

EUROPÄISCHE UNION DER HÖRAKUSTIKER e. V.

Förderpreis 2020

Zeitliche Schwellenbestimmung zur subjektiven Sprachdetektion

Masterarbeit

Verfasserin: Simone Wollermann

Erstprüfer: Dr. Hendrik Husstedt

Zweitprüfer: Prof. Dr. rer. nat. Jürgen Tchorz

Datum der Abgabe: 4. Oktober 2019

EUHA

Europäische Union der
Hörakustiker e.V.

Masterarbeit im Studiengang Hörakustik und Audiologische Technik
Institut für Signalverarbeitung
Universität zu Lübeck

Herausgeber: Europäische Union der Hörakustiker e. V.
Neubrunnenstraße 3, 55116 Mainz, Deutschland
Tel. +49 (0)6131 28 30-0
Fax +49 (0)6131 28 30-30
E-Mail: info@euha.org
Internet: www.euha.org

Alle hier vorhandenen Daten, Texte und Grafiken sind urheberrechtlich geschützt. Eine Verwertung über den eigenen privaten Bereich hinaus ist grundsätzlich genehmigungspflichtig.

© EUHA 2020

Zusammenfassung

Bei der akustischen Szenenanalyse trennt und erkennt der Mensch die Umgebungsgeräusche, welche an seinen Ohren ankommen. Er ist hierdurch in der Lage zu identifizieren, welche Signale wichtig und welche vernachlässigbar sind. Dies erfolgt mit Hilfe von den akustischen Merkmalen der Signale, welche sich grob in spektrale und zeitliche Eigenschaften unterteilen lassen. Da sich die akustische Szene im Alltag in Sekundenbruchteilen ändert, ist es wichtig, dass die Detektion verschiedener Signale möglichst schnell stattfindet. So kann z.B. ein plötzlich herannahendes Auto im Straßenverkehr rechtzeitig wahrgenommen werden. Für das soziale Leben ist besonders die Detektion von Sprachsignalen wichtig. In dieser Arbeit wurden daher die zeitlichen Schwellen zur subjektiven Sprachdetektion untersucht. Zeitliche Detektionsschwellen eines Sprachsignals definieren die minimale Länge des Signals, welche benötigt wird, damit es als Sprachsignal erkannt wird.

Es wurde eine Studie mit 30 älteren Probanden durchgeführt, deren Hörvermögen sich in die WHO Klassen 0 bis 3 einteilt. In einem 2-AFC Verfahren wurden zehn Sprachsignale mit unterschiedlichen Initialphonemen und daraus erzeugte Rauschsignale dargeboten. Die Sprach- und Rauschsignale wiesen die gleichen spektralen Eigenschaften auf, unterschieden sich jedoch in der zeitlichen Feinstruktur und in der Einhüllenden. Die verschiedenen Hörverluste der Probanden ermöglichten es zu untersuchen, welchen Einfluss das Hörvermögen auf die zeitlichen Detektionsschwellen hat. Eine Vergleichsgruppe aus fünf normalhörenden jungen Probanden ermöglichte zudem eine tendenzielle Untersuchung des Alterseinflusses.

Die Ergebnisse zeigten einen signifikanten Einfluss des Initialphonems auf die zeitlichen Detektionsschwellen. Hingegen ergab sich keine allgemeine Korrelation zwischen dem Hörvermögen und den Detektionsschwellen. Da Schwerhörige die zeitliche Feinstruktur schlechter verarbeiten können, lassen die Ergebnisse vermuten, dass die Einhüllende eines Sprachsignals mehr von Bedeutung für die Sprachdetektion ist als die zeitliche Feinstruktur. Zudem wurde ein Zusammenhang von Alter und Kognition mit den Detektionsschwellen gefunden, was auf einen Alterungsprozess der auditorischen Verarbeitung schließen lässt.

Abstract

In acoustic scene analysis, the human ear separates and recognizes the ambient sounds that arrive at the ears. This allows the identification of important and negligible signals. This identification is done with the help of the acoustic characteristics of the signals, which can be roughly divided into spectral and temporal characteristics. Since the acoustic scene is constantly changing, it is important that the detection of different signals takes place as quickly as possible. Thus, for example, a suddenly approaching car can be detected in time in road traffic. For social life, the detection of speech signals is particularly important. Therefore, in this thesis, the temporal detection thresholds for subjective speech detection were examined. Temporal detection thresholds for speech signals define the minimum length of a signal that is required for correctly recognizing it as a speech signal.

A study was conducted with 30 elderly volunteers whose hearing is classified in WHO classes 0 to 3. In a 2-AFC procedure, ten speech signals with different initial phonemes and corresponding noise signals were presented. The speech and noise signals had the same spectral characteristics, but differed in the temporal fine structure and in the envelope. The different grades of hearing loss of the test subjects made it possible to investigate the influence of hearing ability on the temporal detection thresholds. A comparison group of five young normal hearing volunteers also allowed a tendency analysis of the influence of age on the detection thresholds.

The results showed a significant influence of the initial phoneme on the temporal detection thresholds. However, there was no general correlation between hearing loss and detection thresholds. Since hearing impaired people are less able to process the temporal fine structure, the results suggest that the envelope of a speech signal is used for the speech detection rather than the temporal fine structure. In addition, a correlation between age and cognition with the detection thresholds was found, suggesting an aging process of auditory processing.

Abkürzungsverzeichnis

AFC	Alternative Forced Choice
FFT	Fast Fourier Transformation
GÖSA	Göttinger Satztest
IPA	International Phonetic Association
MoCA	Montreal Cognitive Assessment
OEG	Open Ear Gain (Eigenverstärkung des Gehörgangs)
PTA	Pure Tone Average (Mittelwert der Reintonaudiometrie): Mittelwert der Hörschwelle bei 0.5, 1, 2 und 4 kHz
RMS	Root Mean Square (Quadratisches Mittel)
SNR	Signal-to-Noise Ratio (Signal-Rausch-Abstand)
SPL	Sound Pressure Level (Schalldruckpegel) bezogen auf 20 μPa
SRT	Speech Reception Threshold (Sprachverständlichkeitsschwelle): SNR-Wert, bei dem ein Sprachverstehen von 50 % erreicht wird
VOT	Voice Onset Time
WHO	World Health Organization (Weltgesundheitsorganisation)

Abbildungsverzeichnis

2.1.	Menschliches Ansatzrohr zur Sprachbildung	5
2.2.	Übersicht der deutschen Vokale	6
2.3.	Übersicht der deutschen Konsonanten	6
2.4.	Einhüllende und zeitliche Feinstruktur eines Sprachsignals	10
2.5.	Typischer Verlauf einer psychometrischen Funktion	13
2.6.	Beobachtete und ratekorrigierte Wahrscheinlichkeit	13
2.7.	Reiz-Reaktionsschema eines Ja/Nein-Experiments	14
2.8.	Wahrscheinlichkeitsdichten für Stimuli mit und ohne Zielreiz	15
2.9.	Einfluss des internen Kriteriums	16
3.1.	Hörverlust der älteren Probanden	21
3.2.	PTA der Probanden in Abhängigkeit vom Alter	21
3.3.	Freifeldbezogener Frequenzgang des HDA 280 Kopfhörers	22
3.4.	Onset der Sprachsignale	24
3.5.	Exemplarische Zeitsignale von Sprache und Rauschen	25
3.6.	Einstellung der Lautstärke für die Schwellenbestimmung	26
3.7.	Versuchsaufbau und Ablauf der zeitlichen Schwellenbestimmung	26
4.1.	Test/Retest des SRT	29
4.2.	Test/Retest der zeitlichen Detektionsschwellen	29
4.3.	Schwellwerte in Abhängigkeit von der Darbietungsreihenfolge	30
4.4.	Verteilung der Detektionsschwelle pro Signal	32
4.5.	Zeitsignal /Ab/ und zugehörige Detektionsschwellen	34
4.6.	Zeitsignal /Ba/ und zugehörige Detektionsschwellen	34
4.7.	Zeitsignal /Ka/ und zugehörige Detektionsschwellen	35
4.8.	Zeitsignal /Na/ und zugehörige Detektionsschwellen	35
4.9.	Zeitsignal /Ra/ und zugehörige Detektionsschwellen	36
4.10.	Korrelation von Hörvermögen und zeitlichen Detektionsschwellen	37
4.11.	Korrelation von PTA und Detektionsschwellen pro Signal	38
4.12.	Einfluss der Frequenzaufösung	41
4.13.	Vergleich der Schwellwerte von Jung und Alt	42
4.14.	Schwellen in Abhängigkeit von Alter und Kognition	43
4.15.	Korrelation von Alter und MoCA Punktzahl pro Proband	44
4.16.	Relative Häufigkeit der MoCA Punktzahlen für Erinnerung und Visuospatial/Exekutiv	44

Tabellenverzeichnis

I. Grad des Hörverlustes nach WHO	20
II. Ausgewählte Worte für das Sprachmaterial	23
III. Signifikanz der Unterschiede zwischen den Schwellwerten in Abhängigkeit vom Initial- phonem	33

Inhaltsverzeichnis

1. Einleitung	1
2. Grundlagen	3
2.1. Signalklassifikation	3
2.2. Sprachwissenschaftliche Grundlagen	5
2.2.1. Spracherzeugung	5
2.2.2. Sprachwahrnehmung	7
2.2.3. Einfluss einer Schwerhörigkeit auf die Sprachwahrnehmung	10
2.3. Psychophysikalische Grundlagen	12
2.3.1. Einführung	12
2.3.2. Psychometrische Funktion	12
2.3.3. Signalentdeckungstheorie	14
2.3.4. Adaptive Messverfahren	17
3. Material und Methoden	19
3.1. Studiendesign	19
3.2. Probanden	20
3.3. Pegelbezug und Frequenzgang des Kopfhörers	22
3.4. Testmaterial	23
3.4.1. Sprachsignale	23
3.4.2. Generierung des Störgeräusches	24
3.4.3. Lautstärke der Testsignale	25
3.5. Zeitliche Schwellenbestimmung	26
4. Ergebnisse und Diskussion	28
4.1. Test-ReTest Reliabilität und Robustheit des Messverfahrens	28
4.2. Einfluss des Initialphonems auf die Detektionsschwellen	32
4.3. Einfluss des Hörvermögens auf die Detektionsschwellen	37
4.4. Einfluss von Alter und Kognition auf die Detektionsschwellen	41
5. Zusammenfassung und Ausblick	46
Literatur	48
A. Anhang	52
A.1. Methoden	52
A.2. Messergebnisse	58

1. Einleitung

Um sich in seinem Alltag sicher zurecht zu finden, nutzt der Mensch seine sensorischen Empfindungen zur Identifikation und Klassifikation von Objekten in seiner Umgebung. Die Fähigkeit, verschiedene Geräusche durch das Hören zu erkennen und zu identifizieren, wird als akustische Gnosis bezeichnet [1]. Um sich ein vollständiges Bild von den Geräuschen, welche einen umgeben, machen zu können, findet mit Hilfe des auditorischen Systems eine akustische Szenenanalyse statt [2]. Diese basiert zum einen auf kontextbezogenen Top-Down Prozessen, also auf der willkürlichen Wahrnehmung von Geräuschen, welche wir in bestimmten Situationen erwarten. Hinzu kommen sensorische Bottom-Up Prozesse, welche die unwillkürliche Wahrnehmung von Signalen anhand ihrer akustischen Merkmale beschreiben [2]. Da sich die akustische Szene in einem Sekundenbruchteil ändern kann, ist es von Vorteil, bereits sehr kurz wahrgenommene Signale richtig klassifizieren zu können. Hierfür wurden in den letzten Jahren verschiedene Studien durchgeführt, welche sich mit der zeitlichen Schwellenbestimmung zur Geräuschklassifikation beschäftigen. Obert und Tchorz (2018) untersuchten, wie kurz ein Signal sein darf, damit es von normal hörenden jungen Probanden korrekt einer von vier Signalklassen (Sprache, Tiergeräusch, Rauschen, Musik) zugeordnet werden kann [3]. Je nach Geräuschkategorie lagen die zeitlichen Schwellen zwischen 20 und 55 ms. Auf Basis dieser Studie wurden in zwei weiteren Untersuchungen ähnliche Messungen mit jungen normalhörenden sowie älteren schwerhörigen Probanden durchgeführt [4], [5]. Die Studien zeigten, dass die Schwellen bei älteren Probanden mit Hörverlust höher waren als bei jungen normalhörenden Probanden. Die zeitliche Schwellenbestimmung könnte somit ein hilfreicher und sinnvoller Test sein, um die auditorische Verarbeitung zu analysieren und die Hörfähigkeit eines Menschen zu evaluieren. Da jedoch der Hörverlust in beiden Studien mit dem Alter korrelierte, konnte keine präzise Aussage darüber getroffen werden, ob die erhöhten Schwellen aufgrund des Alters oder des Hörverlustes zustande kamen.

In allen drei Studien stellte sich zudem heraus, dass die Schwellen für Sprachsignale niedriger waren, als für die anderen Signalklassen. Sprache dient im gesellschaftlichen Leben als ein Mittel zur Verständigung [6]. Die Fähigkeit zu Hören und Sprache zu verstehen, ist demnach eine wichtige Eigenschaft für das soziale Leben des Menschen. Schinnerl et al. (2019) beschränkten sich in ihrer Untersuchung zur Geräuschklassifikation auf Sprachsignale und untersuchten die zeitliche Klassifikationsschwelle für normalhörende Probanden. Die Probanden stellten sich hierfür die zeitliche Schwelle, ab welcher sie die Signale als Sprache klassifizieren konnten, selbst ein. Im Mittel lag diese Schwelle bei ca. 52 ms. Es konnte weiterhin gezeigt werden, dass die Wahl des Initialphonems einen Einfluss auf die zeitlichen Schwellen aufweist [7].

In den bisherigen Studien fand die zeitliche Klassifikation der Signale sowohl anhand von spektralen als auch von den zeitlichen Eigenschaften der Signale statt. Es konnte folglich keine Aussage darüber getroffen werden, welche Eigenschaften der Signale für die korrekte Klassifikation benötigt werden. In dieser Arbeit soll daher untersucht werden, welchen Einfluss die zeitliche Struktur auf die zeitlichen Detektionsschwellen von Sprachsignalen hat. Des Weiteren konnten die in den bisherigen Studien aufgezeigten Unterschiede in den zeitlichen Schwellen nicht eindeutig auf den Hörverlust oder das Alter zurückgeführt werden, da beide Einflussfaktoren miteinander korrelierten. Es wird in dieser Arbeit

daher die Abhängigkeit der Detektionsschwellen von dem Hörvermögen analysiert, indem die Detektionsschwellen für Probanden in einer kleinen Altersspanne mit Hörverlusten der WHO Klassen 0 bis 3 miteinander verglichen werden. Hierbei wird darüber hinaus der Einfluss des Initialphonems auf die Schwellen untersucht. Zusätzlich wird durch einen Vergleich mit den Detektionsschwellen von jüngeren Normalhörenden zu älteren Normalhörenden der Einfluss des Alters analysiert und weiterhin der Zusammenhang der Detektionsschwellen mit den kognitiven Fähigkeiten überprüft.

Einige Grundlagen zur Signalklassifikation sowie die wichtigsten sprachwissenschaftlichen und psychoakustischen Grundlagen für die Sprachdetektion sind in Kapitel 2 beschrieben. Kapitel 3 beinhaltet das vollständige Studiendesign, inklusive Signalerstellung, Messaufbauten und verwendeter Messverfahren. In Kapitel 4 werden die Ergebnisse der Studie vorgestellt und weitergehend diskutiert. Eine Zusammenfassung sowie ein Ausblick auf zukünftige Untersuchungen ist Kapitel 5 zu entnehmen.

2. Grundlagen

2.1. Signalklassifikation

Die Erkennung von alltäglichen Objekten mittels Integration sensorischer Empfindungen wird als Gnosis bezeichnet [8]. In unserem Alltag sind wir von einer Vielzahl an verschiedensten Geräuschen umgeben. Die akustische Gnosis ist die Fähigkeit, diese Umgebungsgeräusche durch das Hören zu identifizieren und zu verarbeiten [1]. So kann z.B. in einem Büro das Klappern der Tastatur eines Kollegen zeitgleich zu Schritten auf dem Flur und einem Gespräch aus dem Nachbarraum wahrgenommen werden. Obwohl die Signale als einzelne Schallquellen identifiziert werden, kommt an den Ohren lediglich ein Summensignal aus der Überlagerung der verschiedenen Schallereignisse an. Um also wahrzunehmen, was in der Umgebung passiert und passend darauf reagieren zu können, müssen die einzelnen Schallquellen getrennt und ihre Eigenschaften separat verarbeitet werden. Dies erlaubt z.B., sich auf eine bestimmte Schallquelle zu konzentrieren und die restlichen Geräusche weitestgehend auszublenden [9].

Die Interpretation des komplexen Summensignals als Kombination einzelner Schallquellen wird nach Bregman (1990) als akustische Szenenanalyse bezeichnet [2]. Hierfür nutzt das auditorische System verschiedene Techniken, um gleichzeitig stattfindende Schallereignisse zu trennen, neu auftretende Signale als solche zu detektieren, oder den Fokus auf eine bestimmte Schallquelle zu lenken. Zudem können Signale klassifiziert, deren übermittelte Informationen verstanden und bisheriges Wissen über bestimmte Schallquellen genutzt werden, um diese leichter extrahieren zu können [9].

Neben dem elementaren, sensorischen Prozess des Hörens erfordert die auditorische Verarbeitung also auch den kognitiven Prozess der Interpretation des Gehörten [10]. Die Interpretation von Signalen geschieht kontextbezogen und gilt als Top-Down Verarbeitung. Sie erfolgt mit bewusster Aufmerksamkeit und auf Basis von bisherigen Erfahrungen. So können z.B. im Straßenverkehr Autos als solche wahrgenommen werden, weil willkürlich die Aufmerksamkeit darauf gelenkt wird, dass Autos in den Straßenverkehr gehören. Das Hauptaugenmerk bei der akustischen Szenenanalyse liegt jedoch bei den elementaren Prozessen, auch Bottom-Up Verarbeitung genannt. Diese beruhen auf der unbewussten Wahrnehmung der akustischen Merkmale des sensorischen Inputs, wie z.B. der spektralen oder zeitlichen Struktur eines Schallsignals [2].

Nach Bregman (1990) erfolgt die Interpretation des sensorischen Inputs zur Analyse der akustischen Szene anhand der Gruppierung von sequentiellen und simultanen Merkmalen des Summensignals [2]. Die sequentielle Gruppierung verknüpft den sensorischen Input über die Zeit. Aufeinanderfolgende Schallereignisse werden anhand der Ähnlichkeit ihres Spektrums entweder zu sogenannten Streams verbunden, oder auf verschiedene Geräuschquellen zurückgeführt [11]. Je langsamer der Wechsel im Spektrum erfolgt, desto eher wird ein einzelnes, sich zeitlich änderndes Signal wahrgenommen. Erfolgt dagegen ein plötzlicher Wechsel im Frequenzspektrum, wird dies als ein neu auftretendes Signal interpretiert [2].

Die simultane Gruppierung erfolgt anhand der zeitgleich auftretenden Komponenten des Summensignals. Auf Basis der physikalischen Eigenschaften, wie z.B. der Periodizität oder der räumlichen Herkunft, werden all jene Signalanteile formiert, welche entweder zu einer einzelnen, oder zu mehreren

Schallquellen gehören [11]. Auch die Harmonizität und die zeitliche Asynchronität von Signalen zählen zu den simultanen Merkmalen eines Summensignals. In einem harmonischen Signal entsprechen alle Frequenzen den Vielfachen einer Grundfrequenz. Wird eine solche Frequenzmischung in einem Signal wahrgenommen, wird sie als zusammengehörig zu einer Schallquelle interpretiert. Bei der Anwesenheit von zwei oder mehr Sprechern z.B. kann somit eine Differenzierung anhand der Grundfrequenz getroffen werden [2]. Der Effekt der zeitlichen Asynchronität bezieht sich hingegen auf die Tatsache, dass voneinander unabhängige Signale selten zu exakt dem gleichen Zeitpunkt starten oder enden. Eine zeitliche Asynchronität führt somit zu der Wahrnehmung unterschiedlicher Schallquellen, während eine Zuordnung zu der gleichen Schallquelle erfolgt, wenn Frequenzkomponenten als zeitlich synchron wahrgenommen werden [12]. Bereits die zeitliche Asynchronität einer einzelnen Frequenz von 35 ms erzeugt in einem komplexen Signal den Eindruck eines weiteren Schallereignisses [9].

Die Klassifikation von Umgebungsgeräuschen erfolgt somit anhand der spektralen sowie der zeitlichen Merkmale der Signale. Shafiro (2003) zeigte, dass eine höhere spektrale Auflösung von Signalen zu einer verbesserten Identifikation führt. Gygi et al. (2004) verwiesen zudem auf die zeitlichen Merkmale von Signalen, welche die Einhüllende, die Periodizität und die Konsistenz der zeitlichen Veränderung über die Frequenzen hinweg beinhalten. Des Weiteren verbessert sich die Diskriminationsfähigkeit für verschiedene Signale bis zu einer bestimmten Länge mit länger werdendem Signal [13]. Um sich im Alltag einen akkuraten und sicheren Eindruck von der sich ständig wechselnden akustischen Hörumgebung machen zu können, muss die Klassifikation von Signalen jedoch möglichst schnell erfolgen. So können geschulte Versuchspersonen Vokale erkennen, wenn der dazugehörige Sprachlaut weniger als 10 ms lang ist [14]. Jüngste Studien untersuchten die Fragestellung, wie kurz ein Signalanteil sein darf, um vom Menschen der richtigen Kategorie zugeordnet werden zu können [3]-[5], [7]. Da sich besonders die Detektion von Sprachsignalen als spannende Untersuchung gezeigt hat, werden im nachfolgenden einige Grundlagen zu Spracherzeugung und Sprachwahrnehmung dargelegt.

2.2. Sprachwissenschaftliche Grundlagen

2.2.1. Spracherzeugung

Sprachlaute, sogenannte Phone, werden durch die Artikulationsorgane gebildet, welche sich aus Lunge, Luftröhre inkl. Stimmlippen, Kehlkopf, Rachen und Mund zusammensetzen. Das Ansatzrohr, dargestellt in Abbildung 2.1, fasst die Hohlräume oberhalb der Stimmlippen zusammen, welche für die Formung und Verstärkung des Grundtons der menschlichen Stimme verantwortlich sind. Je nach Position der einzelnen Artikulationsorgane können verschiedene Sprachlaute gebildet werden.

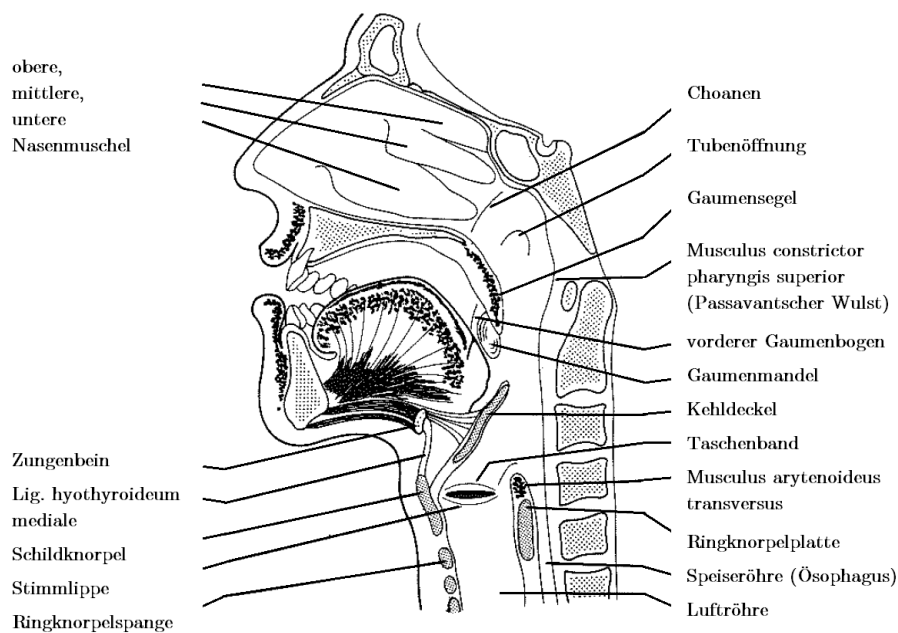


Abb. 2.1: Menschliches Ansatzrohr zur Sprachbildung. Darstellung angelehnt an Petursson (1996).

Jede Sprache hat ihr eigenes Lautinventar, bzw. Phoninventar, welches alle in dieser Sprache verwendeten Laute enthält. Allgemein lassen sich diese Sprachlaute in Vokale und Konsonanten unterteilen. Vokale sind, im Gegensatz zu Konsonanten, grundsätzlich stimmhaft, was bedeutet, dass sich die Stimmbänder bei der Artikulation eines Lautes schnell hintereinander öffnen und schließen, wodurch sie in Schwingung geraten und so einen Ton erzeugen. Vokale werden zudem ohne wesentliche Engstellen im Vokaltrakt und mit hauptsächlichlicher Schallabstrahlung über die Mundöffnung gebildet [15]. Die Kategorisierung für Vokale und Konsonanten unterscheidet sich je nach Literatur. Die IPA (engl.: International Phonetic Association) hat einen weltweiten Standard zur Notation der phonetischen Darstellung aller Sprachen erstellt. Dieser Standard unterteilt die Vokale nach Zungenlage und Mundöffnung. Die Zungenlage wird dabei in vorne, zentral und hinten unterteilt. Die Mundöffnung kann geschlossen, halb-geschlossen, halb-offen oder offen sein. Eine Übersicht der deutschen Vokale nach IPA ist in Abbildung 2.2 dargestellt.

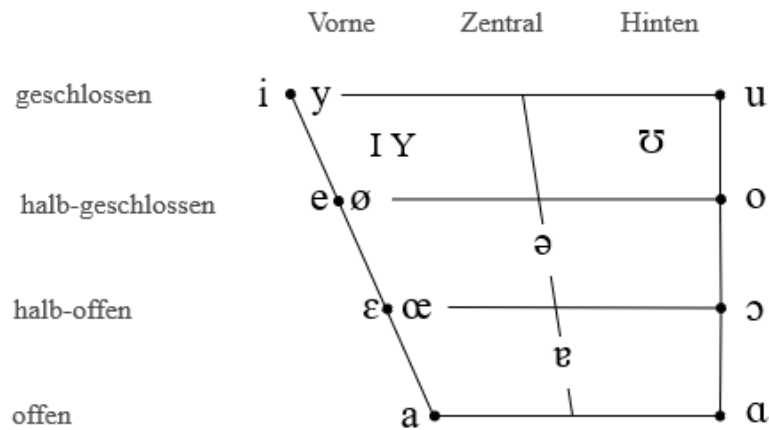


Abb. 2.2: Übersicht der deutschen Vokale. Darstellung angelehnt an IPA (2015).

Für die Erzeugung von Konsonanten muss eine Verengung im Vokaltrakt erfolgen. Die Unterteilung der Konsonanten erfolgt in die Art der Artikulation und den Ort im Vokaltrakt, an dem die Engstelle gebildet wird [15]. Auch hierfür hat die IPA eine Übersicht veröffentlicht, welche angepasst auf die deutschen Konsonanten in Abbildung 2.3 zu sehen ist.

		Artikulationsort												
		bilabial		labiodental		alveolar		postalveolar	palatal	velar	uvular	glottal		
Artikulationsart	Plosiv	p	b			t	d			k	g			
	Nasal		m				n				ŋ			
	Vibrant													
	Frikativ			f	v	s	z		ʃ	ç	x		ʁ	h
	Affrikate													
	Approximant									j				
	Lateraler Aproximant						l							

Abb. 2.3: Übersicht der deutschen Konsonanten. Darstellung angelehnt an IPA (2015).

