Hochschule Aalen

Bachelorarbeit des Studiengangs

Augenoptik / Augenoptik und Hörakustik

Alexander Thal

VUST - Ein videounterstützter Sprachtest

Prüfer: Prof. Dr. S. Kreikemeier

Zweitprüfer: Prof. Dr. A. Limberger

VUST - Ein videounterstützter Sprachtest

Zugelassene Abschlussarbeit des Studiengangs
Augenoptik/Augenoptik und Hörakustik
zur Erlangung des akademischen Grades Bachelor of Science

vorgelegt von:

Alexander Thal

Tag der Einreichung:
13.04.2016

Fakultät Optik und Mechatronik
Hochschule Aalen

Inhaltsverzeichnis

Zusammenfassung ..... **III**

Liste der Abkürzungen ..... **V**

VUST - Ein videounterstützter Sprachtest **1**

1 Einleitung **1**

2 Grundlagen **3**

2.1 Audiovisuelle Wahrnehmung in der Audiologie 3

2.2 Angewandte Sprachtests 6

2.3 Unterstützung durch computeranimierte Sprachbilder 8

2.4 Höranstrengung bei der Sprachprüfung 9

3 Ziel der Arbeit **10**

4 Versuchsablauf, Material und Methoden **11**

4.1 Verwendeter Sprachtest 11

4.2 Entwicklung des Videomaterials 11

4.3 Implementierung der Videos und Kalibrierung 12

4.4 Fragebogen 13

4.5 Probanden 13

4.6 Versuchsaufbau 15

4.7 Versuchsdurchführung 16

5 Ergebnisse **17**

5.1 Datenanalyse 17

5.2 Vergleichende Darstellung der Ergebnisse 17

5.3 Fragebögen 22

6 Diskussion **24**

6.1 Auswahl des Sprachtests.....24

6.2 Verbesserung durch visuellen Reiz 25

6.3 Anstrengung und Konzentration während der Messung.....28

6.4 Verbesserungen und Ausblick auf weitere Arbeiten.....30

6.5 Schlussfolgerung 31

Erklärung VI

Literaturverzeichnis.....VII

Anhang.....XI

Zusammenfassung

Neben der subjektiven Bewertung der Hörsystemanpassung durch die Kunden und die Verifikation durch die Akustikerin / den Akustiker, in der Regel insitu, ist auch die Validierung einer Hörsystemanpassung nötig bzw. vorgeschrieben. Der in der Hörgeräteakustik verbreitetste Test, um das Sprachverstehen zu testen, ist auch weiterhin der Freiburger Sprachtest. Dieser findet nicht nur Anwendung bei Akustikerinnen und Akustikern, sondern auch bei HNO Ärztinnen und Ärzten. Somit ist er aus der Praxis nicht wegzudenken. Neben den bekannten Problemen der ungleichen Listen (Bangert, 1980; Alich, 1985; Winkler & Holube 2014) und der mangelnden Eignung des Freiburger Sprachtests für Messungen im Störschall (Kießling et al., 2008), haben jedoch alle Sprachtests gemein, dass sie nur akustisch und nicht visuell dargeboten werden (Meister et al., 2011). Viele Schwerhörige, auch HörgeräteträgerInnen, nutzen jedoch im realen Leben neben der reinen akustischen auch die visuelle Information über das Mundbild der Sprecherin / des Sprechers. Diese Zusatzinformation kann die subjektive Höranstrengung deutlich reduzieren (Schreitmüller et al., 2014). Aus diesem Grund wurde ein Videomaterial zum Freiburger Sprachtest, sowohl für Einsilber als auch für Mehrsilber, entwickelt. Hierbei handelt es sich um einen realen Sprecher, welcher die Ein- und Mehrsilber des Freiburger Sprachtests spricht.

Die Tonspur des Videos wurde durch den Originalton des Freiburger Sprachmaterials ersetzt und mit den Artikulationsbewegungen des Sprechers exakt synchronisiert. So entstand eine natürliche Aussprache für die Zahlen und Wörter mit Videomundbild und originalem Ton. Das Video

wurde dem Probanden auf einem Kundenbildschirm präsentiert. Der Test wurde im ACAM 5 Messsystem implementiert und konnte auf gleiche Weise wie der Freiburger Sprachtest ohne Videomaterial gesteuert werden.

In dieser Studie wurde die Sprachverständlichkeit (SV) im Freiburger Sprachtest im Freifeld mit und ohne Videomaterial bei insgesamt 22 Personen (16 Probanden mit Hörverlust [67 ± 7,6 J.] und 6 Normalhörenden [27 ± 3,5 J.]) untersucht. Alle Teilnehmer verbesserten ihre SV durch das zusätzliche Videobild. Signifikante Verbesserungen der SV konnten bei den Hörbeeinträchtigten (ohne HG) bei leisem (50 dB, 22 % höhere SV) und mittlerem (65 dB, 20 % höhere SV) Pegel festgestellt werden, jedoch nicht bei lautem (80 dB, 10 % höhere SV) Pegel. Trugen die Probanden ihre Hörsysteme zeigte sich nur bei leisem Pegel ein signifikanter Unterschied (25 % höhere SV). Normalhörende erreichten nur bei sehr leisem Pegel (35 dB) eine signifikante Verbesserung von 20 %. Insgesamt profitierten Hörbeeinträchtigte signifikant mehr als Normalhörende vom Videobild.

Sowohl die Messergebnisse als auch die subjektiven Bewertungen der Probanden mit Hörverlust, ermittelt durch einen Fragebogen, zeigen Vorteile der zusätzlichen Videodarbietung. Durch das Video des Sprechers kann eine nahezu realistische Alltagssituation im Labor erreicht werden und bietet für den Kunden einen Bezug zum Test. Das Video eliminiert zudem die Problematik, dass Kunden häufig von der Darbietung der Worte überrascht werden, da das Video bereits eine Sekunde vor der Sprachwiedergabe sichtbar wird.

Liste der Abkürzungen

Abb.	Abbildung
ANOVA	Analysis of varianz
CI	Cochlear-Implantat
ER	Erkennungsrate
HG	Hörgerät
HV	Hörverlust
NH	Normalhörende
RMS	root mean square
SNR	signal noise ratio
SV	Sprachverständlichkeit
v.	Version

VUST - Ein videounterstützter Sprachtest

1 Einleitung

Der Mensch nimmt seine Umwelt mit unterschiedlichen Sinnen war. Dazu gehören Auge und Ohr inklusive Gleichgewichtsorgan, Nase und Mund sowie die Haut. Die größte Quelle für Informationen ist das Auge. Licht und Schatten, Stillstand und Bewegung, Formen und Farben werden erkannt. Das Ohr bietet durch die zusätzliche akustische Information Hilfe bei der Orientierung, beurteilt die eigene räumliche Lage durch das Gleichgewichtsorgan, ermöglicht zusammen mit dem Sprechorgan das verbale Kommunizieren und verursacht Emotionen bei Musik.
Durch Nase, Mund und Haut schließlich erfahren wir jede Menge weitere Eindrücke aus unserer Umwelt und komplettieren unsere Wahrnehmung.
Jedes Sinnesorgan ist durch bestimmte Eigenschaften und Fähigkeiten charakterisiert.
Unser Umfeld wird dabei nie durch nur einen einzelnen Sinn erlebt. Zu jeder Zeit fließen unzählige Informationen aus allen Bereichen ein und geben uns Rückantwort über unser Befinden in Raum und Zeit. Hierbei treten jedoch diejenige Sinnesempfindung in den Vordergrund, durch welche die meiste Information kommt. Sehen wir uns im Kino zum Beispiel einen Film an, arbeiten hauptsächlich Auge und Ohr. Hier spielen Bilder sowie die gesprochene Sprache, also verbale Kommunikation, die bedeutendste Rolle.
Auch im Zeitalter von Smartphone und geschriebenen Kurznachrichten ist die allgemeine verbale Kommunikation und somit die Verständlichkeit von Sprache eine zentrale und grundlegende Fähigkeit. Durch viele verschiedene Einflüsse kann diese Fähigkeit des Sprachverstehens eingeschränkt sein. Diese Einflüsse können plötzliche, von der Außenwelt kommende Ereignisse sein, oder eine krankhafte Veränderung des Hörsinns. In den meisten Fällen jedoch ist das Altern des Hörorgans für eine verminderte Wahrnehmung und Verständlichkeit von Sprache der Grund. Haben wir Blickkontakt zu dem Sprecher, sei es eine reale Person oder aber auch ein Sprecher in einem Film, können wir uns Hilfe beim Verstehen durch das

Mundbild holen und somit das nicht Gehörte durch das Gesehene kompensieren. Wir konzentrieren uns auf die Aussprache und das Mundbild.
Mit zunehmendem Hörverlust gewinnt das Mundbild immer mehr an Bedeutung. Oft erst wenn der Spagat zwischen „nicht hören“ bzw. „schlecht verstehen“ und „kompensieren durch Mundbild“ zu groß oder zu belastend wird, entschiedet sich der Betroffenen zu einem Hörgerät.
In der Anpassung von Hörsystemen ist das Ziel einer jeden Versorgung laut Hilfsmittelrichtlinie der Ausgleich des Hörverlustes sowie vor allem das Verbessern des Sprachverstehens. Hierbei muss in den meisten Fällen der Kunde somit zu jenem Zeitpunkt versorgt werden, bei welchem er seinen persönlichen Spagat nicht mehr schafft. Der Prozess des Angewöhnens an eine neue Hörsituation bis eine optimale Sprachverständlichkeit erreicht wird, kann oft ein sehr langer sein.
Es gilt also dem Betroffenen diesen Weg einfach und effizient zu gestalten, damit dieser so schnell wie möglich wieder eine für ihn angenehme Hörempfindung erlebt. Dazu gehört es auch, die nötigen Messungen des Gehörs oder des vorhandenen Sprachverstehens so realistisch wie möglich zu gestallten, um eine exakte Analyse und Einschätzung des Betroffenen zu erhalten und ihm so eine angemessene und angepasste Versorgung gewährleisten zu können.

2 Grundlagen

2.1 Audiovisuelle Wahrnehmung in der Audiologie

Doch auch ohne eine Hörminderung liefert das Auge dem Ohr Unterstützung bei dem Verstehen von Sprache. In großen Gesellschaften beispielsweise hilft der Blickkontakt zur sprechenden Person, sich auf sie zu konzentrieren und gezielt zu hören. Dieses Phänomen der bewussten Selektion wird in der Literatur als Cocktailparty-Effekt (Cherry, 1953, McDermott, 2009) beschrieben. Neuere Untersuchungen zu diesem Phänomen stellten Golumbic et al. (2013) mit Videobild an. Sie untersuchten die Sprachverständlichkeit bei Normalhörenden zum einen mit einem Sprecher und zum anderen mit einem zweiten, konkurrierenden Sprecher jeweils mit und ohne Videodarbietung. Auch hier zeigte sich, dass ein zusätzlicher visueller Reiz hilfreich beim Verstehen und Selektieren von Sprache ist, besonders bei mehreren Sprechern.

Schon seit den 50er Jahren und bis heute also wird der positive Einfluss des Mundbildes und der Lippenbewegungen auf die Verständlichkeit von Sprache untersucht.

