Geräte und Verfahren, wie den hier genannten Messungen von Unterschiedsschwellen, gemacht werden. Daher dürfen diese Parameter beim Vergleich von Ergebnissen verschiedener Studien nicht außer Acht gelassen werden.

Für die klinische Diagnostik bei Kindern ab 4 Jahren steht nach Kenntnis des Autors der vorliegenden Dissertation derzeit kein anderes kommerziell erhältliches Messsystem zur Verfügung, als das im Kapitel 3 vorgestellte. In der Literatur wird zwar von einigen wenigen, einander recht ähnlichen Systemen berichtet (im folgenden Gerät *A* (Minning, 2009), *B* (Warnke, 2015) *F* (Fischer & Hartnegg, 2004), *P* (Thie, 2001) und *T* (Heath, 2014) genannt). Davon war jedoch lediglich von Gerät *P* der bestimmungsgemäße Gebrauch psychoakustische Untersuchungen zum Zweck der Diagnose (Thie, 2001) mit Normdaten für Kinder ab 7 Jahren. Derzeit ist ein einziges der oben genannten Systeme (Gerät *B*) einschließlich Normdaten und Zulassung als Medizinprodukt allgemein verfügbar. Dessen Zweckbestimmung ist das Training verschiedener Bausteine der zentralen Verarbeitung und Wahrnehmung im menschlichen Gehirn bei Kindern ab 5 Jahren (Warnke, 2015). Das Gerät *A* ist grundsätzlich als Trainingsgerät für den Heimgebrauch verfügbar (Minning, 2009). Wie eine englische Kopie von Gerät *A* wirkt Gerät *T*. Diese ist ausschließlich im Rahmen einer Kombination aus Hörtraining und Klangtherapie als Hilfsmittel bei der Beratung verfügbar, allerdings ohne eigene Normwerte (Heath, 2014). Das Gerät *F* kommt in einem darauf spezialisierten Labor bei der Diagnose und dem Training von Kindern mit Lese-Rechtschreibstörungen ab einem Alter von 7 Jahren zum Einsatz; es ist nicht kommerziell verfügbar, aber es sind Normwerte vorhanden (Fischer & Hartnegg, 2004). Manche Forschungsgruppen setzen auch selbstgebaute Testausrüstung ein (Bungert-Kahl et al., 2004; Ludwig, 2008). Im weiteren Verlauf werden nur die Geräte *P* (Anfang der 2000er Jahre verfügbares diagnostischen Messsystem) und *F* (stellvertretend für Kombinationsgeräte, die auch in Trainings zum Einsatz kommen) betrachtet.

Mit Gerät *P* können Unterschiedsschwellen für die Frequenz- und Hörpegel-Differenz gemessen werden. Dabei kommt ein Zwei-Intervall-Zwangswahlverfahren zum Einsatz. Es erfolgt eine automatische Präsentation der nächsten beiden Stimuli direkt nach der Eingabe, sodass eine Wiederholung der Präsentation nicht möglich ist. Die Stimulusdauer beträgt 300 ms mit einer Pause von 600 ms zwischen den beiden Stimuli. Die Messung der Frequenz-Differenzschwelle findet mit einem 1 kHz-Ton und einer Startdifferenz von 100 Hz statt. Als Präsentationspegel werden 75 dB genannt. Hörpegel-Differenz-Schwellen werden mit einem nicht näher spezifiziertem Rauschen gemessen. Der Referenzpegel hierbei ist mit 65 dB und die Startdifferenz ist mit 8 dB angegeben. Während der Messung erhält die Testperson nach jeder Eingabe eine Rückmeldung, ob die Eingabe richtig war. Da die Auswahl des Ohres nicht vorgesehen ist, wird angenommen, dass die Messungen binaural (vgl. Abb. 4.18) durchgeführt werden. Als Verfahren zur Schwellenbestimmung wird lediglich genannt, dass es sich

um ein 2-up 1-down Verfahren mit adaptiver Schrittweite handelt, das nach höchstens 60 Versuchen beendet ist. (Ptok & Meisen, 2008; Thie, 2001)

Das Gerät *P* wurde in mehreren Studien eingesetzt und zum Teil im Rahmen von Evaluationen zur Prüfung, inwiefern die damit erzielten Ergebnisse zur Diagnose von AVWS herangezogen werden können, verwendet. Im Folgenden werden die Unterschiedsschwellen \pm SD von Hörpegel- und Frequenz-Differenz (ΔL und Δf) der Kontrollgruppe beschrieben, sofern auch Kinder mit AVWS-Diagnose an der Studie teilnahmen. Ptok und Meisen (2008) führten eine Studie an Schulkindern der 3. und 4. Jahrgangsstufe im Alter zwischen 7 und 11 Jahren durch, mit dem Ziel Normwerte zu erheben. Sie gaben für die von ihnen gemessenen Schulkindern der 3. Jahrgangsstufe ΔL mit $5,5 \pm 4,1$ dB und Δf mit $92,2 \pm 51,8$ Hz an. Die Ergebnisse der Schulkindern der 4. Jahrgangsstufe sind: $\Delta L = 4,0 \pm 3,3$ dB und $\Delta f = 77,9 \pm 65,6$ Hz (Ptok & Meisen, 2008). Nickisch und Kiese-Himmel (2009) untersuchten als Kontrollgruppe Schulkindern der 3. und 4. Jahrgangsstufe im Alter von 8–10 Jahren und verglichen die Ergebnisse mit denen von altersgleichen Kindern mit AVWS-Diagnose. Sie ermittelten die Werte $\Delta L = 6,12 \pm 15,64$ dB und $\Delta f = 49,65 \pm 60,23$ Hz (Nickisch & Kiese-Himmel, 2009). Für die AVWS-Gruppe wird $\Delta L = 4,38 \pm 6,58$ dB angegeben (Nickisch & Kiese-Himmel, 2009). Hier ist nicht auszuschließen, dass eine Vertauschung der Ergebnisse vorliegen könnte. Nickisch et al. (2013) führten eine Studie analog zu Nickisch und Kiese-Himmel (2009) durch, allerdings mit Schulkindern der 2. Jahrgangsstufe (Altersdurchschnitt 7,6 Jahre) und altersgleichen Kindern mit AVWS-Diagnose. Folgende Ergebnisse wurden dabei erzielt: $\Delta L = 3,5 \pm 1,8$ dB und $\Delta f = 59,7 \pm 27,6$ Hz (Nickisch et al., 2013). Entsprechend der Vorgehensweise der in den beiden zuletzt beschriebenen Studien untersuchten Kiese-Himmel et al. (2018) Schulkindern der 1. Jahrgangsstufe (Durchschnittsalter 6,9 Jahre) mit folgenden Ergebnissen: $\Delta L = 3,52 \pm 1,92$ dB und $\Delta f = 77,77 \pm 53,35$ Hz (Kiese-Himmel et al., 2018). Abgesehen von der oben genannten Unregelmäßigkeit bei ΔL in den Daten von Nickisch und Kiese-Himmel (2009) fügen sich ΔL und Δf in ein Muster ein, bei dem die Schwellen mit zunehmendem Alter niedriger werden. Die Werte von Ptok und Meisen (2008) sind höher als diejenigen, die von Kiese-Himmel et al. (2018), Nickisch und Kiese-Himmel (2009) und Nickisch et al. (2013) genannt werden. Allgemein fällt eine verhältnismäßig hohe Standardabweichung auf, die ungefähr in der Größenordnung der jeweiligen Messwerte liegt. In der Kontrollgruppe der 1. Jahrgangsstufe konnten nur 89 % der Kinder die Tests erfolgreich abschließen (Kiese-Himmel et al., 2018).

Gerät *F* ist ebenfalls in der Lage, Unterschiedsschwellen für die Frequenz- und Hörpegel-Differenz zu messen. Dabei kommt ein Zwei-Intervall-Zwangswahlverfahren zum Einsatz. Es erfolgt eine automatische Präsentation der nächsten beiden Stimuli direkt nach der Eingabe, sodass eine Wiederholung der Präsentation nicht möglich ist. Die Stimulusdauer beträgt 300 ms mit einer Pause von 150 ms zwischen den beiden Stimuli. Die Messung der Frequenz-Differenz-Schwelle findet mit einem 1 kHz-Ton und einer Startdifferenz von 100 Hz statt. Als Präsentationspegel werden 65 dB(A) genannt. Die Messung der Hörpegel-Differenz-Schwelle erfolgt mit weißem Rauschen. Der Referenzpegel ist hierbei mit 55 dB(A) und die Startdif-

ferenz ist mit 8 dB angegeben. Da keine Angaben zum gemessenen Ohr gemacht werden, wird angenommen, dass die Messungen binaural (vgl. Abb. 4.18) durchgeführt werden. Während der Messung erhält die Testperson keinerlei Rückmeldung, ob die Eingabe richtig war. Die Unterschiede werden bei jedem Versuch um 10 % verkleinert. Die Messung wird nach 40 Versuchen, oder wenn 3 von 7 hintereinander folgenden Eingaben falsch sind, beendet. Als Schwelle ist der Wert der letzten richtigen Eingabe vor dem ersten Fehler in der Sequenz definiert. (Fischer & Hartnegg, 2004)

Studien, die Normwerte für Gerät *F* beinhalten, liegen von Fischer (2007) und Fischer und Hartnegg (2004) vor. Die Ergebnisse wurden ausschließlich in Form von Diagrammen angegeben (Fischer & Hartnegg, 2004, S. 114, Abb. 3) und (Fischer, 2007, S. 16f, Abb. 2.4 und 2.5), sodass sämtliche Zahlen, die hier wiedergegeben werden aus den Diagrammen abgelesen wurden. Fischer und Hartnegg (2004, S. 114, Abb. 3) zeigen ausschließlich den Median. Aus den beiden genannten Veröffentlichungen ist nicht zweifelsfrei erkennbar, ob die von Fischer (2007, S. 17, Abb. 2.5) genannten Mittelwerte aus denselben Daten berechnet wurden wie die Mediane. Für 8-Jährige konnten die folgenden Werte aus den Abbildungen abgelesen werden: Median $\Delta L = 3,8$ dB und $\Delta f = 35$ Hz; Mittelwert \pm SD $\Delta L = 4,4 \pm 0,2$ dB und $\Delta f = 68 \pm 8$ Hz. Für 20-Jährige lassen sich folgende Werte aus den oben genannten Diagrammen ablesen: Median $\Delta L = 1,8$ dB und $\Delta f = 5$ Hz; Mittelwert \pm SD $\Delta L = 1,4 \pm 0,2$ dB und $\Delta f = 8,5 \pm 1,5$ Hz. Aus der Abbildung (Fischer, 2007, S. 16, Abb. 2.4) geht hervor, dass 53 % der Kinder und 92 % der Erwachsenen den Δf -Test erfolgreich abschließen konnten. Für den ΔL -Test sind die Werte 76 % (Kinder) und 100 % (Erwachsene). Bemerkenswert ist, dass die SD in den Abbildungen mit sehr niedrigen Werten dargestellt wird.

