

Aus der Hals-Nasen-Ohren-Klinik,

Universitätsklinikum des Saarlandes, Homburg/Saar

Direktor: Prof. Dr. med. Alessandro Bozzato

In Zusammenarbeit mit der Hochschule Trier und Prof. Dr. -Ing. Jörg Lohscheller

**Erhebung und Analyse von Normdaten zum telemedizinischen Monitoring von
Stimmstörungen**

**Dissertation zur Erlangung des Grades eines Doktors der Medizin
der Medizinischen Fakultät**

der UNIVERSITÄT DES SAARLANDES

2020

vorgelegt von: Irina Maria Fraundorfer

geb. am: 20.03.1992 in Roding

Tag der Promotion: 21.01.2021

Dekan: Univ.-Prof. Dr. med. M. D. Menger

Berichterstatter:

Professor Dr. J. Lohscheller

Professor Dr. A. Langenbacher

Inhaltsverzeichnis

Inhaltsverzeichnis..... I

Zusammenfassung..... V

1 Einleitung und Zielsetzung der Arbeit..... 1

2 Grundlagen..... 4

 2.1 Anatomie des Larynx 4

 2.2 Grundlagen der Stimmproduktion 8

 2.3 Pathologien der Stimme13

 2.4 Bildgebende Diagnostik des Larynx17

 2.4.1 Stroboskopie18

 2.4.2 Videokymographie.....19

 2.4.3 Hochgeschwindigkeitsvideoendoskopie19

 2.5 Subjektive Stimmbewertung.....19

 2.5.1 Selbstbewertung der Stimme mit VHI-Test.....20

 2.5.2 Fremdbewertung21

 2.6 Apparative Stimmdiagnostik.....21

 2.6.1 Stimmfeldmessung.....21

 2.6.1.1 Sprechstimme22

 2.6.1.2 Singstimme23

 2.6.2 Akustische Irregularitätsparameter23

 2.6.3 Dysphonia Severity Index (DSI).....24

 2.6.4 Die Grundfrequenz F_025

 2.7 Akustische Parameter für Rauschen25

 2.7.1 Harmonics-to-noise ratio (HNR)26

 2.7.2 Noise-to-harmonic ratio (NHR)27

 2.7.3 Normalized-noise energy (NNE).....27

 2.8 Spektrale Stimmqualitätsparameter28

 2.8.1 H_1H_2 28

 2.8.2 Spektrale Neigung.....29

 2.8.3 Smoothed cepstral peak prominence (CPPs)30

3 Material und Methoden32

 3.1 Ethikantrag.....32

 3.2 Probanden32

 3.3 Einschluss in die Studie32

 3.3.1 Einverständniserklärung und Anamnese33

 3.3.2 VHI-12 Test.....33

Inhaltsverzeichnis

3.3.3	Stroboskopie	34
3.3.4	Stimmanalyse.....	34
3.3.5	Stimmmonitoring und Aufbau der Handyapp Voicediary.....	36
3.3.6	Einstellung und Bedienung der Handyapplikation.....	36
3.3.7	Abschlussuntersuchung der Studie	39
3.4	Statistische Auswertung.....	39
4	Ergebnisse	41
4.1	Probanden	41
4.2	Zusammenfassung der Anamnesedaten.....	41
4.3	VHI-12 Auswertung.....	42
4.4	Stroboskopiebefunde	42
4.5	Ergebnisse der Stimmfeldanalyse.....	43
4.5.1	Sprechstimmprofil.....	43
4.5.2	Singstimmprofil.....	45
4.6	Ergebnisse von „Voice Diary“	47
4.6.1	Stimmbelastung	48
4.6.2	Stimmqualität	52
4.6.3	Missempfindungen	53
4.6.4	Ermüdungsgrade.....	54
4.7	Ergebnisse des telemedizinischen Monitorings.....	56
4.7.1	Grundfrequenz F_0	56
4.7.2	Irregularitätsparameter: Jitter & Shimmer	58
4.7.3	HNR	61
4.7.4	NHR	63
4.7.5	NNE	65
4.7.6	H_1H_2	67
4.7.7	Spektrale Neigung: Tilt & Slope.....	70
4.7.8	Smoothed cepstral peak prominence (CPPs)	73
4.7.9	t-SNE	75
5	Diskussion	81
5.1	Einschluss in die Studie	81
5.2	Auswertung des Fragebogens in „Voice Diary“	84
5.3	Parameter aus dem telemedizinischen Monitoring	86
5.3.1	F_0 und Irregularitätsparameter	87
5.3.2	Akustische Parameter für Rauschen	88
5.3.3	Spektrale Parameter	90
5.3.4	Verteilung aller Werte in t-SNE Grafik	93
6	Anhang.....	94