Durch viele Untersuchungen ist bekannt, dass die Verknüpfung von visuellem und auditivem Reiz bereits im sogenannten colliculus superior geschieht (Stein und Meredith, 1994). Dieser Teil des Gehirns, auch Vierhügelplatte genannt, befindet sich im Mittelhirn und bildet das akustische und optische Reflexzentrum (Huch und Jürgens, 2015). Damit liegt es, von der Nervenlaufbahn vom Auge in Richtung Gehirn aus gesehen, zwischen Sehnervenkreuzung und Sehzentrum. Akustische Signale werden getrennt über den Thalamus sowohl an das Mittelhirn zum Reflexzentrum wie auch in das Hörzentrum gleitet. Dies bedeutet im Umkehrschluss, dass die Verknüpfung von auditiven und visuellen Wahrnehmung zum Beispiel noch vor der eigentlichen Objektwahrnehmung oder Sprachbildung stattfindet. Abbildung 1 zeigt die einzelnen Funktionszentren im Hirnstamm von der Seitenansicht. Die Augen befinden sich hier links. Farblich hervorgehoben sind die für die audio-visuelle Verknüpfung wichtigen Komponenten wie die Sehnervenkreuzung, das Mittelhirn, das Zwischenhirn, in dem sich der Thalamus befindet, sowie die Vierhügelplatte.

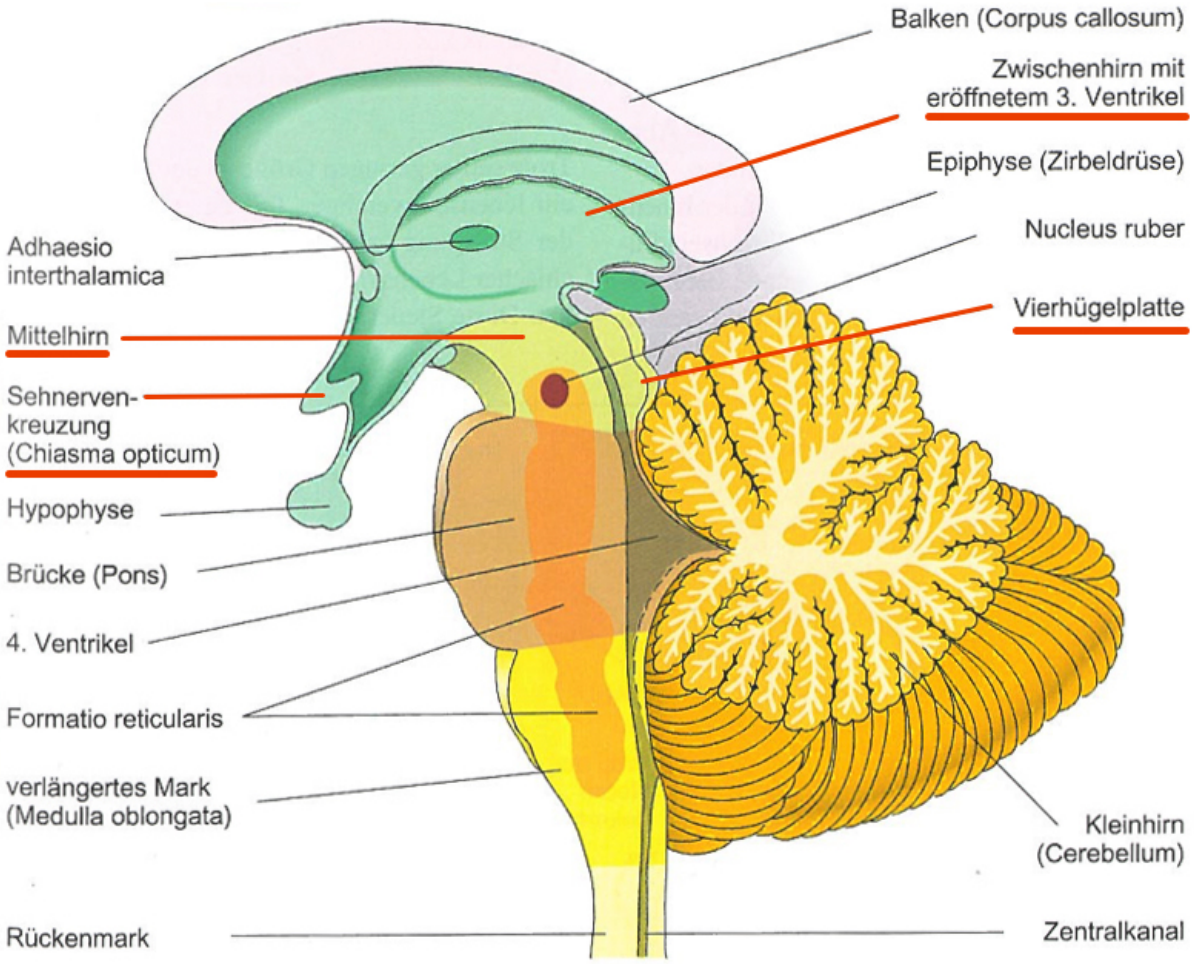

Abb. 1: Funktionszentren im Hirnstamm (Huch und Jürgens, 2015), eigene Bearbeitung

Dabei kann man bei den unterschiedlichen Nervenfasern verschiedene Übertragungsgeschwindigkeiten beobachten. So sind zum Beispiel die Reaktionszeiten nach einem akustischen Reiz kürzer als nach einem optischen (Andreassi und Greco, 1975). Ein kombinierter Reiz von Bild (Lichtblitz) und Ton führt zu einer noch schnelleren Reaktion. Außerdem ruft gleichzeitige audio-visuelle Wahrnehmung wesentlich höhere Amplituden bei evozierten Potenzialen hervor.

Schnellere Reaktionen, wie beispielsweise Ausgleichsbewegungen für den Erhalt des Gleichgewichts, können unter anderem auch durch Nervenfasern ermöglicht werden, welche gemeinsame Informationen aus verschiedenen Sinnesorganen tragen. So liefern einige Fasern bereits zusammengefasste Auskünfte aus den visuellen, okolumotorischen und vestibularen Informationen (Stein und Meredith, 1994).

Eine verbundene Wahrnehmung von Gesehenem und Gehörtem wird deutlich, wenn die Aktivität der verschiedenen Areale im Gehirn beobachtet wird. Sehen wir zusätzlich zu einem akustischen Signal ein bewegtes Mundbild, so wird die Aktivität im auditiven Kortex modifiziert (Sams et al., 1991). Die auditive Wahrnehmung wird somit verändert.
Auch in Situationen, in denen es durch viele Umgebungsgeräusche schwieriger wird, das Gesprochene zu verstehen, wird der Einfluss des Sehzentrums größer. Bei sinkendem SNR verschiebt sich die Aktivität der Hirnareale weiter nach hinten vom Hör- zum Sehzentrum (Sekiyama et al., 2003). Dies scheint an der gesteigerten visuellen Aufmerksamkeit oder an einer größeren Interaktion zwischen Sehen und Hören zu liegen.

Die Funktion des Wahrnehmungsapparates eines Jeden ist somit ein wichtiger Faktor bei dem Verstehen von Sprache. Die Sprachverständlichkeit hängt eng mit der Fähigkeit zusammen, Gehörtes und Gesehenes richtig zu unterscheiden bzw. zu kombinieren und erfordert einen intakte Neurophysiologie.
Offensichtlicher als der Einfluss der Verarbeitungsfähigkeit ist dagegen wahrscheinlich das Vorhandensein eines Hörverlustes und der Art der Hörminderung. Die Sprachanteile, welche wichtig für das Verstehen sind, befinden sich in einem Frequenzbereich von 250 Hz bis 3000 Hz (Kompis, 2013). Dabei liegen die wichtigeren Sprachanteile im höheren Frequenzbereich. Diese Tatsche betrifft vor allem Menschen mit Hörbeeinträchtigungen durch das Alter, welche sich durch eine symmetrische, hochtonbetonte Innenohrschwerhörigkeit auszeichnet. Dabei nimmt der Hörverlust ab 1000 Hz bis zu 60 dB bei 8000 Hz zu (ISO 7029). Die Betroffenen tun sich somit schwerer, Sprache zu verstehen.
Dem gegenüber stehen jedoch Untersuchungen, die zeigen, dass ein Hochtonverlust sich nicht sehr stark auf eine „face-to-face“ Konversation auswirkt und den alltäglichen Gesprächsfluss nur wenig beeinflusst (Erber, 2003). Die visuelle Komponente scheint das Nichtgehörte somit zu kompensieren.

Neuere Untersuchungen zum Einfluss einer zusätzlichen visuellen Darbietung bei Sprachtests machten unter anderen Rouger et al. (2007) in einer Langzeitstudie über

acht Jahre mit CI-Trägern. Sie entdeckten ein erhöhtes Sprachverstehen bei audio-visueller Präsentation der Testreihen. Hierbei profitierten CI-Träger mehr als normalhörende Probanden vom dargebotenen Mundbild. Als Grund vermuten sie eine Umstrukturierung der für die Spracherkennung verantwortlichen Areale im Gehirn bei den CI-Trägern, die eine größere Mitwirkung der visuellen Areale begünstigt.

Aber auch als Normalhörender ohne Einschränkung des Hörsinns kann man erkennen, wie stark das Auge sich beim Sprachverstehen einmischt. Der McGurk Effekt (McGurk und MacDonald, 1976) zeigt eindrucksvoll die Verknüpfung von gehörtem Laut und dem dazu gesehenen Sprachbild. In ihren Versuchen zeigten sie ein Video, in dem eine Frau die Silbe [ga] wiederholt. Später wurde nur die Tonspur durch die Silbe [ba] ersetzt. Die Probanden gaben nun an, den Laut [da] zu hören. Sie wiederholten diese Versuche mit weiteren sinnlosen Lautverbindungen und beschrieben jeweils das selbe Phänomen.
Auch wenn dies zunächst einer optischen Täuschung gleicht, finden diese Silbenverbindungen ohne Sinn, auch Logatome genannt, in der Audiologie ebenfalls Anwendung, etwa zur Überprüfung der Silbenverständlichkeit.

2.2 Angewandte Sprachtests

Neben diesen Testverfahren zur Messung der Silbenverständlichkeit werden im Laufe einer Hörsystemanpassung vor allem Tests zur Prüfung der Sprachverständlichkeit angewendet. So kann der Akustiker feststellen, welchen Nutzen das Gerät für den Kunden im Bezug auf die Verständlichkeit von Sprache liefert. Hierbei stehen dem Prüfer unter anderem mehrere Verfahren zur Auswahl.
So findet der Freiburger Einsilbertest nach Halbrock zum Beispiel im deutschsprachigen Raum die häufigste Anwendung (Ulrich und Hoffmann, 2011). Er wird unter anderen Tests von den Kostenträgern bei der Anpassung von Hörgeräten laut Hilfsmittelrichtilinie verlangt. Der sowohl aus Zahlen und einsilbigen Wörtern bestehende Test ist in den Normen DIN 45621-1 und DIN 45626-1 festgelegt. Durch die Messung mit Mehrsilbern (Zahlen von 13 bis 99) wird der Hörverlust für Sprache

bestimmt, die Prüfung mit Einsilbern ermittelt den Diskriminationsverlust bzw. die Verständlichkeit von Sprache.