Der Artikulationsort lässt sich aufteilen in:

- Labial: An den Lippen
- Dental: Mit der Zungenspitze an den Zähnen
- Alveolar: Mit der Zungenspitze an der oberen Zahnwurzel
- Palatal: Mit dem Zungenrücken gegen den harten Gaumen
- Velar: Mit dem Zungenrücken gegen den weichen Gaumen
- Uvular: Mit dem Zungenrücken im Bereich des Zäpfchens
- Pharyngal: Im Rachenraum
- Glottal: An der Glottis (Stimmorgan im Kehlkopf)

Die Artikulationsart beschreibt die Geste, welche von den Artikulationsorganen ausgeführt werden muss, um den jeweiligen Sprachlaut zu erzeugen. Konsonanten lassen sich hierdurch unterteilen in:

- Plosiv: Durch das momentane Unterbrechen des Luftstroms am Artikulationsort entsteht ein Verschlusslaut. Je nach zeitlicher Länge der Unterbrechung wird die akustische Energie mehr oder weniger reduziert und im Anschluss abrupt wiederhergestellt.
- Frikativ: Am Artikulationsort wird eine schmale Engstelle gebildet. Aufgrund des Drucks, welcher sich aufbaut, entstehen Luftwirbel und der Konsonant ist als Reibelaut zu hören.
- Approximant: Wird ähnlich gebildet wie ein Frikativ, ist jedoch durch weniger Druck stark abgeschwächt.
- Nasal: Der Mund wird am Artikulationsort durch einen Verschluss blockiert, sodass die Luft über die Nase entweichen muss. Hierdurch können zusätzliche Resonanzen oder Anti-Resonanzen entstehen.
- Lateraler Approximant: Die Zunge bindet einen zentralen Verschluss, sodass die Luft über die Seiten entweichen muss.
- Vibrant: Vibration von Lippen, Zungenspitze oder Zäpfchen durch schnell aufeinanderfolgendes Öffnen und Verschließen im Vokaltrakt.

Unabhängig von Artikulationsort und Artikulationsgeste lassen sich die Sprachlaute in die abstrakten Klassen der Phoneme einteilen. Zu einem Phonem gehören all jene Sprachlaute, welche die gleiche bedeutungsunterscheidende Funktion haben. Phoneme sind somit die kleinsten bedeutungsunterscheidenden Elemente von gesprochener Sprache, d.h. der Wechsel eines einzelnen Phonems innerhalb eines Wortes führt zu einer neuen Wortbedeutung (z.B. Maus - Haus). Fast immer ist ein Phon daher auch ein Phonem. Zur Unterscheidung werden Phone in eckigen Klammern und Phoneme durch Schrägstriche notiert (z.B. [a] = Phon und /a/ = Phonem). Einige Phone führen durch Austausch allerdings zu keinem wahrnehmbaren Unterschied in der Wortbedeutung. Diese werden als Allophone bezeichnet [16]. Es ist somit sowohl von Bedeutung wie ein Sprachlaut erzeugt wird, als auch wie Sprache vom Zuhörer wahrgenommen wird.

2.2.2. Sprachwahrnehmung

Die perzeptive Phonetik beschäftigt sich mit der Aufnahme und Verarbeitung von Sprachlauten. Dies beinhaltet die Schallaufnahme durch das Außenohr, die Schallweiterleitung durch das Mittel- und Innenohr, sowie die Verarbeitung im Gehirn. Betrachtet wird häufig die subjektive Wahrnehmung von Sprachlauten in Abhängigkeit von den physikalisch messbaren Eigenschaften des Signals, wie z.B. Tonhöhe oder Lautstärke. Des Weiteren gehören zur Thematik der Sprachwahrnehmung auch semantische Eigenschaften, grammatikalische Regeln oder die Identifikation von einzelnen Sprechern. Im Hinblick auf die Untersuchungen in dieser Arbeit wird die Sprachwahrnehmung im Nachfolgenden jedoch nur dahingehend beschrieben, anhand welcher Eigenschaften eines Signals der Mensch in der Lage ist, Sprachlaute als solche wahrzunehmen und zu erkennen.

In den letzten Jahrzehnten wurden durch unterschiedlichste Forschungen verschiedene Eigenschaften von Sprachlauten und deren Wahrnehmung aufgedeckt. So werden z.B. die akustischen Eigenschaften eines Sprachlautes durch vorherige oder nachfolgende Sprachlaute beeinflusst. Der zugehörige Begriff der Co-Artikulation entstand durch eine Untersuchung von Liberman (1967), bei welcher sich zeigte, dass sich die Frequenz des zweiten Formanten eines /d/ bei der Aussprache von /di/ und /du/ deutlich unterschied. Die Untersuchung zeigte weiterhin, dass es sich bei der Wahrnehmung von gewissen Sprachlauten um eine kategorische Wahrnehmung handelt. Minimale Änderungen in der Frequenz des zweiten Formanten eines einzelnen Phonems werden von dem Menschen nicht als solche erkannt. Ab einer gewissen Frequenzänderung wird jedoch plötzlich ein anderes Phonem wahrgenommen. Bei Musik hingegen ist der Mensch in der Lage, selbst feinste Frequenzunterschiede wahrzunehmen [17]. In einer Studie von Remez (1981) hörten Versuchspersonen lediglich drei Sinustöne, welche mit der Amplitude und der Frequenz der ersten drei Formanten eines Sprachsignals moduliert waren [18]. Der Klang wurde zunächst als unnatürlich und Science-Fiction mäßig beschrieben. Erhielten die Versuchspersonen jedoch den Auftrag, ein Sprachsignal aus dem Geräusch heraus zu hören, waren sie dazu in der Lage. Ist man also darauf bedacht ein Sprachsignal zu hören, wird das Gehörte anders interpretiert. Ein weiterer Effekt bei der Sprachwahrnehmung ist die audiovisuelle Integration, auch bekannt als McGurk-Effekt [19]. Die gleichzeitige Darbietung eines akustischen Sprachlautes und des visuellen Lippenbildes eines anderen Sprachlautes führt zu der Wahrnehmung eines dritten Sprachlautes.

Alle zuvor genannten Effekte treten vor allem für Sprachsignale auf, lassen sich jedoch auch bei verschiedenen Nicht-Sprach-Signalen feststellen. Um gewisse Effekte besser erklären zu können, wurden verschiedene Theorien entwickelt, welche die Sprachwahrnehmung im Allgemeinen beschreiben sollen. Die Motor Theorie erklärt die Wahrnehmung eines Sprachlautes als Wahrnehmung der artikulatorischen Geste des Sprechers [20]. Die Produktion und die Wahrnehmung von Sprache sind demnach eng miteinander verknüpft und von Natur aus spezifiziert. Die Motor Theorie bietet damit auch Erklärungsansätze für die Effekte der kategorischen Wahrnehmung und der audiovisuellen Integration.

Das Trace-Modell nutzt einen Erklärungsansatz auf Basis von neuronalen Netzen [21]. Ein Sprachsignal wird demnach zunächst in grobe Kategorien und davon ausgehend in immer feinere Kategorien unterteilt. Die Wahrscheinlichkeit für jeden Pfad hin zu der Wahrnehmung eines bestimmten Lautes wird verglichen und der Laut mit der größten Wahrscheinlichkeit wird wahrgenommen.

Eine dritte Theorie basiert auf der Aussage von Bloomfield, welcher bereits 1933 darauf verwies, dass Sprache distinktive Merkmale enthält, welche für die Wahrnehmung und die Identifikation von Sprachlauten notwendig sind [22]. Die Theorie der invarianten Merkmale führt an, dass ein wahrgenommenes Sprachsignal in eine diskrete Segmentfolge zerlegt wird [23]. Jedes Segment kann dann durch eine Reihe binärer Unterscheidungsmerkmale beschrieben werden. Die Verarbeitung eines Signals erfolgt dabei in drei Schritten. Zunächst werden akustische Orientierungspunkte anhand von Signalspitzen und Tälern, sowie Unterbrechungen in bestimmten Frequenzbereichen identifiziert. Aus Segmenten nahe dieser Orientierungspunkte werden verschiedene distinktive Merkmale extrahiert und im Anschluss miteinander kombiniert. In einem mentalen Lexikon sind Worte als Segmentfolgen gespeichert, welche

durch eine Gruppe besonderer Merkmale beschrieben werden können. Bei der Wahrnehmung eines Signals wird dann diejenige Segmentfolge ausgewählt, welche am besten zu den akustisch wahrgenommenen Eigenschaften des Signals passen [24].

Die Analyse zur Wahrnehmung eines Sprachsignals erfolgt nach Cole (1974) unter Verwendung von mindestens drei verschiedenen Merkmalen. Invariante Merkmale, kontextabhängige Merkmale und Merkmale, welche aus der Einhüllenden der Wellenform extrahiert werden [25]. Es konnte gezeigt werden, dass jede Silbe sowohl aus unveränderlichen als auch aus kontextabhängigen Merkmalen besteht. Weisen die unveränderlichen Merkmale auf mehrere Phoneme hin, können die kontextabhängigen Merkmale zu einer klaren Zuordnung verhelfen.

In mehreren Untersuchungen wurden verschiedene akustische Eigenschaften von Sprachlauten aufgezeigt, welche bei der Unterscheidung von Phonemen Anwendung finden [26], [27]. Affrikate und Nasale weisen demnach sehr schnelle Wechsel im Frequenzspektrum auf, wohingegen die Wechsel bei Approximanten deutlich langsamer sind. Plosivlaute zeichnen sich neben einem schnell wechselnden Frequenzspektrum auch durch schnelle Wechsel in der Amplitude aus. Um einzelne Plosive voneinander zu unterscheiden wird die Voice Onset Time (VOT) genutzt. Diese beschreibt die Zeit zwischen der Freisetzung des Konsonanten, also der Verschlussöffnung, und dem Einsatz der Stimme.

Sprache ist somit zum einen durch schnelle Wechsel in der Frequenz über die Zeit charakterisiert, was die Form des Vokaltraktes reflektiert, wenn die Artikulationsgeste von einem Laut zum nächsten wechselt. Des Weiteren entsprechen die schnellen oder langsamen Wechsel in der Amplitude eines Sprachlautes den zeitlichen Informationen eines Signals. Die Sprachverarbeitung in der Cochlea wird über eine Reihe an Bandpassfiltern realisiert, deren Mittenfrequenzen bestimmten Positionen auf der Basilarmembran entsprechen. Die zeitlichen Strukturen, welche sich aus den gefilterten Signalanteilen ergeben, sind in Abbildung 2.4 dargestellt. Es wird zwischen Einhüllender und zeitlicher Feinstruktur unterschieden. Als Einhüllende wird die langsame Änderung der Amplitude über die Zeit bezeichnet. Die zeitliche Feinstruktur beschreibt die schnellen Schwankungen nahe der Mittenfrequenz des schmalbandigen Signalanteils [28].

Als zeitliches Auflösungsvermögen wird die Fähigkeit bezeichnet, Änderungen in einem Stimulus zu detektieren, die sich z.B. aus der Modulation oder einer Lücke im Signal ergeben und bezieht sich eher auf die Einhüllende eines Signals als auf die zeitliche Feinstruktur [24]. Die Modulation von Rauschen oder Sinustönen mittels der Einhüllenden einzelner Frequenzbänder eines Sprachsignals kann bereits ab vier Frequenzbändern zu einem adäquaten Sprachverstehen in Ruhe führen [24]. Sobald jedoch Hintergrundgeräusche hinzukommen, wird das Sprachverstehen deutlich schlechter. Die Einhüllende scheint als zeitliche Information nicht auszureichen um die Sprache aus dem Störgeräusch heraus zu hören. Des Weiteren konnte gezeigt werden, dass eine zusätzliche Präsentation der zeitlichen Feinstruktur das Sprachverstehen deutlich verbessert [29].

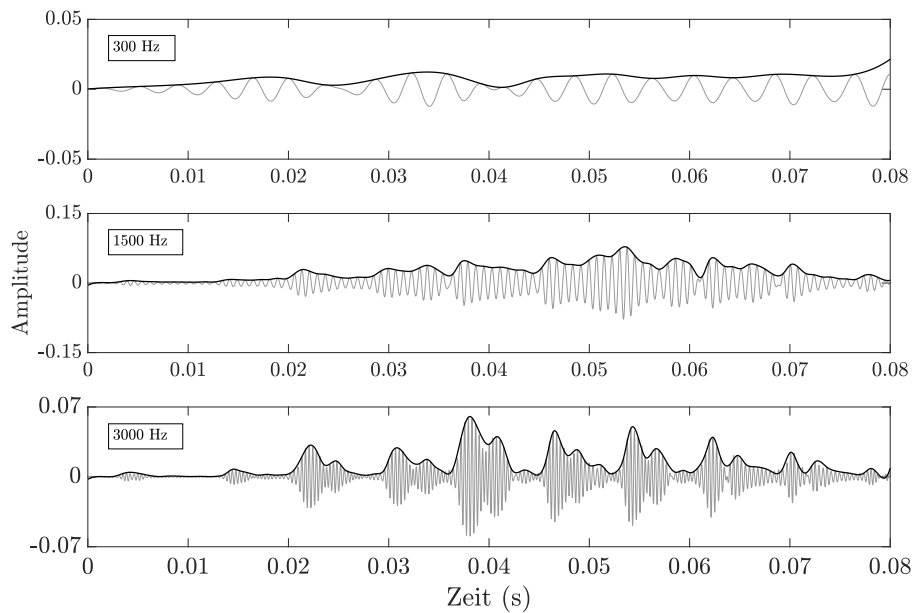


Abb. 2.4: Einhüllende und zeitliche Feinstruktur eines terzbandgefilterten Sprachsignals mit den Mittenfrequenzen 300, 1500 und 3000 Hz. Die grauen Kurven entsprechen der zeitlichen Feinstruktur des Signals und die schwarzen Kurven stellen die Einhüllende dar.

Die Fähigkeit Sprache im Störgeräusch zu verstehen, wird mit der Fähigkeit der Lückenerkennung, also der gerade noch wahrnehmbaren zeitlichen Lücke in einem Rauschsignal, in Verbindung gebracht [28]. Für das Sprachverstehen im Störgeräusch werden Informationen des Sprachsignals aus diesen kurzen Bereichen extrahiert, in welchen das Sprachsignal lauter als das Störgeräusch ist, um somit ein adäquates Sprachverstehen zu erzielen. Die minimale Länge, welche benötigt wird, um eine Lücke in einem Breitbandrauschen zu detektieren, liegt auf Basis verschiedener Untersuchungen bei knapp 2-3 ms [24].

2.2.3. Einfluss einer Schwerhörigkeit auf die Sprachwahrnehmung

In Deutschland sind nach einer aktuellen Studie von Gablenz (2017) ca. 16 % der Bevölkerung schwerhörig [30]. Die Ursachen für Hörverluste sind vielfältig und reichen von Lärmbelastung über ototoxische Medikamente zu erblichen Faktoren. Ein Hörverlust, welcher bedingt durch das Alter auftritt, wird als Presbyakusis bezeichnet [24]. Eine Schwerhörigkeit kann allgemein in drei Hauptgruppen unterteilt werden. Schallleitungsschwerhörigkeit, sensorineurale Schwerhörigkeit und eine kombinierte Schwerhörigkeit, welche sich aus einer Kombination der beiden zuvor genannten zusammensetzt.

Als Schallleitungsschwerhörigkeit wird eine reduzierte Schallübertragung bis zur Cochlea bezeichnet, welche z.B. akut aufgrund von Fremdkörpern im Gehörgang oder chronisch aufgrund von Anomalien im Außen- oder Mittelohr entstehen kann [31]. Eine Schallleitungsschwerhörigkeit wirkt sich meist über alle Frequenzen gleichermaßen aus und äußert sich dadurch, dass Schallsignale allgemein als leiser wahrgenommen werden. Bei entsprechender Anhebung der Lautstärke in den hörbaren Bereich kann daher ein normales Sprachverstehen in Ruhe und im Störgeräusch erreicht werden.

Die häufigste Art einer Schwerhörigkeit ist die sensorineurale Schwerhörigkeit [32]. Diese entsteht entweder durch Störungen im Bereich der Cochlea, im Hörnerv oder in den zentralen Anteilen der Hör-

bahn. Defekte in der Cochlea basieren auf der Degeneration der äußeren und inneren Haarsinneszellen. Das Ausmaß einer solchen Degeneration ist, besonders bei älteren Menschen, meist frequenzabhängig und für höhere Frequenzen stärker ausgeprägt als für tiefere Frequenzen [24]. Die Wahrnehmung von Sprache wird hierdurch verzerrt. Da vor allem leise hochfrequente Konsonanten (Frikative) gar nicht oder nur vermindert gehört werden, ist die Verständlichkeit von Sprache eingeschränkt und es kann unter anderem zu Verwechslungen von Sprachlauten kommen [32]. Anders als bei der Schalleitungsschwerhörigkeit führt eine reine Anhebung der Lautstärke jedoch nicht automatisch zu einem besseren Sprachverstehen. Während leise Signale durch die Degeneration der Haarzellen nicht mehr ausreichend verstärkt werden können, fehlt darüber hinaus die dämpfende Eigenschaft der Haarzellen für laute Signale. Es ergibt sich somit ein eingeschränkter Dynamikbereich, da die Hörschwelle zu höheren Pegeln verschoben ist, laute Signale aber schnell als unangenehm laut empfunden werden. Dies resultiert in einem schnelleren Lautheitsanstieg im verbliebenen Hörbereich als bei Normalhörenden, d.h. kleine Pegelunterschiede werden stärker wahrgenommen.

Durch entsprechende Anhebung und Komprimierung der Signale, z.B. durch Hörgeräte, in den Resthörbereich einer Person mit sensorineuraler Schwerhörigkeit ist häufig ein normales Sprachverstehen in Ruhe möglich. Das Sprachverstehen im Störgeräusch stellt jedoch weiterhin eine Herausforderung dar und ist oftmals, obwohl alle Signalkomponenten im hörbaren Bereich liegen, deutlich eingeschränkt. Diese Schwierigkeit ergibt sich daraus, dass bei einer sensorineuralen Schwerhörigkeit, anders als bei der reinen Schalleitungsschwerhörigkeit, das spektrale und das zeitliche Auflösungsvermögen zusätzlich eingeschränkt sind [32].

Einzelne Komponenten aus einem komplexen Signal zu erkennen und voneinander zu trennen wird als Frequenzselektivität bezeichnet. Als Frequenzauflösungsvermögen versteht man zudem die Fähigkeit, Frequenzen voneinander zu unterscheiden [24]. Hierfür ist jedem Ort auf der Basilarmembran im Innenohr eine bestimmte Frequenz zugeordnet. Aufgrund einer Schädigung der Haarzellen an bestimmten Bereichen auf der Basilarmembran ist diese Selektivität bei der sensorineuralen Schwerhörigkeit eingeschränkt. Frequenzen, welche in benachbarten Frequenzbereichen angesiedelt sind, können hierdurch schwerer getrennt werden. So stellt z.B. die Detektion eines Sprachsignals bei gleichzeitiger Darbietung eines spektral ähnlichen Störgeräusches eine Herausforderung dar [32].

Bezogen auf das zeitliche Auflösungsvermögen führt eine sensorineurale Schwerhörigkeit zu verschiedenen Einschränkungen. Die maskierenden Effekte von lauten Signalen, welche direkt vor leisen Signalen auftreten, sind z.B. stärker ausgeprägt als bei Normalhörenden. Dies resultiert durch den Verlust von Informationen, die gar nicht erst wahrgenommen werden, in einer Einschränkung im Sprachverstehen [32]. Auch die Fähigkeit der Lückenerkennung ist mit zunehmendem Hörverlust deutlich reduziert [12]. Da die Lückenerkennung jedoch das Heraushören von Sprache aus einem Störgeräusch unterstützt, ist durch die sensorineurale Schwerhörigkeit das Sprachverstehen im Störgeräusch erschwert. Zudem profitieren schwerhörige Personen bedeutend weniger von den Informationen aus der zeitlichen Feinstruktur eines Signals [29], [33], wodurch sich ein zusätzlicher Einfluss auf das Sprachverstehen im Störgeräusch ergibt.

2.3. Psychophysikalische Grundlagen

2.3.1. Einführung

Die Psychophysik, begründet von Gustav Theodor Fechner (1860), beschäftigt sich mit der „Untersuchung der Beziehungen zwischen den physikalischen Merkmalen von Reizen, z. B. Reizintensität, und dem psychischen Erleben dieser Reize“ [34]. Häufig ist das Ziel von Untersuchungen die Ermittlung der Detektionsschwelle oder der Unterscheidungsschwelle. Während die Detektionsschwelle die kleinste Reizintensität ist, welche zu einer Empfindung führt, beschreibt die Unterscheidungsschwelle die minimale Differenz zwischen zwei Stimuli, welche nötig ist, um einen Unterschied in der Empfindung herbeizuführen [35]. Zur Ermittlung der jeweiligen Schwellen werden einer Versuchsperson Stimuli mit unterschiedlichen Reizintensitäten präsentiert und das Antwortverhalten untersucht. Wird die Reizintensität im Verlauf einer Messung erhöht, gibt es jedoch keinen definierten Punkt, ab welchem plötzlich eine Empfindung stattfindet. Es handelt sich vielmehr um einen Intensitäts-Bereich, in welchem der Zielreiz bei einigen Darbietungen detektiert wird und bei anderen nicht [24]. Daher wird diejenige Reizintensität, welche in 50 % der Darbietungen zu einer Wahrnehmung führt, als Schwellwert bezeichnet. Die Abhängigkeit der Empfindung von der Reizintensität lässt sich als psychometrische Funktion darstellen.

2.3.2. Psychometrische Funktion

Eine psychometrische Funktion ist die graphische Darstellung einer stetigen eindimensionalen Wahrscheinlichkeitsverteilung, welche die Wahrnehmung als einen stationären stochastischen Prozess modelliert [36]. Die Form der psychometrischen Funktion lässt sich durch eine Basisfunktion beschreiben. Hierfür kann z.B. die logistische Funktion f_{sig} verwendet werden

$$f_{sig}(c_\theta, c_\sigma, \Phi) = \frac{1}{1 + e^{-\frac{c_\sigma}{c_\theta}(\Phi - c_\theta)}}, \quad (1)$$

wobei Φ die Reizintensität definiert und die Steigung c_σ an der Wahrnehmungsschwelle c_θ die Empfindlichkeit der Versuchsperson gegenüber einer Veränderung der Reizintensität definiert [35]. Der typische Verlauf einer psychometrischen Funktion ist in Abbildung 2.5 dargestellt. Während der Wertebereich der psychometrischen Funktion im Idealfall zwischen 0 und 1 liegt, wird er in der Praxis durch die Ratewahrscheinlichkeit und den Lapsus-Fehler eingeschränkt. Die Ratewahrscheinlichkeit γ ergibt sich durch Situationen, in denen die Versuchsperson unsicher in ihrer Antwort ist und durch Raten eine richtige Entscheidung trifft. Für n verschiedene Antwortmöglichkeiten liegt die Ratewahrscheinlichkeit bei $\frac{1}{n}$. Der Lapsus-Fehler λ hingegen beschreibt die Wahrscheinlichkeit, dass eine Versuchsperson im überschwelligen Bereich einen Fehler, z.B. durch Unaufmerksamkeit, macht und ist meistens vernachlässigbar [36].

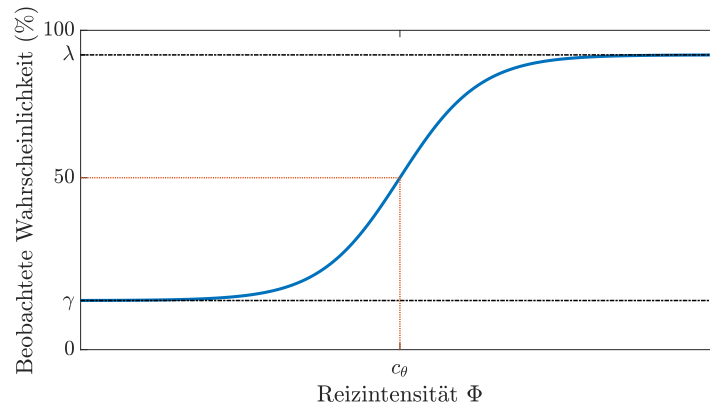


Abb. 2.5: Typischer Verlauf einer psychometrischen Funktion.

Die gesuchte Wahrscheinlichkeit einer Reizempfindung, auch Konvergenzniveau genannt, muss aufgrund der Ratewahrscheinlichkeit und des Lapsus-Fehlers korrigiert werden [37]. Die ratekorrigierte Wahrscheinlichkeit Ψ^* ergibt sich zu

$$\Psi^*(\Phi) = \frac{\Psi(\Phi) - \gamma}{1 - \gamma - \lambda}, \tag{2}$$

mit Ψ als tatsächlich beobachteter Wahrscheinlichkeit eines Ergebnisses. Gleichung 2 kann weiterhin nach Ψ umgestellt werden zu

$$\Psi(\Phi) = \gamma + (1 - \gamma - \lambda) \cdot \Psi^*(\Phi), \tag{3}$$

um die zu beobachtende Wahrscheinlichkeit für ein gewünschtes ratekorrigiertes Konvergenzniveau zu ermitteln. In Abbildung 2.6 ist das Verhältnis der beobachteten und der ratekorrigierten Wahrscheinlichkeit ohne Lapsus-Fehler und für eine Ratewahrscheinlichkeit von 50 % dargestellt.

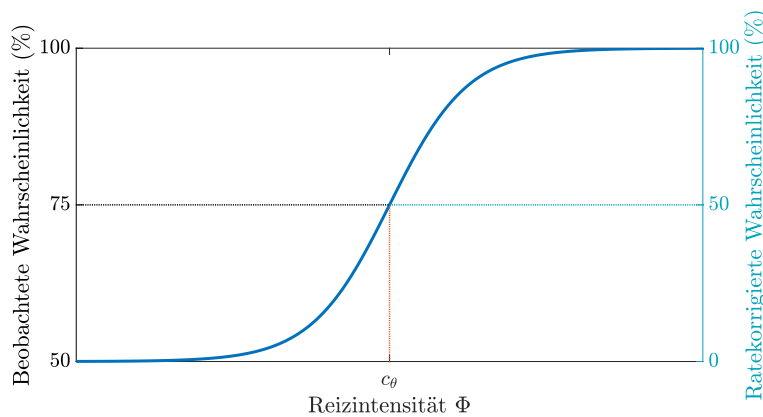


Abb. 2.6: Beobachtete und ratekorrigierte Wahrscheinlichkeit in Abhängigkeit von der Reizintensität.

Vor allem im Bereich der Detektionsschwelle c_θ ist die Empfindung einer Versuchsperson immer mit einer gewissen Unsicherheit behaftet. Das Antwortverhalten, welches sich daraus ergibt, lässt sich anhand der Signalentdeckungstheorie darlegen.

2.3.3. Signalentdeckungstheorie

Das Antwortverhalten einer Versuchsperson bei Entscheidungsprozessen, welchen eine gewisse Unsicherheit zugrunde liegt, kann durch die Signalentdeckungstheorie beschrieben werden [38]. Sie beruht auf der Annahme, dass sich das Antwortverhalten zum einen aus dem sensorischen Prozess und zum anderen aus einem Entscheidungsprozess zusammensetzt. Im sensorischen Prozess erzeugt die Darbietung eines Stimulus eine interne Antwort, welche zusätzlich von äußeren und inneren Faktoren abhängig ist. Im Entscheidungsprozess wird die interne Antwort mit einem internen Kriterium verglichen, um sich für oder gegen die Wahrnehmung eines Reizes zu entscheiden [39].