Inhaltsverzeichnis

7	Literaturverzeichnis.....	102
8	Abbildungsverzeichnis	109
9	Danksagung	115
10	Lebenslauf	116

Abkürzungsverzeichnis

CPPs	Smoothed Cepstral Peak Prominence
DSI	Dysphonia Severity Index
ELS	European Laryngological Society
H ₁ H ₂	Amplitudendifferenz der ersten und zweiten harmonischen Schwingung
HNR	Harmonics-to-Noise Ratio
Jitter	Frequenzänderungen
LTAS	Long-Term Average Spectrum
mSsl	mittlere Sprechstimmlage
NHR	Noise-to-Harmonic Ratio
NNE	Normalized-Noise Energy
RBH	Rauigkeit, Behauchtheit, Heiserkeit
Shimmer	Amplitudenschwankung
Slope	Parameter der spektralen Neigung in einem Energiespektrum
Tilt	Parameter der spektralen Neigung in einem Energiespektrum
t-SNE	t-distributed Stochastic Neighbor Embedding
VHI	Voice Handicap Index

Zusammenfassung

Einleitung: In unserer kommunikativen Gesellschaft können Stimmstörungen, wie Heiserkeit, zu Einschränkungen im privaten und öffentlichen Leben und zur Isolation der Betroffenen führen. Daher ist ein frühzeitiges Diagnostizieren von Stimmstörungen entscheidend für die Lebensqualität der Stimmkranken. Um einen größeren Informationsgehalt über die Stimme zu erhalten, als bisher in punktuellen Untersuchungen möglich war, wurden in dieser Studie Stimmdaten von Stimmgesunden mittels telemedizinischen Monitorings via Handy-App erhoben und mit bestehenden Referenzwerten aus der Literatur verglichen.

Material und Methoden: Die Teilnehmer dieser Studie waren 20 deutschsprachige Studenten im Alter von 22 bis 33 Jahren, die nach einer unauffälligen Voruntersuchung anamnestisch, stimmqualitativ (Stimmfeldmessung) und durch eine visuelle Begutachtung der Stimmlippen mittels Stroboskopie als stimmgesund eingestuft wurden. In einem 14-tägigen Untersuchungszeitraum sollte in sechs festgelegten Zeitintervallen mindestens je eine Aufnahme eines ponierten Vokals über die Handyapplikation „Voice Diary“, die extra von der Hochschule Trier entwickelt und auf HUAWEI Handys installiert wurde, aufgezeichnet werden. Jede Stimmaufnahme wurde danach mit einem Fragebogen bewertet. Aus den erhobenen Daten des telemedizinischen Monitorings wurden ausgewählte akustische Parameter mit MATLAB berechnet und in verschiedenen Grafiken dargestellt. Zu den ausgewählten Parametern zählten die Grundfrequenz (F_0) und die Irregularitätsparameter Jitter und Shimmer, mit welchen sich Frequenz- und Amplitudenschwankungen bestimmen lassen. Zudem erfolgte die Berechnung der Harmonics-to-noise ratio (HNR) und der Noise-to-harmonics ratio (NHR), um das Verhältnis von harmonischen Anteilen und Geräuschanteilen in einem Stimmsignal zu beschreiben. Mit der Normalized-noise energy (NNE) wurde die gesamte Geräuschenergie eines Stimmsignal berechnet. Neben den Rauschparametern wurden spektrale Stimmqualitätsparameter wie der H_1H_2 , als akustisches Korrelat für das Verhältnis von Offen- und Schlussphase der Stimmlippen verwendet. Des Weiteren wurde die spektrale Neigung des Stimmsignals mit dem Tilt, dem Slope und der Smoothed cepstral peak prominence (CPPs) ermittelt. Zuletzt erfolgte durch ein t-distributed stochastic neighbor Embedding (t-SNE) eine Dimensionsreduktion aller Parameter und der Vergleich in einem gemeinsamen Plot.

Ergebnisse: Insgesamt konnten die Daten von 17 Probanden in dieser rein deskriptiven Studie verwendet werden. Die subjektive Stimmbewertung der meisten Aufnahmen war sehr gut oder gut. Die Frage nach der Stimmqualität ist in dieser Studie mit den anderen Qualitätsfragen nicht vergleichbar.