Da der Freiburger Sprachtest zahlreiche, anerkannte Kritikpunkte besitzt, entwickelten Hahn und Limberger (2013) den Freiburger Reloaded. Die in dieser Arbeit angegangen Probleme des originalen Freiburgers sind vor allem Einsilber mit zu positiver oder negativer Assoziation, ein veralteter Wortschatz (Alich, 1985) sowie ein ungleicher Schwierigkeitsgrad der Gruppen (Winkler und Holube, 2014). Zudem wurde zwischen den Wortdarbietungen ein Rauschen eingefügt. Dadurch soll die Aufmerksamkeit des Prüfling erhöht werden und das Hörgerät in einen eingeschwungenen Zustand versetzen (Hahn und Limberger, 2013).

Neben den Ein- bzw. Mehrsilbertests gibt es Verfahren zur Überprüfung des Sprachverstehens mittels Sätzen. Der Oldenburger Satztest (Wagener et al., 1999), kurz OlSa, bietet dem Probanden Sätze mit immer gleichem Satzaufbau dar:

Name - Verb - Zahl - Adjektiv - Substantiv

Dabei werden die Sätze für jede Darbietung zufällig aus einem Kollektiv von 50 Wörtern (jeweils zehn pro Satzglied) gewählt. Die dabei entstehenden Sätze ergeben nur eingeschränkt Sinn. Aus einer Reihe von mehreren Sätzen und einem zusätzlichen Rauschen, welches als Störgeräusch dient, wird der SNR für den Probanden durch ein automatisiertes Einpendelungsverfahren ermittelt.

Die meisten Sprach- und Satztests bestehen dabei lediglich aus Tonspuren, welche dem Probanden beziehungsweise Kunden durch einen Kopfhörer oder ein Lautsprechersystem vorgespielt werden. Der optische Reiz wird in diesen Situationen nicht angesprochen.

Einige Testverfahren in der Sprachaudiologie bieten allerdings auch eine optische Unterstützung. So ist zum Beispiel für den Oldenburger Satztest in der geschlossenen Version ein Auswahlfenster der möglichen Wörter gegeben, mit dem der Proband durch das Anklicken der verstandenen Satzglieder per Computermaus eine Antwort geben kann.

Der Phonemtest der Firma Phonak zeigt ebenfalls eine optische Oberfläche für den Kunden auf einem Bildschirm. Der Kunde kann mit der Maus aktiv am Test teilnehmen und seine Antworten ebenfalls durch Anklicken auswählen.

Diese Testverfahren bieten allerdings nur eine erweiterte Benutzeroberfläche für die zu prüfende Person, die damit eine Auswahlmöglichkeit von Antworten bekommt. Keine der bisherigen Messvarianten unterstützt jedoch das Testsignal an sich durch ein zusätzliches Mundbild oder Ähnlichem.

2.3 Unterstützung durch computeranimierte Sprachbilder

Seit den Anfängen der Animation von Gesicht und Kopf eines Menschen in den 70er Jahren, erlebt die computergestützte Animation eine stetige Entwicklung und findet ein immer größeres Anwendungsfeld (Ping et al., 2013). Nicht nur in der Filmindustrie, sondern auch in Bereichen von Computerspielen oder der Kommunikation zwischen Mensch und Maschine, werden animierte Gesichter eingesetzt.
Die Interaktion zwischen Mensch und Maschine wird in der Literatur auch im Zusammenhang mit der Schulbildung genannt und findet so beispielsweise Anwendung als Sprachtherapie für den Unterricht mit hörgeschädigten Kindern (Cole et al., 1999). So beschreiben Cole et al. (1999) ein fiktives Szenario einer hochgradig schwerhörigen Schülerin, die mit Hilfe eines computeranimierten Lehrers ihre Aussprache und Wortbildung lernt und verbessert. Hierbei wählt die Schülerin selbst ihr Lerntempo und kann schwere Lektionen beliebig oft wiederholen.
Kühnel et al. (2010) sprechen von der heutigen Anwendung von animierten Kopfmodellen in der Logopädie oder beim Erlernen neuer Fremdsprachen.
In der Hörakustik und Hörsystemanpassung hat diese Technik zur visuellen Unterstützung von Sprach- oder Satztests jedoch noch keine feste Anwendung gefunden.
Einen ersten Ansatz zur Einbeziehung lieferten Meister et al. (2011). Sie untersuchten die Sprachverständlichkeit mit einem in der Hörsystemanpassung gebräuchlichen Satztest bei Normalhörenden sowie hörgeschädigten CI-Trägern und entdeckten ein erhöhtes Sprachverstehen bei audio-visueller Präsentation der Testreihen. Sie nutzten dabei ein bereits 2004 entwickeltes Kopfmodell, welches Daten der natürlichen Artikulation für die Erstellung des Mundbildes verwendet (Fagel und Clemens, 2004).

Weitere Untersuchungen mit dem selben Kopfmodell wurden später zur audio-visuellen Sprachwahrnehmung von Konsonanten in Störgeräusch durchgeführt, welche zu ähnlichen Ergebnissen kamen (Lipski et al., 2013).
Neueste Untersuchungen dazu zeigten neben einer verbesserten Sprachwahrnehmung durch ein zusätzliches Mundbild zudem, dass jüngere Probanden deutlich mehr Wörter alleine nur durch das Mundbild absehen konnten als ältere (Schreitmüller et al., 2015).

2.4 Höranstrengung bei der Sprachprüfung

Ein weiterer Begriff, der oft und immer häufiger in Verbindung mit dem Verstehen von Sprache und somit auch dem Testen des Sprachverstehens gebracht wird, ist die Höranstrengung. Sie wird unter anderem beschrieben durch die kognitiven Anforderungen beim Verstehen von Sprache (Hicks und Tharpe, 2002). Wie schwerfällig und ermüdend ist also Zuhören und Verstehen? Hier liegt die Vermutung nahe, dass Menschen mit einem Hörverlust eine wesentlich höhere Anstrengung aufbringen müssen und somit schneller ermüden. Der Begriff der Höranstrengung ist an dieser Stelle jedoch zu unterscheiden von der Hörermüdung. Hier findet eine Gewöhnung an ein Tonsignal statt, welches eine Verschiebung der individuellen Hörschwelle mit sich führt (Lehnhardt und Laszig, 2009). Dies ist bei der Höranstrengung nicht der Fall.
Neben verschiedenen audiologischen Messverfahren kommen zur Beurteilung der Höranstrengung auch Fragebögen zum Einsatz.

3 Ziel der Arbeit

Trotz der Tatsache, dass ein Mundbild die Sprachverständlichkeit verbessert, haben bewegte Bilder in der Anpassung von Hörgeräten noch keine feste Anwendung gefunden.

Diese Arbeit soll einen Ansatz liefern, die Prüfung des Sprachverstehens in der Hörsystemanpassung mit einem visuellen Reiz zu ergänzen. Dies beinhaltet die Entwicklung eines Videomaterials, welches in einen anerkannten Sprachtest eingebunden wird. Des Weiteren soll getestet werden, inwieweit das Mundbild das Ergebnis einer Sprachmessung bei Hörbeeinträchtigten und Normalhörenden beeinflusst. Durch einen kurzen Fragebogen wird ebenfalls die Komponente Höranstrengung mit und ohne Videobild erfragt.

Dabei werden im Wesentlichen zwei Hypothesen gesetzt:

1) Das Mundbild verbessert die Sprachverständlichkeit des Probanden. Hierbei profitieren Hörgeschädigte mehr als Normalhörende. Gerade im Bereich der Hörschwelle liefert das Mundbild eine merkliche Unterstützung. Dabei verfälscht das Video das Ergebnis nicht, sondern hebt es auf ein verbessertes Niveau. Die Verständlichkeit von Sprache ist rein visuell ohne Ton geringer als das auditive Sprachverstehen.
2) Durch ein Video ist der Zuhörer aufmerksamer und somit motivierter ein gutes Ergebnis zu erzielen. Er bekommt den Eindruck einer faireren Messbedingung und sieht so dem Ergebnis positiver entgegen. Dies wirkt sich positiv auf die Höranstrengung aus. Der Proband ist wacher bei der Messung und verfolgt die Anpassung seiner Hörsysteme mit größerem Interesse.

Des Weiteren besteht der Gedanke, die Hörgeräteversorgung in der Messkabine näher an die Alltagssituation anzugleichen und den klinischen, rein testbasierten Charakter einer Sprachmessung zu minimieren. Durch das Einbringen eines realen „Sprechpartners“ erhält der Kunde einen Bezug zum Testverfahren.

4 Versuchsablauf, Material und Methoden

4.1 Verwendeter Sprachtest

Für die Umsetzung eines videounterstützten Sprachtests wurde der Freiburger Sprachtest nach Halbrock von 1953 mit der Neuaufnahme von 1969 verwendet. Der Test ist standardisiert durch die Normen DIN 45621-1 und DIN 45626-1 und wird von den Kostenträgern bei der Anpassung von Hörgeräten laut Hilfsmittelrichtlinie verlangt.

Die Wortlisten des Tests wurden nach Bangert (1980) und Alich (1985) aussortiert. Sie beschreiben Listen, die mit besonders gutem bzw. schlechtem Ergebnis vom Mittelwert abweichen. Somit wurden nur die Listen 6, 7, 8, 10, 11, 17, 18, 19 zu Messungen verwendet.

4.2 Entwicklung des Videomaterials

Das Video wurde mit Hilfe des Medienzentrums der Hochschule Aalen erstellt. Es verfügt über ein professionell ausgestattetes Studio und bietet somit eine korrekte Ausleuchtung des Sprechers und stellt eine angemessene Aufnahme von Ton und Bild sicher.

Als Sprecher wurde ein Mitarbeiter der Hochschule Aalen gewählt, der keine Verbindung zur Hörakustik hat. Vor der Aufnahme hatte er die Möglichkeit die Audiodateien des Freiburger Tests anzuhören und sich die Aussprache einzuprägen, um ein möglichst gleiches Mundbild zu bekommen.

Zur Aufnahme wurde ein Panasonic AJ-HPX3100 P2HD Camcorder verwendet. Für die spätere Synchronisierung und Bearbeitung der Audiospur wurde ebenfalls der Ton des Sprechers mit einem Sennheiser MKH 416 P48 Kondensator-Mikrofon aufgezeichnet. Das Material wurde mit Hilfe des Programs Adobe Premier Pro (CS6, v. 12.0, Adobe Systems Incorporated) aufgenommen und gespeichert. Auch zur späteren Bearbeitung der Videosequenzen wurde dieses Program herangezogen.

Aufgezeichnet wurden insgesamt 373 Wörter und die zehn Zahlengruppen des Freiburgers. Jedes Wort und jede Zahl wurde einzeln aus dem Videomaterial getrennt und die entsprechende Audiodatei unterlegt. Vor und nach der Tonspur wurde jeweils eine Sekunde (entspricht vier Video-Frames á 0.25 Sekunden)

eingefügt, um einen Vor- und Nachlauf zu erhalten. Nach der Bearbeitung der Videos wurden diese im 4:3 Format als H264 Datei exportiert. Die Auflösung beträgt nach dem Export 784 x 576 Pixel mit einer Größe zwischen 700 bis 900 KB. Abbildung 2 zeigt ein vom Bildschirm aufgenommenes Standbild des fertigen Videomaterials.