Bei den beiden Geräten *P* und *F* fehlt die Angabe, wie die Darbietungspegel bestimmt wurden und somit bleibt unklar, welchen Hörpegeln die angegebenen Pegel entsprechen. Ein 2-up 1-down-Verfahren, wie es für Gerät *P* angegeben wurde (Ptok & Meisen, 2008), bedeutet nach Levitt (1971), dass der 29 %-Punkt der psychometrischen Funktion bestimmt wird. Daher besteht die Möglichkeit, dass Gerät *P* gegebenenfalls ein 1-up 2-down-Verfahren einsetzt, das den 71 %-Punkt und damit einen Wert misst, der oberhalb der Ratewahrscheinlichkeit von 50 % liegt.

Bungert-Kahl et al. (2004) und Ludwig (2008) verwendeten ein selbstgebautes Messsystem mit einem Drei-Intervall-Zwangswahlverfahren und einer separaten Antwortbox. Es werden keine Angaben gemacht, ob es eine Rückmeldung nach der Eingabe gab. Die Reizdarbietung erfolgte mit 250 ms langen Sinustönen und einer Pause von 750 ms zwischen jeweils zwei Tönen. Vor jeder Untersuchung wurde mit dem System die individuelle Hörschwelle der Versuchspersonen bestimmt. Die Messung der Unterschiedsschwellen erfolgte bei 30 bis 35 dB SL (Bungert-Kahl et al., 2004; Ludwig, 2008). Im Detail gab es Unterschiede zwischen Bungert-Kahl et al. (2004) und Ludwig (2008): Bungert-Kahl et al. (2004) arbeiteten mit der Version II der zugrundeliegenden freiprogrammierbaren Tucker-Davis-Messstation und Ludwig (2008) mit der Version III. Ein weiterer Unterschied ist, dass Bungert-Kahl et al. (2004)

ein 1-up 1-down-Verfahren zur Schwellenbestimmung einsetzen während (Ludwig, 2008) die Schwelle mittels Maximum-Likelihood-Schätzung der Parameter einer psychometrischen Funktion ermittelte.

Ludwig (2008) hat Versuchspersonen im Alter zwischen 6 und 25 Jahren gemessen. Die Ergebnisse liegen als Median sowie unteres und oberes Quartil vor. Sie werden hier für die Stimulusfrequenz 1 kHz wiedergegeben. Für die Gruppe der 6–8-Jährigen und die Darbietungsart monaural sind die folgenden Werte (jeweils Median (unteres – oberes Quartil)): $\Delta f = 21,3$ Hz (13,4–42,1 Hz); $\Delta L = 4,3$ dB (3,0–5,6 dB); $\Delta t = 48,0$ ms (38,2–63,2 ms). Bei Erwachsenen wurden die Darbietungsarten dichotisch und interaural an Personen im Alter zwischen 18 und 19 Jahren gemessen. Monaurale Messungen fanden an 21–25-Jährigen statt. Monaurale Messungen führten zu den folgenden Ergebnissen: $\Delta f = 14,2$ Hz (11,4–17,6 Hz); $\Delta L = 2,7$ dB (1,7–4,0 dB); $\Delta t = 29,5$ ms (24,5–31,8 ms). Für interaurale Messungen wurden folgende Werte genannt: $\Delta f = 1,0$ Hz (0,6–1,2 Hz); $\Delta L = 2,2$ dB (1,6–2,9 dB); $\Delta t = 11,1$ ms (6,7–14,9 ms). Bei dichotischen Messungen wurden folgende Ergebnisse erzielt: $\Delta f = 12,0$ Hz (7,6–15,5 Hz); $\Delta L = 2,5$ dB (1,3–3,2 dB); $\Delta t = 28,2$ ms (22,9–35,8 ms). (Ludwig, 2008)

Von Bungert-Kahl et al. (2004) liegen die Ergebnisse der Erwachsenengruppe im Alter von 20 bis 29 Jahren als Mittelwert \pm SD vor. Die monauralen Messungen führten zu den folgenden Ergebnissen: $\Delta f = 6,0$ Hz \pm 4,1 Hz; $\Delta L = 1,3$ dB \pm 0,4 dB; $\Delta t = 13,2$ ms \pm 6,4 ms. Für interaurale Messungen wurden folgende Werte genannt: $\Delta f = 5,7$ Hz \pm 3,9 Hz; $\Delta L = 2,1$ dB \pm 0,9 dB; $\Delta t = 9,0$ ms \pm 4,2 ms. Bei den dichotischen Messungen wurden folgende Ergebnisse erzielt: $\Delta f = 4,5$ Hz \pm 3,2 Hz; $\Delta L = 1,8$ dB \pm 0,8 dB; $\Delta t = 17,9$ ms \pm 7,3 ms. (Bungert-Kahl et al., 2004)

Thompson et al. (1999) untersuchten die Schwellen für die Frequenz-Differenz Unterscheidungsfähigkeiten von 5-Jährigen mit einem Zwei-Alternativen-Zwei-Intervall-Zwangswahlverfahren. Dabei enthielt jedes Intervall zwei Töne mit einer Pause von 200 ms, wobei in einem Intervall ein Ton eine höhere Frequenz hatte. Die Pause zwischen den beiden Intervallen betrug 600 ms. In der Studie bestand die Aufgabe darin, mündlich das Intervall mit den unterschiedlichen Tönen zu nennen, was lediglich 5 von 16 Kindern gelang, d.h. 69% der Kinder konnten nicht getestet werden. Die Unterschiedsschwelle \pm SD der 5 Kinder für einen 200 ms langen 1 kHz-Ton betrug 22,8 Hz \pm 17,5 Hz (Thompson et al., 1999). Olsho et al. (1987) führten eine Studie zur Erforschung der Frequenz-Differenz-Unterschiedsschwellen an Säuglingen durch, die jeweils im Alter 3, 6 und 12 Monaten getestet wurden. Um das zu bewerkstelligen war ein aufwendiges Verfahren der Verhaltensaudiometrie mit 2 Untersuchungspersonen und einem umfangreichen Training notwendig, bei Bedarf auch während der Messung. Aus den Messdaten wurden psychometrische Funktionen berechnet und daraus die Schwellen abgelesen. Im Alter von 12 Monaten und für 1 kHz-Töne bei einem Darbietungspegel von 40 dB SL beträgt Δf ungefähr 27 Hz. Zusätzlich fanden Messungen zum Vergleich an Erwachsenen statt, mit dem Ergebnis von Δf ist ungefähr 6 Hz für 1 kHz-Töne bei einem Darbietungspegel von 40 dB SL (Olsho et al., 1987).

Der in dieser Arbeit neu entwickelte Test (s. Abschnitt 4.3) verwendet ein Drei-Intervall-Zwangswahlverfahren. Er wird vollständig per Touchscreen gesteuert. Während des Tests sind Wiederholungen der Darbietung möglich und nach jeder Eingabe erfolgt eine Rückmeldung, ob die Eingabe richtig war. Es werden ausschließlich Sinustöne verwendet mit einer Länge von 400 ms und einer Pause zwischen zwei Signalen von 600 ms. In Schwellennähe kommt ein 1-up 2-down-Verfahren mit variabler, leicht randomisierter Schrittweite zur Anwendung. Zur Berechnung der Schwelle werden die Umkehrpunkte in Anlehnung an ISO 8253-1 (2011) verwendet. Das gesamte Messsystem ist als Medizinprodukt zertifiziert und zugelassen. Dementsprechend sind alle Kopfhörerausgabepegel gemäß der Norm für Reintonaudiometer (DIN 60645-1, 2018) kalibriert.