Bei sogenannten Ja/Nein-Experimenten werden einer Versuchsperson Stimuli mit und ohne Zielreiz dargeboten. Die Versuchsperson soll nach der Darbietung jedes Stimulus entscheiden, ob der gesuchte Zielreiz vorhanden war oder nicht. Die vier verschiedenen Möglichkeiten, welche sich aus dem dargebotenen Stimulus und der Antwort der Versuchsperson ergeben, lassen sich, wie in Abbildung 2.7, als Reiz-Reaktionsschema in einer 2D Kontingenztabelle darstellen.

		Reaktion	
		ja	nein
Reiz	vorhanden	Treffer (engl. Hit)	Verpasser (engl. Miss)
	nicht vorhanden	falscher Alarm (engl. False Alarm)	korrekte Ablehnung (engl. Correct Rejection)

Abb. 2.7: Reiz-Reaktionsschema eines Ja/Nein-Experiments als 2D Kontingenztabelle.

Die Performance der Versuchsperson ist neben den äußeren Einflüssen wie Beleuchtung, Versuchsaufbau oder Umgebungsgeräuschen maßgeblich von den inneren Faktoren abhängig. Diese lassen sich wiederum in eine Leistungskomponente (Sensitivität) und die Reaktionsneigung (internes Kriterium) unterteilen. Ein Bestandteil der Signalentdeckungstheorie ist die getrennte Bewertung dieser beiden Komponenten [39]. In die Leistungskomponente fließen neben der Konzentration und Aufmerksamkeit einer Versuchsperson auch die Empfindlichkeit des Sinnessystems mit ein. Da die Neurone im Gehirn eine gewisse Spontanaktivität zeigen, wird die neuronale, bzw. interne Antwort, welche auf die Darbietung eines Stimulus folgt, zusätzlich verrauscht. Dies bedeutet, dass die interne Antwort einer Versuchsperson bei der mehrfachen identischen Darbietung eines Stimulus unterschiedlich ausfallen kann, unabhängig davon, ob ein Zielreiz vorhanden ist oder nicht [40]. Es ergeben sich somit zwei Verteilungen der internen Antwort, eine für Stimuli ohne Zielreiz und eine für Stimuli mit Zielreiz, welche gewöhnlich als normalverteilt angenommen werden [41].

In Abbildung 2.8 sind diese Verteilungen beispielhaft dargestellt. Die Mittelwerte der zwei Verteilungen und der daraus resultierende Abstand zwischen den Kurven wird durch die Intensität des Zielreizes bestimmt. Je stärker die Intensität, desto leichter ist der Zielreiz zu erkennen und die Wahrscheinlichkeitsdichte für Stimuli mit Zielreiz verschiebt sich nach rechts, sodass sich die beiden Dichten weniger überlagern.

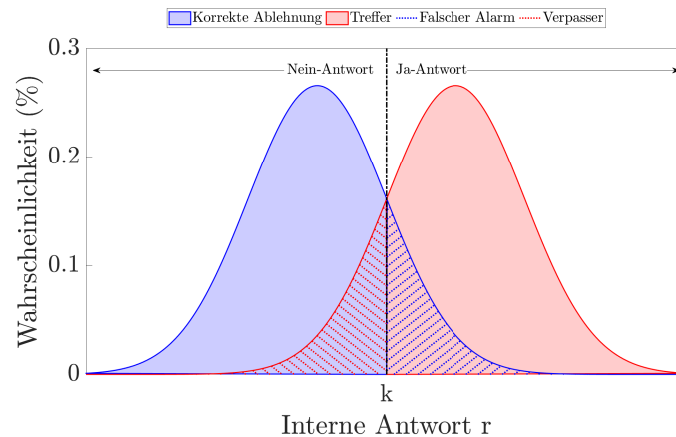


Abb. 2.8: Wahrscheinlichkeitsdichten für Stimuli mit Zielreiz (rote Kurve) und ohne Zielreiz (blaue Kurve) in Abhängigkeit von der internen Antwort.

Die Breite der Kurven liegt den internen und externen Einflüssen zugrunde. Je weniger störende Einflüsse vorhanden sind, desto besser kann ein Stimulus wahrgenommen werden. Hierdurch verringert sich die Streuung, was in schmaleren Kurven resultiert. Die Diskriminationsfähigkeit eines Zielreizes hängt somit sowohl vom Abstand der Mittelwerte als auch von der Streuung ab.

Ein Maß, um die Diskriminierbarkeit zu beschreiben, ist der Diskriminationsindex d' . Geht man davon aus, dass die Streuung, und die daraus resultierende Standardabweichung σ für beide Kurven gleich ist, ergibt sich d' aus

$$d' = \frac{\mu_+ - \mu_-}{\sigma}, \quad (4)$$

mit den Mittelwerten μ_+ für die Wahrscheinlichkeitsdichte von Stimuli mit Zielreiz und μ_- für Darbietungen ohne Zielreiz.

Der Diskriminationsindex beschreibt zwar die interne Antwort einer Versuchsperson, steht aber nicht mit dem tatsächlichen Entscheidungsprozess in Verbindung. All jene Reizintensitäten, welche in der Nähe der Wahrnehmungsschwelle liegen, befinden sich in dem Bereich, in welchem sich die beiden Dichten überlagern. Diese Reize sind somit schwer zu trennen und die Versuchsperson muss ein Kriterium k festlegen, anhand dessen sie sich für eine Antwort entscheidet. Dieses interne Kriterium beschreibt dann einen Schwellwert der internen Antwort, ab welchem die Versuchsperson mit Ja antwortet. Interne Antworten unterhalb des Schwellwertes werden mit Nein bewertet [38].

Die Wahl des internen Kriteriums hängt als interner Faktor von der Reaktionsneigung, also der Antworttendenz (engl. response bias) der Versuchsperson ab. Diese ergibt sich unter anderem aus der Persönlichkeit, den Erwartungen oder der Kosten-Nutzen-Abwägung. Abbildung 2.9 stellt den Einfluss des internen Kriteriums auf das Antwortverhalten dar. Bei einer neutralen Wahl des Kriteriums liegt dieses genau auf der Schnittstelle der beiden Verteilungsdichten (Abbildung 2.9b). Wählt eine Versuchsperson mit der liberalen Strategie ein eher niedriges Kriterium, wird eine hohe Anzahl von Treffern erzielt, was jedoch auch zu einer hohen Anzahl an falschen Alarmen führt (Abbildung 2.9a). Um die Anzahl falscher Alarme zu reduzieren kann mit der konservativen Strategie ein hohes Kriterium

angesetzt werden, jedoch erhöht sich dadurch zeitgleich die Anzahl an Verpassern (Abbildung 2.9c). Bei Entscheidungen, denen eine Unsicherheit zugrunde liegt, sind Fehler daher unvermeidbar. Durch die Wahl des internen Kriteriums kann die Versuchsperson jedoch die Art der Fehler modifizieren.

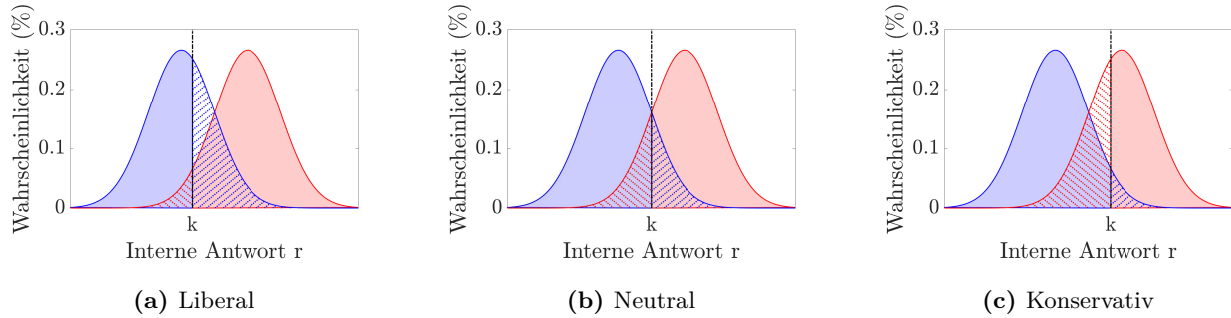


Abb. 2.9: Einfluss des internen Kriteriums auf das Antwortverhalten der Versuchsperson.

Bei der Untersuchung von Wahrnehmungsschwellen ist das interne Kriterium dennoch als Nachteil zu nennen, da Zielreize nur dann auch als solche wahrgenommen werden, wenn die zugehörigen Stimuli zu internen Antworten oberhalb des Kriteriums führen.

Um dem Einfluss des internen Kriteriums vorzubeugen, wurden sogenannte Alternative Forced-Choice-Methoden (AFC-Methoden) entwickelt [40]. Bei diesen Methoden werden der Versuchsperson zwei oder mehr Stimuli präsentiert, deren Darbietung räumlich oder zeitlich getrennt ist. Die Versuchsperson hat im Anschluss die Aufgabe auszuwählen, welcher präsentierte Stimulus den gesuchten Reiz enthalten hat. Die ermittelten Schwellwerte sind häufig niedriger als bei Ja/Nein-Paradigmen [36], da sich die Versuchsperson auch bei Unsicherheit für einen bestimmten Stimulus entscheiden muss. Auch bei Reizintensitäten, welche zu internen Antworten unterhalb eines internen Kriteriums führen würden, könnte sich die Versuchsperson trotzdem für den richtigen Stimulus entscheiden.

Bei einem AFC Verfahren hängt die Entscheidung der Versuchsperson davon ab, welcher der präsentierten Stimuli die größte interne Antwort geliefert hat, ohne dass ein internes Kriterium gesetzt werden muss. Für ein 2-AFC Verfahren (zwei Stimuli werden präsentiert) ergibt sich die Wahrscheinlichkeit richtig zu antworten somit aus der Wahrscheinlichkeit, dass der Stimulus, welcher den Zielreiz enthalten hat, eine größere interne Antwort liefert als der Stimulus ohne Zielreiz [40]. Die Wahrscheinlichkeit für eine richtige Antwort folgt somit aus

$$P(\text{korrekt}) = P[r_+(\Phi) > r_-(\Phi)], \quad (5)$$

mit r_+ als interner Antwort für den gesuchten Stimulus, r_- als interner Antwort für den falschen Stimulus sowie Φ als Reizintensität.

Unabhängig davon, ob als Antwortparadigma ein Ja/Nein Experiment oder ein AFC-Verfahren genutzt wird, finden für die Bestimmung eines Schwellwertes in Abhängigkeit von der Reizintensität adaptive Messverfahren Anwendung.

2.3.4. Adaptive Messverfahren

Bei Untersuchungen mit adaptiven Messverfahren wird die Reizintensität in Abhängigkeit von der vorherigen Antwort der Versuchsperson angepasst. Im Vergleich zu klassischen Messverfahren führt dies in Bezug auf die Versuchsdauer und die Genauigkeit der Schwellwertbestimmung zu einer höheren Leistungsfähigkeit [35]. Adaptive Messverfahren können entweder nicht-parametrisch oder parametrisch sein. Parametrische Messverfahren werden angewandt, wenn die grundsätzliche Form der psychometrischen Funktion, welche den Messdaten zugrunde liegt, bekannt ist. Andernfalls finden die nicht-parametrischen Messverfahren Anwendung, auf welche im Rahmen dieser Arbeit Bezug genommen wird.

Nach Hatzfeld (2014) müssen für ein nicht-parametrisches, adaptives Messverfahren vier Kriterien festgelegt werden:

- Verlaufsregel, also wann und wie sich die Reizintensität ändert,
- Startbedingung, also mit welcher Reizintensität begonnen wird,
- Stoppbedingung, also wann die Messung beendet wird und
- Berechnungsvorschrift, also wie der Schwellwert berechnet wird.

Als Verlaufsregel haben sich verschiedene Varianten der Staircase-Methode, auch Up/Down-Methode genannt, etabliert [37]. Bei der einfachen Up/Down-Methode ergibt sich die Reizintensität direkt aus der vorherigen Antwort der Versuchsperson. Für die $n + 1$ -te Darbietung ergibt sich eine Reizintensität von

$$\Phi_{n+1} = \Phi_n - \delta(2 \cdot Z_n - 1), \quad (6)$$

mit Φ_n als Reizintensität der n -ten Darbietung und Z_n als zugehöriger Antwort des Probanden. Die Schrittweite δ ist dabei in beide Richtungen gleich groß. Antwortet die Versuchsperson richtig ($Z_n = 1$) wird die Reizintensität um die Schrittweite verringert, bei einer falschen Antwort ($Z_n = 0$) wird sie um die Schrittweite erhöht. Diese Methode konvergiert gegen diejenige Reizintensität, bei welcher mit einer Wahrscheinlichkeit von 50 % eine Detektion des Zielreizes stattfindet [35].

Um auch andere Detektionswahrscheinlichkeiten und ratekorrigierte Konvergenzniveaus abzufragen, kann eine Änderung der Verlaufsregel erfolgen. Neben der stochastischen Approximation [42] und der beschleunigten stochastischen Approximation [43] können modifizierte Up/Down-Methoden, wie die transformierte Up/Down-Methode [44] oder die nicht-parametrische Up/Down-Methode [45] angewandt werden.

Eine weitere Variante der simplen Up/Down-Methode, welche jedes gewünschte Konvergenzniveau abfragen kann, ist die gewichtete Up/Down-Methode nach Kaernbach (1991), bei welcher die Schrittweite für falsche und richtige Antworten unterschiedlich gewählt wird [46]. Das Verhältnis der Schrittweite für richtige Antworten δ_{\downarrow} und der Schrittweite für falsche Antworten δ_{\uparrow} lässt sich berechnen aus

$$\delta_{\uparrow} \cdot (1 - \Psi) = -\delta_{\downarrow} \cdot \Psi, \quad (7)$$

für ein gewünschtes Konvergenzniveau Ψ . Als Beispiel liefert ein 2-AFC Verfahren eine Ratewahrscheinlichkeit von 50 % und erfordert nach Gleichung 3, unter Vernachlässigung des Lapsus-Fehlers, ein ratekorrigiertes Konvergenzniveau von 75 %. Aus Gleichung 7 folgt für die gewichtete Up/Down-Methode somit ein Verhältnis der Schrittweite von $3 \cdot \delta_{\downarrow} = -1 \cdot \delta_{\uparrow}$, was bedeutet, dass die Schrittweite für falsche Antworten dreimal so groß ist wie die Schrittweite für richtige Antworten.

Als Startbedingung für die jeweilige Verlaufsregel kann entweder eine Reizintensität gewählt werden, die überschwellig, also oberhalb der vermuteten Detektionsschwelle, oder unterschwellig liegt. In beiden Fällen führt das adaptive Messverfahren dazu, dass die Reizintensität im Verlauf der Messung an die wahre Detektionsschwelle angenähert wird. Die Wahl der Startbedingung kann jedoch einen Einfluss auf die Motivation der Versuchsperson haben. Ein überschwelliger Startwert kann z.B. sinnvoll sein, um der Versuchsperson mehr Sicherheit in der Bearbeitung der Aufgabe zu geben, da sie mit den Teststimuli vertraut gemacht wird, wodurch sich die Lernphase verkürzen kann.

Die Stoppbedingung sowie die Berechnungsvorschrift können auf verschiedene Weise getroffen werden. Bei der stochastischen Approximation wird eine minimale Grenze für die Schrittgröße festgelegt. Ist diese erreicht, so ist die Messung beendet. Die zuletzt getestete Reizintensität wird dann als Schwellwert verwendet [37]. Die häufigsten Stoppbedingungen für AFC-Verfahren mit Up/Down-Methoden sind das Erreichen einer bestimmten Anzahl an Darbietungen, oder nachdem eine bestimmte Anzahl an Wendepunkten erreicht wurde. Der Schwellwert wird aus den Reizintensitäten an einer gewissen Anzahl an letzten Wendepunkten oder letzten Darbietungen gemittelt. Diese Mittelung kann sowohl linear als auch geometrisch erfolgen [47].

3. Material und Methoden

3.1. Studiendesign

Diese Studie befasst sich mit der zeitlichen Schwellenbestimmung zur subjektiven Sprachdetektion. Die ermittelten Schwellwerte, also die minimale Signallänge eines Sprachsignals, welche benötigt wird um es als Sprache wahrzunehmen, werden im Nachfolgenden als zeitliche Detektionsschwellen bezeichnet. Folgende Fragestellungen werden untersucht:

- Welchen Einfluss haben die Hörschwelle und das Sprachverstehen im Störgeräusch auf die Detektionsschwellen?
- Welchen Einfluss hat das Initialphonem des Testsignals auf die Detektionsschwellen?
- Welchen Einfluss hat das Alter auf die Detektionsschwellen?

An der Studie haben 30 ältere Probanden mit und ohne Hörverlust, sowie fünf normalhörende junge Probanden teilgenommen. Die ersten 20 älteren Probanden wurden an zwei Terminen mit einer Woche Abstand eingeladen, um die Test-Retest-Reliabilität zu überprüfen. Die jüngeren Probanden wurden eingeladen, um den Einfluss des Alters auf die Detektionsschwellen untersuchen zu können.

Alle Messungen fanden in einem akustisch abgeschirmten Raum des Deutschen Hörgeräte Instituts statt. Nach einem Anamnese-Gespräch und einer Otoskopie wurde bei allen Probanden eine Reintonaudiometrie am ACAM 5 Audiometer durchgeführt, um die individuelle Hörschwelle zu ermitteln. Aus den audiometrischen Daten wurde anschließend der PTA (engl.: Pure Tone Average) berechnet, welcher den Mittelwert der Hörschwelle bei 0.5, 1, 2 und 4 kHz beschreibt.

Zur Bestimmung des Sprachverstehens im Störgeräusch wurde die Sprachverständlichkeitsschwelle (SRT, engl.: Speech Reception Threshold) gemessen. Hierfür sollte ein Test angewandt werden, welcher in möglichst kurzer Zeit ein aussagekräftiges Ergebnis liefern kann. Die Wahl fiel auf den Göttinger Satztest (GÖSA), da dieser aufgrund der verwendeten Alltagssätze keine Darbietung einer separaten Trainingsliste erfordert [48].

Der GÖSA wurde im Freifeld mit den Oldenburger Messprogrammen, Version 1.3.5.1, durchgeführt. Die Darbietung der Signale erfolgte über einen Lautsprecher des Typen „8020C“ des Herstellers Genelec. Die Sitzposition der Probanden war auf einen Abstand von einem Meter frontal vor dem Lautsprecher festgelegt. Der Pegel des Störgeräusches war, unabhängig vom Hörverlust der Probanden, fest auf 65 dB SPL eingestellt und die Darbietung erfolgte nicht-kontinuierlich (kein Störgeräusch während der Antwortphase). Der Pegel vom Sprachsignal wurde adaptiv je nach Antwort der Probanden angepasst und war zu Beginn der Messung für einen Signal-Rausch-Abstand (SNR) von 0 dB auf 65 dB SPL eingestellt. Jeder Proband absolvierte eine Testliste mit 30 Sätzen.

Es wurde vermutet, dass die kognitive Leistung einen Einfluss auf die Detektionsschwellen haben könnte. Um mögliche Auffälligkeiten in der kognitiven Verarbeitung der Probanden aufzudecken, wurde daher als kognitiver Screening-Test der Montreal Cognitive Assessment (MoCA) Test durchge-

führt. Der MoCA beinhaltet Aufgaben aus acht unterschiedlichen Bereichen (Visuospatial/Exekutiv, Benennen, Gedächtnis, Aufmerksamkeit, Sprache, Abstraktion, Erinnerung und Orientierung) und wird mittels Punktesystem ausgewertet. Die Einweisung sowie die Punkteverteilung sind beim MoCA streng vorgegeben und wurden bei allen Probanden gleichermaßen durchgeführt [49]. Das verwendete Aufgabenblatt ist Anhang A.1 zu entnehmen.

Für den Test zur Bestimmung der Detektionsschwellen wurden die Probanden vor einen Touchscreen platziert. Die zeitlichen Detektionsschwellen wurden für zehn Signale mit unterschiedlichen Initialphonen ermittelt, welche über alle Probanden nach einem lateinischen Quadrat randomisiert dargeboten wurden. Die Wiedergabe der Signale erfolgte über den „HDA 280“ Kopfhörer der Firma Sennheiser, welcher an einer mit dem Laptop verbundenen „Fireface 802“ Soundkarte von RME angeschlossen war. Das Messprogramm für die zeitliche Schwellenbestimmung wurde in Matlab R2018b realisiert. In einem 2-AFC Verfahren lag die Aufgabe der Probanden darin, nach der Darbietung eines Sprach- und eines spektral angeglichenen Störsignals, dasjenige Signal auszuwählen, welches sie als Sprache wahrgenommen haben. Die Länge der dargebotenen Signale wurde je nach Antwort der Probanden adaptiv angepasst. Eine detailliertere Beschreibung des Ablaufs folgt in Abschnitt 3.5.

3.2. Probanden

Als Probanden sollten ältere Erwachsene zwischen 60 und 85 Jahren mit verschiedenen Hörverlusten eingeladen werden. Die Hörverluste wurden, wie von der Weltgesundheitsorganisation (WHO) vorgeschlagen und in Tabelle I zu sehen, in vier Kategorien eingeteilt. Hierbei ist jeweils der PTA des besseren Ohres ausschlaggebend. Um möglichst symmetrische Hörverluste miteinander zu vergleichen, sollte sich der PTA zwischen den Ohren um weniger als 10 dB unterscheiden. Da es sich bei dem Testmaterial um deutsche Sätze und Wörter handelte, wurde zudem die deutsche Muttersprache als Einschlusskriterium vorausgesetzt.

Tabelle I: Grad des Hörverlustes nach WHO

Grad der Schwerhörigkeit	PTA
WHO 0 - keine Schädigung	0 - 25 dB
WHO 1 - leichte Schwerhörigkeit	26 - 40 dB
WHO 2 - mittelgradige Schwerhörigkeit	41 - 60 dB
WHO 3 - hochgradige Schwerhörigkeit	61 - 80 dB

Die Auswahl der Probanden erfolgte über die Probandendatenbanken der Akademie für Hörakustik und der Arbeitsgruppe Auditive Kognition der Universität zu Lübeck. Insgesamt wurden 31 ältere Probanden eingeladen. Die Daten einer Probandin wurden für die Auswertung nicht berücksichtigt, da aufgrund von Handhabungsschwierigkeiten und Sekundenschlaf keine robusten Ergebnisse ermittelt werden konnten. Somit wurden die Ergebnisse von 30 älteren Probanden (15 männlich, 15 weiblich) analysiert. Das Alter der Probanden lag im Mittel bei 72,7 Jahren mit einer Standardabweichung von 6,3 Jahren. Der mittlere Hörverlust über alle älteren Probanden ist in Abbildung 3.1 dargestellt.

Die große Streuung lässt sich dadurch erklären, dass in der Darstellung die Hörverluste der WHO

Klassen 0 bis 3 vertreten sind. Die individuellen Hörschwellen der Probanden können Anhang A.2 entnommen werden.

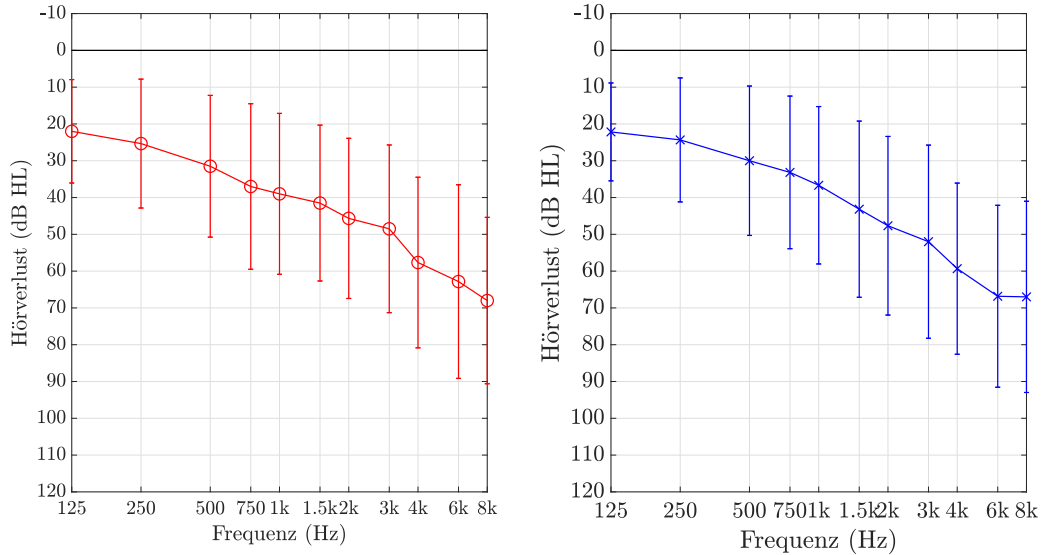


Abb. 3.1: Mittlerer Hörverlust mit Standardabweichung über alle 30 älteren Probanden.

Um die Effekte des Hörverlustes getrennt von den Effekten des Alters zu untersuchen, wurden die Probanden gezielt so ausgewählt, dass keine Korrelation zwischen Hörverlust und Alter besteht. Wie in Abbildung 3.2 zu sehen, liefert der Spearman Korrelationskoeffizient keine signifikante Korrelation zwischen Alter und Hörverlust ($r^2 = 0,06$ mit $p = 0,19$).

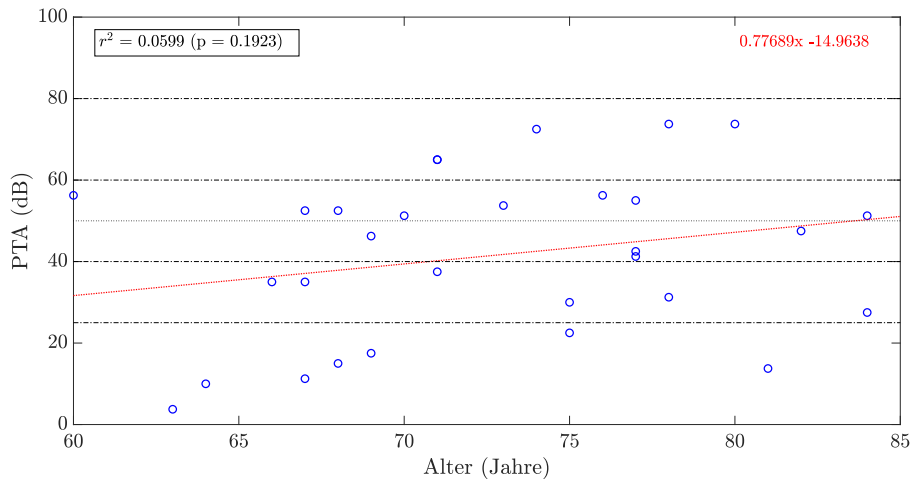


Abb. 3.2: PTA der Probanden in Abhängigkeit vom Alter. Die gestrichelte rote Linie symbolisiert den monotonen Zusammenhang, welcher sich aus der Spearman Korrelation ergibt, jedoch nicht signifikant ist.

Um zusätzlich den Einfluss des Alters auf die zeitlichen Detektionsschwellen zu überprüfen, wurden des Weiteren fünf normalhörende junge Probanden (drei weiblich, zwei männlich) mit einem mittleren Alter von $23,6$ Jahren ± 2 Jahren eingeladen. Die Audiogramme der jungen Probanden sind Anhang

A.2 zu entnehmen. Um vergleichbare Ergebnisse zu den älteren Probanden zu erzielen, wurden neben der Tonaudiometrie auch die Sprachverständlichkeitsmessung im Störgeräusch und der MoCA Test durchgeführt, bevor die Messung zur Schwellenbestimmung stattfand. Die Ergebnisse der zeitlichen Detektionsschwellen konnten anschließend mit den Ergebnissen der normalhörenden (WHO 0) älteren Probanden verglichen werden.