Abb. 2: Bildschirmfotos des Videomaterials

4.3 Implementierung der Videos und Kalibrierung

Als Messbox für die Testdurchführungen wurde die ACAM 5 (v. 5.14.0, Acousticon Hörsysteme GmbH) verwendet. Zur einfacheren Einbindung in den Sprachtest wurden die Videodateien in das Videoformat .avi umgewandelt. Vor der Einfügung wurden die Videodateien und original Tonspuren auf ihre Lautstärke verglichen. Dazu wurde das kostenfreie Programm Audacity (v. 2.1.0 für Mac OSX) verwendet. Mit Hilfe des Plug-Ins FFmpeg (v. 2.2.2 für Mac OSX) können die Audiospuren einer Videodatei innerhalb von Audacity bearbeiten werden (https://www.ffmpeg.org). Mit einem weiteren Hilfsprogramm WaveStats (v. 0.3 für Mac OSX) für Audacity werden einige Funktionen zur Analyse einer Tonspur ermöglicht.

Der Vergleich der beiden Tonspuren wurde mittels RMS durchgeführt. Hierbei ergab sich ein konstanter Unterschied von 3.1 dB, um die die Videosequenzen leiser waren. Dies wurde bei der Installation berücksichtigt und die Lautstärke der Testeinheit um 3.1 dB erhöht, um eine korrekte Kalibrierung zu erhalten.
Die Installation und Implementierung in die ACAM wurde von der Firma Acousticon Hörsysteme GmbH selbst übernommen.
Des Weiteren wurde bei ersten Versuchen die richtige Kalibrierung durch vergleichende Messungen mit einem Schallpegelmesser in einem Meter Abstand sicher gestellt.

4.4 Fragebogen

Um zu sehen, inwieweit das Videobild sich auf die Höranstrengung auswirkt, wurde ein kurzer Fragebogen mit insgesamt drei Fragen entwickelt (s. Anhang 1). Die Fragen behandelten somit im Wesentlichen den Aspekt der Anstrengung und geforderten Konzentration. Hier musste zunächst beantwortet werden, ob ein Unterschied zwischen beiden Varianten der Prüfung (mit oder ohne Video) bemerkt wurde. Der Proband konnte zudem auf einem Zahlenstrahl von 1 bis 10 angeben, wie hoch er die jeweilige Anstrengung beurteilt. Dabei bedeutete 1 eine geringe, 10 eine sehr hohe Konzentration. Durch eine sehr feine Einteilung der Skala konnten Werte mit einer Dezimalstelle angestrichen werden. Ein weiterer Aspekt, der mit dem Fragebogen verfolgt werden sollte, war die Frage nach dem subjektiv empfundenen Ergebnis. Bei welcher Variante hatte der Proband also ein besseres Gefühl im Bezug auf das Resultat. Welcher Test, ob mit oder ohne Video, für weitere Messungen in der Zukunft vorgezogen würde, konnte der Proband in der letzten Frage beantworten. Nach dem Frageteil hatte der Teilnehmer Platz für persönliche Anmerkungen.

4.5 Probanden

Insgesamt nahmen 22 Probanden teil. Sechs davon normalhörend (2 weiblich, 4 männlich) und 16 mit einem HV (3 weiblich, 13 männlich). Abbildung 3 zeigt den mittleren HV der linken und rechten Seite inklusive Standardabweichung getrennt

sowie den gesamten Bereich, in dem Hörverluste gemessen wurden (weiblich und männlich zusammen).

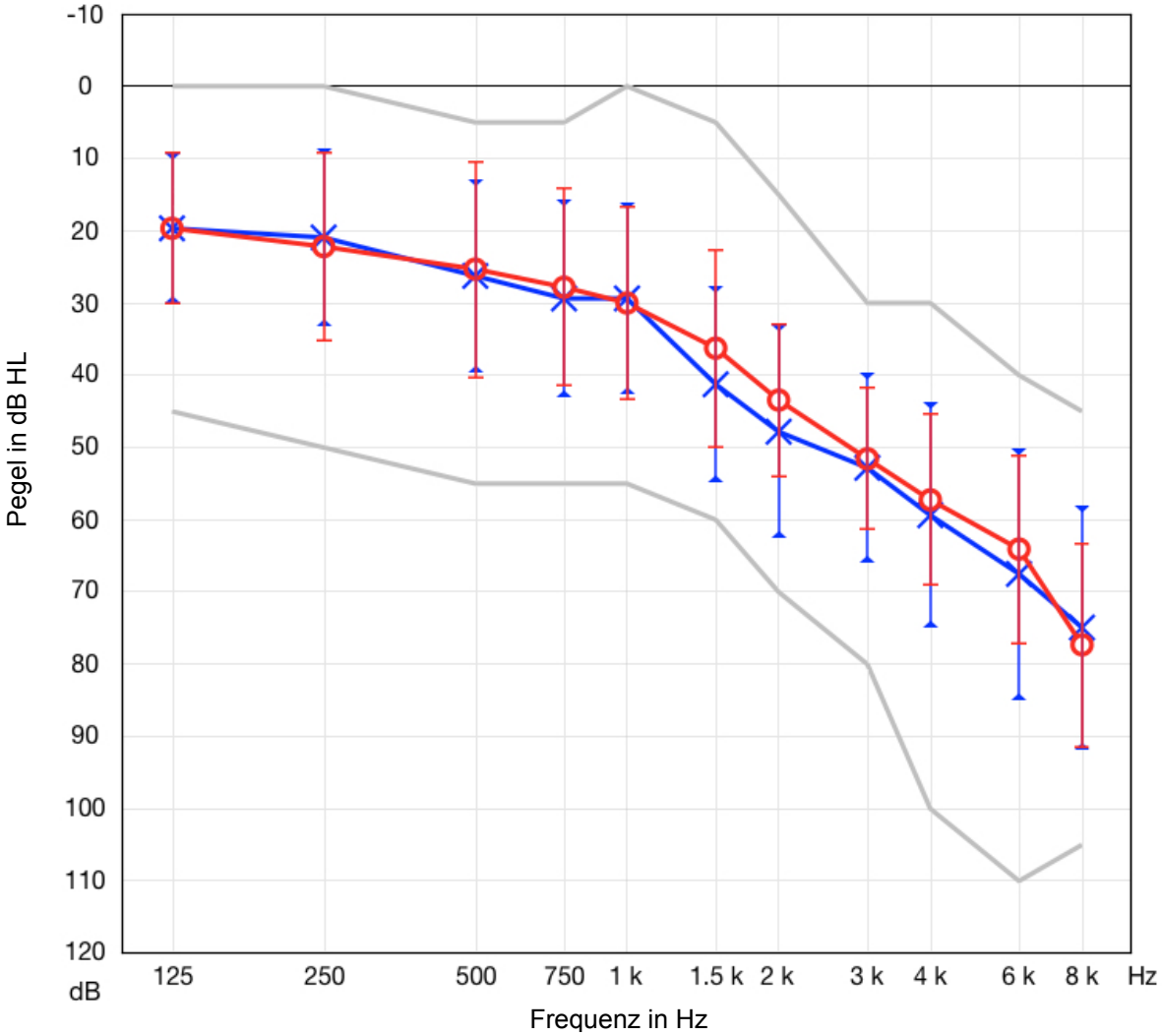

Abb. 3: Mittlerer Hörverlust mit Standardabweichung sowie gesamter Bereich der Hörverluste, n = 16
(— rechts, — links, — gesamter Bereich)

Das durchschnittliche Alter der Normalhörenden betrug 27 ± 3,5 Jahre, das der hörbeeinträchtigten Probanden 67 ± 7,6 Jahre. Im Mittel verwendeten die Hörgeschädigten ihre Hörsysteme seit 4,6 Jahren. Der eigentliche Beginn des Hörverlusts war im Durchschnitt vor elf Jahren.

4.6 Versuchsaufbau

Für die Sprachmessungen mit Video wurde ein Lautsprecher vor einen Bildschirm für den Kunden gestellt, sodass eine vollständige Einsicht des gesamten Bildschirms gegeben war. Der Bildschirm besitzt eine Auflösung von 1680 x 1050 Pixeln mit einer Bildschirmaktualisierungsrate von 60 Hz. Der Prüfraum sowie die Sprachprüfungen wurden gemäß der Normung ISO 8253-3 aufgebaut bzw. durchgeführt. Abbildung 4 zeigt den Versuchsaufbau aus Sicht des Probanden.

Abb. 4: Versuchsaufbau aus Sicht des Probanden

4.7 Versuchsdurchführung

Die Probanden wurden jeweils an zwei Terminen getestet. An einem Termin sollten alle Messungen ohne Video durchgeführt werden, an dem anderen sämtliche Messungen mit Video. Dabei wurde immer randomisiert, ob der erste Termin mit oder ohne Video-Messung begonnen wurde. Die Probanden mit Hörverlust führten die Messungen jeweils mit und ohne Hörgerät durch. Auch hier wurde randomisiert, ob der Proband innerhalb eines Termins zunächst mit Hörgerät oder ohne Hörgerät den Test begann. Dadurch konnte sicher gestellt werden, dass ein eventuell entstehender Lerneffekt ausgeschlossen wird.

Vor Beginn wurde notiert, wann der Hörverlust bei dem Probanden eingesetzt hat und ab welchem Zeitpunkt eine Versorgung mit Hörgeräten stattfand.

Für jeden Probanden wurde mit vorheriger Otoskopie ein Hörtest durchgeführt, sofern das aktuellste Audiogram älter als sechs Monate war. Die Messungen erfolgten mit der ACAM 5 (v. 5.14.0, Acousticon Hörsysteme GmbH) und dem Kopfhörer DT 48 der Firma Beyerdynamics. Getestet wurde die Luftleitung rechts und links in den Frequenzen von 125 Hz bis 8000 Hz.

Für die Sprachmessungen mit und ohne Video wurden die Lautstärke, bei der 50 % der Wörter verstanden wurden, sowie die Verständlichkeit bei den Lautstärken 50 dB, 65 dB und 80 dB notiert. Bei normalhörenden Probanden wurde die Diskrimination anstatt bei 80 dB bei 35 dB gemessen. Um zu erkennen, wie groß das Verstehen von Ein- und Mehrsilber durch den rein visuellen Reiz ist, wurde zusätzlich jeweils Zahlen und Wörter eine Testreihe mit 0 dB dargeboten. Alle Messungen erfolgten im Freifeld.

Nach den Messungen am zweiten Termin sollten die Probanden den Fragebogen selbstständig ausfüllen. Bei Verständnisfragen wurde geholfen.

5 Ergebnisse

5.1 Datenanalyse

Die erste Analyse und Einsichtung der Daten erfolgte mit Excel 2011 (Microsoft Office) für Apple Macintosh. Für eine weiterführende Statistik wurde R verwendet (RStudio für Mac OSX, v. 0.99.879, © 2009-2016 RStudio, Inc.).

Zur weiteren statistischen Auswertung und der Überprüfung von Signifikanzen innerhalb der einzelnen Messungen mit und ohne Video wurde eine mehrfaktorielle Varianzprüfung für abhängige Messreihen durchgeführt (ANOVA). Zur Einsicht der Signifikanzen innerhalb der einzelnen Lautstärken wurde der TukeyHSD Test ausgeführt. Für die Beurteilung auf signifikante Unterschiede der Skaleneinträge zur Einschätzung der Höranstrengung wurde der t-Test für abhängige Stichproben eingesetzt.