Die Ergebnisse der binaural durchgeführten Messungen an Kindern im Alter von 5;5 bis 5;11 Jahren im Rahmen der für diese Arbeit durchgeführten Studien (s. Abschnitt 5.3.2) sind wie folgt (Median (Interquartilsabstand)): $\Delta f = 40,0$ Hz (41,4 Hz); $\Delta L = 6,0$ dB (9,0 dB); $\Delta t = 65,3$ ms (104,0 ms). Im Vergleich mit den Daten aus der Literatur liegen die Schwellen und die Streuung im selben Bereich oder niedriger als diejenigen, die mit den Geräten *P* (Kiese-Himmel et al., 2018; Nickisch & Kiese-Himmel, 2009; Nickisch et al., 2013; Ptok & Meisen, 2008) oder *F* (Fischer, 2007; Fischer & Hartnegg, 2004) zumeist in Schulen an Kindern über 5;11 Jahren erhoben wurden. Im Vergleich mit den Studien, die in Kliniken oder Forschungseinrichtungen stattfanden (Ludwig, 2008; Olsho et al., 1987; Thompson et al., 1999) weisen die eigenen Daten, die in Kindertagesstätten erhoben wurden, höhere Werte auf. Bei Messungen in Schulen oder Kindertagesstätten ist es nicht auszuschließen, dass die Kinder kurz vor den Tests Lärm ausgesetzt waren und auch der Störschall kann hier schlechter kontrolliert werden als in Kliniken oder Forschungseinrichtungen. Die Unterschiedsschwellen und Streuungen bei den Kindern unter 5;5 Jahren nehmen mit abnehmendem Alter zu, was plausibel ist. So zeigten auch die Ergebnisse von Ludwig (2008) einen deutlichen Anstieg bei der jüngsten Gruppe im Vergleich zur Gruppe der 8 bis 9-Jährigen. Dieser Trend setzt sich in den eigenen Daten der Kinder unter 5;5 Jahren fort. Der Anteil an Kindern, mit denen keine Ergebnisse für Δf erzielt werden können, lag in der Gruppe der 5;5 bis 5;11 Jahre alten Kinder beim BASD bei 9%. Im Vergleich zu 11% (Kiese-Himmel et al., 2018), oder 47% (Fischer, 2007) für im Mittel 7- oder 8-Jährige erscheint der Anteil akzeptabel. Bei Thompson et al. (1999) konnten 69% der 5-Jährigen den Test nicht erfolgreich durchführen. Für den BASD traf das nur bei 18% der 5;0 bis 5;5 Jahre alten Kinder zu.

Die mit dem sprachfreien Teil der BASD-Testbatterie an Erwachsenen erzielten Ergebnisse lassen sich zum Vergleich mit der Literatur folgendermaßen zusammenfassen (Median (unteres – oberes Quartil)): Binaural $\Delta f = 5,0$ Hz (4,0–7,0 Hz); $\Delta L = 2,2$ dB (1,8–2,8 dB); $\Delta t = 43,4$ ms (30,0–58,5 ms). Interaural $\Delta f = 2,0$ Hz (1,6–2,5 Hz); $\Delta L = 2,8$ dB (2,1–3,4 dB); $\Delta t = 20,9$ ms (14,2–41,5 ms). Dichotisch $\Delta f = 3,8$ Hz (2,3–4,8 Hz); $\Delta L = 2,4$ dB (1,6–2,9 dB); $\Delta t = 54,3$ ms (41,0–73,1 ms). In den eigenen Daten unterscheiden sich die Ergebnisse der binaural und dichotisch (s. Abb. 4.18) durchgeführten Tests nicht statistisch signifikant. Das ist grundsätzlich im Einklang mit den Feststellungen von Ludwig (2008) und Bungert-Kahl

et al. (2004) bezüglich monaural und dichotisch durchgeführter Tests. Auch die Rangordnung der Ergebnisse in Abhängigkeit der Präsentationsart ist gleich. Mit Ausnahme von Δt , sind die Werte der eigenen Daten eher in Übereinstimmung mit den Daten von Bungert-Kahl et al. (2004), die insgesamt niedriger ausfallen als bei Ludwig (2008). Allerdings weisen die eigenen Daten beim Subtest Tondauer-Differenz deutlich höhere Werte auf, was möglicherweise an Unterschieden der Stimulus- und Pausenlänge und der Dauer für das Ein- und Ausblenden des Tones liegen könnte.

Bei Untersuchungen im Rahmen einer AVWS-Diagnose bieten sprachhaltige Tests, mit leichten altersabhängigen Unterschieden, eine besonders gute Trennschärfe bei Schulkindern zwischen Kindern mit und ohne AVWS (Kiese-Himmel et al., 2018; Nickisch & Kiese-Himmel, 2009; Nickisch et al., 2013). Bei Kindern zwischen 8 und 10 Jahren stellten sich das Wiederholen von immer länger werdenden, aus Nonsense-Silben bestehenden Kunstwörtern, Sprachverstehen im Störgeräusch und Phonemdifferenzierung als am besten geeignet zur Differenzierung heraus (Nickisch & Kiese-Himmel, 2009). Die Sprachaudiometrie im Störgeräusch und das Nachsprechen von Kunstwörtern finden sich auch in der Liste der leistungsfähigsten Tests für die Gruppe von Kindern im 2. Schuljahr wieder, sowie ein Test zum Zahlenfolge-Gedächtnis und einer zum Verstehen und Nachsprechen von beidohrig gleichzeitig dargebotenen unterschiedlichen Wörtern (z.B. links „die Schlittenfahrt“ und rechts „der Birnenbaum“) (Nickisch et al., 2013). Abgesehen von Sprachverstehen im Störgeräusch eignen sich diese Testverfahren auch für Kinder im 1. Schuljahr am besten (Kiese-Himmel et al., 2018). Auf Grundlage der mit dem Gerät *P* gemachten Erfahrungen, insbesondere bedingt durch die Quote an nicht erfolgreich absolvierten Tests in der Kontrollgruppe der Schulkinder der 1. Jahrgangsstufe, konnte keine Empfehlung für die Verwendung von sprachfreien psychoakustischen Tests in der AVWS-Diagnostik ausgesprochen werden (Kiese-Himmel et al., 2018). Das Gleiche wird auch von Ludwig (2008) anhand einer Gruppe von 84 Kindern im Alter von 6 bis 17 Jahren berichtet, die mit dem Verdacht auf AVWS in der Klinik vorstellig wurden. Die von ihr verwendete sprachfreie psychoakustische Testbatterie wies bei der Klassifizierung in dieser Gruppe eine Sensitivität von 0,64 und eine Spezifität von 0,49 auf. Sie verglich die auffälligen Kinderergebnisse der psychoakustischen Tests mit den Diagnosen der Klinik, in denen der Verdacht bei 39 Kindern bestätigt wurde und bei 45 Kindern AVWS nicht nachgewiesen werden konnte (Ludwig, 2008).

Um dem Rechnung zu tragen und die Testperformance zu verbessern, beinhaltet die neu entwickelte BASD-Testbatterie neben sprachfreien psychoakustischen Tests auch zwei sprachhaltige Tests zur Phonemdifferenzierung. Wie von der Studienleitung berichtet wurde, trägt das neu entwickelte Testdesign in Form eines per Touchscreen gesteuerten „Memoryspiels“ zur Steigerung der Motivation und Senkung des Anteils an nicht erfolgreichen Testdurchführungen bei. Die verhältnismäßig hohe Streuung wurde bei ähnlichen psychoakustischen Testverfahren auch von anderen Forschungsgruppen bei jungen Kindern beobachtet (Kiese-Himmel et al., 2018; Ludwig, 2008; Nickisch & Kiese-Himmel, 2009; Nickisch et al., 2013). Es gibt mehrere Faktoren, die zur Streuung bei den BASD-Messungen beigetragen haben könn-

ten. Die Leistung der jungen Kinder kann sich unterscheiden bedingt durch ihren individuellen Entwicklungsstand. Auch eine dem Test möglicherweise vorangegangene Lärmbelastung in der Kindertagesstätte kann einen negativen Einfluss auf die Ergebnisse haben. Es ist darüber hinaus nicht auszuschließen, dass die Aufgabenstellung, insbesondere das genaue Zuhören in Schwellennähe, manchmal nicht vollständig verstanden und umgesetzt wurde und sich somit die Aufmerksamkeit im Verlauf einer Messung vom Hören zum Sehen und Memory spielen verschoben hatte. Darüber hinaus sind große Trainingseffekte vorhanden, entweder durch gezieltes Üben oder wiederholtes Messen, besonders bei der Unterscheidung von Tönen unterschiedlicher Frequenz (Berwanger & Suchodoletz, 2003; Michalski & Tewes, 2001; Schäffler et al., 2004). Die sehr guten Ergebnisse der 25 % besten 4;0 bis 4;5 jährigen Kinder – unteres Quartil: $\Delta f = 20,8$ Hz; $\Delta L = 4,9$ dB; $\Delta t = 79,1$ ms – könnten möglicherweise auf derartige Übungs- und Trainingseffekte zurückzuführen sein.

Die Untersuchung des Zusammenhangs zwischen der Sprachentwicklung sowie AVWS und den ASD-Subtests hat gezeigt, dass eine Testbatterie entwickelt wurde, die in den Untertests Frequenz-Differenz, Diskrimination von /ba/ und /ga/ und Diskrimination von /ga/ und /ka/ signifikant mit auditiven Verarbeitungs- und Wahrnehmungsleistungen und den Sprachfähigkeiten von Kindern zwischen 4;0 und 4;5 Jahren korreliert und zwischen sprachauffälligen und -unauffälligen Kindern differenzieren kann. Der Untertest Hörpegel-Differenz zeigte keine signifikante Korrelation mit der Beurteilung hinsichtlich eines Verdachtes auf AVWS oder der Zuordnung des Sprachentwicklungsstands zu einer der 3 Kategorien (sprachlich regelrecht entwickelt, sprachförderbedürftig, sprachentwicklungsgestört), was übereinstimmend im Kontext der AVWS-Diagnose berichtet wurde (Kiese-Himmel et al., 2018; Nickisch & Kiese-Himmel, 2009; Nickisch et al., 2013).

7. Fazit und Ausblick

Die verbesserte und frühere Diagnose von Hörstörungen im Rahmen des Neugeborenen-Hörscreenings hat den Bedarf nach zusätzlichen leistungsfähigen Untersuchungsmethoden für Gehörprüfungen bei Kleinkindern und Vorschulkindern in den Mittelpunkt gerückt. Im Rahmen dieser Arbeit wurden drei neue Verfahren entwickelt, implementiert und in Studien erprobt, die über die periphere Schallverarbeitung hinaus auch die zentrale Verarbeitung mit einbeziehen und die bei der audiologischen Diagnostik sowie bei der Anpassung und Überprüfung von Hörgeräteeinstellungen zum Einsatz kommen können.