3.3. Pegelbezug und Frequenzgang des Kopfhörers

Um einen absoluten Pegelbezug zum Kopfhörer herzustellen, erfolgte eine Ausmessung des Dynamikbereichs mit Hilfe eines Kunstkopfes und kalibrierten Ohrsimulatoren. Hierfür wurde der „KEMAR G.R.A.S 45BC“ mit den Ohrsimulatoren des Typs RA0045 und den Ohren KB0065 und KB0066 verwendet. Die Ansteuerung des Kopfhörers sollte über eine „Fireface 802“ Soundkarte von RME so erfolgen, dass alle Darbietungen ohne Verzerrungen wiedergegeben werden konnten. Über Matlab R2018b wurde hierfür ein 1000 Hz Sinuston mit einem Pegel von 80 dB SPL über den Kopfhörer an den KEMAR abgegeben. Die Lautstärke-Steller der Soundkarte waren hierbei auf null eingestellt. Da der Gehörgang des KEMAR eine zusätzliche Verstärkung von 3,5 dB bei 1000 Hz liefert, wurde der maximale Ausgangspegel des Kopfhörers in Matlab so lange angeglichen, bis ein Pegel von 83,5 dB SPL am Ohrsimulator anlag.

Abschließend wurde der Frequenzgang des Kopfhörers ausgemessen und ein Freifeldbezug hergestellt. Dieser ist in Abbildung 3.3 dargestellt.

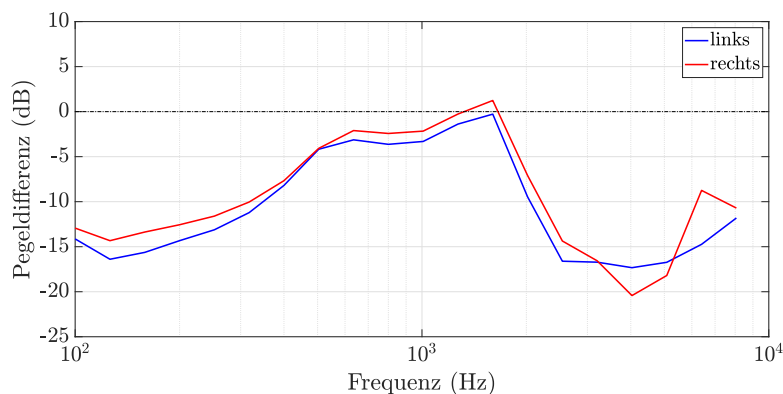


Abb. 3.3: Freifeldbezogener Frequenzgang des HDA 280 Kopfhörers.

Über den Kopfhörer wurde hierfür ein Rosa Rauschen, bandbegrenzt zwischen 50 und 16.000 Hz, mit einem Summenpegel von 70 dB SPL dargeboten. Nachfolgend wurden die Open Ear Gain (OEG) des KEMAR, entnommen aus der DIN IEC TS 60318-7 Tabelle 3, S.29, sowie das ursprüngliche Eingangssignal von der Aufnahme der Ohrsimulatoren abgezogen. Es zeigt sich, dass für tiefe und hohe Frequenzen eine Dämpfung durch den Kopfhörer erfolgt, während das Übertragungsverhalten im mittleren Frequenzbereich nahezu linear ist. Auf eine Frequenzglättung wurde verzichtet.

3.4. Testmaterial

3.4.1. Sprachsignale

Als Sprachmaterial wurden institutsintern zehn verschiedene Worte, die Tabelle II entnommen werden können, aufgenommen. Um zu untersuchen, welchen Einfluss unterschiedliche Initialphoneme auf die Detektionsschwellen haben, wurden der Vokal /a/ und fünf verschiedene Konsonanten als Initialphoneme der Worte ausgewählt. Für die Konsonanten war es wichtig, verschiedene Kategorien wie z.B. Nasale oder Frikative zu vergleichen. Die Auswahl fiel hierfür auf /b/, /k/, /l/, /n/ und /r/. Das zweite Phonem dieser fünf Worte war stets das /a/, um eine bessere Vergleichbarkeit zwischen den Signalen zu haben. Die anderen fünf Worte begannen jeweils mit dem Vokal /a/, gefolgt von den zuvor genannten Konsonanten als zweites Phonem.

Die Auswahl passender Worte erfolgte auf Grundlage des Wörterbuches „A Frequency Dictionary of German“ welches die 4037 häufigsten deutschen Worte nach ihrem Häufigkeitsrang auflistet. Die Einträge basieren auf dem Leipzig/BYU Corpus of Contemporary German, welcher 4,2 Millionen gesprochene sowie geschriebene Worte beinhaltet [50]. Pro Phonem-Kombination wurden alle Worte aufgelistet, die mindestens dreisilbig waren, um Aufnahmen mit einer Wortlänge von mindestens einer Sekunde zu ermöglichen. Im Anschluss wurden die drei häufigsten Substantive, Verben und sonstigen Worte pro Phonem-Kombination selektiert (siehe Anhang A.1) und letztlich jeweils eines der drei häufigsten Substantive als Sprachsignal ausgewählt.

Tabelle II: Ausgewählte Worte für das Sprachmaterial

Wortinitial	Vokal	Wortinitial	Konsonant
Ab	Abteilungen	Ba	Bakterien
Ak	Aktivitäten	Ka	Kandidaten
Al	Alternativen	La	Landwirtschaft
An	Angebote	Na	Nachmittage
Ar	Arbeitsplätze	Ra	Rahmenbedingungen

Die Aufnahme der Sprachsignale erfolgte mit einem einzelnen männlichen Sprecher im schallarmen Raum des Deutschen Hörgeräte Instituts. Es wurde nur ein einziger Sprecher gewählt, da die Variabilität des Sprechers nicht in die Detektionsschwellen einfließen sollte. Der Sprecher achtete auf eine möglichst gleichbleibende Sprechweise ohne besondere Betonungen. Um den Klang der Aufnahmen dennoch so natürlich wie möglich zu halten, wurden die Worte jeweils in einen Satz eingebettet: „Ich spreche jetzt über...“ oder „Ich mag ...“. Für die Aufnahmen wurde ein „Type 4190“ Freifeldmikrofon mit Windschutz von Bruel & Kjaer, sowie eine „Fireface 802“ Soundkarte von RME verwendet. Die Ansteuerung erfolgte über Matlab R2018b.

Die Vorverarbeitung der Signale erfolgte ebenfalls in Matlab R2018b. Zunächst wurden die Worte grob aus den Aufnahmen ausgeschnitten und im Pegel angeglichen. Hierbei wurde darauf geachtet, dass mindestens drei Millisekunden Ruhe vor und nach dem Wort vorhanden waren. Um einen ge-

meinsamen Onset für alle Worte zu ermöglichen, wurden die Signale mit gleitendem quadratischem Mittelwert (Moving RMS) mit einer Fensterbreite von 10 ms betrachtet. In Abbildung 3.4 ist das Verfahren der Schwellwertbetrachtung beispielhaft für die Signale mit Vokal-Initialphonem dargestellt.

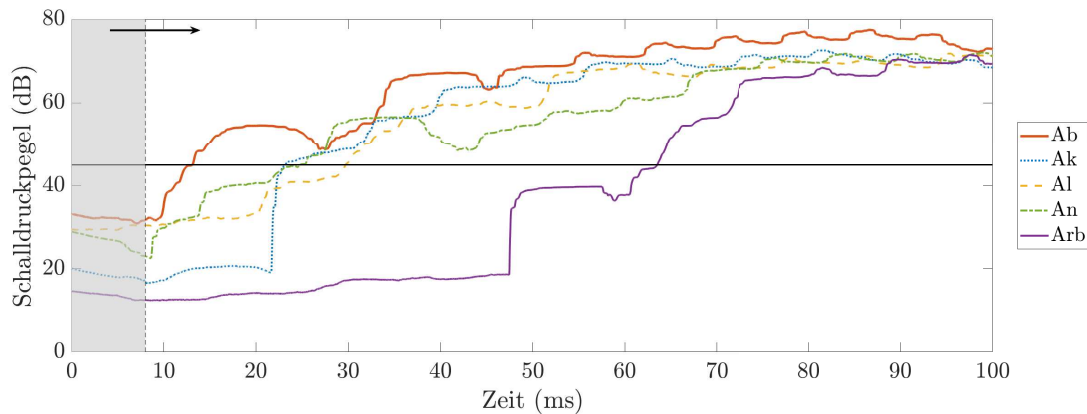


Abb. 3.4: Bestimmung des Onsets der Sprachsignale mit Vokal-Initialphonem. Der Schwellwert wurde auf 45 dB SPL festgelegt.

Für den Onset der Signale wurde jeweils diejenige Position ausgewählt, an welcher der Moving RMS einen bestimmten Schwellwert erreichte. Dieser Schwellwert wurde einheitlich auf 45 dB SPL gesetzt und liegt deutlich über dem Eigenrauschen des Mikrofons und den Hintergrundgeräuschen des Raumes. Im Anschluss wurde das Ende der Signale, unabhängig von der Verteilung der Silben, so gekürzt, dass alle Signale eine einheitliche Länge von einer Sekunde aufwiesen. Abschließend erfolgte eine Bandbegrenzung der Signale auf einen Frequenzbereich von 100 bis 10.000 Hz. Um den Wertebereich der Signallänge von einer Sekunde bis zu wenigen Millisekunden adäquat einteilen zu können, wurde eine logarithmische Zeiteinteilung gewählt. Hierfür wurde die Signallänge in Dezibel relativ zu 1 ms angegeben. Die Länge der Signale wurde von 60 dB re 1 ms in Schritten von 1 dB bis auf 15 dB re 1 ms, was einer minimalen Länge von ca. 5,6 ms entspricht, reduziert. Somit ergaben sich für jedes der zehn Signale 46 Signalstücke.

3.4.2. Generierung des Störgeräusches

Als Störgeräusch sollte ein Rauschen dienen, welches das gleiche Spektrum aufweist wie das zugehörige Sprachsignal. Da sich das Frequenzspektrum von Sprachsignalen verändert, wenn deren Länge variiert wird, wurde das Störgeräusch für jedes Signalstück einzeln erzeugt. Hierfür wurde die Fast Fourier Transformation (FFT) jedes einzelnen Signalstücks berechnet. Im Anschluss erfolgten eine Randomisierung der Phase, entsprechend einer Gleichverteilung, im Frequenzbereich und eine abschließende inverse FFT zurück in den Zeitbereich. Eine beispielhafte Darstellung von jeweils zwei Sprach- und Störsignalen ist in Abbildung 3.5 abgebildet.

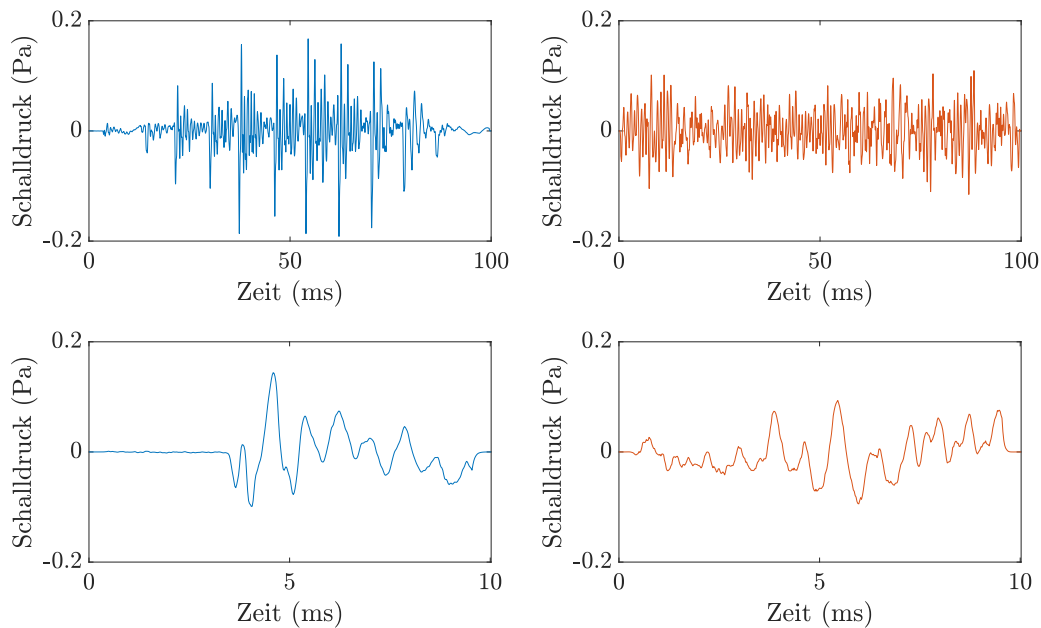


Abb. 3.5: Exemplarische Zeitsignale von Sprache (links) und Rauschen (rechts) für einen Pegel von 65 dB SPL. Oben: Signale der Länge 100 ms. Unten: Signale der Länge 10 ms.

Es ergab sich für jedes Signalstück ein Störgeräusch, welches ein identisches Frequenzspektrum aufweist und lediglich in der zeitlichen Feinstruktur und in der Einhüllenden modifiziert ist. Dies ist an Abbildung 3.5 deutlich zu erkennen. Während der Onset der Sprachsignale sowie deren periodischer Verlauf deutlich zu erkennen sind, sind die daraus resultierende Einhüllende und die zeitliche Feinstruktur in der Darstellung der Störsignale stark modifiziert. Abschließend wurden die Signale in einem Zeitbereich von 0,5 ms ein- und ausgeblendet, um ein Knacken des Kopfhörers zu verhindern, welches durch den plötzlichen Sprung der am Kopfhörer anliegenden Spannung entstehen könnte.

3.4.3. Lautstärke der Testsignale

Für die Messung der zeitlichen Detektionsschwellen sollte gewährleistet sein, dass allen Probanden, unabhängig vom Hörverlust, eine überschwellige Darbietung der Testsignale möglich ist. Als Sicherheitsfaktor wurde der Ausgangspegel auf eine maximale mittlere Leistung von 100 dB SPL begrenzt. Die Darbietung der Signale erfolgte über den Kopfhörer „HDA 280“ der Firma Sennheiser. Vor der eigentlichen Schwellenbestimmung stellten die Probanden die Lautstärke für die Testsignale selbst so ein, dass sie als angenehm laut empfunden wurden. Hierfür konnten jeweils ein langes und ein kurzes Sprach-, sowie Rauschsignal als Beispiele angehört und in der Lautstärke angepasst werden. Abbildung 3.6 stellt die zugehörige graphische Benutzeroberfläche dar. Der eingestellte Schalldruckpegel wurde anschließend auf alle Sprach- und Rauschsignale angewandt.

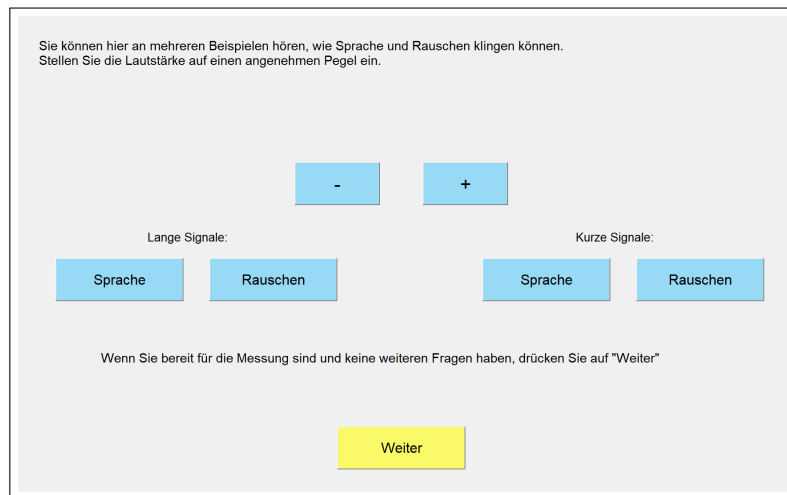


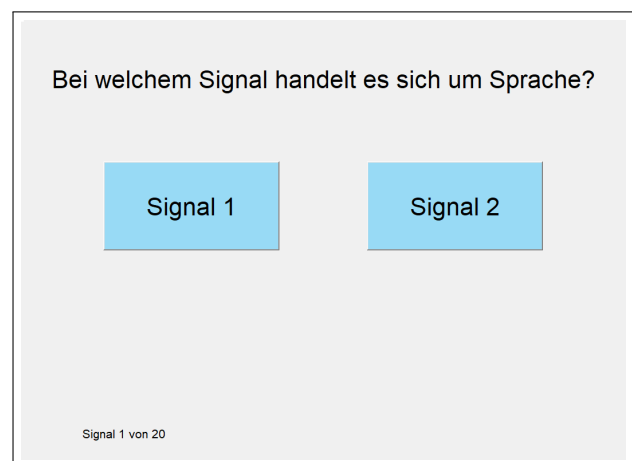
Abb. 3.6: Einstellung der Lautstärke für die Schwellenbestimmung.

3.5. Zeitliche Schwellenbestimmung

Die Aufgabe für die Probanden war ein 2-AFC Verfahren, d.h. sie hörten hintereinander zwei Signale, von denen eines ein Sprachsignal und das andere das zugehörige Rauschen war. Die Darbietung der Signale erfolgte hierbei doppelblind in randomisierter Reihenfolge. Nach der Präsentation mussten die Probanden entscheiden, ob es sich bei dem ersten oder bei dem zweiten Signal um das Sprachsignal handelte. Je nach Antwort der Probanden wurde die Länge für die nachfolgend dargebotenen Signale entweder gekürzt oder verlängert. Der Versuchsaufbau sowie die Benutzeroberfläche für die Probanden sind in Abbildung 3.7 dargestellt. Die Schwellenbestimmung erfolgte für alle zehn Signale in randomisierter Reihenfolge.



(a) Versuchsaufbau



(b) Graphische Benutzeroberfläche

Abb. 3.7: Versuchsaufbau und Ablauf der zeitlichen Schwellenbestimmung.

Es wurde ein adaptives Messverfahren mit der gewichteten Staircase-Methode nach Kaernbach (1991) als Entscheidungsregel angewandt. Die initiale Signallänge lag bei 46 dB re 1 ms, was ca. 200 ms entspricht. Da für das 2-AFC Verfahren eine ratekorrigierte Wahrscheinlichkeit von 50 % einer beobachteten Wahrscheinlichkeit von 75 % entspricht, musste nach Gleichung 7 die Schrittweite für falsche Antworten dreimal so groß sein wie für richtige Antworten. Um die Signallänge schnell in den Bereich der gesuchten Schwelle zu bringen, erfolgte zunächst eine Trainingsphase mit einer Schrittweite von +6 dB für falsche Antworten und -2 dB für richtige Antworten. Nach dem vierten Wendepunkt begann anschließend die Testphase. Hierfür wurde die Schrittweite auf +3 dB für falsche Antworten und -1 dB für richtige Antworten reduziert. Die Messung wurde beendet, sobald der achte Wendepunkt innerhalb der Testphase erreicht war. Die zeitliche Schwelle wurde abschließend aus den Signallängen an den letzten vier Wendepunkten gemittelt. Erreichte ein Proband die minimale Signallänge von 15 dB re 1 ms und entschied sich in mehr als 20 aufeinanderfolgenden Darbietungen richtig für das Sprachsignal, wurde die Messung manuell beendet und der Schwellwert für das zugehörige Signal auf 15 dB re 1 ms gesetzt. Die Messung wurde dann für das nächste Signal fortgesetzt. Um die Höranstrengung sowie Ermüdungseffekte möglichst gering zu halten, erfolgte nach jedem zweiten Durchlauf eine Pause von mindestens 60 Sekunden.

4. Ergebnisse und Diskussion

4.1. Test-Retest Reliabilität und Robustheit des Messverfahrens

Die ersten 20 Probanden wurden für zwei Termine mit einem Abstand von einer Woche eingeladen, um die Test-Retest Reliabilität und die Robustheit der Messmethode zu überprüfen. Die Messungen fanden an beiden Terminen im gleichen Umfang und in der gleichen Reihenfolge statt. Es sei an dieser Stelle angemerkt, dass der Untersucher beim zweiten Termin für elf Probanden ein anderer war als beim ersten Termin. Die Einweisung der Probanden in die einzelnen Aufgaben erfolgte jedoch an beiden Terminen gleichermaßen.

Die ermittelte Hörschwelle für die beiden Termine zeigt mit einem Maximum von ± 10 dB die für die Tonaudiometrie übliche Streuung [51]. Der Vergleich des PTA zwischen dem ersten und zweiten Termin ergab hierdurch eine Streuung von $-0,2 \text{ dB} \pm 2,4 \text{ dB}$. Die zugehörigen Abbildungen sind Anhang A.2 zu entnehmen.

Der MoCA Test, welcher als erste Aufgabe beim jeweiligen Termin durchgeführt wurde, war zu Beginn für alle Probanden unbekannt und für eine audiologische Studie eher unüblich. Hierdurch ergab sich beim ersten Termin eine gewisse Unsicherheit der Probanden. Da der Test an beiden Terminen die gleichen Aufgaben beinhaltete, wurde vermutet, dass aufgrund der Wiedererkennung bessere Ergebnisse beim zweiten Termin erzielt werden als beim ersten Termin. Zudem erzählten einige Probanden, dass sie bestimmte Aufgaben zu Hause geübt haben, um besser vorbereitet zu sein. Trotz der erwarteten Verbesserungen wurde der MoCA auch beim zweiten Termin durchgeführt, um die Messbedingungen und den Ablauf für beide Termine gleich zu halten. Wie Anhang A.2 zu entnehmen ist, bestätigte sich die vermutete Verbesserung des MoCA Testergebnisses beim zweiten Termin mit einer Erhöhung der Punktzahl von $1,5 \pm 2,1$ Punkten.

Zur Messung der SRTs wurde an beiden Terminen der GÖSA mit je einer Testliste von 30 Wörtern durchgeführt. Die Testliste war beim zweiten Termin eine andere als beim ersten, damit die Probanden nicht durch das Wiedererkennen bekannter Sätze ein besseres Ergebnis erzielen konnten. Für zwei Probanden aus der WHO 3 Klasse konnte an beiden Terminen kein SRT bestimmt werden, da der Pegel, welcher für ein Sprachverstehen benötigt wurde, über die Leistungsgrenze des Lautsprechers hinaus ging. Somit sind in Abbildung 4.1 die Ergebnisse der Sprachverständlichkeitsmessung von Test und Retest für 18 Probanden dargestellt.

Um zu untersuchen, ob es einen Unterschied zwischen den SRTs des ersten und zweiten Termins gibt, wurde der t-Test für verbundene Stichproben angewandt. Das Signifikanzniveau wurde, wie auch für alle nachfolgenden Signifikanztests, auf $\alpha = 0,05$ festgelegt. Der t-Test liefert mit einem p-Wert von 0,0065 einen signifikanten Unterschied zwischen den SRTs für Test und Retest. Der SRT verbesserte sich beim zweiten Termin im Mittel um $0,9 \text{ dB} \pm 1,2 \text{ dB}$, was sich jedoch nicht mit den Erwartungen deckt, da der GÖSA „über keinen Lerneffekt verfügt, solange keine Testlisten wiederholt werden“ [52], und somit keine signifikanten Unterschiede zwischen den beiden Terminen erwartet wurden.

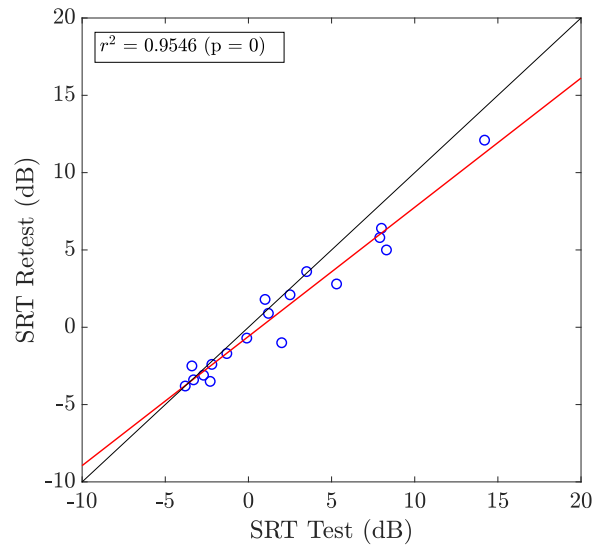


Abb. 4.1: Test und Retest des SRT pro Proband.

Da die Messung des SRT an beiden Terminen nach dem MoCA Test stattfand, könnte sich die Unsicherheit, welche sich beim ersten Termin durch den MoCA ergab, auf die Konzentrationsfähigkeit der Probanden ausgewirkt haben. Während sie sich beim ersten Termin zunächst mit den ihnen gestellten Aufgaben vertraut machen mussten, wussten sie beim zweiten Termin bereits vorab, was sie erwartet. Zudem könnte eine Interrater-Reliabilität eingeflossen sein, welche zu einem untersucherabhängigen Testergebnis führt [48].

Die Verteilung der zeitlichen Detektionsschwellen wurde für Test und Retest über die mittlere Schwelle pro Proband betrachtet. Die Ergebnisse sind in Abbildung 4.2 dargestellt.

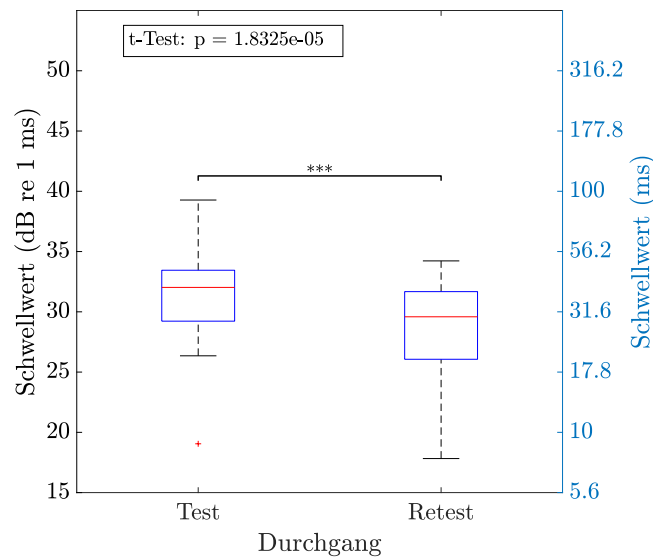


Abb. 4.2: Test und Retest der zeitlichen Detektionsschwellen pro Proband.

Für die Verteilung der mittleren Schwelle pro Proband zeigen sich für den zweiten Termin signifikant niedrigere Schwellwerte (p-Wert des t-Test: $1,8e-05$). Der Median verringerte sich von 32 dB re 1 ms auf 29,6 dB re 1 ms. Die Streuung ist für beide Durchgänge ähnlich groß, verschiebt sich für den Retest jedoch zu niedrigeren Werten. Die höheren Schwellen beim ersten Termin könnten auch für diese Messung durch die Unsicherheit der Probanden zustande gekommen sein, da auch die Schwellenbestimmung eine bisher unbekannte Aufgabe war. Beim Retest waren die Probanden hingegen mit der Aufgabe vertraut und konnten diese mit mehr Sicherheit und mehr Konzentration durchführen als beim ersten Termin.

Um zu untersuchen, ob innerhalb eines Termins ein Lerneffekt bei der Schwellenbestimmung stattfand, wurde die Verteilung der Schwellwerte in der Reihenfolge der dargebotenen Signale ermittelt (siehe Abbildung 4.3). Da die Schwellenmessung in beiden Terminen als letzte Messung stattfand, sollte zusätzlich ein möglicher Ermüdungseffekte aufgedeckt werden.