5.2 Vergleichende Darstellung der Ergebnisse

Die Messergebnisse wurden für die Probanden mit Hörbeeinträchtigung jeweils mit und ohne Hörgerät getrennt ausgewertet. Abbildung 5 zeigt vergleichend die erreichte Sprachverständlichkeit mit und ohne Videotest bei Probanden mit Hörverlust mit und ohne Hörgeräte in Form eines Boxplots. Dabei zu sehen sind der Median, 25. und 75. Quartil, kleinster und größter nicht-extremer Wert sowie Ausreißer (gekennzeichnet als Kreis).

Die auf Seite 19 gezeigte Abbildung 6 veranschaulicht die Ergebnisse für die Probanden ohne Hörbeeinträchtigung, ebenfalls als Boxplot.

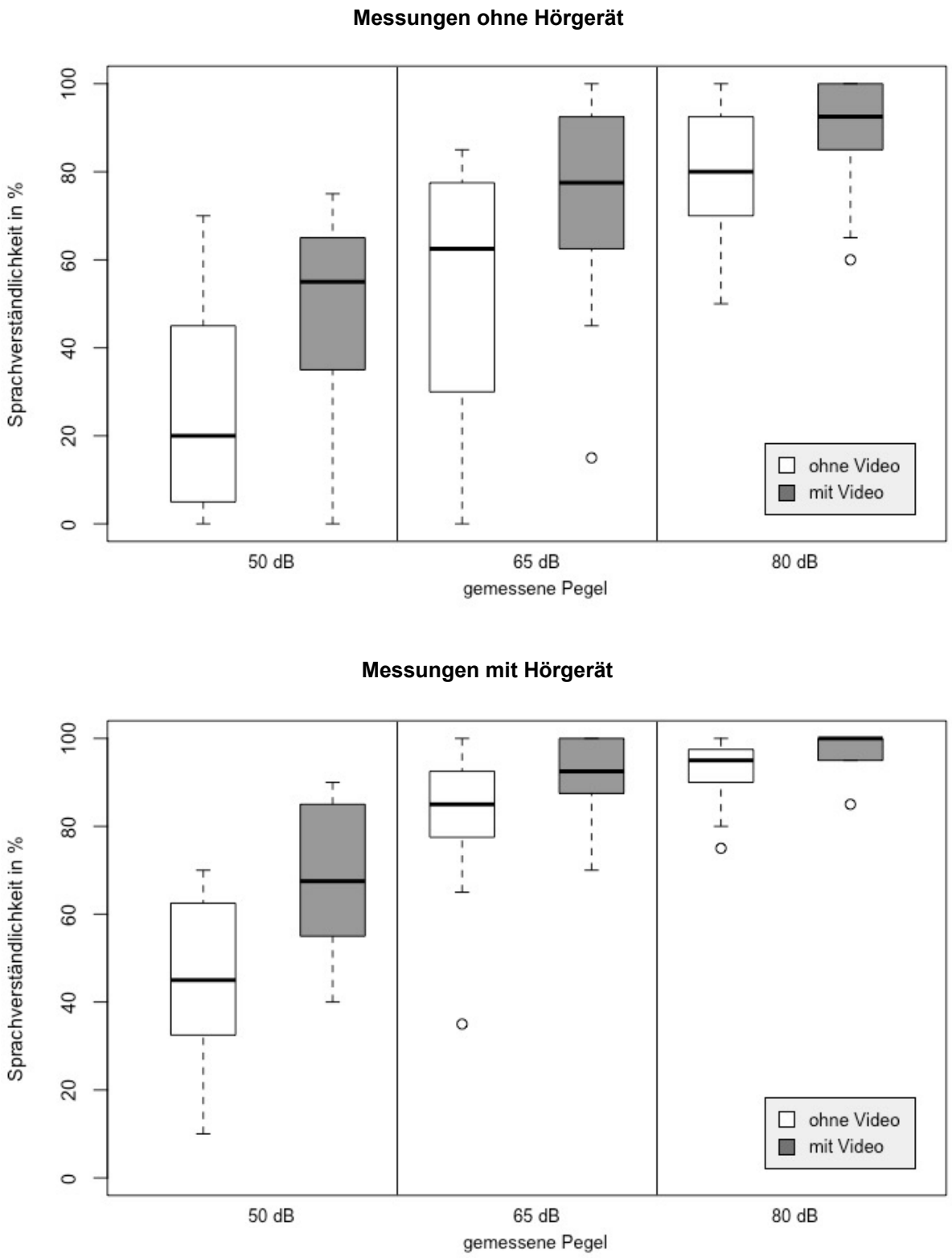

Abb. 5: Sprachverständlichkeiten bei Probanden mit HV (n = 16)

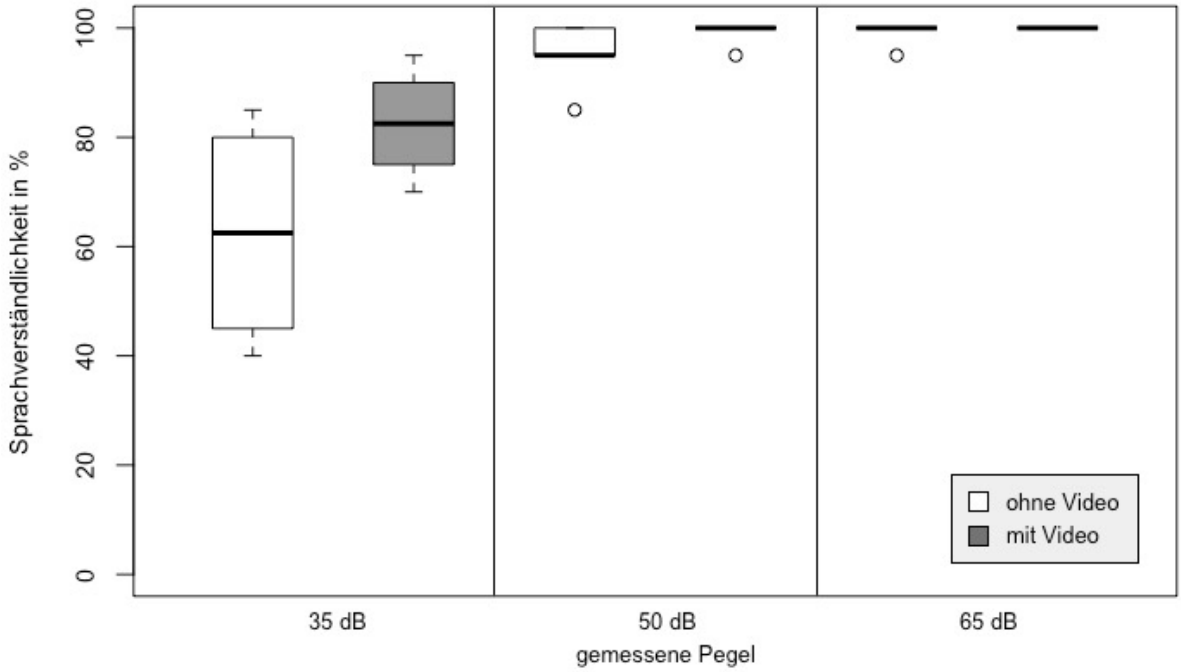

Abb. 6: Sprachverständlichkeiten der normalhörenden Probanden (n = 6)

Alle errechneten, arithmetischen Mittel inklusive Standardabweichung der Sprachverständlichkeit sind in Tabelle 1 zusammengefasst. Die Werte sind auf ganze Zahlen gerundet.

Messung		SV ohne Video in %	SV mit Video in %	Δ in %
ohne HG	50 dB	28 ± 24	49 ± 22	22
	65 dB	53 ± 29	73 ± 23	20
	80 dB	79 ± 17	90 ± 12	10
mit HG	50 dB	44 ± 19	69 ± 16	25
	65 dB	82 ± 16	92 ± 9	10
	80 dB	92 ± 8	98 ± 4	6
NH	35 dB	63 ± 19	83 ± 9	20
	50 dB	95 ± 5	99 ± 2	4
	65 dB	99 ± 2	100 ± 0	1

Tabelle 1: Mittlere Sprachverständlichkeiten der Probanden

Die nachstehende Tabelle 2 zeigt eine Auflistung der Signifikanzen bei den jeweiligen Testvarianten im Bezug auf einen signifikanten Unterschied mit und ohne Video-einblendung bei den gemessenen Lautstärken.

Messung		p-Wert	signifikanter Unterschied (ohne / mit Video)
ohne HG	50 dB	0.0402	ja mit α = 0.05
	65 dB	0.0710	gering mit α = 0.1
	80 dB	0.9000	nein
mit HG	50 dB	0.0080	sehr mit α = 0.01
	65 dB	0.9176	nein
	80 dB	0.9987	nein
NH	35 dB	0.0061	sehr mit α = 0.01
	50 dB	0.9629	nein
	65 dB	0.9999	nein

Tabelle 2: Signifikante Unterschiede bei jeweiligen Messungen mit und ohne Video

Die Verbesserung der Einsilberverständlichkeit durch das Tragen der Hörgeräte unterscheidet sich außerdem nicht signifikant, wenn ohne oder mit zusätzlichem Mundbild durch das Video getestet wird. Abbildung 7 veranschaulicht die gemittelten Sprachverständlichkeit für die Probanden mit HV.

Zusätzlich wurde gemessen, wie viele Einsilber nur mit Videobild und ohne Ton verstanden wurde (Erkennungsrate, ER). So konnten die Probanden mit Hörverlust alleine durch das Mundbild des Sprechers 17,5 ± 8,8 % erkennen, Normalhörende wiederholten im Mittel 16 ± 6,5 % richtig. Diese Werte entsprechen drei bis vier Wörtern, die richtig benannt wurden.

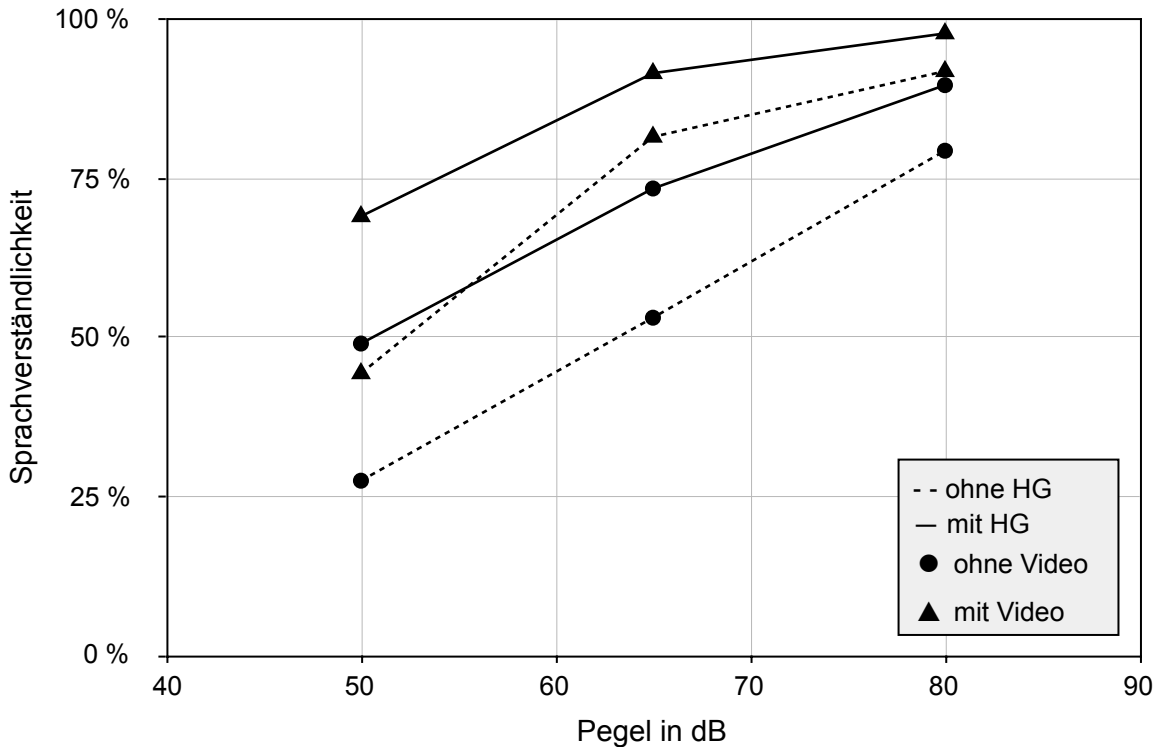

Abb. 7: Mittlere Sprachverständlichkeiten mit / ohne HG und mit / ohne Video

In Tabelle 3 sind die jeweiligen Verbesserungen mit kritischem p-Wert zur Überprüfung der Signifikanz eingetragen.