Die Bestimmung der Reintonschwelle spielt eine zentrale Rolle in der audiologischen Diagnostik. Daher wurde ein neuer Selbsttest namens „Multiple-choice Auditory Graphical Interactive Check“ (MAGIC) für Kinder ab 3 Jahren entwickelt. Die wesentliche Neuerung gegenüber Verfahren, die meist zum Einsatz kommen, ist, dass das Kind die Stimulusdarbietung selbst starten und wiederholen und die Eingabe, ob ein Ton gehört wurde, machen darf. Dadurch sinkt der mögliche Einfluss der Vigilanz auf die Messung und die Motivation steigt. Darüber hinaus muss der Messalgorithmus keine Darbietungspegel wiederholen, weil sichere Eingaben gemacht werden. Durch Pendeln um die Hörschwelle in kleiner werdenden Schrittweiten wird die Schwelle gewöhnlich in 3–5 Schritten ermittelt. Die Messung wird per Touchscreen über altersgerechte Bilder gesteuert. Damit das zuverlässig funktioniert, ist eine Eingewöhnung zu Beginn der Messung notwendig, bei der das Kind mit der Messung vertraut gemacht wird. Dabei lernt das Kind auf leise Töne zu achten und das Konzept der „gehört“ und „nicht gehört“ Eingabe richtig umzusetzen.

Im Rahmen der Validierung des Tests wurden insgesamt 156 Kinder mit Hörverlusten von -5 bis 95 dB HL im Alter zwischen 3;6 und 10;3 Jahren mit dem MAGIC und konventioneller Spielaudiometrie (PTA) im Frequenzbereich von 250 Hz bis 8 kHz untersucht. Test-Retest-Messungen wurden an 49 Kindern aus dem Testkollektiv durchgeführt. Zusätzlich fand eine Befragung statt, welches Testverfahren mehr Spaß gemacht hatte und welche Methode als anstrengender empfunden wurde. Beim Vergleich der Ergebnisse von MAGIC und PTA liegen 71,1 % der Daten innerhalb des Bereichs von -5 dB bis 5 dB Differenz zwischen den beiden Messmethoden und 89,5 % befinden sich zwischen -10 dB und 10 dB. Zwischen MAGIC und PTA besteht eine starke positive Korrelation 0,79 ($p < 0,001$). Die mittlere Differenz zwischen den MAGIC- und PTA-Ergebnissen beträgt 0,5 dB bei einer SD von 7,6 dB. Beim Vergleich der Ergebnisse von Test und Retest liegen 68,2 % der Daten innerhalb des Bereichs von -5 dB bis 5 dB Differenz zwischen den beiden Tests und 86,1 % befinden sich zwischen -10 dB und 10 dB. Test und Retest sind stark positiv korreliert ($r = 0,81$; $p < 0,001$). Test und Retest unterscheiden sich im Mittel um 1,4 dB bei einer SD von 8,0 dB. Die Kinderbefragung ergab, dass einer deutlichen Mehrheit der MAGIC mehr Spaß machte und sie ihn weniger anstrengend fanden als PTA. Es ist daher davon auszugehen, dass der MAGIC insgesamt eine gute Ergänzung zu den etablierten Testverfahren in der Audiologie darstellt.

Des Weiteren wurde ein neuer, deutscher, altersgerechter Sprachaudiometrietest „Mainzer Audiometric Test for Children“ (MATCH) entwickelt, der auf 26 Nomen basiert, die mit großer Wahrscheinlichkeit Teil des Wortschatzes von 2-jährigen Kindern sind. Als Antwort können die Kinder ein Item aus 4 alternativen Bildern oder „nichts gehört“ auswählen. Der MATCH bietet im Vergleich zu bisher in der Praxis verwendeten Tests ein erweitertes Vokabular und ermöglicht die Messung des Sprachverständnisses unabhängig von der Fähigkeit der Sprachproduktion des Kindes sowie der Hörleistung und Konzentration des Untersuchungspersonals. Im Rahmen der Normierung und Validierung des Tests wurden insgesamt 179 Kinder im Alter von 2;11 bis 6;9 Jahren untersucht. Innerhalb dieser Testgruppe wiesen 51 Kinder auf beiden Ohren einen Hörverlust von bis zu 90 dB auf. Die normalhörenden Kinder wurden altersgemäß in 3 Gruppen aufgeteilt. Für jede Altersgruppe wurden die Sprachverständlichkeitsschwelle (SVS) und die psychometrische Funktion des Sprachverstehens bestimmt. Zur Validierung wurden Test-Retest-Messungen an 85 Ohren durchgeführt, sowie an 91 Ohren die Korrelation zwischen dem mittleren Hörverlust (mHV PTA) bei 500 Hz, 1 kHz, 2 kHz und 4 kHz und der SVS berechnet.

Das entwickelte Testmaterial weist eine relative Phonemhäufigkeit auf, die im Wesentlichen der Verteilung der deutschen Spontansprache entspricht. Zudem zeigt das Langzeitspektrum der Sprachaufnahmen eine gute Übereinstimmung mit dem internationalen Langzeit-Sprachspektrum. Die Normdaten wurden bestimmt, um in der klinischen Diagnostik individuelle Testergebnisse in einen Kontext setzen zu können und damit mögliche Hörbeeinträchtigungen bei Kindern zu erkennen und zu beurteilen. Der verwendete adaptive Algorithmus führt zu einer SVS bei 71,4% Sprachverständlichkeit. Die zugehörigen Werte für die 3 Altersgruppen sind 38,3 dB, 34,0 dB und 30,0 dB. An dieser Stelle hat die Diskriminationsfunktion der jeweiligen Altersgruppe eine Steigung von 2,1 %/dB, 2,6 %/dB und 3,5 %/dB. Der Test wurde durch Korrelationsanalysen zwischen MATCH und Tonaudiometrie und mittels Test-Retest-Vergleich validiert. Die Korrelation zwischen MATCH-SVS und mHV PTA beträgt $r = 0,84$ ($p < 0,001$). Die mittlere Differenz zwischen Test und Retest ist $1,7 \pm 4,6$ dB bei einer Korrelation von $r = 0,89$ ($p < 0,001$)

Die mittlere Testdauer beim MATCH beträgt je nach Altersgruppe zwischen 3 und 4 Minuten pro Ohr und variiert statistisch signifikant zwischen den Gruppen. Die kürzeste Messzeit liegt deutlich unter 2 Minuten, während die längste Messzeit mit 8,5 Minuten in der Gruppe der jüngsten Kinder auftritt. Die Unterschiede zwischen Normalhörenden und Schwerhörenden bezüglich der Testdauer sind gering, und es zeigt sich eine Abnahme der Testdauer mit zunehmendem Alter um etwa 1 Minute. Die Reaktionszeit wird als Zeitspanne zwischen dem Ende der akustischen Signalpräsentation und der Eingabe per Touchscreen erfasst. Sie ist niedrig bei deutlich überschwellig gehörten Wörtern oder wenn nichts gehört wurde und sie ist verlängert in Schwellennähe. Dort kann eine verlängerte Reaktionszeit auf erhöhte Höranstrengung hindeuten und somit auf konzentrierte Mitarbeit.

Die neu entwickelte Testbatterie zur Prüfung der zentralen Schallverarbeitung, „Bochum Auditory Speech Discrimination Test“ (BASD), beinhaltet sprachfreie psychoakustische Verfahren (Messung der Unterschiedsschwellen für Frequenz (Δf), Pegel (ΔL), Tondauer (Δt) und Amplitudenmodulationsfrequenz) und sprachhaltige Untertests (Diskrimination der Silben /ba/ und /ga/ sowie /ga/ und /ka/). Alle Untertests können bei den Testfrequenzen 500 Hz und 1 kHz und den Darbietungsarten monaural, binaural, interaural und dichotisch durchgeführt werden. Der BASD ist ein Selbsttest der ein Drei-Alternativen-Zwangswahlverfahren in Form eines Memoryspiels zur Schwellenbestimmung verwendet. In Schwellennähe wird ein 1-up 2-down-Verfahren mit variabler, leicht randomisierter Schrittweite eingesetzt, das in der Lage ist, mit wenigen Schritten die Schwelle zu finden, selbst wenn die Startdifferenz weit davon entfernt ist.

Der BASD wurde an 30 Erwachsenen im Alter zwischen 18;9 und 26;9 Jahren erprobt, an 293 Kindern im Alter zwischen 4;0 und 5;11 Jahren normiert und an 158 Kindern im Alter zwischen 4;0 bis 4;5 Jahre hinsichtlich seiner Fähigkeit die auditiven Verarbeitungs- und Wahrnehmungsleistungen und Sprachfähigkeiten der Kinder zu klassifizieren untersucht. Die Ergebnisse bei den Erwachsenen sind wie folgt (Median (unteres – oberes Quartil)): Binaural $\Delta f = 5,0$ Hz (4,0–7,0 Hz); $\Delta L = 2,2$ dB (1,8–2,8 dB); $\Delta t = 43,4$ ms (30,0–58,5 ms). Interaural $\Delta f = 2,0$ Hz (1,6–2,5 Hz); $\Delta L = 2,8$ dB (2,1–3,4 dB); $\Delta t = 20,9$ ms (14,2–41,5 ms). Dichotisch $\Delta f = 3,8$ Hz (2,3–4,8 Hz); $\Delta L = 2,4$ dB (1,6–2,9 dB); $\Delta t = 54,3$ ms (41,0–73,1 ms). Die Normdaten der 4;0 bis 5;11 Jahre alten Kindern sind in den Tabellen 5.11 und 5.12 enthalten. Ihre Unterschiedsschwellen sind deutlich größer als die der Erwachsenen und steigen mit abnehmendem Alter. In der Studie zur Validierung wurde mittels t-Test ein signifikanter Unterschied ($p = 0,028$) zwischen Kindern mit diagnostizierter AVWS und den Kontrollen im Untertest /ga/ und /ka/ gefunden. Hinsichtlich der Sprachfähigkeiten bestehen statistisch signifikante Spearman-Korrelationen (ρ ; p_S) und t-Test-Gruppenunterschiede (p_t) bei den Untertests Frequenz-Differenz ($\rho = 0,24$; $p_S = 0,014$; $p_t = 0,023$), Diskrimination von /ba/ und /ga/ ($\rho = 0,27$; $p_S = 0,001$; $p_t = 0,002$) und Diskrimination von /ga/ und /ka/ ($\rho = 0,27$; $p_S = 0,001$; $p_t = 0,001$).