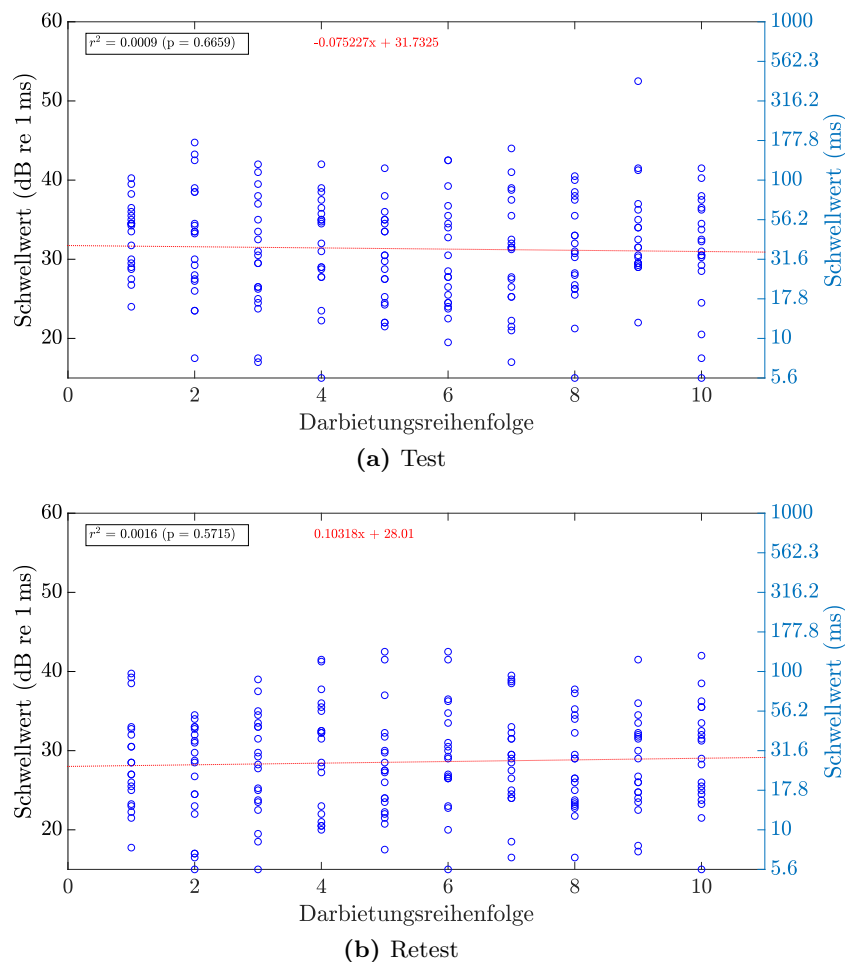


Abb. 4.3: Schwellwerte in Abhängigkeit von der Darbietungsreihenfolge.

Abbildung 4.3a zeigt die Verteilung der Schwellwerte in Abhängigkeit von der Darbietungsreihenfolge für den ersten Termin. Es ist kein konstanter Abfall in den zeitlichen Detektionsschwellen zu sehen, was dafürspricht, dass kein Lerneffekt vorliegt. Ebenso ist kein konstanter Anstieg zu entdecken, so dass ein Ermüdungseffekt über den gesamten Verlauf der Messung zur Schwellenbestimmung ebenfalls unwahrscheinlich ist. Eine Berechnung des Spearman-Korrelationskoeffizienten unterstützt diese Vermutung mit einem r^2 von 0,0009 (p-Wert $\approx 0,67$) und die lineare Regression liefert nur eine Steigung von -0,075.

Für den Retest ergaben sich sehr ähnliche Ergebnisse. Die Spearman Korrelation liefert mit einem p-Wert von 0,57 ein r^2 von 0,0016 und die Steigung der linearen Regression liegt bei 0,103. Somit scheint die Schwellenbestimmung auch beim zweiten Termin weder durch einen Lerneffekt, noch durch einen Ermüdungseffekt beeinflusst worden zu sein.

Für die nachfolgenden Abschnitte wurden die Ergebnisse von allen 30 älteren Probanden ausgewertet. Von den ersten 20 Probanden wurden hierfür nur die Ergebnisse vom ersten Termin herangezogen.

4.2. Einfluss des Initialphonems auf die Detektionsschwellen

Um einen ersten Vergleich zwischen den Initialphonemen machen zu können, wurde die Verteilung der Detektionsschwellen pro Signal ermittelt. Die Ergebnisse sind als Boxplots in Abbildung 4.4 dargestellt. Um Signifikanztests bezüglich der Unterschiede zwischen den Detektionsschwellen der einzelnen Signale durchführen zu können, wurde vorab überprüft, ob die ermittelten Detektionsschwellen normalverteilt sind. Der Shapiro-Wilk-Test ergab, dass bei der Hälfte der Signale keine Normalverteilung der Daten vorliegt. Zur Überprüfung auf signifikante Unterschiede wurde daher der Friedman Test mit anschließender Tukey Kramer Methode durchgeführt. Hierdurch entstanden insgesamt 45 Paarvergleiche. Aus Tabelle III können die Wahrscheinlichkeitswerte der Paarvergleiche entnommen werden.

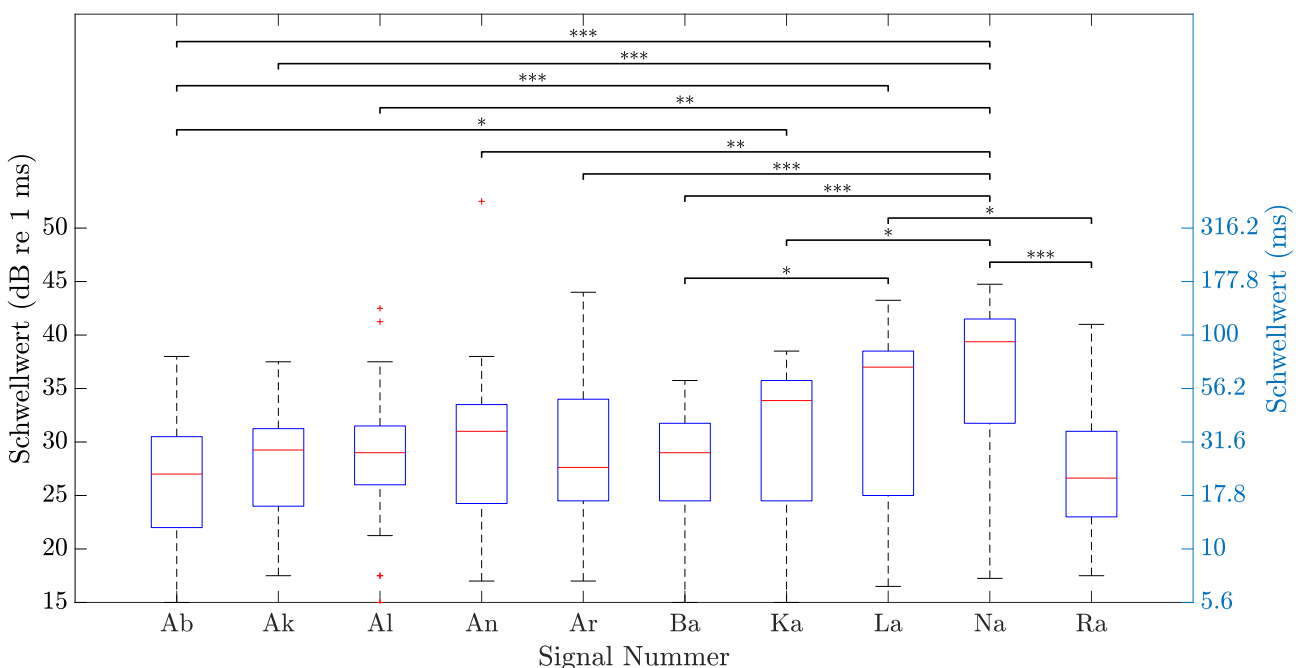


Abb. 4.4: Verteilung der Detektionsschwelle pro Signal. Für Signifikanz wurde der Friedman-Test verwendet. Die Sternchen kennzeichnen die Signifikanz mit * für $p < 0.05$, ** für $p < 0.01$ und *** für $p < 0.001$

Die Detektionsschwellen für jedes einzelne Signal weisen hohe interindividuelle Unterschiede auf, da die Streuung, wie in Abbildung 4.4 dargestellt, für alle Verteilungen sehr groß ist.

Zudem fällt auf, dass die Verteilungen für /Ab/, /Al/, /Ba/ und /Ka/ die untere Grenze von 15 dB re 1 ms erreichen. Hierbei handelte es sich um die Ergebnisse zweier Probanden der WHO Klasse 0, sowie einer Probandin der WHO Klasse 2 (breitbandiger Hörverlust), welche für diese Signale bis zum kürzesten Signalstück einen spektralen Unterschied zwischen Sprach- und Rauschsignal wahrgenommen haben. Dies wird in Kapitel 4.4 näher untersucht.

Die Mediane der Schwellwerte aller fünf Signale welche mit /a/ beginnen, sowie der Signale mit den Initialen /Ba/, /Ka/ und /Ra/ sind signifikant niedriger als der mediane Schwellwert des Signals mit dem Initial /Na/. Wie Tabelle III entnommen werden kann, ist dieser Unterschied für /Ab/, /Ak/, /Ar/, /Ba/ und /Ka/ sogar höchst signifikant. Bei dem /n/ handelt es sich um ein Nasal, bei welchem

Tabelle III: Signifikanz der Unterschiede zwischen den Schwellwerten in Abhängigkeit vom Initialphonem nach dem Friedman Test mit Tukey Kramer Korrektur.

	Ab	Ak	Al	An	Ar	Ba	Ka	La	Na	Ra	p < 0.05
Ab	1,0000	0,5613	0,2558	0,1092	0,4255	0,8251	0,0313	0,0001	0,0000	0,8137	p < 0.01
Ak		1,0000	1,0000	0,9981	1,0000	1,0000	0,9619	0,1432	0,0002	1,0000	p < 0.001
Al			1,0000	1,0000	1,0000	0,9971	0,9987	0,3827	0,0017	0,9976	
An				1,0000	0,9998	0,9659	1,0000	0,6373	0,0070	0,9696	
Ar					1,0000	0,9999	0,9876	0,2234	0,0005	0,9999	
Ba						1,0000	0,8137	0,0467	0,0000	1,0000	
Ka							1,0000	0,8853	0,0313	0,8251	
La								1,0000	0,7379	0,0499	
Na									1,0000	0,0000	
Ra										1,0000	

die Luft bei der Aussprache, wie in Kapitel 2.2.3 beschrieben, über die Nase entweichen muss. Dies lässt vermuten, dass der Informationsgehalt, welcher für die Sprachdetektion benötigt wird, erst nach einem längeren zeitlichen Verlauf zur Verfügung steht als für die anderen Phoneme.

Die Schwellwerte für das Initial /La/ waren ebenfalls signifikant höher als die Schwellwerte für die Initiale /Ab/, /Ba/ und /Ra/. Bei der Aussprache des lateralen Approximanten /l/ muss aufgrund des zentralen Verschlusses der Zunge ein Teil der Luft über die Seiten entweichen. Ähnlich zum /n/ ist der Informationsgehalt vermutlich auch beim /l/ zunächst zu gering, um die zeitlichen Merkmale für die Sprachdetektion zu nutzen.

Auch zwischen /Ab/ und /Ka/ zeigt sich ein signifikanter Unterschied. Das /k/ ist der einzige ausgewählte Konsonant, welcher stimmlos ausgesprochen wird. Womöglich ist durch die Stimmlosigkeit ohne den Einsatz der Stimmbänder weniger Information in der zeitlichen Feinstruktur vorhanden, die für die Sprachdetektion genutzt werden kann.

Der Sprecher hat während der Aufnahmen, wie in Kapitel 3.4.1 beschrieben, auf eine bewusst monotone Aussprache ohne spezielle Betonungen geachtet. Unter der Annahme, dass für die Sprachdetektion nur das Initialphonem benötigt wird, wurde daher vermutet, dass es keinen Unterschied zwischen den Schwellwerten der Signale mit Vokal-Initialphonem gibt. Diese Vermutung bestätigt sich, da, wie Abbildung 4.4 und Tabelle III zu entnehmen, keine signifikanten Unterschiede zwischen den Verteilungen bestehen.

Um die Annahme zu überprüfen, ob für jedes Signal das Initialphonem für die Sprachdetektion ausreichend war, wurde untersucht, in welchem Bereich des jeweiligen Zeitsignals sich die zugehörigen Detektionsschwellen befanden. In den Abbildungen 4.5 bis 4.9 sind daher die ersten 200 ms von fünf exemplarischen Zeitsignalen für einen Darbietungspegel von 65 dB SPL dargestellt. Die übrigen Zeitsignale sind Anhang A.2 zu entnehmen. Der Übergang vom initialen zum zweiten Phonem lässt sich nur schwer bestimmen, da sich die Artikulationsorgane beim Sprechen fließend von der Position für das erste Phonem zur Position für das zweite Phonem bewegen [24]. Es wurde daher eine subjektive

Einteilung erstellt. Der rote Bereich definiert hierbei das Vokal-Phonem und der gelbe Bereich das Konsonant-Phonem, unabhängig davon, ob das jeweilige Phonem initial oder an zweiter Position liegt. Der blaue Bereich stellt das 1. bis 3. Quartil der Schwellwertverteilung dar und die vertikale blaue Linie definiert den Median.

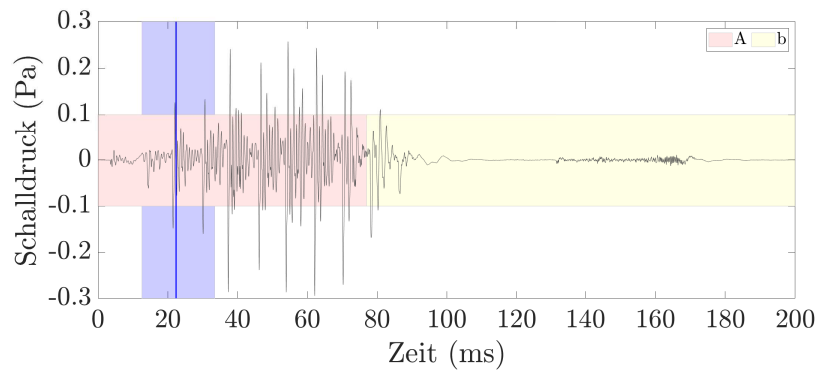


Abb. 4.5: Zeitsignal /Ab/ und zugehörige Detektionsschwellen

In Abbildung 4.5 ist zu sehen, dass der initiale Vokal /a/ bis ca. 80 ms andauert und anschließend der Konsonant gesprochen wird. Der typisch periodische Verlauf des Vokals ist deutlich zu erkennen. Die zeitlichen Detektionsschwellen liegen im 1. bis 3. Quartil zwischen 12,6 und 33,5 ms, mit einem Median von 22,4 ms. Es reicht somit das Initialphonem für die Sprachdetektion aus. Die Ergebnisse für die vier weiteren Signale, welche ebenfalls mit dem Vokal beginnen, weisen hohe Ähnlichkeit auf. Der Median der Detektionsschwellen liegt für alle Vokal-Initialphoneme zwischen 20 und 40 ms.

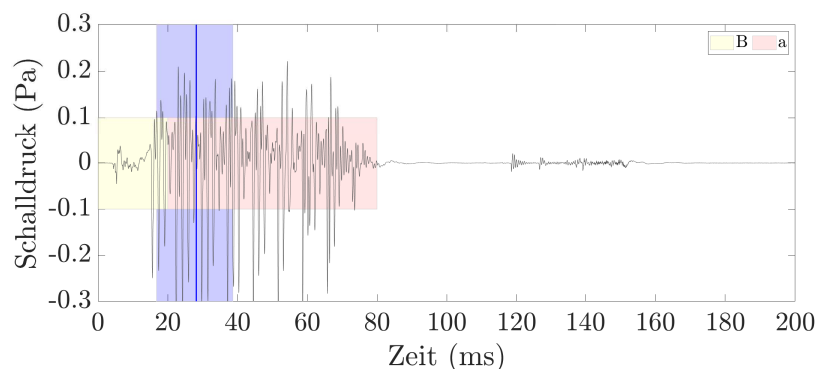


Abb. 4.6: Zeitsignal /Ba/ und zugehörige Detektionsschwellen

Die Verteilung der Schwellwerte für /Ba/ ist in Abbildung 4.6 dargestellt. Zunächst fällt auf, dass das Initialphonem /b/ mit ca. 20 ms deutlich kürzer ist als der nachfolgende Vokal, da das /b/ ein bilabialer stimmhafter Plosivlaut ist, für dessen Artikulation die Verengung im Vokaltrakt erst direkt an den Lippen gebildet wird. Die Schwellwerte liegen im Vergleich zu Abbildung 4.5 mit einem Median von 28,2 ms zwar nur um wenige Millisekunden zu höheren Werten verschoben, jedoch befindet sich ein Großteil der Detektionsschwellen in dem zeitlichen Bereich, in welchem bereits das zweite Phonem

anteilig dargeboten wird. Dies lässt vermuten, dass der Plosivlaut /b/ aufgrund seiner Kürze keinen ausreichenden Informationsgehalt aufweist, um eine Sprachdetektion herbeizuführen.

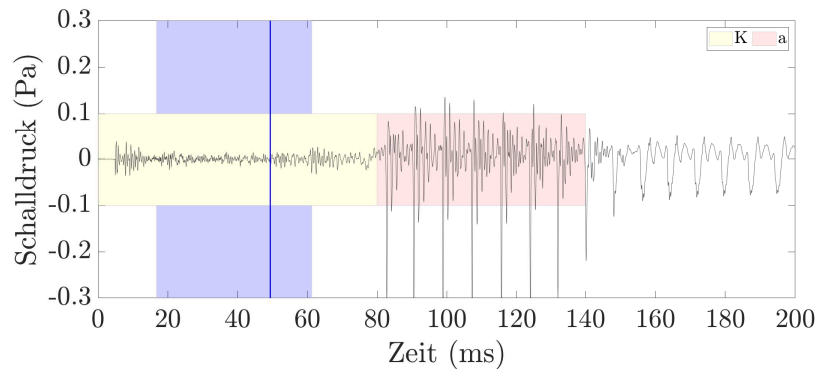


Abb. 4.7: Zeitsignal /Ka/ und zugehörige Detektionsschwellen

Für das Initial /Ka/ liegen die zeitlichen Detektionsschwellen deutlich innerhalb des ersten Phonems, was in Abbildung 4.7 zu sehen ist. Die Schwellen für /Ka/ sind jedoch mit einem Median von 49,4 ms deutlich höher als für /Ab/ oder /Ba/. Zudem fällt der rauschähnliche Verlauf des /k/ auf, welcher sich aus der stimmlosen Aussprache ergibt. In diesem rauschähnlichen zeitlichen Verlauf scheint der für die Sprachdetektion benötigte Informationsgehalt, im Vergleich zum /a/, erst nach einer längeren Signaldauer zur Verfügung zu stehen.

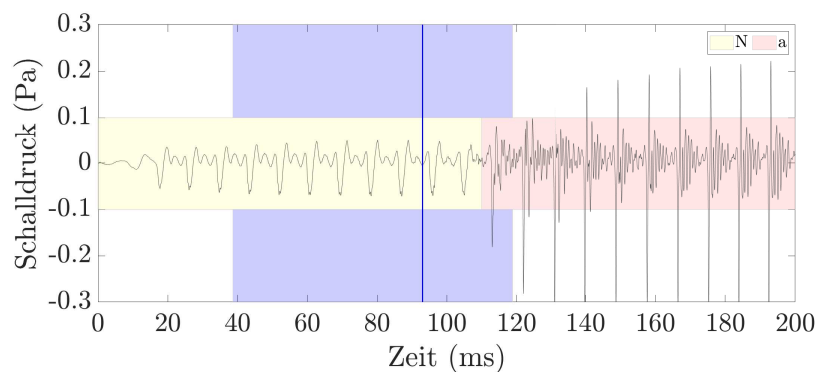


Abb. 4.8: Zeitsignal /Na/ und zugehörige Detektionsschwellen

Die Schwellwerte in Abhängigkeit vom Zeitsignal für /Na/ sind in Abbildung 4.8 dargestellt und sind ebenfalls repräsentativ für die Ergebnisse des Initials /La/. Die Aussprache des /n/ ist mit ca. 110 ms sehr lang und es ist deutlich der periodische Verlauf durch die stimmhafte Aussprache zu sehen. Die Sprachdetektion erfolgt bereits anhand des Initialphonems, jedoch sind die zeitlichen Detektionsschwellen mit einem Median von 93,1 ms am höchsten von allen Signalen. Die hohen Detektionsschwellen könnten erneut anhand des Informationsgehalts erklärt werden, welcher sich für nasale Sprachlaute aufgrund der langsamen Schwingungen erst im Verlauf der Aussprache aufbaut, während er für Vokale sehr schnell ansteigt.

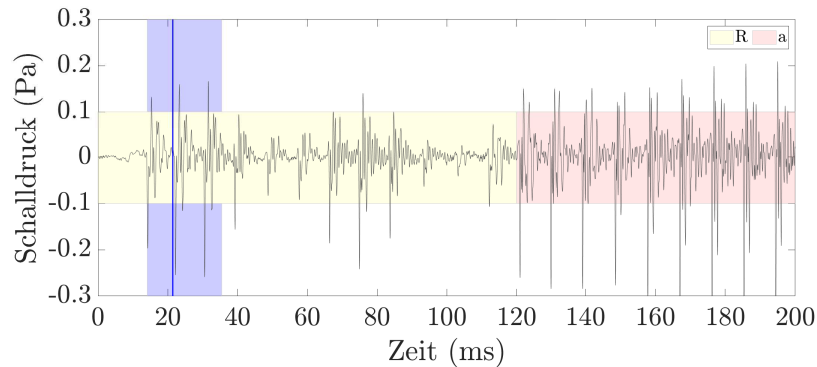


Abb. 4.9: Zeitsignal /Ra/ und zugehörige Detektionsschwellen

In Abbildung 4.9 ist abschließend die Verteilung der Schwellwerte für /Ra/ zu sehen. Mit einem Median von 21,4ms ergaben sich für /Ra/ die kürzesten Detektionsschwellen. Bei der Aussprache des stimmhaften uvularen Frikativs entstehen Luftwirbel aufgrund des aufgebauten Drucks am Artikulationsort. Es lässt sich demnach vermuten, dass hierdurch bereits nach einer kurzen Signaldauer genügend Informationen aus den zeitlichen Merkmalen für die Sprachdetektion zur Verfügung stehen.

Die Ergebnisse aus den Abbildungen 4.4 bis 4.9 legen dar, dass die Eigenschaften der Initialphoneme einen bedeutenden Einfluss auf die zeitlichen Detektionsschwellen für Sprache aufweisen. Mit einem Median von 21,4ms waren die Schwellwerte für /Ra/ die niedrigsten, während die Schwellen für /Na/ mit einem Median von 93,1 ms am längsten waren. Es lässt sich also vermuten, dass die Art der Artikulation, welche für die einzelnen Phoneme unterschiedlich stattfindet, einen starken Einfluss auf den für die Sprachdetektion benötigten Informationsgehalt hat.

Des Weiteren lag der Median bei neun von zehn Signalen im Bereich des ersten Phonems. Das Initialphonem scheint also genügend zeitliche Informationen für die Sprachdetektion zu liefern. Eine Ausnahme macht hierbei das Signal mit den Initialen /Ba/, bei welchem ein Teil des zweiten Phonems zur Sprachdetektion benötigt wurde. Eine Erklärung hierfür könnte die vergleichsweise kurze zeitliche Länge des Plosivlautes /b/ und dessen zeitlicher Verlauf sein, welche anscheinend nicht ausreicht, um genügend zeitliche Merkmale für die Sprachdetektion zur Verfügung zu stellen.

Ob neben den unterschiedlichen Initialphonemen auch das Hörvermögen der Probanden einen Einfluss auf die zeitlichen Detektionsschwellen aufweist wurde im nachfolgenden Abschnitt untersucht.

4.3. Einfluss des Hörvermögens auf die Detektionsschwellen

In vorherigen Studien wurde vermutet, dass der Hörverlust einen Einfluss auf die zeitlichen Detektionsschwellen aufweist [4], [5]. Diese Vermutung wurde im Rahmen dieser Arbeit ebenfalls überprüft. Hierfür wurde zum einen die Hörschwelle gemessen, um daraus den PTA zu bilden, und zum anderen das Sprachverstehen im Störgeräusch, bzw. der SRT ermittelt. Zunächst wurde der Zusammenhang zwischen PTA und SRT überprüft. Die Ergebnisse sind in Anhang A.2 dargestellt. Es zeigt sich eine signifikante Korrelation zwischen dem PTA und dem SRT, welche sich mit den Erwartungen deckt, da das Sprachverstehen durch das reduzierte spektrale sowie zeitliche Auflösungsvermögen eingeschränkt ist [53].

Im Anschluss wurde der Zusammenhang vom PTA und vom SRT mit den zeitlichen Detektionsschwellen untersucht (siehe Abbildung 4.10). Hierfür wurden zunächst die Detektionsschwellen pro Proband über alle zehn Signale gemittelt.

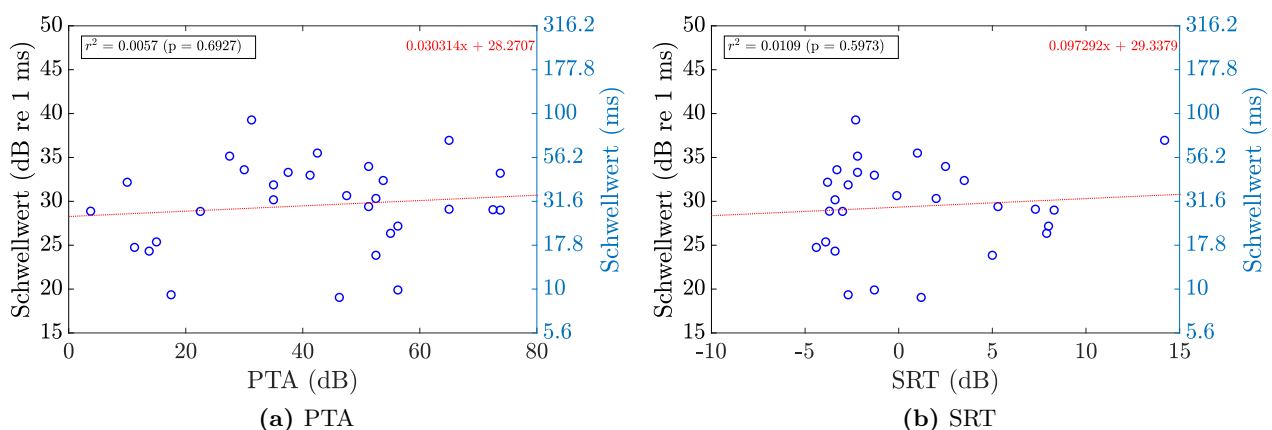


Abb. 4.10: Korrelation des Hörvermögens mit den zeitlichen Detektionsschwellen.

In Abbildung 4.10a ist zu sehen, dass kein Zusammenhang zwischen dem Hörverlust und den zeitlichen Detektionsschwellen vorliegt. Der Spearman-Korrelationskoeffizient lieferte ein r^2 von 0,0057 (mit $p = 0,6927$). Auch zwischen dem SRT und den Detektionsschwellen liegt keine signifikante Korrelation vor. Das Verhältnis ist in Abbildung 4.10b dargestellt und ergibt ein r^2 von 0,0109 (mit $p = 0,5973$). Es scheint somit keinen allgemeinen Zusammenhang zwischen dem Hörvermögen und den zeitlichen Detektionsschwellen zu geben. In Kapitel 4.2 hat sich jedoch gezeigt, dass die Schwellwerte eine Abhängigkeit zum Initialphonem aufweisen. Um zu untersuchen, ob sich das Hörvermögen unterschiedlich auf die Detektionsschwellen der verschiedenen Signale auswirkt, wurde daher weitergehend der Zusammenhang zwischen dem Hörvermögen und den Detektionsschwellen für jedes einzelne der zehn Signale betrachtet. In der nachfolgenden Abbildung 4.11 sind die Korrelationen von den Schwellwerten der zehn verschiedenen Signale in Abhängigkeit vom PTA dargestellt.