Bis auf den NNE verteilen sich alle anderen akustischen Parameter bei den einzelnen Probanden und in Anbetracht des Gesamtkollektivs regelmäßig. Dies zeigt auch die gleichmäßige Verteilung innerhalb eines t-SNE Plots nach Dimensionsreduktion und Ausschluss des NNE. Zudem stimmen sie mit bisher beschriebenen Referenzwerten aus der Literatur überein. Einzelne Ausreißer lassen sich durch unregelmäßige Messsituationen im häuslichen und freien Umfeld durch Hintergrundgeräusche erklären.

Diskussion: Das telemedizinische Monitoring mit Mobiltelefonen ergab unter Ausschluss des NNE stabile akustische Parameter, die mit Referenzwerten von Stimmgesunden aus der Literatur übereinstimmen. Das Stimmmonitoring via Handyapp ist daher geeignet, Stimmverläufe von gesunden Probanden valide aufzuzeichnen und kann bei Stimmgesunden in der Stimmdiagnostik verwendet werden.

Summary

Introduction: In a communicative society voice disorders, like hoarseness, can lead to limitations in a person's private and public life and to isolation of the affected individual. An early diagnosis of voice disorders is essential for improving the patient's quality of life.

In this telemedical monitoring study a mobile phone application was used to capture the voice parameters of individuals with healthy voices. This served to create a useful data set which was then compared to pre-existing reference values in the literature.

Material and Methods: In this study 20 German speaking students from the ages of 22 to 33 years volunteered. They each were given a non-pathological preexamination of voice quality, in addition to the anamnestic and a normal visual appearance of the vocal cords during stroboscopy. During the research period of 14 days, the phonations of vocalization during 6 different time periods each day were recorded with the application Voice Diary. This application was invented at the technical faculty in Trier and downloaded and installed on Huawei phones. Each record was evaluated afterwards using a standardized questionnaire. Through calculation of the recording of the telemedical monitoring certain acoustic parameters were computed with MATLAB and presented in different charts. The selected parameters included the fundamental frequency (F_0) and the parameters of irregularity like jitter and shimmer, which describe the variation of frequency and the amplitude. Furthermore, the Harmonics-to-noise ratio (HNR) and the Noise-to-harmonic ratio (NHR) were calculated to describe the ratio of the harmonic and noisy components in a voice signal. The Normalized-noise energy (NNE) was used to calculate the complete amount of noise energy in a voice signal. Besides the noise parameters, spectral voice parameters like the H_1H_2 as an acoustic correlation of the ratio between opening and closing of the vocal folds were determined. For the calculation of the spectral tilt of a voice signal, parameters like the Tilt, the Slope and the Smoothed cepstral peak prominence (CPPs) were used. In a final step the parameters were compared in a common Plot after the dimension reduction with t-distributed stochastic neighbor Embedding (t-SNE).

Results: On the whole the recordings and results of 17 participants were able to be compared in this descriptive study. The subjective voice evaluation of most recordings was good or very good. In the study the results of the voice quality were not comparable with the other questions of the standardized questionnaire.

Except the NNE, the spreading of the elected acoustic parameters of all test subjects was normal compared to the average. The same result was shown in the t-SNE plot after the reduce of dimension and exclusion of the NNE. Furthermore, the results were conforming to preexisting reference values in the literature. Single spikes in the data are explained by the subjects experiencing disturbed recording situations as well as background noises in private and public areas.

Discussion: The telemedical monitoring of voices using mobile phones is able to record stable acoustic parameters, which are comparable to pre-existing reference values of individuals with healthy voices. Therefore, the telemedical monitoring using mobile phones is a valid method for recording voice changes of healthy people and can be used as an instrument in the diagnostic of healthy voices.