Ausblick

Auf der Basis der in dieser Arbeit vorgestellten und diskutierten Ergebnisse und aufbauend auf den Möglichkeiten, die das neu entwickelte Messsystem bietet, werden abschließend Anregungen für zusätzliche neue Funktionen und weiterführende Studien skizziert.

Der MAGIC wurde mit Sinustönen als Stimulus durchgeführt, allerdings besteht ebenso die Option frequenzmodulierte Töne zu verwenden. Entsprechend der Literatur sind keine Unterschiede in Bezug auf die Genauigkeit und Messunsicherheit bei Verwendung von frequenzmodulierten Tönen zu erwarten (Arlinger & Jerlvall, 1987; Orchik & Rintelmann, 1978). Da diese Signale für Kinder interessanter als Sinustöne sind (Orchik & Rintelmann, 1978), könnten die Messungen bevorzugt mit frequenzmodulierten Tönen erfolgen. Im Zuge einer Weiterentwicklung könnten die Reaktionszeiten, ähnlich wie das beim MATCH der Fall ist, an-

gezeigt werden. Auch die Anzahl an Wiederholungen des Stimulus als weiterer Parameter könnte Hinweis über die Zuverlässigkeit der Messungen geben.

Der MATCH wurde von seiner Architektur und Implementierung her darauf ausgelegt, dass der Test nach dem gleichen Konzept in verschiedenen Sprachen durchgeführt werden kann. Allerdings sind die Ergebnisse hinsichtlich Tonaufnahme und Sprache unterschiedlich, so dass jeweils vergleichbare Studien in den fremdsprachlichen Versionen notwendig sind. Insbesondere für Kinder ab 6 Jahren könnte im Zuge einer Weiterentwicklung die Möglichkeit zur Messung mit kleineren Schrittweiten (z.B. 5 dB und 2 dB statt 10 dB und 4 dB) zur Erhöhung der Messgenauigkeit in Erwägung gezogen werden.

Mit dem BASD liegt eine Testbatterie vor, die leicht um weitere, insbesondere sprachfreie psychoakustische Testverfahren ergänzt werden kann. Im Zuge einer Weiterentwicklung könnten auch weitere von den Fachgesellschaften empfohlene Verfahren (DGPP, 2019) implementiert werden, wie z.B. die Lückenerkennung oder die Messung der zeitlichen und räumlichen Ordnungsschwelle. Auch die Auswertung und Anzeige von Reaktionszeiten sind denkbare Erweiterungen, die im Zuge einer Weiterentwicklung mit aufgenommen werden sollten. Sie könnten Hinweise über die Zuverlässigkeit der Messungen geben und möglicherweise in einem Zusammenhang bezüglich der Sprachentwicklung oder AVWS stehen. Bisher wurden alle Kindermessungen in der Präsentationsart binaural durchgeführt. Es wären daher weitere Studien an Kindern wünschenswert, bei denen auch die Präsentationsarten interaural und dichotisch zum Einsatz kommen.

Literatur

- American Speech-Language-Hearing Association. (1988). Determining threshold level for speech [Guidelines]. Available from www.asha.org/policy. <https://doi.org/10.1044/policy.GL1988-00008>
- Arlinger, S. D., & Jerlvall, L. B. (1987). Reliability in warble-tone sound field audiometry. *Scand Audiol*, *16*, 21–7.
- Baumann, R. (2006). *Der Würzburger Kindersprachtest: Entwicklung, Standardisierung und dessen Erprobung bei normalhörenden und cochleaimplantierten Kindern* [Dissertation]. Bayerische Julius-Maximilians-Universität zu Würzburg.
- Berwanger, D., & Suchodoletz, W. v. (2003). Evaluation eines Trainings von Ordnungsschwelle und Richtungshören. *Forum Logopädie*, *6*(17), 12–19.
- Biesalski, P., Leitner, H., Leitner, E., & Gangel, D. (1974). Der Mainzer Kindersprachtest. Sprachaudiometrie im Vorschulalter. *HNO*, *22*, 160–1.
- Billich, P. (1981). *Heidelberger CVC-Audiometrie – Entwicklung und Erprobung*. Geers-Stiftung: Schriftenreihe, Bd. 2, median-Verlag.
- Brand, T., Achtzehn, J., & Kollmeier, B. (1999). Erstellung von Testlisten für den Oldenburger Kinder-Reimtest (K. Schorn, Hrsg.). *Zeitschrift für Audiologie, Supplementum II*, 50–51.
- Brinkmann, K., & Richter, U. (1989). Free-field sensitivity level of audiometric earphones to be used for speech audiometer calibration. *Scand Audiol*, *18*(2), 75–81. <https://doi.org/10.3109/01050398909070726>
- Brinkmann, K., & Richter, U. (1997). Ensuring reliability and comparability of speech audiometry in Germany. In M. Martin (Hrsg.), *Speech audiometry* (2. Aufl., xii, 368 p.). Whurr Publishers.
- Bundesministerium für Gesundheit. (2008). *Bekanntmachung eines Beschlusses des Gemeinsamen Bundesausschusses über eine Änderung der Kinder-Richtlinien: Einführung eines Neugeborenen-Hörscreenings vom 19. Juni 2008*. Verfügbar 30. Juni 2023 unter http://www.g-ba.de/downloads/39-261-681/2008-06-19-Kinder-%C3%B6rscreening_BAnz.pdf
- Bungert-Kahl, P., Biedermann, F., Dörrscheidt, G. J., von Cramon, D. Y., & Rübsamen, R. (2004). Psychoacoustic Test Tools for the Detection of Deficits in Central Auditory Processing: Normative Data. *Zeitschrift für Audiologie*, *43*(2), 48–71.
- Byrne, D., Dillon, H., Tran, K., Arlinger, S., Wilbraham, K., Cox, R., Hagerman, B., Hetu, R., Kei, J., Lui, C., Kiessling, J., Kotby, M. N., Nasser, N. H. A., El Kholly, W. A. H., Naka-

- nishi, Y., Oyer, H., Powell, R., Stephens, D., Meredith, R., Sirimanna, T., Tavartkiladze, G., Frolenkov, G. I., Westerman, S., & Ludvigsen, C. (1994). An international comparison of long-term average speech spectra. *The Journal of the Acoustical Society of America*, *96*, 2108–2120.
- Chilla, R., Gabriel, P., Kozielski, P., Bansch, D., & Kabas, M. (1976). Der Göttinger Kindersprachverständnistest. I. Sprachaudiometrie des Kindergarten- und retardierten Kindes mit einem Einsilber-Bildtest. *HNO*, *24*, 342–6.
- Coninx, F. (2005). Construction and norming of the Adaptive Auditory Speech Test (AAST). *22. Wissenschaftliche Jahrestagung der Deutschen Gesellschaft für Phoniatrie und Pädaudiologie*. <https://www.egms.de/static/en/meetings/dgpp2005/05dgpp045.shtml>
- Coninx, F. (2006). Entwicklung und Erprobung des Adaptiven Auditiven SprachTests (AAST). *9. Jahrestagung der Deutschen Gesellschaft für Audiologie*. <http://www.uzh.ch/orl/dga2006/programm/wissprog/Coninx.pdf>
- Deutsche Gesellschaft für Phoniatrie und Pädaudiologie e.V. (DGPP). (2013). *S2k-Leitlinie Periphere Hörstörungen im Kindesalter*. Verfügbar 30. Juni 2023 unter <https://dgpp.de/de/download/48/>
- Deutsche Gesellschaft für Phoniatrie und Pädaudiologie e.V. (DGPP). (2019). *S1-Leitlinie Auditive Verarbeitungs- und Wahrnehmungsstörungen (AVWS)*. Verfügbar 2. Juli 2023 unter <https://dgpp.de/de/download/46/>
- DIN 45621-3. (1985). DIN 45621-3:1985-04, Sprache für Gehörprüfung; Wörter für die Gehörprüfung bei Kindern. <https://doi.org/10.31030/1234555>
- DIN 45626-1. (1995). DIN 45626-1:1995-08, Tonträger mit Sprache für Gehörprüfung – Teil 1: Tonträger mit Wörtern nach DIN 45621-1 (Aufnahme 1969). <https://doi.org/10.31030/2791997>
- DIN 60318-1. (2010). DIN EN 60318-1:2010-07, Akustik – Simulatoren des menschlichen Kopfes und Ohres – Teil 1: Ohrsimulator zur Kalibrierung von supra-auralen und circumauralen Kopfhörern. (*IEC 60318-1:2009*), Deutsche Fassung EN 60318–1:2009. <https://doi.org/10.31030/1629722>
- DIN 60645-1. (2018). DIN EN 60645-1:2018-08, Akustik – Audiometer – Teil 1: Reinton-Audiometer. (*IEC 60645-1:2017*), Deutsche Fassung EN 60645–1:2017. <https://doi.org/10.31030/2847334>
- Dobie, R. A. (1983). Reliability and validity of industrial audiometry: implications for hearing conservation program design. *Laryngoscope*, *93*, 906–27.
- Downs, D. W. (1982). Effects of hearing and use on speech discrimination and listening effort. *J Speech Hear Disord*, *47*(2), 189–193. <https://doi.org/10.1044/jshd.4702.189>