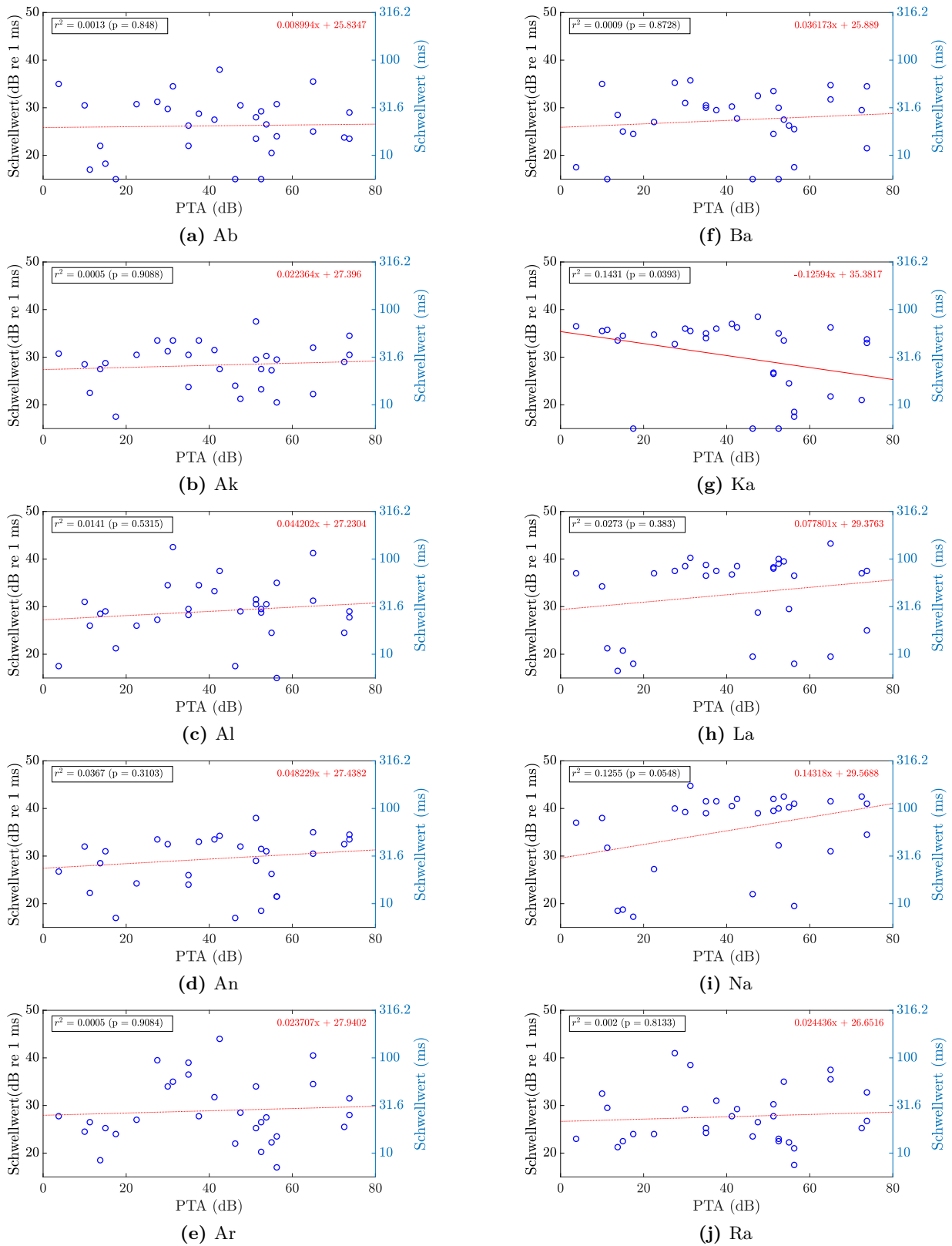


Abb. 4.11: Korrelation von PTA und zeitlichen Detektionsschwellen pro Signal. Ist die rote Regressionsgerade durchgezogen liegt ein signifikanter Zusammenhang vor.

Die Auswertung der in Abbildung 4.11 dargestellten Daten erfolgte zunächst auf der Annahme, dass es sich um einzelne Datensätze handelt. Für jeden Datensatz wurde einzeln die Hypothese, dass es keinen Zusammenhang zwischen dem PTA und den Detektionsschwellen des jeweiligen Signales gibt, überprüft.

Die Abbildungen 4.11a - 4.11e zeigen die Ergebnisse für diejenigen Signale, welche mit dem Vokal beginnen. Die Verteilung der Schwellwerte weist für alle fünf Signale eine hohe Streuung auf, sodass sich kein linearer Zusammenhang erschließen lässt. Die Spearman-Korrelation liefert Korrelationskoeffizienten nahe Null mit p-Werten zwischen 0,31 und 0,91 und deutet somit ebenfalls auf keine signifikanten Korrelationen hin.

Ähnlich sieht die Verteilung der Schwellwerte für die Signale der Initiale /Ba/ und /Ra/ aus (Abbildung 4.11f, 4.11j). Auch hier ergeben sich hohe Streuungen und Spearman-Korrelationskoeffizienten nahe Null.

Für /La/ und /Na/, zu sehen in Abbildung 4.11h und 4.11i, zeigt sich die Tendenz eines linearen Zusammenhangs. Der Spearman-Korrelationskoeffizient liefert mit einem r^2 von 0,027 ($p = 0,28$) für /La/ und $r^2 \approx 0,126$ ($p = 0,055$) für /Na/ zwar für beide Signale keinen signifikanten Zusammenhang, jedoch lässt sich anhand der Steigung der linearen Regressionsgerade von 0,078 für /La/ und 0,14 für /Na/ eine Tendenz erkennen, dass sich für größer werdende PTAs höhere Detektionsschwellen ergeben. Diese Tendenz ist für /Na/ stärker ausgeprägt als für /La/.

Für das Initialphonem /Ka/ (Abbildung 4.11g) zeigt sich ein signifikanter Zusammenhang, dass Probanden mit höherem PTA das Sprachsignal auch bei kürzeren Signallängen richtig als Sprache zuordnen konnten als Probanden mit niedrigerem PTA. Zudem ist zu sehen, dass die Streuung für größere PTAs genauso hoch ist wie bei anderen Phonemen. Bei niedrigen PTAs (WHO Klasse 0) ist die Streuung jedoch sehr gering. Eine mögliche Erklärung für den unerwarteten Zusammenhang könnte sein, dass das /k/ ein hochfrequenter Konsonant ist und möglicherweise ein Teil der für Sprache markanten Merkmale für die nahezu normalhörenden Probanden nicht hörbar war, wodurch die Detektion erschwert war. Denn auch die Probanden der WHO 0 Klasse wiesen einen leichten Hörverlust in den hohen Frequenzen (ab 2 kHz) auf, stellten jedoch die Lautstärke der Testsignale deutlich niedriger ein als die schwerhörigeren Probanden.

Betrachtet man die Ergebnisse aus Abbildung 4.11 unter der Annahme, dass es sich um einen gemeinsamen Datensatz handelt, für den zehn unabhängige Hypothesen (kein Zusammenhang zwischen PTA und Detektionsschwellen für Signal x) geprüft werden, führt die notwendige Bonferroni-Korrektur dazu, dass es für keines der zehn Signale einen signifikanten Zusammenhang zwischen PTA und Detektionsschwellen gibt.

Es sei an dieser Stelle angemerkt, dass, wie in Kapitel 3.2 beschrieben, die Einteilung der Hörverluste anhand des PTA in die einzelnen WHO Klassen erfolgte. Es wurde jedoch nicht näher betrachtet, in welchem Frequenzbereich die jeweiligen Hörverluste angesiedelt waren. Ein großer Teil der Probanden hatte einen Hochtonhörverlust, welcher typisch für eine altersbedingte Schwerhörigkeit ist [54]. Ob sich jedoch je nach Art des Hörverlustes womöglich unterschiedliche Korrelationen zu den Detektionsschwellen ergeben würden, konnte aufgrund der kleinen Stichprobe nicht näher untersucht werden.

Für eine spezifischere Untersuchung zu den Auswirkungen verschiedener Hörverlust-Typen auf die Detektionsschwellen müssten die Probanden in nachfolgenden Untersuchungen speziell nach der Art ihres Hörverlustes ausgewählt werden.

Aufgrund der hohen Korrelation zwischen PTA und SRT ergaben sich für den Zusammenhang zwischen dem SRT und den Detektionsschwellen ähnliche Verläufe für die einzelnen Signale wie sie in Abbildung 4.11 zu sehen sind. Auf Basis der Spearman-Korrelation sind diese jedoch für keines der Signale signifikant. Die zugehörige Abbildung kann Anhang A.2 entnommen werden.

Bei der Erstellung der Rauschsignale wurde darauf geachtet, dass sich Sprache und Rauschen nicht spektral, sondern nur in der zeitlichen Einhüllenden und der zeitlichen Feinstruktur unterscheiden. Da schwerhörige Menschen nachweislich Schwierigkeiten damit haben, die zeitliche Feinstruktur von Signalen auszuwerten [29], [33], wurde vermutet, dass sich im Rahmen dieser Studie ein linearer Zusammenhang zwischen dem Hörvermögen und den Detektionsschwellen ergibt. Dies konnte jedoch nicht bestätigt werden. Dass eine Schwerhörigkeit einen negativen Einfluss auf die Verarbeitung der zeitlichen Feinstruktur bei der Sprachwahrnehmung hat, sich jedoch kein allgemeiner Einfluss auf die Detektionsschwellen ergab, lässt vermuten, dass die zeitliche Einhüllende des Sprachsignals unabhängig vom Hörverlust verarbeitet wird und für die Sprachdetektion wichtiger ist als die zeitliche Feinstruktur.

In vorangegangenen Studien zur Signalklassifikation in Abhängigkeit vom Hörverlust bestand häufig eine Korrelation zwischen Alter und Hörverlust, sodass nicht eindeutig differenziert werden konnte, ob gefundene Einflüsse durch das Alter oder durch die Schwerhörigkeit entstanden. In der vorliegenden Arbeit wurde die Probandengruppe jedoch gezielt so ausgewählt, dass sich keine Korrelation zwischen Alter und Hörvermögen ergab. Da sich nur für ein Signal eine signifikante Korrelation zwischen Hörvermögen und Detektionsschwellen zeigte, wurde in einem weiteren Schritt der Einfluss des Alters auf die Detektionsschwellen untersucht.

4.4. Einfluss von Alter und Kognition auf die Detektionsschwellen

Da sich zwischen dem Hörvermögen und den zeitlichen Detektionsschwellen der älteren Probanden lediglich ein signifikanter Zusammenhang zu einem Signal gezeigt hat, wurden die Schwellwerte für fünf junge normalhörende Probanden ermittelt, um diese mit den Schwellwerten der älteren normalhörenden (WHO 0) Probanden zu vergleichen und einen möglichen Einfluss des Alters aufzudecken. Bei der Ermittlung der Schwellwerte erreichten die jungen Probanden in 19 von insgesamt 50 Durchläufen unabhängig des Signals die untere Begrenzung der Signallänge. Die Entscheidung für das Sprachsignal erklärten sie in allen Fällen anhand eines deutlich hörbaren Frequenzunterschiedes. Aufgrund dieser Aussagen wurden die Sprach- und Rauschsignale noch einmal spektral miteinander verglichen. Hierbei ergaben sich zwei verschiedene Faktoren, welche womöglich zu dem hörbaren Frequenzunterschied beigetragen haben.

Es zeigte sich, dass sich das Ein- und Ausblenden bei der Erstellung der Signale unterschiedlich auf die Sprach- und Rauschsignale ausgewirkt hat, wodurch sich Differenzen zwischen den Fourier-Spektren ergaben. Für lange Signale hatte das Ein- und Ausblenden nahezu keinen Einfluss. Je kürzer die Signale jedoch wurden, desto größer wurde der spektrale Unterschied zwischen Sprach- und Rauschsignal. Angelehnt an die Verarbeitung in der Cochlea wurden die Signale zusätzlich in einem Terzspektrum miteinander verglichen. In Abbildung 4.12 sind beispielhaft für zwei Signalstücke unterschiedlicher Länge die Fourier- und Terzspektren der Sprach- und Rauschsignale gegenübergestellt.

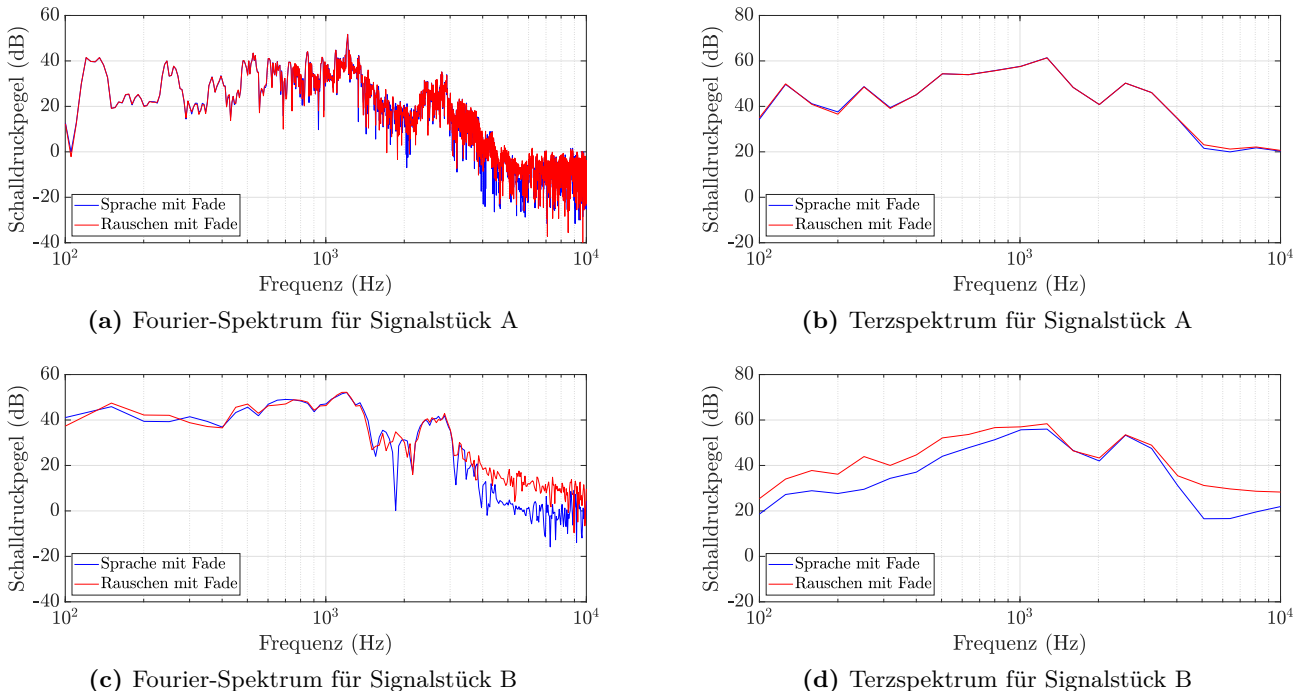


Abb. 4.12: Einfluss der Frequenzauflösung auf den Frequenzabgleich für zwei beispielhafte Signalstücke unterschiedlicher Länge bei einem Schalldruckpegel von 65 dB. Signalstück A hat eine zeitliche Dauer von 200 ms und Signalstück B von 20 ms.

In Abbildung 4.12a sind die zuvor genannten Frequenzunterschiede, welche sich auf Basis des Ein- und Ausblendens ergeben, für ein Signalstück A der Länge 200 ms dargestellt. Die zugehörige Darstellung im Terzspektrum ist in Abbildung 4.12b gezeigt. Es ist zu sehen, dass der Unterschied zwischen Sprach- und Rauschsignal für das Signalstück A im Fourier- sowie Terzspektrum sehr ähnlich ist. Für Signalstück B ist das Fourier-Spektrum für Sprache und Rauschen in Abbildung 4.12c zu sehen. Für dieses Signalstück mit einer zeitlichen Dauer von 20 ms ist der Unterschied im Terzspektrum, dargestellt in Abbildung 4.12d, jedoch deutlich ausgeprägter als im Fourier-Spektrum. Die dargestellten Unterschiede stehen exemplarisch für eines der zehn verschiedenen Signale. Es zeigten sich jedoch für alle zehn Signale ähnliche Verläufe, teilweise stärker und teilweise schwächer ausgeprägt. Die Frequenzunterschiede waren für alle jungen normalhörenden Probanden bei einigen Signalen deutlich hörbar. Die ermittelten Detektionsschwellen sind demnach nicht weiter auswertbar, da nicht festgestellt werden kann, in welchem Maße die spektralen Unterschiede die Sprachdetektion beeinflusst haben. Es erfolgte jedoch ein Vergleich der Detektionsschwellen der jungen und der älteren normalhörenden Probanden. Die Verteilungen sind in Abbildung 4.13 dargestellt.

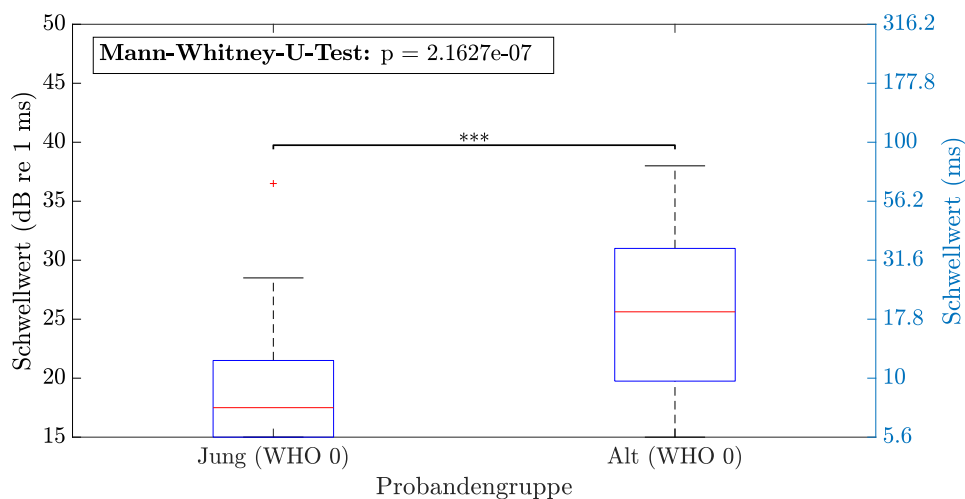


Abb. 4.13: Vergleich der zeitlichen Detektionsschwellen der jungen und älteren normalhörenden Probanden.

Die geringe Streuung von den Detektionsschwellen der jungen Probanden, sowie ein Median von 17,5 dB re 1 ms lassen vermuten, dass die unterschiedlichen Frequenzspektren einen Einfluss auf die Detektionsschwellen besitzen. Auch die Verteilung der Detektionsschwellen der älteren normalhörenden Probanden reicht bis zur unteren Begrenzung, da zwei der sechs Probanden ebenfalls für einzelne Signale einen Frequenzunterschied hörten, welchen sie vermutlich als Hinweis zur Sprachdetektion nutzen konnten. Die Streuung der Detektionsschwellen ist jedoch für die älteren normalhörenden Probanden sichtbar größer als für die jüngeren Probanden. Das 1. Quartil liegt zudem oberhalb des Medians der jungen Probanden und auf Basis des Mann-Whitney-U Tests besteht mit einem p-Wert von $2,16e-7$ ein höchst signifikanter Unterschied zwischen den Medianen. Dieser Unterschied deutet darauf hin, dass die spektralen Unterschiede für die älteren Probanden nicht ausschlaggebend für die Sprachdetektion waren.

Hopkins und Moore (2012) zeigten auf, dass die Fähigkeit, die zeitliche Feinstruktur eines Signals zu verarbeiten, mit dem Alter abnimmt, auch wenn das Hörvermögen normal ist [55]. Diese These unterstützt die Vermutung, dass das Alter, bzw. die damit verbundenen kognitiven Fähigkeiten einen Einfluss auf die zeitlichen Detektionsschwellen haben. Um diese Vermutung näher zu untersuchen, wurde innerhalb der Altersspanne der 30 älteren Probanden ein möglicher Einfluss des Alters auf die mittleren Detektionsschwellen betrachtet. Zudem wurden die Ergebnisse des MoCA Tests herangezogen, um zu untersuchen, ob ein Zusammenhang zwischen der kognitiven Verarbeitung und den Schwellwerten besteht. Die Ergebnisse sind in Abbildung 4.14 dargestellt.

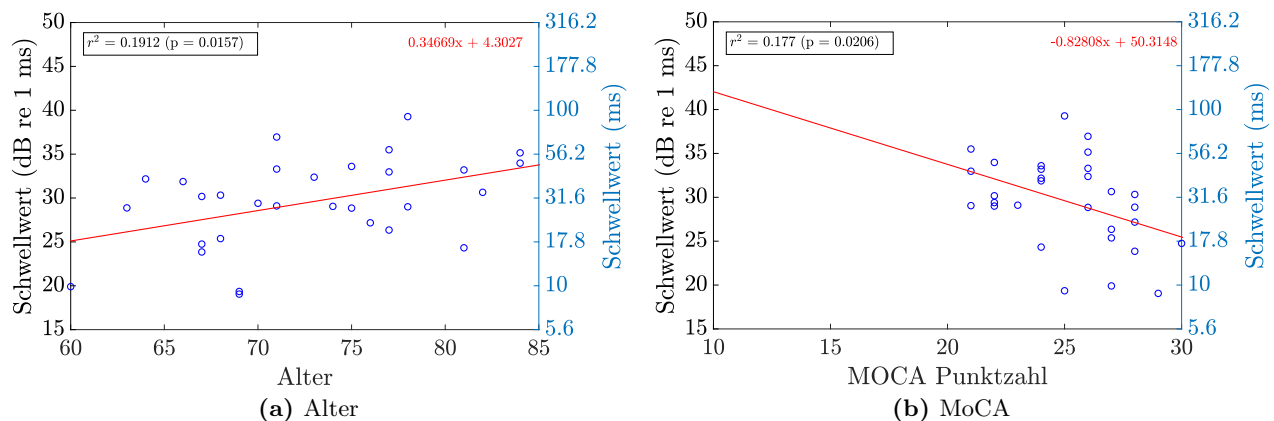


Abb. 4.14: Zeitliche Detektionsschwellen in Abhängigkeit von Alter und Kognition.

In Abbildung 4.14a ist die Korrelation vom Alter der älteren Probanden und der Detektionsschwellen zu sehen. Der Pearson-Korrelationskoeffizient liefert mit einem r^2 von 0,19 mit einem p-Wert von $\approx 0,016$ einen signifikanten Zusammenhang. Die lineare Regression liefert eine Steigung von ca. 0,35 dB pro zusätzlichem Lebensjahr. Je älter also die Probanden sind, desto höher liegen die Schwellwerte. Auch die Kognition hat einen signifikanten Einfluss auf die zeitlichen Detektionsschwellen, was in Abbildung 4.14b zu sehen ist. Der Pearson-Korrelationskoeffizient liefert ein r^2 von 0,177 (p-Wert = 0,02) und die lineare Regression eine Steigung von -0,828 dB pro zusätzlichem MoCA Punkt. Je höher also die erreichte Punktzahl im MoCA Test, desto niedriger liegen die Schwellwerte. Dies unterstützt die Vermutung, dass die Kognition einen Einfluss auf die Verarbeitung der zeitlichen Informationen von Sprachsignalen hat und war zudem zu erwarten, da die kognitive Leistung allgemein stark mit dem Alter korreliert [56].

Eine Korrelation vom Alter der älteren Probanden und den zugehörigen MoCA Testergebnissen ist in Abbildung 4.15 dargestellt.

Mit einem r^2 von 0,1 und einem zugehörigen p-Wert von 0,077 liefert die Pearson-Korrelation zwar keinen signifikanten Zusammenhang, jedoch eine starke Tendenz dahingehend, dass die kognitive Leistung mit höherem Alter abnimmt. Die Nicht-Signifikanz könnte ein Anzeichen dafür sein, dass der MoCA Test als Screening Test nicht sensitiv genug ist, um detaillierte Aussagen über die kognitive Verarbeitung der Probanden treffen zu können.

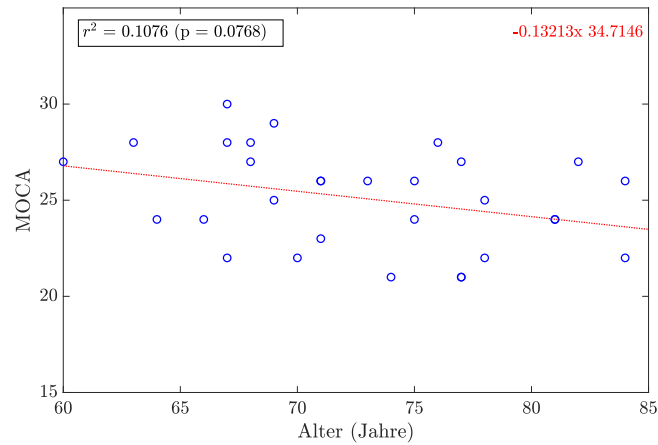


Abb. 4.15: Korrelation von Alter und MoCA Punktzahl pro Proband.

Da der MoCA Test verschiedene Kategorien beinhaltet, wurden abschließend die erreichten Punktzahlen pro Kategorie ausgewertet, um zu untersuchen, welche kognitiven Bereiche einen Einfluss auf die Detektionsschwellen haben könnten.

Einige Kategorien wiesen keine oder nur geringe Streuungen in der Verteilung der Punktzahlen auf, sodass die zugehörigen kognitiven Bereiche demnach keinen oder nur einen geringen Einfluss auf die Detektionsschwellen zu haben scheinen. Die entsprechenden Ergebnisse sind Anhang A.2 zu entnehmen. Die Ergebnisse für die beiden Kategorien „Erinnerung“ und „Visuospatial/Exekutiv“ wiesen jedoch eine deutliche Streuung in den Punktzahlen auf. Die relativen Häufigkeiten der jeweils erreichten Punktzahlen sind in Abbildung 4.16 dargestellt.

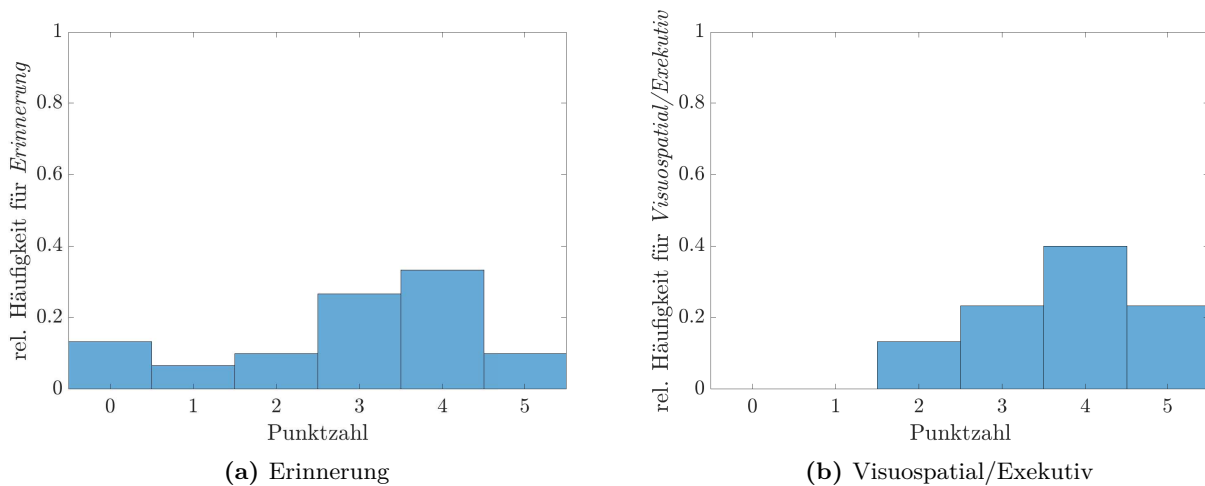


Abb. 4.16: Relative Häufigkeit der MoCA Punktzahlen für Erinnerung und Visuospatial/Exekutiv.