Verbesserung ohne - mit HG	ohne Video	mit Video	p-Wert	signifikanter Unterschied
50 dB	17 %	20 %	0.9873	nein
65 dB	28 %	18 %	0.4224	nein
80 dB	13 %	8 %	0.9553	nein

Tabelle 3: Mittlere Verbesserungen mit HG verglichen ohne und mit Video

Nur für fünf Probanden mit HV konnte ein Hörverlust für Sprache (50 %) bei der Messung mit Mehrsilbern gemessen werden. Alle anderen Probanden konnten auch bei einer Darbietung der Zahlen ohne Ton und nur mit Video mehr als die Hälfte richtig wiedergeben.

Tabelle 4 zeigt die im Mittel gemessenen Lautstärken, bei denen die Probanden die Hälfte der Mehrsilber richtig verstanden. Es liegt bei keiner der Messungen eine signifikante Verbesserung durch das Videobild vor.

	ohne Video	mit Video	Verbesserung	p-Wert	signifikanter Unterschied
ohne HG	49 ± 7,8 dB	44 ± 6,5 dB	5 dB	0.7780	nein
mit HG	38 ± 4,9 dB	35 ± 5,7 dB	3 dB	0.6046	nein

Tabelle 4: Mittlere Lautstärke für 50 % Mehrsilberverstehen (n = 5)

Dabei wurde bei den weiteren Probanden (n = 11) eine mittlere Erkennungsrate von 65,5 ± 8,2 % bei Messung ohne Ton und mit Videobild ermittelt.
Die normalhörenden Teilnehmer verstanden bis auf eine Person ebenfalls mehr als die Hälfte der Wörter, im Mittel 74 ± 8,9 %. Ein Proband konnte sein Mehrsilberverstehen von 50 % bei 27 dB ohne Video auf 22 dB mit Video verbessern. Ohne Ton wurden nur 40 % der Zahlen richtig wiederholt.

5.3 Fragebögen

Die Auswertung der Fragebögen erfolgte für Normalhörende und Probanden mit Hörverlust getrennt. Abbildung 8 zeigt jeweils das Ergebnis der ersten Frage, ob eine unterschiedliche Konzentration und somit Anstrengung für die Testvariante mit oder ohne Video erforderlich war.

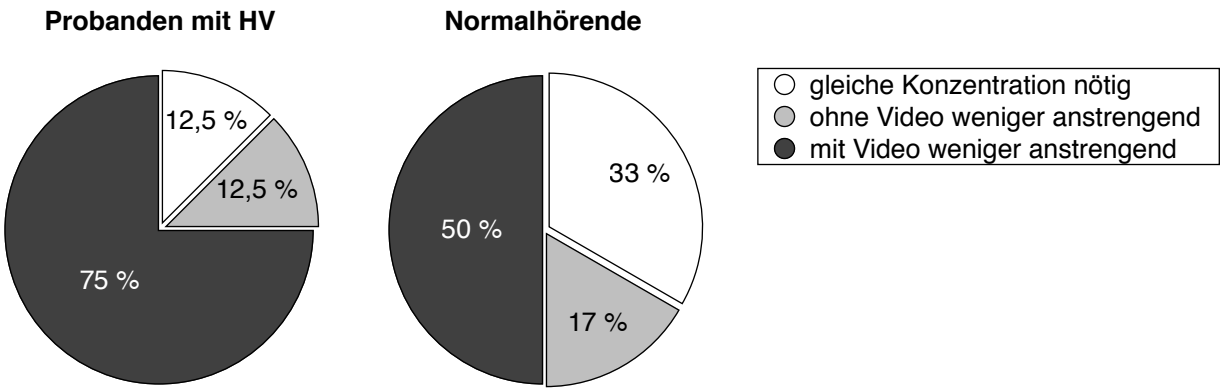

Abb. 8: Auswertung der Frage 1 nach einer unterschiedlichen Konzentration

Zusätzlich sollten die Probanden ihre Anstrengung auf einer Zahlenskala von 1 bis 10 eintragen, wobei 10 eine hohe und 1 eine niedrige Konzentration bedeuteten. Die Hörgeräteträger trugen für die Variante ohne Video im Durchschnitt 6,8 ± 2,3 ein, mit Video konnte 4,5 ± 2,3 im Mittel ausgewertet werden. Die Daten zeigen einen sehr signifikanten Unterschied ($p < 0.005$). Normalhörende kreuzten im Mittel 3,5 ± 1,4 für den Test ohne Video und 2,6 ± 1,4 für die Testvariante mit Videobild an, womit sich die Bewertungen nicht signifikant unterscheiden ($p > 0.1$).

Auf die Frage, bei welcher Testversion das Ergebnis subjektiv besser empfunden wurde (unabhängig vom tatsächlichen Resultat) gaben alle Probanden, sowohl normalhörend als auch mit HV, an, dass sie durch das zusätzliche Video ein besseres Resultat erwarteten.

Welche Testvariante für weitere Messungen einer Hörgeräteversorgung beim Akustiker oder HNO-Arzt bevorzugt wird, wurde in Frage 3 beantwortet. 13 der insgesamt 16 Probanden mit HV (81,3 %) bzw. 100 % der Personen ohne HV gaben an, auch für zukünftige Messungen eine Videounterstützung zu begrüßen.

6 Diskussion

Die Ziele dieser Arbeit bestanden aus der Entwicklung eines Videomaterials, welches in einen anerkannten Sprachtest eingebunden werden und zur visuellen Erweiterung dienen sollte. Des Weiteren sollte in einer Pilotstudie untersucht werden, inwiefern das bewegte Sprachmundbild das Sprachverstehen innerhalb dieses Tests verbessert. Dabei wurde auch der Gedanke gefasst, dass eine Messung mit einer zusätzlichen visuellen Informationsquelle weniger Konzentration benötigt und somit zu einer geringeren Anstrengung führt.

6.1 Auswahl des Sprachtests

Für die Erweiterung eines Sprach- bzw. Satztests mit Video wurden zu Beginn der Arbeit mehrere Testverfahren zur Überlegung herangezogen.

Zunächst lag die Entscheidung nahe, den Oldenburger Satztest zu nehmen. In den zuvor beschriebenen Studien von Meister et al. (2011) oder Schreitmüller et al. (2015) wurde ebenfalls dieser Satztest eingesetzt. Bei Verwendung des OlSa wäre somit eine direkte Vergleichbarkeit zwischen animierten Mundbild und Video eines realen Sprechers gegeben.

Hierbei lag jedoch das Problem bei der Einbindung eines Videos. Da die Testreihen im OlSa bei jeder Messung zufällig gewählt werden, muss das Video für einen Testsatz aus einzelnen Videosequenzen bestehen. Für jedes Wort ein Video. Somit wäre ein flüssiges und realistisches Mund- bzw. Sprachbild nicht realisierbar gewesen. Auf eine feste Auswahl und Zusammenstellung der Satzglieder, um ein komplettes Video pro zu prüfenden Satz erstellen zu können, wurde ebenfalls bewusst verzichtet, da vor allem die Randomisierung der einzelnen Wortglieder eines des größten Argumente für den Oldenburger Satztest ist.

Somit führten die Überlegungen auf einen Sprachtest mit Einsilbern oder zumindest wenigen Mehrsilbern, die zufällig angespielt werden müssen. Der ebenfalls eingangs beschriebene Freiburger Reloaded von Hahn und Limberger (2014), abgeleitet aus dem Freiburger Sprachtest, erforderte die geringste technische Umsetzung. Zudem bietet er einige Verbesserungen gegenüber des originalen Tests wie einen aussortierten Wortschatz oder eine Neuverteilung der Wortgruppen.

Da der Freiburger Reloaded jedoch nicht in ausreichendem Maße validiert ist und er somit keine genügende Vergleichbarkeit bietet, wurde für die Entwicklung und Umsetzung eines videounterstützten Sprachtests der Freiburger Einsilbertest nach Hahlbrock (1953) mit der Neuaufsprache von 1969 verwendet. Er ist ausreichend validiert und wird nach Verlangen der Krankenkassen für die Hörsystemanpassung herangezogen. Daher liefert dieser Test eine große Vergleichbarkeit mit bereits vorangegangenen Messungen und Erfahrungen aus der Anpassung von Hörgeräten.
Da die einzelnen Wortgruppen und Zahlen in der Norm DIN 45621-1 „Sprache für Gehörprüfung“ einlesbar sind und zur Verwendung stehen, sind rechtliche Einschränkungen nicht vorhanden. Dies ist bei anderen Sprach- und Satztests nicht der Fall. Durch die Eingrenzung der verwendeten Wortgruppen konnte eine gleichmäßige Schwierigkeit erreicht werden.

6.2 Verbesserung durch visuellen Reiz

An den Ergebnissen der Messungen zur Verständlichkeit ist zu erkennen, dass in jedem der Fälle eine Verbesserung durch das zusätzliche Mundbild über Video erreicht wurde. Dabei unterscheiden sich die Resultate bei Personen mit Hörverlust signifikant ($p < 0.05$) bei leiser (50 dB) und gering signifikant ($p < 0.1$) bei mittlerer Sprache (65 dB), sofern die Probanden keine Hörhilfe tragen. Wird das Sprachverstehen durch die Hörgeräte unterstützt, lassen sich nur noch im leisen Sprachbereich deutliche Unterschiede erkennen ($p < 0.01$) . Ebenfalls zu sehen ist, dass die Probanden mit Unterstützung der visuellen Informationsquelle insgesamt mehr oder gleich viele Wörter verstehen konnten. So wurden im Test bei lauter Sprache (80 dB) ohne Hörsystem, dafür mit Video, im Durchschnitt 90 % verstanden und mit Hörgeräten, dafür ohne Video, 92 %, also annähernd gleich viel. Deutlicher zeigt sich dies bei den Messungen mit leisem Sprachpegel (50 dB). Im Mittel verstanden die Probanden beim Test mit Mundbild, allerdings ohne Hörgeräte, 5 % mehr als beim Test mit Hörsystemen und ohne Video. Dies verdeutlicht, dass die visuelle Information gerade bei leiser Sprache hilfreich ist. Dabei ist jedoch ebenfalls zu bedenken, dass durch das Starten eines Videos auf dem Kundenbildschirm die Aufmerksamkeit des Probanden automatisch gesteigert wird und er somit weiß, wann bzw. dass ein Testwort erscheint. Ohne Video wird das Wort ohne Ankündigung

abgespielt. Gerade bei leisen Pegel kann es somit vorkommen, dass der Proband von der Darbietung des Wortes überrascht wird und es somit nicht korrekt wiederholen kann.