- Downs, D. W., & Crum, M. A. (1978). Processing demands during auditory learning under degraded listening conditions. *J Speech Hear Res*, 21(4), 702–714. <https://doi.org/10.1044/jshr.2104.702>
- Efron, B. (1979). Bootstrap Methods: Another Look at the Jackknife. *The Annals of Statistics*, 7(1), 1–26. <https://doi.org/10.1214/aos/1176344552>
- Eikelboom, R. H., Swanepoel, D. W., Motakef, S., & Upson, G. S. (2013). Clinical validation of the AMTAS automated audiometer. *International Journal of Audiology*, 52(5), 342–349. <https://doi.org/10.3109/14992027.2013.769065>
- Elliott, L. L., Connors, S., Kille, E., Levin, S., Ball, K., & Katz, D. (1979). Children's understanding of monosyllabic nouns in quiet and in noise. *The Journal of the Acoustical Society of America*, 66(1), 12–21. <https://doi.org/10.1121/1.383065>
- Fastl, H., & Zwicker, E. (1983). A free-field equalizer for TDH 39 earphones. *J Acoust Soc Am*, 73, 312–4.
- Feuerstein, J. F. (1992). Monaural versus binaural hearing: ease of listening, word recognition, and attentional effort. *Ear Hear*, 13(2), 80–86.
- Fischer, B. (2007). *Looking for Learning: Auditory, Visual & Optomotor Processing of Children with Learning Problems*. Nova Science Publishers Inc.
- Fischer, B., & Hartnegg, K. (2004). On the development of low-level auditory discrimination and deficits in dyslexia. *Dyslexia*, 10(2), 105–118. <https://doi.org/10.1002/dys.268>
- Foster, J. R., & Haggard, M. P. (1987). The Four Alternative Auditory Feature Test (FAAF) – linguistic and psychometric properties of the material with normative data in noise. *British Journal of Audiology*, 21(3), 165–174. <https://doi.org/10.3109/03005368709076402>
- Fraser, S., Gagné, J.-P., Alepins, M., & Dubois, P. (2010). Evaluating the effort expended to understand speech in noise using a dual-task paradigm: the effects of providing visual speech cues. *J Speech Lang Hear Res*, 53(1), 18–33. [https://doi.org/10.1044/1092-4388\(2009/08-0140\)](https://doi.org/10.1044/1092-4388(2009/08-0140))
- Gosselin, P. A., & Gagné, J.-P. (2010). Use of a Dual-Task Paradigm to Measure Listening Effort Utilisation d'un paradigme de double tâche pour mesurer l'attention auditive. *Revue canadienne d'orthophonie et d'audiologie-Vol*, 34(1).
- Gravel, J. S. (2001). Potential Pitfalls in the Audiological Assessment of Infants and Young Children. In R. C. Seewald & J. S. Gravel (Hrsg.), *A Sound Foundation Through Early Amplification 2001 – Proceedings of the Second International Conference* (S. 85–101). Phonak AG.

- Hahlbrock, K.-H. (1953). Über Sprachaudiometrie und neue Wörterteste. *Archiv für Ohren-, Nasen- und Kehlkopfheilkunde*, 162(5), 394–431. <https://doi.org/10.1007/BF02105664>
- Heath, A. (2014). *Test of Auditory and Visual Skills: TAVS User Guide* (English edition: May 2014). Version TAVS version 5.3. Advanced Brain Technologies. Centerville, Utah 84014, United States.
- Heller, O., Boretzki, M., Fichtl, E., Knoblach, W., May, B., Nowak, T., & Stock, A. (1995). *Entwicklung eines integrativen Verfahrens zur audiologischen Diagnostik, ausgerichtet auf die Prüfung und Bewertung technischer Hörhilfen für Hörgeschädigte. Abschlußbericht (01VJ8903)* (Techn. Ber.). Psychologisches Inst. der Univ. Würzburg, Lehrstuhl III.
- Heller, O., & Opp-Enzinger, A. H. A. (2006, 22. August). *Patentanmeldung, Self-paced In-Situ Audiometry, WO 2007/009287 A2, PCT/CH2006/00450*. Phonak AG, Stäfa.
- Heyn, J. (2019). *Kurz- und langfristige Lern- und Habituationseffekte bei der Anwendung des Oldenburger Satztests an jungen, normalhörenden Probanden* [Dissertation]. Julius-Maximilians-Universität Würzburg.
- Hick, C. B., & Tharpe, A. M. (2002). Listening effort and fatigue in school-age children with and without hearing loss. *J Speech Lang Hear Res*, 45(3), 573–584. [https://doi.org/10.1044/1092-4388\(2002/046\)](https://doi.org/10.1044/1092-4388(2002/046))
- Hickling, S. (1966). Studies on the reliability of auditory threshold values. *Journal of Auditory Research*, 6(1), 39–46.
- Hnath-Chisolm, T. E., Laipply, E., & Boothroyd, A. (1998). Age-related changes on a children's test of sensory-level speech perception capacity. *Journal of speech, language, and hearing research: JSLHR*, 41(1), 94–106. <https://doi.org/10.1044/jslhr.4101.94>
- Hoffmann, V., & Schäfer, K. (2020). *Kindliche Hörstörungen. Diagnostik - Versorgung - Therapie*. Springer. <https://doi.org/10.1007/978-3-662-61126-5>
- Hoth, S., Mühler, R., Neumann, K., & Walger, M. (2015). *Objektive Audiometrie im Kindesalter*. Springer. <https://doi.org/10.1007/978-3-642-44936-9>
- Hoth, S., & Steffens, T. (2016). Audiometrie – Die Untersuchung des Gehörs und seine technische Versorgung. In R. Kramme (Hrsg.), *Medizintechnik: Verfahren - Systeme - Informationsverarbeitung* (S. 1–61). Springer. https://doi.org/10.1007/978-3-662-45538-8_15-1
- Institut für Phonetik und Sprachverarbeitung. (2023). *Bavarian Archive for Speech Signals (BAS) Web Services* [Ludwig-Maximilians-Universität, München]. Verfügbar 19. Juni 2023 unter <https://hdl.handle.net/11858/00-1779-0000-0028-421B-4>

- International Bureau for Audiophonology. (1996). Audiometric Classification of Hearing Impairments. *BIAP Recommendation 02/1*. <https://www.biap.org/de/recommandations/recommandations/tc-02-classification/213-rec-02-1-en-audiometric-classification-of-hearing-impairments/file>
- ISO 389-1. (2018). DIN EN ISO 389-1:2018-06, Akustik – Standard-Bezugspegel für die Kalibrierung audiometrischer Geräte – Teil 1: Äquivalente Bezugs-Schwellenschalldruckpegel für reine Töne und supra-aurale Kopfhörer. (*ISO 389-1:2017*), Deutsche Fassung EN ISO 389–1:2018. <https://doi.org/10.31030/2757944>
- ISO 389-9. (2009). DIN EN ISO 389-9:2009-12, Akustik – Standard-Bezugspegel für die Kalibrierung audiometrischer Geräte – Teil 9: Vorzugs-Messbedingungen zur Bestimmung von Bezugs-Hörschwellenpegeln. (*ISO 389-9:2009*), Deutsche Fassung EN ISO 389–9:2009. <https://doi.org/10.31030/1505366>
- ISO 8253-1. (2011). DIN EN ISO 8253-1:2011-04, Akustik – Audiometrische Prüfverfahren – Teil 1: Grundlegende Verfahren der Luft- und Knochenleitungs-Schwellenaudiometrie mit reinen Tönen. (*ISO 8253-1:2010*), Deutsche Fassung EN ISO 8253–1:2010. <https://doi.org/10.31030/1705319>
- ISO 8253-3. (2022). DIN EN ISO 8253-3:2022-11, Akustik – Audiometrische Prüfverfahren – Teil 3: Sprachaudiometrie. (*ISO 8253-3:2022*), Deutsche Fassung EN ISO 8253–3:2022. <https://doi.org/10.31030/3296501>
- Jerlvall, L., & Arlinger, S. (1986). A Comparison of 2-dB and 5-dB Step Size in Pure-Tone Audiometry. *Scandinavian Audiology*, 15(1), 51–56. <https://doi.org/10.3109/01050398609045954>
- Kaernbach, C. (2001). Adaptive threshold estimation with unforced-choice tasks. *Percept Psychophys*, 63, 1377–88.
- Kandzia, F. (2005). *Automatisierte, binaurale und multifrequente Erfassung von DPOAE und AMFR auf einer DSP-basierten Messplattform* [Diplomarbeit]. Lehrstuhl für Realzeit-Computersysteme, TU München (Betreuer: J. Oswald).
- Kandzia, F. (2011). *Knochenleitungsevozierte otoakustische Emissionen in der audiologischen Diagnostik* [Dissertation]. Technische Universität München. Dr. Hut.
- Katz, J., Chasin, M., English, K., Hood, L. J., & Tillery, K. L. (Hrsg.). (2014). *Handbook of Clinical Audiology* (7. Aufl.). Lippincott Williams and Wilkins.
- Kiese-Himmel, C., Nickisch, A., & Werner, F. (2018). Auditive Verarbeitungs- und Wahrnehmungsstörung: Ist die Diagnose schon bei Erstklässlern möglich? *Laryngorhinootologie*, 97(01), 37–43. <https://doi.org/10.1055/s-0043-120580>
- Kisler, T., Reichel, U., & Schiel, F. (2017). Multilingual processing of speech via web services. *Computer Speech & Language*, 45, 326–347. <https://doi.org/10.1016/j.csl.2017.01.005>