In Abbildung 4.16a sind die relativen Häufigkeiten der erreichten Punktzahlen für die Kategorie „Erinnerung“ zu sehen. Die Aufgabe bestand darin, sich fünf Worte über einen Zeitraum von wenigen Minuten zu merken und anschließend zu wiederholen. Die von den Probanden erreichten Punktzahlen sind zwischen null und fünf von fünf möglichen Punkten verteilt. Dies lässt vermuten, dass das Erin-

nerungsvermögen einen Einfluss auf die Detektionsschwellen haben könnte. Für die Sprachdetektion hörten die Probanden nacheinander das Sprach- und das Rauschsignal und wählten erst anschließend aus, welches der beiden Signale sie als Sprachsignal wahrgenommen haben. Hierfür mussten sich demnach beide Signale bestmöglich gemerkt werden, um sie miteinander vergleichen zu können. Schwierigkeiten mit der Merkfähigkeit könnten demnach möglicherweise die Detektion des Sprachsignals erschwert haben.

Die relativen Häufigkeiten der erreichten Punktzahlen für die Kategorie „Visuospatial/Exekutiv“ sind in Abbildung 4.16b dargestellt. Die zugehörigen Aufgabenstellungen sind dem MoCA Aufgabenblatt aus Anhang A.1 zu entnehmen. Die Punktzahlen der Probanden streuen zwischen zwei und fünf von fünf möglichen Punkten. Bezogen auf die Schwellenbestimmung erfolgte die Detektion der Sprachsignale über die manuelle Eingabe am Touchscreen. Möglicherweise war die Fähigkeit, den linken Button für das zuerst gehörte Signal oder den rechten Button für das zweite Signal zu drücken, aufgrund von Schwierigkeiten bei der visuell-räumlichen Verarbeitung eingeschränkt.

Um jedoch genauere Aussagen darüber treffen zu können, in welchem Maße die verschiedenen kognitiven Fähigkeiten einen Einfluss auf die Sprachdetektion haben, wäre es hilfreich, verschiedene kognitive Tests bei den Messungen einzubeziehen, die jeweils darauf abzielen, verschiedene kognitive Bereiche zu testen und getrennt voneinander auszuwerten.

5. Zusammenfassung und Ausblick

In dieser Arbeit wurden die zeitlichen Schwellen zur subjektiven Sprachdetektion untersucht. Hierbei wurde sowohl der Einfluss der Hörschwelle und des Sprachverstehens im Störgeräusch, als auch der Einfluss des Initialphonems auf die Detektionsschwellen betrachtet. Zudem wurde untersucht, ob das Alter und die Kognition mit den Schwellwerten in Beziehung stehen.

Die Ergebnisse zeigen einen signifikanten Einfluss des Initialphonems auf die zeitlichen Detektionsschwellen. Die kürzesten Schwellwerte ergaben sich für /Ra/ mit einem Median von 21,4 ms. Sämtliche Schwellwerte für die Signale mit Vokal-Initialphonem lagen im Bereich zwischen 20 und 40 ms nah beieinander. Die höchsten Schwellwerte ergaben sich für /Na/ mit einem Median von 93,1 ms. Bis auf ein Signal lagen zudem alle Schwellwerte im Bereich des ersten Phonems. Es lässt sich also sagen, dass, wenn der Informationsgehalt bei der Aussprache eines Phonems groß genug ist, dies scheinbar für die Sprachdetektion ausreicht. Das /L/ und das /N/ scheinen diesen Informationsgehalt als Nasal und lateraler Approximant erst nach einer gewissen Zeit zur Verfügung zu stellen, da diese die höchsten Schwellen im Vergleich zu den anderen Phonemen aufwiesen.

Ein Einfluss der Hörschwelle auf die zeitlichen Detektionsschwellen zeigte sich nur für eines von zehn Signalen mit einer signifikanten Korrelation. Eine Abhängigkeit der Detektionsschwellen vom Sprachverstehen im Störgeräusch konnte nicht nachgewiesen werden. Es wurde ursprünglich vermutet, dass die Detektionsschwellen mit größerer Schwerhörigkeit höher werden, da die Verarbeitung der zeitlichen Feinstruktur von Sprachsignalen durch eine Schwerhörigkeit beeinflusst wird. Dies ließ sich jedoch nicht bestätigen. Da die verwendeten Rauschsignale sowohl in der zeitlichen Feinstruktur, als auch in der Einhüllenden modifiziert waren, lässt sich vermuten, dass die zeitliche Einhüllende eines Sprachsignals, unabhängig vom Hörverlust, mehr von Bedeutung für die Sprachdetektion ist als die zeitliche Feinstruktur.

Zudem wurde ein Zusammenhang von Alter und Kognition mit den Detektionsschwellen gefunden, was auf einen Alterungsprozess der auditorischen Verarbeitung schließen lässt. Dieser zeigte sich sowohl im Vergleich der Detektionsschwellen von Jung und Alt, als auch in der Korrelation der MoCA Testergebnisse und dem Alter der älteren Probanden mit den Detektionsschwellen. Bei den MoCA Testergebnissen zeigte sich, dass die älteren Probanden vor allem bei den Kategorien „Erinnerung“ und „Visuospatial/Exekutiv“ Schwierigkeiten hatten. Der MoCA Test ist jedoch nicht darauf ausgelegt, einzelne Bereiche der kognitiven Verarbeitung präzise zu untersuchen. Demnach kann keine Aussage darüber getroffen werden, welche kognitiven Prozesse tatsächlich die Detektionsschwellen für Sprachsignale beeinflussen.

Es ergab sich außerdem die Schwierigkeit, dass der Frequenzabgleich zwischen den Sprach- und Rauschsignalen nur begrenzt möglich war, wodurch sich ein Frequenzunterschied ergab, welcher von allen jungen Normalhörenden und von drei älteren Probanden für einzelne Signale deutlich hörbar war.

Im Hinblick auf die Ergebnisse dieser Studie lässt sich sagen, dass das gewählte Verfahren zur Bestimmung der zeitlichen Detektionsschwellen für Sprachsignale anwendbar ist, sich jedoch Grenzen

gezeigt haben, welche bei weiterführenden Untersuchungen berücksichtigt werden müssen.

Bezogen auf den Frequenzabgleich zwischen Sprach- und Rauschsignalen sollte in nachfolgenden Untersuchungen darauf geachtet werden, dass sich keine hörbaren Frequenzunterschiede zwischen den Signalen ergeben. Nur so kann sichergestellt werden, dass lediglich die Informationen aus den zeitlichen Merkmalen der Signale für die Sprachdetektion genutzt werden können.

Zudem wurde bei der Erstellung der Rauschsignale neben der zeitlichen Feinstruktur auch die Einhüllende modifiziert. Da sich kein Zusammenhang zwischen Hörvermögen und Detektionsschwellen feststellen ließ, Schwerhörige jedoch die zeitliche Feinstruktur schlechter verarbeiten können als Normalhörende, wäre eine spezifischere Untersuchung von Interesse. Indem sich das Rauschsignal lediglich in der zeitlichen Feinstruktur von dem zugehörigen Sprachsignal unterscheiden würde, könnte untersucht werden, inwiefern eine Schwerhörigkeit die Sprachdetektion anhand der Informationen aus der zeitlichen Feinstruktur beeinflusst. Es wird davon ausgegangen, dass sich die Detektionsschwellen unabhängig vom Hörverlust zu höheren Werten verschieben, da die Detektion ohne die Informationen aus der Einhüllenden grundsätzlich erschwert ist. Zudem sollten sich Schwellen ergeben, welche vom Grad der Schwerhörigkeit abhängen, da dies die Verarbeitung der zeitlichen Feinstruktur beeinflusst. Auch der Einfluss der kognitiven Verarbeitung auf die zeitlichen Detektionsschwellen müsste näher untersucht werden. Hierfür wäre es sinnvoll, kognitive Tests zu verwenden, welche speziell darauf ausgelegt sind, bestimmte Bereiche der Kognition zu untersuchen. So könnten präzisere Aussagen darüber getroffen werden, welche kognitiven Prozesse in die Sprachdetektion involviert sind.

Literatur

- [1] L. Robert Slevc und Alison R. Shell: *Auditory agnosia*. In: *Handbook of Clinical Neurology*, Band 129, Seiten 573–587. Elsevier B.V., 2015. <http://dx.doi.org/10.1016/B978-0-444-62630-1.00032-9>.
- [2] Albert S. Bregman: *Auditory Scene Analysis: The Perceptual Organization of Sound*. Journal of the Acoustical Society of America, 95, 1990.
- [3] Sebastian Obert und Jürgen Tchorz: *Geräuschklassifikation*. Projektarbeit, Technische Hochschule Lübeck, 2018.
- [4] Daniel Bank, Mario Schinnerl, Marlitt Frenz, Felix Gassenmeyer und Hendrik Husstedt: *Duration Threshold for Identifying Sound Samples of Elderly Hearing Impaired*. In: *Studierendentagung BioMedTec*, Lübeck, 2019.
- [5] Durga Lal Budathoki, Greg O’Beirne und Jürgen Tchorz: *Time Required to Classify Different Sound Types and Their Relation to Hearing Ability and Speech Understanding in Noise*. Dissertation, University of Canterbury, 2019.
- [6] Wilhelm Schmidt, Elisabeth Berner, Norbert Richard Wolf und Helmut Langner: *Geschichte der deutschen Sprache - ein Lehrbuch für das germanistische Studium*. Hirzel, Stuttgart, 11. Auflage, 2013, ISBN 978-3-777-62272-9.
- [7] Mario Schinnerl, Marlitt Frenz, Felix Gassenmeyer und Hendrik Husstedt: *Procedure with self Adjustment to measure the minimum Time needed for Identifying Sound Samples*. In: *Studierendentagung BioMedTec*, Lübeck, 2019.
- [8] Russell J. Love und Wanda G. Webb: *Neurology for the Speech-Language Pathologist*. Butterworth-Heinemann, Stoneham, 2. Auflage, 2013, ISBN 1483141993.
- [9] Stephen McAdams und Carolyn Drake: *Auditory Perception and Cognition*. In: Hal Pashler und Steven Yantis (Herausgeber): *Stevens’ Handbook of Experimental Psychology*, Seiten 397–452. Wiley, NewYork, 3. Auflage, 2002.
- [10] Carryl L. Baldwin: *Auditory Cognition and Human Performance - Research and Applications*. CRC Press, Boca Raton, Fla, 2016, ISBN 978-1-466-55354-5.
- [11] Jürgen Hellbrück und Wolfgang Ellermeier: *Hören: Physiologie, Psychologie und Pathologie*. Hogrefe Verlag, Göttingen, 2. Auflage, 2004.
- [12] Vishakha W Rawool: *The Effects of Hearing Loss on Temporal Processing*. Hearing Review, 2006. <http://www.hearingreview.com/2006/06/the-effects-of-hearing-loss-on-temporal-processing/>.

-
- [13] Josh H. McDermott, Michael Schemitsch und Eero P. Simoncelli: *Summary statistics in auditory perception*. Nature Neuroscience, 16(4):493–498, 2013, ISSN 10976256. <http://dx.doi.org/10.1038/nn.3347>.
- [14] Giles Wilkeson Gray: *Phonemic microtomy: The minimum duration of perceptible speech sounds*. Speech Monographs, 9(1):75–90, 1942, ISSN 0038-7169.
- [15] Wolfgang Hess: *Grundlagen der Phonetik - 2. Deskriptive Phonetik*. 2002.
- [16] T. Alan Hall: *Phonologie: Eine Einführung*. Walter de Gruyter, Berlin, 2. Auflage, 2011, ISBN 978-3-110-21587-8.
- [17] Alvin M. Liberman, Franklin S. Cooper, Donald P. Shankweiler und Michael Studdert-Kennedy: *Perception of the speech code*. Psychological Review, 74(6):431–461, 1967.
- [18] Robert E. Remez, Philip E. Rubin, David B. Pisoni und Thomas D. Carrell: *Speech Perception Without Traditional Speech Cues*. Science, 212(4497):947–949, 1981.
- [19] H. McGurk und J. MacDonald: *Hearing lips and seeing voices*. Nature, 264(5588):746–748, 1976.
- [20] Alvin M. Liberman und Ignatius G. Mattingly: *The motor theory of speech perception revised*. Cognition, 21:1–36, 1985, ISSN 10699384.
- [21] James L. McClelland und Jeffrey L. Elman: *The TRACE model of speech perception*. Cognitive psychology, 18(1):1–86, 1986.
- [22] Leonard Bloomfield: *Language*. Holt, Rinehart and Winston., New York, 1933.
- [23] Kenneth N. Stevens: *Toward a model for lexical access based on acoustic landmarks and distinctive features*. The Journal of the Acoustical Society of America, 111(4):1872–1891, 2002, ISSN 0001-4966.
- [24] Brian C. J. Moore: *An Introduction to the Psychology of Hearing*. BRILL, Leiden - Boston, 6. auflage Auflage, 2013.
- [25] Ronald A. Cole und Scott Brian: *Towards a theory of speech perception*. Psych, 81(4):348–374, 1974.
- [26] Sheila E. Blumstein und Kenneth N. Stevens: *Acoustic Invariance In Speech Production: Evidence From Measurements Of The Spectral Characteristics Of Stop Consonants*. Journal of the Acoustical Society of America, 66(4):1001–1017, 1979, ISSN NA.
- [27] Kenneth N. Stevens: *Acoustic correlates of some phonetic categories*. Journal of the Acoustical Society of America, 68(3):836–842, 1980, ISSN NA.
- [28] Il Joon Moon und Sung Hwa Hong: *What is temporal fine structure and why is it important?* Korean Journal of Audiology, 18(1):1–7, 2014, ISSN 20933797.
-

-
- [29] Kathryn Hopkins und Brian C. J. Moore: *The importance of temporal fine structure information in speech at different spectral regions for normal-hearing and hearing-impaired subjects*. The Journal of the Acoustical Society of America, 127(3):1595–1608, 2010, ISSN 0001-4966.
- [30] P. von Gablenz, I. Holube und E. Hoffmann: *Prevalence of hearing loss in Northern and Southern Germany. German version*. HNO, 65(8):663–670, 2017, ISSN 0017-6192.
- [31] Thomas Zahnert: *The Differential Diagnosis of Hearing Loss*. Deutsches Arzteblatt, 108(25):433–444, 2011, ISSN 00121207.
- [32] Harvey Dillon: *Hearing Aids*. Thieme, Stuttgart, 2. Auflage, 2012, ISBN 978-1-604-06811-5.
- [33] Christian Lorenzi, Gaëtan Gilbert, Héloïse Carn, Stéphane Garnier und Brian C.J. Moore: *Speech perception problems of the hearing impaired reflect inability to use temporal fine structure*. Proceedings of the National Academy of Sciences of the United States of America, 103(49):18866–18869, 2006, ISSN 00278424.
- [34] Werner Stangl: *Psychophysik. Online Lexikon für Psychologie und Pädagogik*, 2019. <https://lexikon.stangl.eu/78/psychophysik/>, besucht: 2019-06-01.
- [35] Christian Hatzfeld und Roland Werthschützky: *Simulation und Auswahl von psychometrischen Verfahren zur Ermittlung von Kennwerten menschlicher Wahrnehmung*. Technisches Messen, 81(4):173–181, 2014, ISSN 01718096.
- [36] Manuel Kühner, Heiner Bubb, Klaus Bengler und Jörg Wild: *Adaptive Verfahren in der Psychophysik*. Ergonomie aktuell, 13:26–30, 2012.
- [37] Bernhard Treutwein: *Adaptive psychophysical procedures*. Vision Research, 35(17):2503–2522, 1995.
- [38] Thomas D. Wickens: *Elementary Signal Detection Theory*. OXFORD University Press, Los Angeles, 2010.
- [39] Herbert Hagendorf, Joseph Krummenacher, Hermann Joseph Müller und Torsten Schubert: *Allgemeine Psychologie für Bachelor: Wahrnehmung und Aufmerksamkeit. (Lehrbuch mit Online-Materialien)*. Springer-Verlag, Berlin Heidelberg New York, 1. Auflage, 2011, ISBN 978-3-642-12710-6.
- [40] David Heeger: *Signal Detection Theory*. Technischer Bericht, 1997. papers://c941067e-36da-4589-acf1-2f5738fdb5a1/Paper/p764.
- [41] Jochen Müsseler und Martina Rieger: *Allgemeine Psychologie*. Springer-Verlag, Berlin Heidelberg New York, 3. Auflage, 2016, ISBN 978-3-642-53898-8.
- [42] Herbert Robbins und Sutton Monroe: *A Stochastic Approximation Method*. IEEE Transactions on Systems, Man and Cybernetics, 1(4):338–344, 1971, ISSN 21682909.
-

-
- [43] Harry Kesten: *Accelerated Stochastic Approximation*. The Annals of Mathematical Statistics, 29(1):41–59, 1958, ISSN 0003-4851.
- [44] H. Levitt: *Transformed Up-Down Methods for Psychoacoustics*. The Journal of the Acoustical Society of America, 49(2B):467–477, 1971.
- [45] Cyrus Derman: *Non-parametric up-and-down experimentation*. Annals of Mathematical Statistics, 28:795–797, 1957.
- [46] Christian Kaernbach: *Simple adaptive testing with the weighted up-down method*. Perception, & Psychophysics, 49(3):227–229, 1991, ISSN 00315117.
- [47] Miguel A. García-Pérez: *Forced-choice staircases with fixed step sizes: asymptotic and small-sample properties*. Vision Research, 38(12):1861–1881, 1998, ISSN 00426989.
- [48] Thomas Steffens: *Die systematische Auswahl von sprachaudiometrischen Verfahren*. HNO, 65(3):219–227, 2017, ISSN 00176192.
- [49] *MoCA Montreal Cognitive Assessment*, 2019. <https://www.mocatest.org>, besucht: 2019-06-10.
- [50] Randall L. Jones und Erwin Tschirner: *A Frequency Dictionary of German*. Routledge, New York, 2006.
- [51] Roland Laszig, Ernst Lehnhardt, Gerhard Hesse, Thomas Janssen und Martin Kinkel: *Praxis der Audiometrie*. Georg Thieme Verlag, Stuttgart, 9. Auflage, 2009.
- [52] Thomas Brand und Birger Kollmeier: *Eignung des Göttinger und des Oldenburger Satztests für die Bestimmung von Verständlichkeitsfunktionen*. Fortschritte der Akustik - DAGA, Seiten 256–257, 2000.
- [53] Lidia W. Lee und Larry E. Humes: *Evaluating a speech-reception threshold model for hearing-impaired listeners*. Journal of the Acoustical Society of America, 93(5):2879–2885, 1993, ISSN NA.
- [54] Gerhard Hesse: *Innenohrschwerhörigkeit*. Thieme, Stuttgart, 2015.
- [55] B. C. J. Moore: *The importance of temporal fine structure for the intelligibility of speech in complex backgrounds*. ISAAR 2011: Speech Perception and Auditory Disorders, (August 2011):21–32, 2012.
- [56] Paul Verhaeghen und Timothy A. Salthouse: *Meta-analyses of age-cognition relations in adulthood: Estimates of linear and nonlinear age effects and structural models*. Psychological Bulletin, 122(3):231–249, 1997, ISSN 00332909.

A. Anhang

A.1. Methoden

Bestätigung der Ethik Kommission



UNIVERSITÄT ZU LÜBECK

Universität zu Lübeck · Ratzeburger Allee 160 · 23538 Lübeck

Herrn
Dr. Hendrik Husstedt
Deutsches Hörgeräte Institut GmbH
Anschützstraße 1

23562 Lübeck

Ethik-Kommission

Vorsitzender:
Herr Prof. Dr. med. Alexander Katalinic
Stellv. Vorsitzender:
Herr Prof. Dr. med. Frank Gieseler

Ratzeburger Allee 160
23538 Lübeck

Sachbearbeitung: Frau Janine Erdmann
Tel.: +49 451 3101 1008
Fax: +49 451 3101 1024
ethikkommission@uni-luebeck.de

Aktenzeichen: 19-174A
Datum: 10. Mai 2019

Verkürztes Verfahren Anzeige

Eine prospektive Studie für die „zeitliche Schwellenbestimmung zur subjektiven Sprachdetektion“ am Deutschen Hörgeräte Institut im Rahmen einer Masterarbeit des Studiengangs Hörakustik und audilogische Technik der Universität zu Lübeck
Ihr Schreiben vom 08. Mai 2019

Sehr geehrter Herr Dr. Husstedt,


mit Ihrem o.g. Schreiben informieren Sie die Ethik-Kommission über Ihr geplantes Vorhaben.

Es lagen folgende Unterlagen vor:

- Ihr Anschreiben vom 08. Mai 2019
- Studienprotokoll Nr. 1 vom 23. April 2019
- Aufklärung und Einwilligung FB402 V002, gültig ab 06. Mai 2019.

Die Ethik-Kommission nimmt das Vorhaben zustimmend zur Kenntnis.
Eine Behandlung im normalen Antragsverfahren wird nicht für notwendig erachtet.

Mit freundlichen Grüßen


Prof. Dr. med. Alexander Katalinic
Vorsitzender

Probandeninformation (Seite 1)

Deutsches Hörgeräte Institut GmbH

Anschützstr. 1
23562 Lübeck
Tel.: 0451 305051 0
Fax: 0451 305051 99

Sehr geehrte Dame, sehr geehrter Herr,

Vielen Dank für Ihr Interesse an unserer Studie *zeitliche Schwellenbestimmung zur subjektiven Sprachdetektion* – Studienleiter Dr. Hendrik Husstedt.

Im Folgenden erhalten Sie von uns einige grundlegende Informationen zur Studie und den geplanten Messungen. Außerdem informieren wir Sie über den Umgang mit den erhobenen Daten.

Studienziele

In dieser Studie soll untersucht werden, welche minimale Länge ein Sprachsignal haben muss, um von Normalhörenden und Schwerhörigen ohne Hörgeräte eindeutig als Sprache erkannt zu werden.

Vor der eigentlichen Schwellenbestimmung werden eine Reintonaudiometrie sowie eine Messung der Sprachverständlichkeit im Störgeräusch durchgeführt. So kann ein möglicher Zusammenhang von Hörverlust oder Sprachverstehen mit den benötigten Schwellwerten untersucht werden. Des Weiteren erfolgt ein kurzer neuropsychologischer Test.

Die Studie wird im Rahmen der Masterarbeit einer Studentin des Masterstudiengangs „Hörakustik und Audiologische Technik“ der Universität zu Lübeck und im Auftrag des Deutschen Hörgeräte Instituts durchgeführt. Die Studierende ist eine ausgebildete Hörgeräteakustikerin.

Studienumfang und geplanter Ablauf

Die Studie findet an einem Tag, über einen Zeitraum von maximal zwei Stunden, in Gebäude D der Akademie für Hörakustik, Bessemerstr. 3, 23562 Lübeck, statt. Nach einem kurzen Anamnesebogen (Alter, Geschlecht, Beginn des Hörverlustes, Erkrankungen an den Ohren u.ä.) werden eine Begutachtung des äußeren Ohres (Otoskopie) und eine Reintonaudiometrie zur Ermittlung der Hörschwelle durchgeführt, sowie das Sprachverstehen im Störgeräusch ohne Hörgeräte ermittelt. Es folgt ein kurzer neuropsychologischer Test.

Anschließend findet die Bestimmung der zeitlichen Schwelle zur subjektiven Sprachdetektion statt. Hierfür stellen Sie zunächst eigenständig die Lautstärke der folgenden Testsignale auf einen für Sie angenehmen Pegel ein.

Sie bekommen dann zwei Signale dargeboten (einmal Sprache – einmal Störgeräusch) und haben die Aufgabe auszuwählen, bei welchem der beiden Signale es sich um Sprache handelt. Dies wird über einen Touchscreen ausgewählt. Die Präsentationsdauer der einzelnen Signale wird im Verlauf der Messungen

FB402 V002
Gültig ab 06.05.2019

1

Probandeninformation (Seite 2)

immer kürzer. Die Signale werden über Kopfhörer wiedergegeben. Die Lautstärke entspricht hierbei dem von Ihnen eingestellten Pegel.

Die Ergebnisse dieser Studie können als Grundlage zur Hypothesenbildung für weitere Studien dienen. Als Aufwandsentschädigung erhalten Sie 5 Euro Anfahrtspauschale und 7 Euro für jede angefangene Stunde.

Für die Durchführung der Messungen stellen Sie sich vorab selbst die Lautstärke auf einen für Sie angenehmen Pegel ein. Die so eingestellte Lautstärke wird über den gesamten Test konstant gehalten. Sollte die Lautstärke zu einem späteren Zeitpunkt als zu laut empfunden werden, nehmen Sie die Kopfhörer ab oder geben Sie dem Testleiter ein Zeichen, dass dieser die Messung sofort unterbrechen kann.

Das Risiko des Bekanntwerdens der sensiblen Patientendaten wird durch die Pseudonymisierung und Zugriffsbeschränkung minimiert. Es sind keine Medikamenteneinnahmen oder andere invasive Maßnahmen erforderlich.

Datenschutzrechtliche Informationen

Für die Datenverarbeitung verantwortlich ist die Datenschutzbeauftragte des Unternehmens. Die Datenerhebung erfolgt zum Zweck des oben genannten Studienziels durch die durchführende Person der Studie. Die personenbezogenen Daten wie der Name, die Adresse usw. werden streng vertraulich und nach gesetzlichen Bestimmungen behandelt. Alle während der Studie erhobenen Daten werden temporär in pseudonymisierter¹ Form elektronisch gespeichert.

Für die spätere Auswertung werden die Daten aller Probanden vollständig anonymisiert² herangezogen. Sämtliche personenbezogenen Daten, die zum Zweck der Kommunikation zwischen dem Unternehmen und dem Probanden herangezogen werden, werden nach Ende der Studie vollständig gelöscht. Zugriff auf Ihre Daten haben nur Mitarbeiter des Unternehmens. Diese Personen sind zur Verschwiegenheit verpflichtet. Ebenfalls sind die temporär elektronisch gespeicherten Daten vor fremdem Zugriff geschützt.

Sie haben jederzeit das Recht und die Möglichkeit, Ihre Einwilligung an der Studienteilnahme zu widerrufen und/oder die Löschung Ihrer personenbezogenen Daten zu fordern.

¹ Pseudonymisierung: „die Verarbeitung personenbezogener Daten in einer Weise, in der die Daten ohne Hinzuziehung zusätzlicher Informationen nicht mehr einer spezifischen betroffenen Person zugeordnet werden können, sofern diese zusätzlichen Informationen gesondert aufbewahrt werden und technischen und organisatorischen Maßnahmen unterliegen, die gewährleisten, dass die Daten keiner betroffenen Person zugewiesen werden können;“ BDSG §45 Absatz 5

² Anonymisierung: „...Informationen, die sich nicht auf eine identifizierte oder identifizierbare natürliche Person beziehen, oder personen-bezogene Daten, die in einer Weise anonymisiert worden sind, dass die betroffene Person nicht oder nicht mehr identifiziert werden kann.“ Verordnung (EU) 2016/679 (26)

Probandeninformation (Seite 3)



Deutsches Hörgeräte Institut GmbH

Anschützstr. 1
23562 Lübeck
Tel.: 0451 305051 0
Fax: 0451 305051 99

**Einwilligungserklärung zur Teilnahme an der Studie
„zeitliche Schwellenbestimmung zur subjektiven
Sprachdetektion“**

- Ich bestätige hiermit, dass die im Rahmen der Studie erhobenen Daten und Studienbefunde auf Fragebögen und elektronischen Datenträgern aufgezeichnet werden und anonymisiert für wissenschaftliche oder historische Forschungszwecke ohne Namensnennung verwendet werden.
- Ich erkläre mich hiermit einverstanden, dass meine erhobenen Daten in anonymisierter Form für Publikationszwecke in wissenschaftlichen Zeitschriften sowie auf Konferenzen präsentiert werden.
- Ich versichere hiermit, dass ich freiwillig an dieser Studie teilnehme, wobei mir vorbehalten ist, meine Mitwirkung jederzeit ohne Angabe von Gründen zu beenden. In diesem Fall werde ich für meine Teilnahme anteilig bezahlt.
- Ich habe eine schriftliche Information zur oben genannten Studie erhalten, gelesen und verstanden. Ich wurde über den Inhalt, die Vorgehensweise sowie die Risiken der Studie ausführlich und in verständlicher Form aufgeklärt. Meine Fragen wurden ausreichend beantwortet. Ich hatte genügend Zeit, mich gegen eine Teilnahme an der Studie zu entscheiden und willige hiermit in diese ein.