Vergleicht man die erhöhte Verständlichkeit ohne Hörgerät zu mit Hörgerät zum einen gemessen ohne Video und zum anderen gemessen mit Video, so findet man keine signifikanten Unterschiede ($p > 0.1$). Dies bedeutet im Umkehrschluss, dass die relative Verbesserung annähernd gleich ist, egal ob mit oder ohne Video gemessen wird. Das Video verbessert die Sprachverständlichkeit gleichermaßen und hebt sie auf ein erhöhtes Niveau (vgl. Abbildung 7 und Tabelle 3, S. 21). Dies ist eine wichtige Erkenntnis, da sie die Praxistauglichkeit des Videotests zeigt. Eine Anpassung von Hörsystemen muss ausreichend validiert werden. Dies geschieht im Falle des Freiburger Einsilbertests zum einen ohne Hörgeräte und nach der Anpassung mit Hörgeräten. Die Verbesserung an Sprachverständlichkeit, die der Kunde durch die Hörsysteme erreicht, ist unter anderen Kriterien ein wichtiger und unabdingbarer Beweis für eine gute und erfolgreiche Anpassung.

Allerdings weisen die gesammelten Werte eine sehr große Streuung auf. So weicht die Verständlichkeit bei der Messung ohne Hörgeräte bei 50 dB um 24 % vom Mittelwert ab. Bei einer mittleren Lautstärke von 65 dB sogar um 29 %, was im Freiburger Sprachtest einen Unterschied von fünf bis sechs Wörtern entspricht. Tragen die Probanden ihre Hörsysteme nimmt diese Standardabweichung ab.
Diese große Streuung der Daten ist zum einem erklärbar durch die geringe Anzahl an Probanden. Ein weiterer, noch entscheidenderer Faktor sind jedoch die unterschiedlichen Hörverluste. Es wurde versucht, Probanden mit ähnlichen Hörverlusten zu testen. Dennoch variiert die Hörminderung zum Beispiel im Hochtonbereich (6 kHz) von leicht (40 dB) bis fast taub (110 dB) (vgl. Abb. 3, S. 14).
Bei normalhörenden Personen ist diese starke Abweichung nicht zu erkennen. Die Hörkurve aller Probanden befindet sich innerhalb von -5 dB bis 20 dB und liegt somit im nahezu selben Bereich.

Insgesamt profitieren die normalhörende Teilnehmer jedoch signifikant weniger vom Videobild. Nur bei sehr leisen Pegeln (35 dB) konnte eine signifikante Verbesserung erkannt werden ($p < 0.01$). Ein Grund dafür ist sicherlich, dass Personen mit normalem Gehör bei leisem Pegel für die Hörgeräteanpassung (50 dB) bereits annähernd alle Wörter auch ohne Videobild verstehen und somit ihr Verständlichkeit nicht mehr verbessern können. Gleiches gilt für Messungen bei mittlerem Pegel von 65 dB. Versuche mit leiserem Pegel als 35 dB machten in diesem Falle keinen Sinn, da alle Messungen im Freifeld durchgeführt wurden und durch Umgebungsgeräusche wie Lüftung, Computer oder alleine das eigene Atmen die Wortdarbietung nicht gehört wurde.

Zusätzlich zu den üblichen Lautstärken wurde das Verstehen bzw. Sehen von Wörtern ohne Ton getestet, um zu beobachten, wie stark der Einfluss des Videobildes ist. Probanden mit einem Hörverlust konnten im Mittel gleich viele Wörter (drei oder vier) richtig erkennen wie normalhörende Teilnehmer. Dies deutet darauf hin, dass das Lippenlesen in diesem Untersuchungsfeld unabhängig von Hörverlust oder Alter ist. Sieht man sich bei den Probanden mit Hörbeeinträchtigung die individuelle Erkennungsrate in Abhängigkeit von Alter und Zeit mit Hörverlust an, so lässt sich kein Zusammenhang erkennen. Jüngere Probanden mit einem Hörverlust seit ein oder zwei Jahren konnten somit gleich viele Wörter richtig wiederholen wie ältere Teilnehmer, die schon länger mit der Hörbehinderung leben und länger mit Hörgeräten versorgt sind. Diese Beobachtungen stimmen jedoch nicht mit den im Vorfeld genannten Ergebnissen von Schreitmüller et al. (2015) überein. Sie entdeckten, dass jüngere Teilnehmer wesentlich mehr Wörter ohne Ton erkannten als ältere Probanden. Gleichwohl in dieser Studie der Oldenburger Satztest verwendet wurde, ist die Erkennungsrate nur durch das Videobild vergleichbar.
Das gleiche Bild zeigt sich bei den Ergebnissen der Messungen zum Mehrsilberverstehen. Hier ist jedoch zu bemerken, dass die Erkennungsrate wesentlich höher ist. Die Probanden konnten hier zwischen vier und neun Zahlen richtig wiedergeben. Dies liegt zum einen daran, dass der Bereich für die möglichen Antworten für mehrsilbige Zahlen begrenzt ist (13 bis 99). Hinzu kommt, dass die Aussprache wesentlich länger dauert und einem gleichmäßigerem Rhythmus folgt. Die Auswahl

an möglichen Einsilbern hingegen scheint nahezu unbegrenzt für den Probanden und die Wörter erscheinen ohne jeglichen Zusammenhang oder Sinn. Die kurze Dauer der Artikulation erschwert das Erkennen von einsilbigen Wörtern zusätzlich.

Durch die hohe Erkennungsrate der Zahlen bei einer Testdurchführung mit Video, konnte nicht bei allen Probanden eine 50 % Schwelle bestimmt werden. Lediglich bei insgesamt sechs Probanden wurde ohne Ton eine Verständlichkeit von 50 % oder weniger gemessen. Insofern ist hier die Tauglichkeit für die tägliche Praxis in Frage zu stellen.

Die Messungen zeigen somit eine Verbesserung der Sprachverständlichkeit durch ein Videobild, wobei diese bei leiseren Messpegeln größer ist. Dabei verbessern sich die Probanden im Mittel gleich stark bei einer Messung mit Video durch das Hörgerät verglichen mit Messungen ohne Video. Zahlen werden nur durch das gesehen Bild besser erkannt als Wörter, was nicht in allen Fällen eine Messung des Hörverlustes für Zahlen mit Videobild zulässt.

6.3 Anstrengung und Konzentration während der Messung

Durch einen kurzen Fragebogen nach den Messungen konnte die subjektive Empfindung der Probanden vom zusätzlichen Videobild erfasst werden (vgl. Anhang 1). Hierzu mussten die Teilnehmer einfache Fragen zur Anstrengungen und Konzentration beantworten und konnten diese zusätzlich auf einer Skala von 1 bis 10 eintragen.

In den Sozialwissenschaften werden grundsätzlich zur Beurteilung von Intensitäten fünfstufige Ratingskalen mit Antwortmöglichkeiten wie stark, ziemlich oder wenig empfohlen (Rohrmann, 1978). Da diese Art von Beurteilungssystem jedoch keine Aussage über die Streuung der Antworten innerhalb der Probanden gibt, wurde auf diese Art der Intensitätsbeurteilung bewusst verzichtet.

Die Auswertung der Antworten zeigt eine klare Tendenz zu einer größeren Akzeptanz des Sprachtests, wenn mit zusätzlichem Videobild getestet wird. So gaben drei Viertel der untersuchten Teilnehmer mit Hörverlust an, weniger Konzentration und somit Anstrengung zu empfinden, wenn die zusätzliche Videoquelle gezeigt wurde

(vgl. Abb. 8, S. 22). Als einer der Gründe wurde in den Anmerkungen angegeben, dass ein Mundbild eine Kombination von Gehörtem und Gesehenem zulässt und somit eine leichtere Testsituation entsteht, welche sich in einer entspannteren Einstellung des Prüflings zeigt.

Die restlichen Probanden mit Hörminderung beurteilten ihre Konzentration gleich oder besser ohne Video. Hier wurde die Anmerkung gemacht, dass eine zusätzliche Informationsquelle zu sehr ablenkt und somit eine größere Anstrengung verlangt. Bei diesen Probanden konnte beim Test ohne Video beobachtet werden, dass sie ihre Augen schloßen, um sich besser auf die Testwörter zu konzentrieren.

Die optische Information kann also bei manchen Testpersonen die eigentliche, akustische Information überdecken (informational masking). Eine ähnliche Erkenntnis konnten auch Franken et al. (2013) machen. Sie untersuchten die Höranstrengung bei Normalhörenden und CI-Trägern mit und ohne computeranimiertem Mundbild und entdeckten eine erhöhte Anstrengung bei zusätzlicher visueller Modalität.

Eine insgesamt höhere Konzentration zeigt sich in der Auswertung der Zahlenskala. So unterscheiden sich die eingetragenen Werte signifikant ($p < 0.005$), wobei die Testvariante mit Video durchschnittlich mit geringeren Zahlenwerten bewertet wurde. Bei normalhörenden Probanden kann keine Signifikanz unter den Antworten gemessen werden ($p > 0.1$). Im Mittel trugen diese eine geringere Zahl ein als die Hörbeeinträchtigten.

Bei Frage 2 nach dem subjektiv empfundenen Ergebnis wurde von allen Teilnehmern angekreuzt, bei einem Test mit Videobild ein insgesamt positiveres Gefühl im Hinblick auf das Messergebnis zu haben. In der Tat verbesserten sich alle Teilnehmer. Bei drei der Probanden mit Hörverlust sowie drei Teilnehmern ohne Hörverlust konnte jedoch nur eine Verbesserung durch das Video von 15 % oder weniger bei einzelnen Lautstärken gemessen werden. Dennoch gaben die Probanden an, ein wesentlich besseres Ergebnis mit Video zu empfinden. Dies zeigt, dass Probanden dem videounterstützten Test positiver entgegensehen, auch wenn das Resultat ähnlich ist.

Diese Aussage wird bestärkt durch das Ergebnis der dritten Frage, in der die Probanden ihre Präferenz für weiteren Messungen wählen sollten. So bevorzugen die Mehrheit aller Teilnehmer (19 von 22) die Testvariante mit Videounterstützung.

Insgesamt lässt sich also eine Präferenz für einen Test mit Videounterstützung sowohl bei Personen mit als auch ohne Hörverlust erkennen. Die positivere Einstellung und Erwartung zum Test zeigt sich in einer geringeren Konzentration und somit Anstrengung bei der Sprachprüfung.

6.4 Verbesserungen und Ausblick auf weitere Arbeiten

Trotz der zufrieden stellenden Ergebnisse wäre eine Festigung der Ergebnisse durch weitere Messungen mit einem größeren Kollektiv an Probanden wünschenswert. Um in weiteren Untersuchungen große Standardabweichungen zu vermeiden, wäre eine stärkere Eingrenzung der unterschiedlichen Hörminderungen möglich. In diesem Zusammenhang könnte weiter untersucht werden, ob die Verbesserung der Verständlichkeit durch ein Video eine Abhängigkeit vom Hörverlust zeigt.