- Kliem, K., & Kollmeier, B. (1995). Überlegungen zur Entwicklung eines Zweisilber-Kinder-Reimtests für die klinische Audiologie. *Audiologische Akustik*, *34*, 6–10.
- Lenarz, T., & Boenninghaus, H.-G. (2012). *Hals-Nasen-Ohren-Heilkunde* (14. Aufl.). Springer. <https://doi.org/10.1007/978-3-642-21131-7>
- Leurpendeur, A., Weikert, M., Fuhrmann, J., & Schaller, P. (2006). Diagnostik der Auditiven Verarbeitungs- und Wahrnehmungsstörungen – Ein pragmatisches Konzept aus der phoniatriisch-logopädischen Praxis. *Logos interdisziplinär*, *14*(2), 113–121.
- Levitt, H. (1971). Transformed Up-Down Methods in Psychoacoustics. *The Journal of the Acoustical Society of America*, *49*(2B), 467–477. <https://doi.org/10.1121/1.1912375>
- Lowell, E. L., Rushford, G., Hoversten, G., & Stoner, M. (1956). Evaluation Of Pure Tone Audiometry With Preschool Age Children. *Journal of Speech and Hearing Disorders*, *21*(3), 292–302. <https://doi.org/10.1044/jshd.2103.292>
- Ludwig, A. (2008). *Psychoakustische und elektrophysiologische Untersuchungen zu zentral-auditiven Verarbeitungsstörungen während der Kindesentwicklung* [Dissertation]. Universität Leipzig. Leipziger Univ.-Verl.
- Machelett, K. (1994). *Das Lesen von Sonagrammen in der Phonetik* [Magisterarbeit]. Institut für Phonetik und Sprachliche Kommunikation, LMU München.
- Machelett, K. (1996). *Das Lesen von Sonagrammen V1.0* (Begleitendes Hypertext-Dokument zur Vorlesung) (SAMP- Zeichen: <https://www.phonetik.uni-muenchen.de/studium/skripten/SGL/SGLIPATab.html>). Institut für Phonetik und Sprachverarbeitung. Ludwig-Maximilians-Universität München. <https://www.phonetik.uni-muenchen.de/studium/skripten/SGL/SGLHome.html>
- Mahomed, F., Swanepoel, D. W., Eikelboom, R. H., & Soer, M. (2013). Validity of Automated Threshold Audiometry: A Systematic Review and Meta-Analysis. *Ear and Hearing*, *34*(6), 745–752. <https://doi.org/10.1097/01.aud.0000436255.53747.a4>
- Martin, M. (1997). *Speech audiometry* (2. Aufl.). Whurr Publishers.
- McCormick, B. (1993). *Paediatric Audiology – 0-5 Years* (2. Aufl.). Whurr.
- Michalski, S., & Tewes, U. (2001). Zentrale Hörstörungen nachweislich trainierbar? *Hörakustik*, *10*, 98–106.
- Minning, U. (2009). *Bedienungsanleitung BrainFit home* [5. Ausgabe Deutsch, September 2009]. AUDIVA GmbH. Kandern.
- Moser, T., & Zenner, H.-P. (2019a). Peripheres Auditorisches System. In R. Brandes, F. Lang & R. F. Schmidt (Hrsg.), *Physiologie des Menschen: mit Pathophysiologie* (32. Aufl., S. 685–700). Springer. https://doi.org/10.1007/978-3-662-56468-4_52

- Moser, T., & Zenner, H.-P. (2019b). Zentrale auditorische Verarbeitung. In R. Brandes, F. Lang & R. F. Schmidt (Hrsg.), *Physiologie des Menschen: mit Pathophysiologie* (32. Aufl., S. 701–706). Springer. https://doi.org/10.1007/978-3-662-56468-4_53
- Mrowinski, D., Scholz, G., & Steffens, T. (Hrsg.). (2017). *Audiometrie: Eine Anleitung für die praktische Hörprüfung* (5. Aufl.). Georg Thieme Verlag. <https://doi.org/10.1055/b-0037-143208>
- Müller, J. (2002). *Abbildung der subjektiven Lautheitsfunktion durch objektive, nicht kooperative Methoden* [Diplomarbeit]. Lehrstuhl für Realzeit-Computersysteme, TU München.
- Nelson, H. D., Bougatsos, C., & Nygren, P. (2008). Universal newborn hearing screening: systematic review to update the 2001 US Preventive Services Task Force Recommendation. *Pediatrics*, *122*(1), e266–76. <https://doi.org/10.1542/peds.2007-1422>
- Neumann, K., Gross, M., Böttcher, P., Euler, H. A., Spormann-Lagodzinski, M., & Polzer, M. (2006). Effectiveness and efficiency of a universal newborn hearing screening in Germany. *Folia Phoniatr Logop*, *58*, 440–55.
- Neumann, K. (2012). *Ein Screening-Verfahren zur flächendeckenden Erfassung des Sprachstandes vier- bis viereinhalbjähriger Kinder. Optimierung, Validierung, Erweiterung, elektrophysiologische Fundierung* [Förderkennzeichen DLR 01GJ0982, Schlussbericht zum BMBF-Vorhaben]. Universitätsklinikum Bochum.
- Neumann, K., Baumeister, N., Baumann, U., Sick, U., Euler, H. A., & Weißgerber, T. (2012). Speech audiometry in quiet with the Oldenburg Sentence Test for Children. *International Journal of Audiology*, *51*(3), 157–163. <https://doi.org/10.3109/14992027.2011.633935>
- Neumann, K., & Euler, H. A. (2013). Kann ein Sprachstandsscreening zwischen dem Bedarf für Sprachförderung und für Sprachtherapie trennen? In A. Redder & S. Weirner (Hrsg.), *Sprachförderung und Sprachdiagnostik : Interdisziplinäre Perspektiven* (S. 174–198). Waxmann.
- Nickisch, A., Gohde, K., & Kiese-Himmel, C. (2013). AVWS bei Regelschülern im 2. Schuljahr: Welche Tests trennen auffällige von unauffälligen Kindern? *Laryngorhinotologie*, *92*(09), 594–599. <https://doi.org/10.1055/s-0031-1299758>
- Nickisch, A., & Kiese-Himmel, C. (2009). Auditive Verarbeitungs- und Wahrnehmungsleistungen 8- bis 10-Jähriger: Welche Tests trennen auffällige von unauffälligen Kindern? *Laryngo-Rhino-Otologie*, *88*(07), 469–476. <https://doi.org/10.1055/s-0028-1119403>
- Nickisch, A., Kiese-Himmel, C., Wiesner, T., & Schönweiler, R. (2019). Leitlinie „Auditive Verarbeitungs- und Wahrnehmungsstörungen“: Differenzialdiagnose. *HNO*, *67*(8), 576–583. <https://doi.org/10.1007/s00106-019-0645-3>
- Olsho, L. W., Koch, E. G., Carter, E. A., Halpin, C. F., & Spetner, N. B. (1988). Pure-tone sensitivity of human infants. *J Acoust Soc Am*, *84*, 1316–24.

- Olsho, L. W., Koch, E. G., & Halpin, C. F. (1987). Level and age effects in infant frequency discrimination. *The Journal of the Acoustical Society of America*, *82*(2), 454–464. <https://doi.org/10.1121/1.395446>
- Orchik, D. J., & Rintelmann, W. F. (1978). Comparison of pure-tone, warble-tone and narrow-band noise thresholds of young normal-hearing children. *Journal of the American Audiology Society*, *3*(5), 214–220.
- Oswald, J. A. (2005). *Objektive Audiometrie mit otoakustischen Emissionen und akustisch evozierten Potenzialen* [Dissertation]. Technische Universität München.
- Pals, C., Sarampalis, A., van Rijn, H., & Başkent, D. (2015). Validation of a simple response-time measure of listening effort. *The Journal of the Acoustical Society of America*, *138*(3), EL187–EL192. <https://doi.org/10.1121/1.4929614>
- Peterson, G. E., & Barney, H. L. (1952). Control Methods Used in a Study of the Vowels. *The Journal of the Acoustical Society of America*, *24*(2), 175–184. <https://doi.org/10.1121/1.1906875>
- Ptok, M., Kiese-Himmel, C., & Nickisch, A. (2019). Leitlinie „Auditive Verarbeitungs- und Wahrnehmungsstörungen“: Definition. *HNO*, *67*(1), 8–14. <https://doi.org/10.1007/s00106-018-0598-y>
- Ptok, M., & Meisen, R. (2008). Basale auditorische Funktionen bei Schülerinnen und Schülern der 3. und 4. Jahrgangsstufe. *Laryngorhinootologie*, *87*(4), 257–264. <https://doi.org/10.1055/s-2007-995313>
- Ptok, M. (2002). Entscheidungsverhalten in Schwellennähe bei psychoakustischen Untersuchungen: Decision Processes at Threshold Levels During Psychoacoustic Testing. *Laryngorhinootologie*, *81*(10), 696–701. <https://doi.org/10.1055/s-2002-35006>
- Ptok, M. (2011). Early detection of hearing impairment in newborns and infants. *Dtsch Arztebl Int* *2011*, *108*, 426–31. <https://doi.org/10.3238/arztebl.2011.0426>
- Ptok, M. (2014). Subjektive audiometrische Verfahren bei Kindern. *HNO*, *62*(10), 694–701. <https://doi.org/10.1007/s00106-014-2887-4>
- Ptok, M., & Nickisch, A. (2018). Audiometrische Verfahren zur Erkennung kindlicher Hörstörungen. *Monatsschrift Kinderheilkunde*, *166*(3), 218–227. <https://doi.org/10.1007/s00112-018-0442-1>
- Pürner, D. L. (2019). *Änderungen der peripheren und zentralen Schallverarbeitungsleistungen mit dem Alter* [Dissertation]. Technische Universität München.
- Pürner, D. L., Schirkonyer, V., & Janssen, T. (2022). Changes in the peripheral and central auditory performance in the elderly – A cross-sectional study. *Journal of Neuroscience Research*, *100*(9), 1791–1811. <https://doi.org/10.1002/jnr.25068>