Lübeck, den _____

Teilnehmer/in

Studienleiter/in

Die Einwilligung kann jederzeit widerrufen werden. Jedoch werden Verarbeitungen der Daten bis zum Widerruf der Einwilligung nicht berührt.
Eine Einsicht der temporär pseudonymisierten und später anonymisierten Daten ist für die betroffene Person auf schriftlichen Antrag an das Unternehmen jederzeit möglich.
Eine Kopie der Einwilligungserklärung kann jederzeit durch die betroffene Person eingefordert werden.

Sollte die Einwilligung verweigert werden, so ist der Termin zur Studie abzubrechen.
FB402 V002
Gültig ab 06.05.2019

MoCA Aufgabenblatt

NAME :
 Ausbildung :
 Geschlecht :
 Geburtsdatum :
 DATUM :

MONTREAL COGNITIVE ASSESSMENT (MOCA)

VISUOSPATIAL / EXEKUTIV

Ende (E) [A] [B] [C] [D] [5] [1] [2] [3] [4]

Beginn [1]

[] []

Würfel nachzeichnen

[] []

Eine Uhr zeichnen (Zehn nach elf)
(3 Punkte)

[] [] []

Kontur Zahlen Zeiger

___/5

BENENNEN

[]

[]

[]

___/3

GEDÄCHTNIS	GESICHT	SAMT	KIRCHE	TULPE	ROT	Keine Punkte
Wortliste vorlesen, wiederholen lassen. 2 Durchgänge. Nach 5 Minuten überprüfen (s.u.)	1. Versuch					
	2. Versuch					

AUFMERKSAMKEIT Zahlenliste vorlesen (1 Zahl/ Sek.) In der vorgegebenen Reihenfolge wiederholen [] 2 1 8 5 4
 Rückwärts wiederholen [] 7 4 2 ___/2

Buchstabenliste vorlesen (1 Buchst./Sek.). Patient soll bei jedem Buchstaben „A“ mit der Hand klopfen. Keine Punkte bei 2 oder mehr Fehlern
 [] FBACMNAAJKLBAFAKDEAAAJAMOF AAB ___/1

Fortlaufendes Abziehen von 7, mit 100 anfangen [] 93 [] 86 [] 79 [] 72 [] 65 ___/3
 4 oder 5 korrekte Ergebnisse: 3 P., 2 oder 3 korrekt: 2 P., 1 korrekt: 1 P., 0 korrekt: 0 P.

SPRACHE Wiederholen: „Ich weiß lediglich, dass Hans heute an der Reihe ist zu helfen.“ [] ___/2
 „Die Katze versteckte sich immer unter der Couch, wenn die Hunde im Zimmer waren.“ []

Möglichst viele Wörter in einer Minute benennen, die mit dem Buchstaben F beginnen [] _____ (N ≥ 11 Wörter) ___/1

ABSTRAKTION Gemeinsamkeit von z.B. Banane und Apfelsine = Frucht [] Eisenbahn - Fahrrad [] Uhr - Lineal ___/2

ERINNERUNG	GESICHT	SAMT	KIRCHE	TULPE	ROT	Punkte nur bei richtigem Nennen OHNE Hinweis
Worte erinnern OHNE HINWEIS	[]	[]	[]	[]	[]	
Optional Hinweis zu Kategorie						
Mehrfachauswahl						

___/5

ORIENTIERUNG [] Datum [] Monat [] Jahr [] Wochentag [] Ort [] Stadt ___/6

© Z Nasreddine MD Version 7.Nov.2004 deutsche Übersetzung: SM Bartusch, SG Zipper Normal ≥ 26 / 30
 www.mocatest.org Untersucher: _____

TOTAL ___/30
 + 1 Punkt wenn ≤ 12 Jahre Ausbildung

Phonemübersicht - Vokale

	Substantive	Verben	Sonstiges
Ab	Abbildung (439) Abschnitt (1271) Abschluss (1326)	abschließen (974) ablehnen (1113) abhängen (1210)	absolut (993) abhängig (1613) abschließend (3533)
Ak	Aktiengesellschaft (1272) Aktivität (1431) Aktionär (2168)	akzeptieren (1453) aktivieren (3263)	aktuell (777) akademisch (2390)
Al	Alternative (1645) Alkohol (2589) Allianz (3264)		allerdings (221) allgemein (339) allmählich (2052)
An	Angebot (584) Angabe (685) Analyse (1134)	anfangen (373) anbieten (434) ansehen (444)	anderer (59) andererseits (840) angesichts (1146)
Ar	Artikel (903) Arbeitsplatz (1184) Arbeitgeber (1749)	arbeiten (200)	arbeitslos (1647) arabisch (1839)

Phonemübersicht - Konsonanten

	Substantive	Verben	Sonstiges
Ba	Backofen (3405) Bakterie (3406) Batterie (3706)	basieren (3065) bauen (686) backen (3164)	bayrisch (1989) bald (475)
Ka	Kapitel (911) Kandidat (1441) Kategorie (2004)		katholisch (1935) kalt (875)
La	Landwirtschaft (1723) Landesregierung (3992) Landschaft (1442)	lagern (3620) lachen (559) landen (1600)	langfristig (1936) langweilig (2921) landwirtschaftlich (3471)
Na	Nachmittag (1426) Nachfrage (1774) Nachfolger (2843)	nachdenken (1303) nachweisen (1688) nachfolgen (3016)	natürlich (127) national (1007) nahezu (2075)
Ra	Rahmenbedingung (3781) Rahmen (460) Rathaus (3118)	rauchen (2432) raten (2369) rauschen (3640)	radikal (2319) radioaktiv (4003) rasch (1025)

A.2. Messergebnisse

Audiogramme

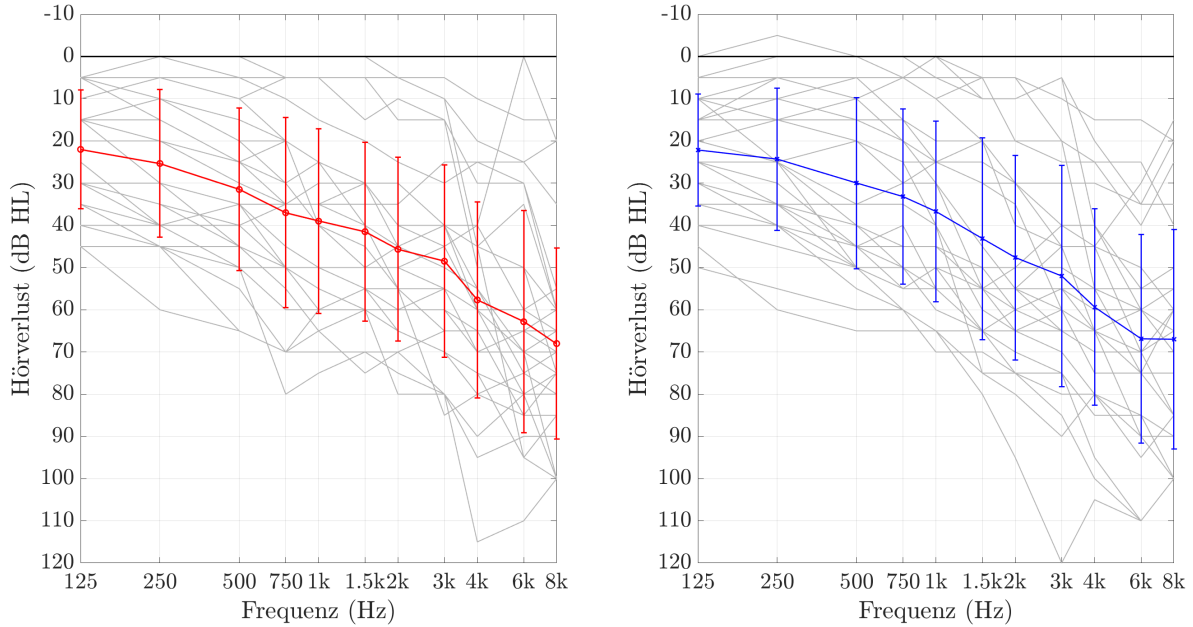


Abb. A.1: Individuelle Tonaudiogramme der 30 älteren Probanden (graue Kurven), sowie der daraus resultierende Mittelwert und die Standardabweichung.

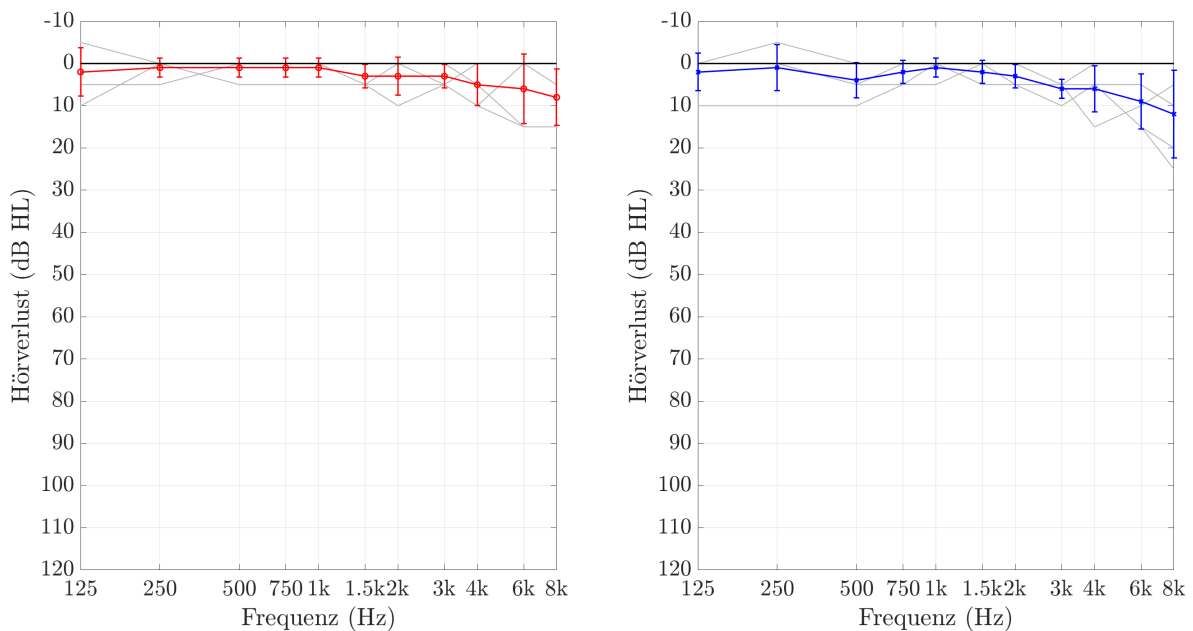


Abb. A.2: Individuelle Tonaudiogramme der fünf jüngeren Probanden (graue Kurven), sowie der daraus resultierende Mittelwert und die Standardabweichung.

Test / Retest

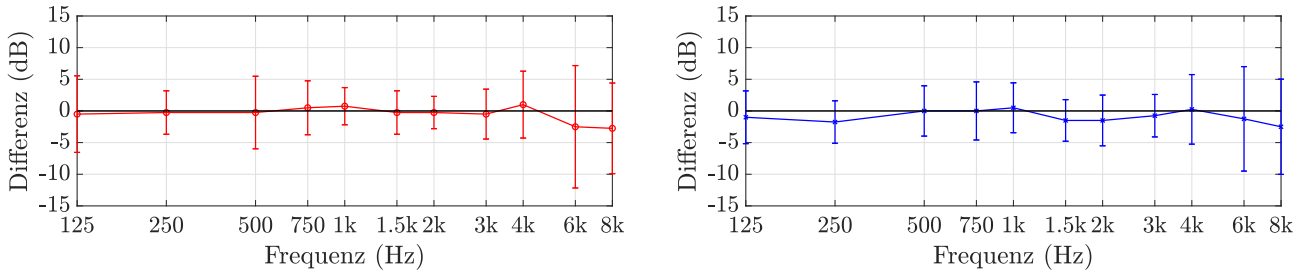
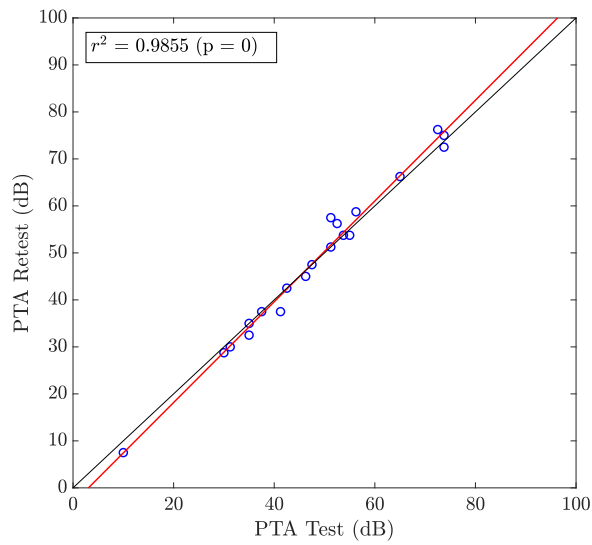
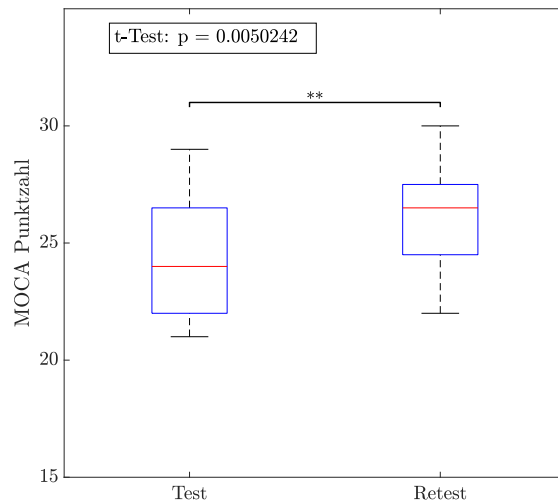


Abb. A.3: Differenz der Tonaudiogramme zwischen Test und Retest.



(a) PTA



(b) MoCA

Abb. A.4: Test und Retest von PTA und MoCA Testergebnissen pro Proband.

Detektionsschwellen in Abhängigkeit vom Zeitsignal

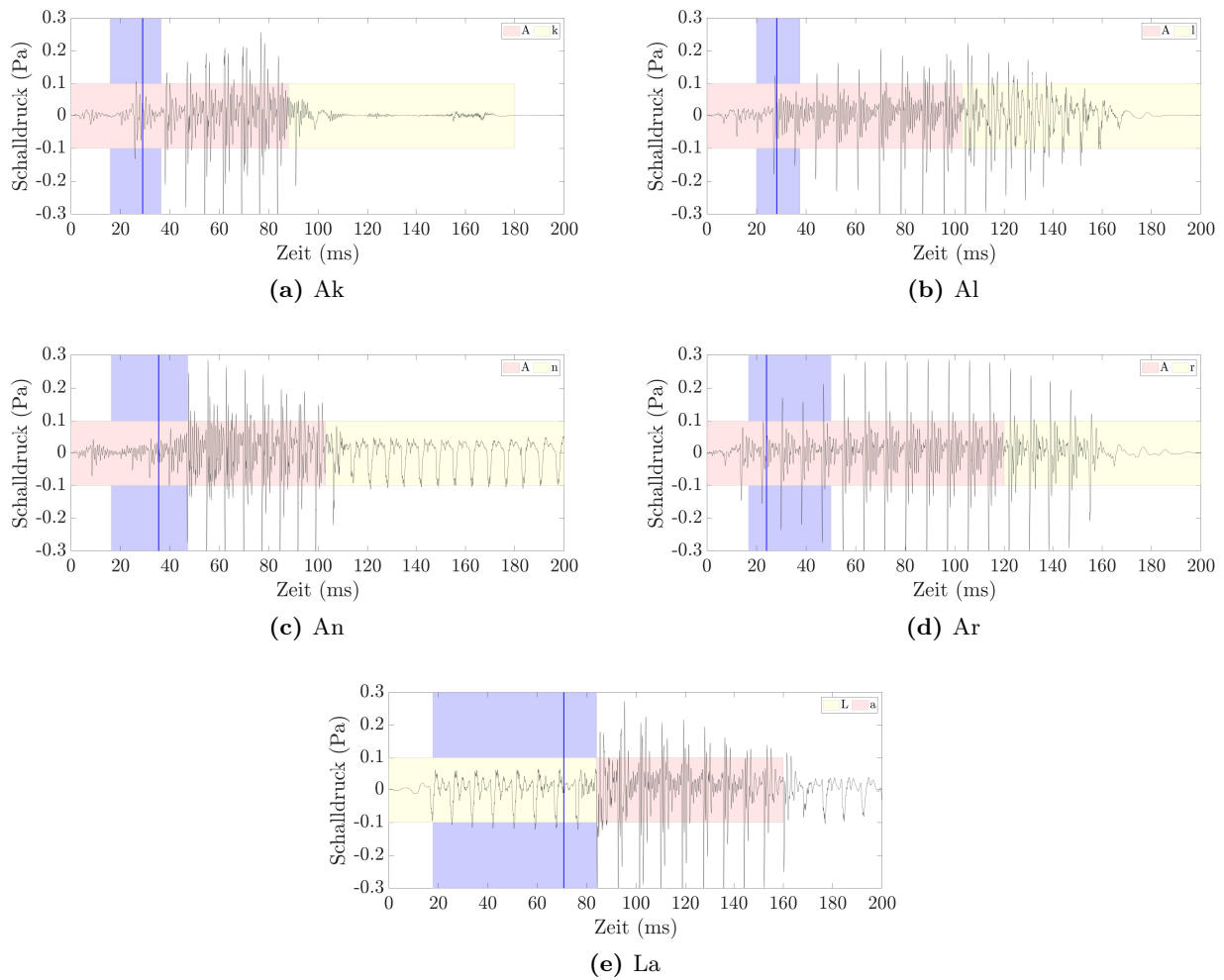


Abb. A.5: Verschiedene Zeitsignale und zugehörige Detektionsschwellen.

Korrelation von PTA und SRT

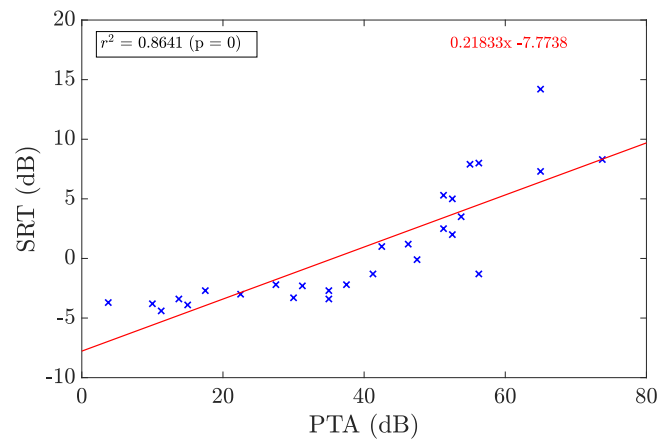


Abb. A.6: Korrelation von PTA und SRT pro Proband.

Schwellen der einzelnen Signale in Abhängigkeit von der SRT

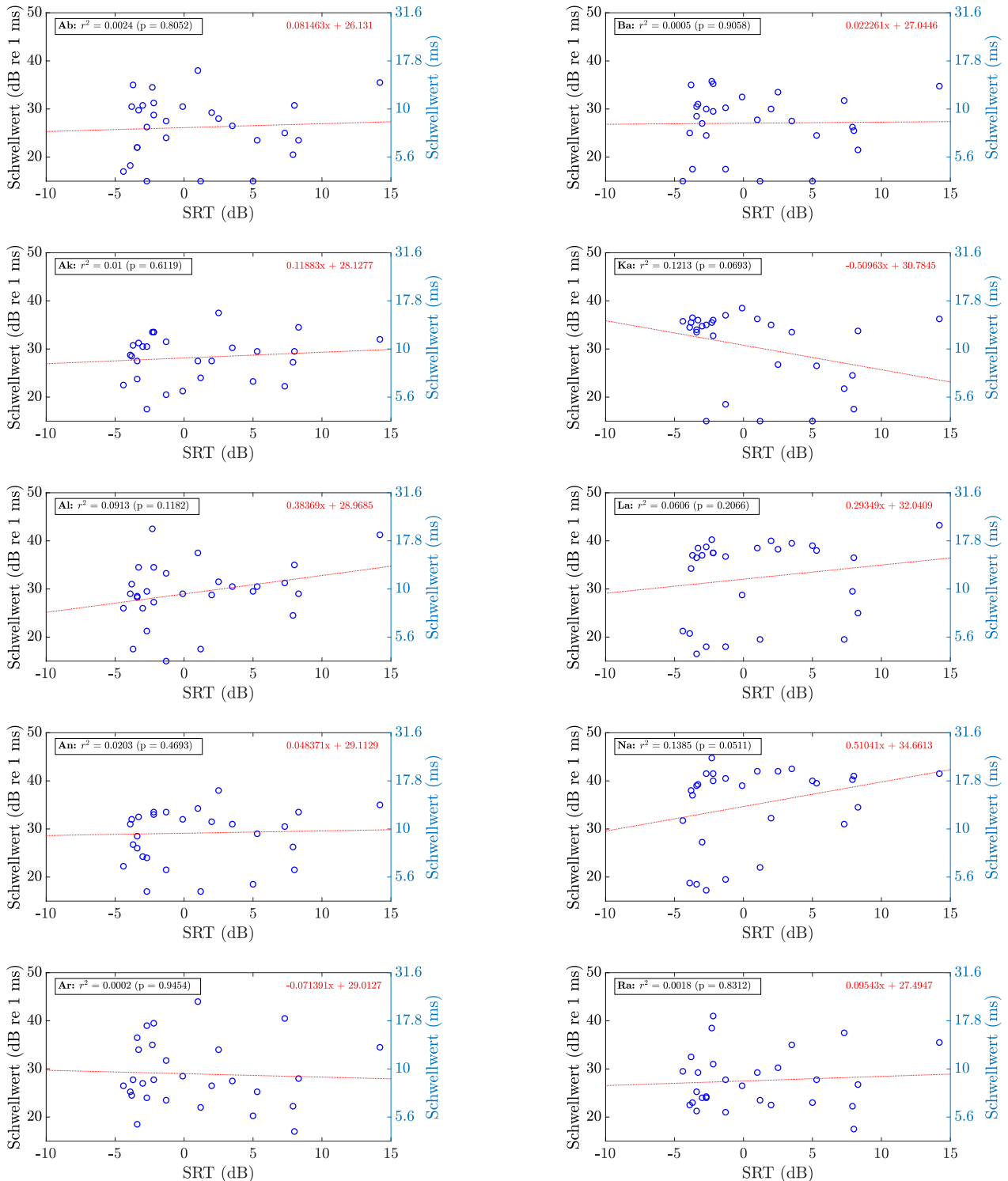


Abb. A.7: Schwellwerte für die einzelnen Signale in Abhängigkeit vom SRT. Die rot gestrichelten Linien besagen, dass keine signifikante Korrelation vorliegt.

Relative Häufigkeit der MoCA Punktzahlen pro Kategorie

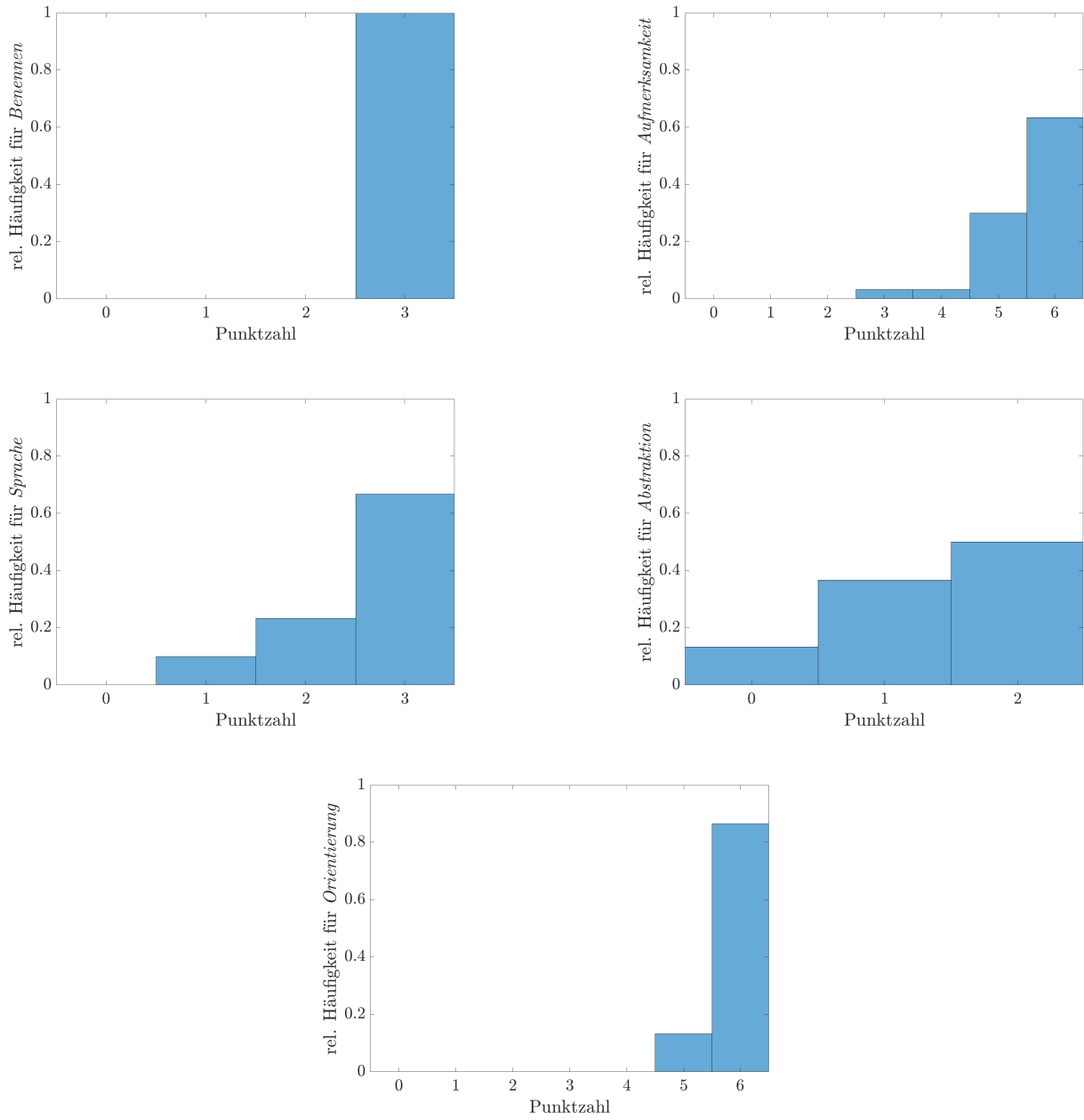


Abb. A.8: Relative Häufigkeit der MoCA Punktzahlen pro Kategorie.