Für eine Bestärkung der hier vorgestellten Ergebnisse wäre ebenfalls eine Untersuchungsreihe mit mehreren Wiederholungen denkbar. Somit könnten Lerneffekte durch eine eventuelle Gewöhnung an das Mundbild erkannt werden.

Ein weitere Ansatz könnte die genauere Messung des Hörverlustes für Sprache mittels Zahlen sein. So wäre eine Prüfungen des Mehrsilberverstehens über Kopfhörer vorstellbar, um in diesem Weg auch leisere Pegel testen zu können und somit eine detaillierte psychometrische Funktion zu erhalten.

Eine mögliche Weiterarbeit könnte die genauere Untersuchung der Höranstrengung bieten. Der in dieser Studie verwendete Fragebogen wurde für eine schnelle und einfache Überprüfung selbst erstellt, da der Schwerpunkt der Arbeit auf der Entwicklung, Einbindung und Prüfung eines Sprachtestes mit Video lag. Es existieren jedoch auch Prüfbögen mit spezifischeren Skalen vor (Meis und Gabriel, 2001).

Diese Pilotarbeit kann somit in einigen Bereichen genauer untersucht werden, um die Ergebnisse zu festigen.

6.5 Schlussfolgerung

Die Messergebnisse aus dieser Studie zeigen deutlich, dass eine Einbindung eines Videos in den Freiburger Sprachtest für die Messung des Einsilberverstehens eine praxistaugliche sowie alltagsrealistische Erweiterung des Tests darstellt. Dabei können zum einen HörgeräteträgerInnen von der geringeren Konzentrations-forderung profitieren und des Weitern bleibt das relevante Ergebnis für die Audiologin / den Audiologen oder die HNO-Ärztin / den HNO-Arzt unverändert. Eine weitere Stärkung der Erbnisse durch weitere Sprachprüfungen mit Video für eine feste Eingliederung in den Sprachtest ist daher zu begrüßen.

Erklärung

Ich versichere hiermit, dass ich die vorliegende Bachelorthesis selbstständig und ohne fremde Hilfe angefertigt und keine andere als die angegebene Literatur benutzt habe. Alle von anderen Autoren wörtlich übernommenen Stellen wie auch die sich an die Gedankengänge anderer Autoren eng anlehnenden Ausführungen meiner Arbeit sind besonders gekennzeichnet. Diese Arbeit wurde bisher in gleicher oder ähnlicher Form keiner anderen Prüfungsbehörde vorgelegt und auch nicht veröffentlicht.

Aalen, 13.04.2016

Ort, Datum

Unterschrift

Literaturverzeichnis

Alich, G. 1985. Anmerkungen zum Freiburger Sprachcerständnistets (FST): Sprache - Stimme - Gehör. *Zeitschrift für Kommunikationsstörungen*, 9, 1-6

Andreassi, J. L., Greco, J. R. 1975. Effects of bisensory stimulation on reaction time and the evoked cortical potential. *Psychological Psychology*, 3 (2), 189-194

Bangert, H. 1980. Probleme bei der Ermittlung des Diskriminationsverlustets nach dem Freiburger Sprachtest. *Audiologische Akustik*, 9, 166-170

Cherry, E.C. 1953. Some experiments on the recognition of speech, with one and two ears. *Journal of the Acoustical Society*, 25 (5), 975-979

Cole, R., Massaro, D. W., de Villiers, J., Rundle, B., Shobaki, K., Wouters, J., Cohen, M., Beskow, J., Stone, P., Connors, P., Tarachow, A., Solcher, D. 1999. New tools for interactive speech and language training: Using animated conversational agents in the classrooms of profoundly deaf children. *Proceedings of ESCA/SOCRATES Workshop on Method and Tool Innovations for Speech Science Education*, 45.52

Deutsches Institut für Normung e.V. (1995) *DIN 45621-1 Sprache für Gehörprüfung Teil 1: Ein- und mehrsilbige Wörter.* Berlin: Beuth Verlag

Deutsches Institut für Normung e.V. (1995) *DIN 45626-1 Tonträger mit Sprache für Gehörprüfung Teil 1: Tonträger mit Wörtern nach DIN 45621-1 .* Berlin: Beuth Verlag

Deutsches Institut für Normung e.V. (2001) *DIN EN ISO 7029 Akustik - Statistische Verteilung von Hörschwellen als eine Funktion des Alters.* Berlin: Beuth Verlag

Deutsches Institut für Normung e.V. (2012) *DIN EN ISO 8253-3 Akustik - Audiometrische Prüfverfahren Teil 3: Sprachaudiometrie .* Berlin: Beuth Verlag

Erber, N. P. 2003. Use of hearing aids by older people: influence of non auditory factors (visual, manual dexterity). *International Journal of Audiology*, 42 (2), 2S21-2S25

Fagel, S. and Clemens, C., 2004. An Articulation Model for Audiovisual Speech Synthesis - Determination, Adjustment, Evaluation. *Speech Communication*, 44, 141-154

Frenken M., Schreitmüller S., Lipski, S., Meister, H. 2013. Audiovisuelle Sprachwahrnehmung II: Höranstrengung - ein wichtiger Faktor bei der Sprachaudiometrie Normalhörender und CI-Träger. *16. Jahrestagung der Deutschen Gesellschaft für Audiologie*. Rostock

Golumbic, E. Z., Cogan, G. B., Schroeder, C. E., Poeppel, D. 2013. Visual Input Enhances Selective Speech Envelope Tracking in Auditory Cortex at a “Cocktail Party”. *Journal of Neuroscience*, 33 (4), 1417-1426

Hahn, F., Prof. Dr. Limberger, A. 2013. *Freiburger reloaded*. Bachelorarbeit, Hochschule für Technik und Wirtschaft Aalen

Hicks, C. B., Tharpe, A. M. 2002. Listening effort and fatigue in school-age children with and without hearing loss. *Journal of Speech, Language, and Hearing Research*, 45 (3), 573-584

Huch, R., Jürgens, K. 2015. *Mensch Körper Krankheit*. München: Elsevier GmbH

Kießling J., 2008. *Versorgung mit Hörgeräten*. In: J. Kießling, B. Kollmeier & G. Diller (eds.) Versorgung und Rehabilitation mit Hörgeräten. Stuttgart New York: Thieme, pp. 120 ff.

Kühnel, C., Weiss, B., Möller, S. 2010. Ein dreistufiger Ansatz zur Evaluation von audio-visueller Systemausgabe. *DAGA Berlin,* 85-86

Lehnhardt, E., Laszig, R. 2009. *Praxis der Audiometrie*. Stuttgart: Georg Thieme

Lipski, S., Schreitmüller, S., Frenken, M., Meister, H. 2013. Audiovisuelle Wahrnehmung von Sprachsegmenten bei CI-Trägern und Normalhörenden. *16. Jahrestagung der Deutschen Gesellschaft für Audiologie*. Rostock

Lozán, J. L., Kausch, H. 1998. *Angewandte Statistik für Naturwissenschaftler: mit 39 Tabellen und 125 Beispielen*. Berlin: Parey

Meis, M. & Gabriel, B., 2001. Entwicklung, Psychometrie und erste Ergebnisse eines Fragebogens zur Höranstrengung. *Zeitschrift für Audiologie*, *Supplementum IV*, 97-99.

Meister, H., Schreitmüller, S., Pyschny, V., Lang-Roth, R., Walger, M., Fagel, S. 2011. Untersuchungen zur audio-visuellen Sprachwahrnehmung bei Personen mit Hörstörungen mittels virtuellem Kopfmodell. *14. Jahrestagung der Deutschen Gesellschaft für Audiologie*. Jena

McDermott, J.H. 2009. The cocktail party problem. *Current Biology*, 19 (22), R1024-R1027

McGurk, H., MacDonald, J. 1976. Hearing lips and seeing voices. *Nature*, 264, 746-748

Ping, H. Y., Abdullah, L. N., Sulaiman, P. S., Halin, A. A. 2013. Computer Facial Animation: A Review. *International Journal od Computer Theory and Engineering*, 5 (4), 658-662

Rohrmann, B. 1978. Empirische Studien zur Entwicklung von Antwortskalen für die sozialwissenschaftliche Forschung. *Zeitschrift für Sozialpsychologie*, 9, 222-245

Kompis, M. 2013. *Audiologie.* Bern: Hans Huber, Hogrefe AG

Rouger, J., Lagleyre, S., Fraysse, B., Deneve, S., Deguine, O., Barone, P. 2007. Evidence that cochlear-implanted deaf patients are better multi sensory integrators. *PNAS*, 104 (17), 7295-7300

Sams, M., Aulanko, R., Hämäläinen, M., Hari, R., Lounasmaal, O. V., Lu., S., Simol, J. 1991. Seeing speech: visual information from lip movements modifies activity in the human auditory cortex. *Neuroscience Letters*, 127, 141-145

Schreitmüller, S., Frenken, M., Ortmann, M., Walger, M., Meister, H. 2015. Audiovisuelle Sprachwahrnehmungsverfassung mit computeranimierten Sprech-bewegungen. *18. Jahrestagung der Deutschen Gesellschaft für Audiologie*. Bochum

Sekiyama, K., Kanno, I., Miura, S., Sugita, Y. Auditory-visual speech perception examined by fMRI and PET. *Neuroscience Research*, 47, 277-287
Stein, B. E., Meredith, M. A. 1994. *The merging of the senses*. Cambridge: MIT Press

Ulrich J., Hoffmann E. 2011. *Hörakustik 2.0 Theorie und Praxis*. Heidelberg: DOZ Verlag

Wagener, K., Kühnel, V., and Kollmeier, B. 1999. Entwicklung und Evaluation eines Satztests für die deutsche Sprache I: Design des Oldenburger Satztests. *Zeitschrift für Audiologie* 38, 4-15.

Winkler, A., Holube, I. 2014. Was wissen wir über den Freiburger Sprachtest? *Zeitschrift für Audiologie*, 53 (4), 146-154

Anhang

1) Fragebogen

Studie: VUST - Alexander Thal

Proband ID: _____ □ C Termin 1: _____ Termin 2: _____

begonnen: □ ohne Video □ mit Video HV seit: _____ HG seit: _____

Fragebogen

1. Können Sie einen Unterschied im Bezug auf die Anstrengung und Konzentration bei den beiden Testvarianten feststellen?

   □ Nein, beide Tests erfordern gleiche Konzentration.

   □ Ja, weniger anstrengend: □ ohne Video □ mit Video

   Wie hoch beurteilen Sie die Anstrengung bei jeweiligem Verfahren auf der Skala von 1 bis 10? (1 = nicht anstrengend, 10 = sehr anstrengend)

   - ohne Video: 1 2 3 4 5 6 7 8 9 10
   - mit Video: 1 2 3 4 5 6 7 8 9 10

2. Können Sie einen Unterschied im Bezug auf das subjektiv empfundene Ergebnis bemerken?

   □ Nein, beide Tests ergaben das gleiche Resultat.

   □ Ja, besseres Ergebnis: □ ohne Video □ mit Video

3. Welche der beiden Varianten des Sprachtestes würden Sie für weitere Messungen bevorzugen?

   □ ohne Video □ mit Video

Anmerkungen:

_____

_____

_____