- Rabbitt, P. (1966). Recognition: Memory for words correctly heard in noise. *Psychonomic Science*, 6(8), 383–384. <https://doi.org/10.3758/BF03330948>
- Rakerd, B., Seitz, P. F., & Whearty, M. (1996). Assessing the cognitive demands of speech listening for people with hearing losses. *Ear Hear*, 17(2), 97–106. <https://doi.org/10.1097/00003446-199604000-00002>
- Reiß, M., Reiß, G., Plontke, S. K., Leuwer, R., & Waldfahrer, F. (2021). Ohr – Anatomie und physiologische Grundlagen. In M. Reiß (Hrsg.), *Facharztwissen HNO-Heilkunde: Differenzierte Diagnostik und Therapie* (2. Aufl., S. 243–266). Springer. https://doi.org/10.1007/978-3-662-58178-0_19
- Robinson, D. W. (1991). Long-term repeatability of the pure-tone hearing threshold and its relation to noise exposure. *British Journal of Audiology*, 25(4), 219–235. <https://doi.org/10.3109/03005369109076594>
- Rosner, T. R. (2004). *Entwicklung und Evaluierung objektiver Analysemethoden der AM-FFR mit einer DSP-basierten Messplattform* [Diplomarbeit]. Lehrstuhl für Realzeit-Computersysteme, TU München.
- Rosner, T. R. (2011). *Zeitoptimierte Hörschwellschätzung mit otoakustischen Emissionen und auditorischen Potentialen* [Dissertation]. Technische Universität München. Dr. Hut.
- Sachse, S., & Buschmann, A. (2016). (Sprach-)Diagnostische Möglichkeiten bei 2-Jährigen. *Sprache · Stimme · Gehör*, 40(02), 68–75. <https://doi.org/10.1055/s-0042-101057>
- Schäffler, T., Sonntag, J., Hartnegg, K., & Fischer, B. (2004). The effect of practice on low-level auditory discrimination, phonological skills, and spelling in dyslexia. *Dyslexia*, 10(2), 119–130. <https://doi.org/https://doi.org/10.1002/dys.267>
- Schiel, F. (1999). Automatic Phonetic Transcription of Non-Prompted Speech. *Proc. of the ICPhS*, 607–610.
- Schiel, F. (2010, Mai). BASat : New Statistical Resources at the Bavarian Archive for Speech Signals. In N. C. (Chair), K. Choukri, B. Maegaard, J. Mariani, J. Odijk, S. Piperidis, M. Rosner & D. Tapias (Hrsg.), *Proceedings of the Seventh International Conference on Language Resources and Evaluation (LREC'10)* (S. 1069–76). European Language Resources Association (ELRA).
- Schirkonyer, V. (2006). *Implementierung eines Sprachverständlichkeitstests auf einer DSP-basierten Plattform* [Studienarbeit]. Lehrstuhl für Realzeit-Computersysteme, Technische Universität München (Betreuer: J. Oswald).
- Schirkonyer, V. (2007). *Entwicklung eines Tympanometriemoduls – Kombination von Tympanometrie und OAE-Messungen* [Diplomarbeit]. Lehrstuhl für Realzeit-Computersysteme, Technische Universität München (Betreuer: F. Kandzia).

- Schirkonyer, V., Keilmann, A., Harmuth, C., Wachtlin, B., Rader, T., & Bohnert, A. (2020). Der neue Mainzer Sprachtest für Kinder 3–7 Jahre (MATCH). *HNO*, *68*(7), 526–533. <https://doi.org/10.1007/s00106-019-00788-x>
- Schlauch, R. S., & Edward, C. (2007). A Multinomial Model for Identifying Significant Pure-Tone Threshold Shifts. *Journal of Speech, Language, and Hearing Research*, *50*(6), 1391–1403. [https://doi.org/10.1044/1092-4388\(2007/097\)](https://doi.org/10.1044/1092-4388(2007/097))
- Schmuziger, N., Probst, R., & Smurzynski, J. (2004). Test-retest reliability of pure-tone thresholds from 0.5 to 16 kHz using Sennheiser HDA 200 and Etymotic Research ER-2 earphones. *Ear Hear*, *25*, 127–32. <https://doi.org/10.1097/01.AUD.0000120361.87401.C8>
- Smith, M. C., Cable, H. R., & Wilmot, J. F. (1988). Pure tone audiometry: comparison of general practice and hospital services. *British Journal of General Practice*, *38*(317), 552–555. <https://bjgp.org/content/38/317/552>
- Steffens, T. (2017). Die systematische Auswahl von sprachaudiometrischen Verfahren. *HNO*, *65*(3), 219–227. <https://doi.org/10.1007/s00106-016-0249-0>
- Steigenberger, C. (2009). *Sprachverstehen im Störlärm mit dem Würzburger Kindersprachtest* [Dissertation]. Julius-Maximilians-Universität Würzburg.
- Steland, A. (2010). *Basiswissen Statistik: Kompaktkurs für Anwender aus Wirtschaft, Informatik und Technik* (2. Aufl.). Springer. <https://doi.org/10.1007/978-3-642-02667-6>
- Stuart, A., Stenstrom, R., Tompkins, C., & Vandenhoff, S. (1991). Test-retest variability in audiometric threshold with supraaural and insert earphones among children and adults. *Audiology*, *30*, 82–90.
- Studebaker, G. A. (1967). Intertest Variability and the Air-Bone Gap. *Journal of Speech and Hearing Disorders*, *32*(1), 82–86. <https://doi.org/10.1044/jshd.3201.82>
- Suchodoletz, W. v., & Sachse, S. (2008a). *Handbuch zum Elternfragebogen zur Früherkennung von Late Talkers (SBE-2-KT)*. Verfügbar 24. Juni 2023 unter <https://www.ph-heidelberg.de/index.php?id=11082>
- Suchodoletz, W. v., & Sachse, S. (2008b). *SBE-2-KT: Sprachbeurteilung durch Eltern – Kurztest für die U7*. Verfügbar 24. Juni 2023 unter <https://www.ph-heidelberg.de/index.php?id=11082>
- Suchodoletz, W. v., & Sachse, S. (2012). *Früherkennung von Sprachentwicklungsstörungen der SBE-2-KT und SBE-3-KT für zwei- bzw. dreijährige Kinder*. Kohlhammer.
- Thie, R. (2001). *Psychoakustisches Testsystem (PATSY basic). Gebrauchsanweisung* [GA 02000.2, Ausgabe: 29.10.2001]. Pilot Blankenfelde medizinisch elektronische Geräte GmbH. Blankenfelde.

- Thompson, N. C., Cranford, J. L., & Hoyer, E. (1999). Brief-Tone Frequency Discrimination by Children. *Journal of Speech, Language, and Hearing Research*, 42(5), 1061–1068. <https://doi.org/10.1044/jslhr.4205.1061>
- Völter, C., Shehata-Dieler, W., Baumann, R., & Helms, J. (2005). Entwicklung eines neuen Kindersprachtests und Erprobung an CI-Kindern. *Laryngorhinootologie*, 84(10), 738–743. <https://doi.org/10.1055/s-2005-861375>
- Wagener, K., Brand, T., & Kollmeier, B. (2006). Evaluation des Oldenburger Kinder-Reimtests in Ruhe und im Störgeräusch. *HNO*, 54, 171–178.
- Wagener, K., Brand, T., & Kollmeier, B. (1999). Entwicklung und Evaluation eines Satztests für die deutsche Sprache I: Design des Oldenburger Satztests (Development and evaluation of a German sentence test part I: Design of the Oldenburg sentence test). *Z. Audiol./Audiological Acoustics*, 38, 4–15.
- Wagener, K., & Kollmeier, B. (2005). Evaluation des Oldenburger Satztests mit Kindern und Oldenburger Kinder-Satztest. *Zeitschrift für Audiologie*, 44, 134–143.
- Warnke, R. (2015). *Brain-Boy® Gebrauchsanweisung* [GBA-BB-V06.5-08.12.2015]. Medi-TECH Electronic GmbH. Wedemark.
- Wechtenbruch, J. (1996). *Bildbenennung zur Wortschatzüberprüfung bei Kindern: Der Einfluss visueller und kognitiver Fähigkeiten auf die Testergebnisse des aktiven Wortschatztests für 3-6jährige* [Diss., Ludwig-Maximilians-Universität München. Wildegger].
- Weißgerber, T., Baumann, U., Brand, T., & Neumann, K. (2012). German Oldenburg Sentence Test for Children: A Useful Speech Audiometry Tool for Hearing-Impaired Children at Kindergarten and School Age. *Folia Phoniatica et Logopaedica*, 64(5), 227–233. <https://doi.org/10.1159/000342414>
- Wiesner, T., Bohnert, A., Limberger, A., Massinger, C., & Nickisch, A. (2019). *Konsenspapier der DGPP zur Hörgeräte-Versorgung bei Kindern, Vers. 4.0*. Verfügbar 2. Juli 2023 unter <https://dgpp.de/de/download/9/>
- World Health Organization. (2021). *World report on hearing* [Licence: CC BY-NC-SA 3.0 IGO]. <https://www.who.int/publications/i/item/9789240020481>
- Yoshinaga-Itano, C., Sedey, A. L., Coulter, D. K., & Mehl, A. L. (1998). Language of early- and later-identified children with hearing loss. *Pediatrics*, 102(5), 1161–71. <https://doi.org/10.1542/peds.102.5.1161>
- Yoshinaga-Itano, C. (2001). The Social-Emotional Ramifications of Universal Newborn Hearing Screening, Early Identification and Intervention of Children who are Deaf or Hard of Hearing. In R. C. Seewald & J. S. Gravel (Hrsg.), *A Sound Foundation Through Early Amplification 2001 – Proceedings of the Second International Conference* (S. 221–231). Phonak AG.

Zwicker, E. (1982). *Psychoakustik* (1. Aufl.). Springer. <https://doi.org/10.1007/978-3-642-68510-1>

Zwicker, E., & Fastl, H. (1999). *Psychoacoustics Facts and Models* (2. Aufl.). Springer. <https://doi.org/10.1007/978-3-662-09562-1>