1 Einleitung und Zielsetzung der Arbeit

Der technische Fortschritt und die Modernisierung erlauben in sämtlichen Bereichen des Lebens neue Möglichkeiten und steigern in vielerlei Hinsicht Effektivität, beispielsweise durch schnelleren Datenaustausch in einer eng vernetzten Gesellschaft und schneller Verarbeitung großer Datenmengen. Im kommunikativen Bereich wurde in den letzten 20 Jahren ein großer Schritt getan. Obwohl durch die Nutzung von Messaging-Diensten und Chatforen die non-verbale Kommunikation gefördert wird, ist die Sprache ein fester Bestandteil im privaten und beruflichen Alltag. Um Sprache zu produzieren, gebrauchen wir Menschen unsere Stimmen. Sie ermöglicht es uns an Kommunikationen aktiv teilzunehmen, unsere Gedanken mitzuteilen und unsere Emotionen auszudrücken. Ein Ausdruck von Trauer beispielsweise können nicht nur Tränen sein, sondern auch eine schwache, leise und stockende Stimme. Zudem können Stimmen einiges über die Persönlichkeit verraten. Ein sehr lauter und wortgewandter Erzähler wird von uns als eher extrovertierte Persönlichkeit wahrgenommen, eine leise, piepsige Stimme schreiben wir dagegen eher einem schüchternen, introvertierten Menschen zu. Der kanadische Psychologe David Feinberg fand heraus, dass tiefe Männerstimmen auf die Zuhörer vertrauenswürdiger und kompetenter wirken [15]. Die Stimme kann beim Zuhörer nicht nur etwas bewirken, sondern ermöglicht es diesem auf das Geschlecht, das Alter und den Gesundheitszustand eines Menschen zu schließen [97]. Infekte der oberen Atemwege können zu einer heiseren und belegten Stimme führen. Stimmstörungen können hingegen auch durch strukturelle und funktionelle Veränderungen im Kehlkopf selbst verursacht werden. Beeinträchtigungen der Funktionstüchtigkeit und Leistungsfähigkeit der Stimme im Alltag und im Beruf können schwerwiegende Konsequenzen haben.

So kann es bei Stimmstörungen im privaten und im beruflichen Umfeld laut einer Studie zu einem erhöhten Depressionsrisiko und zu einer höheren Anzahl an krankheitsbedingten Fehltagen in der Arbeit kommen [16]. Des Weiteren geben Patienten mit Stimmstörungen ein reduziertes Selbstwertgefühl an und Probleme in der Kommunikation bei lauter Umgebung [108]. Vor allem Lehrer sind gegenüber Stimmstörung häufiger betroffen als andere Berufsgruppen [13]. Da die Stimme eine zentrale Stellung im Leben hat, ist das frühzeitige Diagnostizieren von Stimmerkrankungen wichtig.

Mit bisherigen Methoden wie der Stimmfeldmessung oder bildgebenden Untersuchungen, wie der Stroboskopie, im stationären Rahmen konnte der Stimmstatus nur zu einem gewissen Zeitpunkt erhoben werden. Da sich die Stimme im Verlauf des Tages immer wieder ändert, beziehungsweise auch die Belastung der Stimme von Tag zu Tag unterschiedlich ist, ermöglicht ein ambulanter Kontrolltermin durch eine punktuelle Aufnahme der Stimme meist nur eine Aussage über den momentanen Istzustand der Stimme. Da ständige Kontrolltermine

aber eine Kosten- und Zeitfrage für Patient und Klinik sind, wäre ein klinikfernes und mobiles Aufnahmeverfahren im Rahmen eines ambulanten Monitoring hier eine elegante Lösung [67]. Da durch ein telemedizinisches Monitoring bereits in unterschiedlichen Bereichen der Medizin neue diagnostische und auch therapeutische Wege ermöglicht wurden, soll dieses Verfahren auch bei Stimmerkrankungen Anwendung finden. In der Kardiologie beispielsweise lassen sich Hypertonien und Herzinsuffizienzen von Patienten durch ein derartiges Verfahren kontrolliert überwachen [43]. Bereits mehrere Studien in der Vergangenheit haben ambulante Stimmessverfahren getestet und das bestehende Potential darin gezeigt [45, 67, 90]. In einer Studie von Mehta et al. aus dem Jahr 2012 beispielsweise [68] wurden hyperfunktionelle Stimmen im Tagesverlauf vor und nach Therapie aufgezeichnet, um anhand der Daten die Diagnostik und folglich Therapiemöglichkeiten des Krankheitsbildes zu verbessern. Unter Verwendung eines Sensors, der am Hals befestigt wurde, wurden die Vibrationen bei der Sprachproduktion aufgezeichnet und über eine Kabelverbindung an ein Handy, das als Speicherort diente, weitergeleitet. Diese Methode ermöglichte das Erstellen eines Stimmtagesprofil und nicht nur eine punktuelle Datenerhebung. Diese Studie zeigt, dass unter Verwendung von neuen Technologien wie Smartphones, die Möglichkeit besteht, die Stimmanalyse klinikfern und mobil zu gestalten. In Verbindung mit Stimmtagebuch Apps zur Analyse von Stimmstörungen kann zum einen der Untersuchungszeitraum frei gewählt werden und ein größerer Informationsgehalt über die Stimme im Vergleich zu einer einmaligen Untersuchung gewonnen werden. Zum anderen trägt der Patient aktiv zur Datenerhebung bei, er wird mehr in die Untersuchung mit einbezogen und bezüglich seiner Beschwerden sensibilisiert.

Die vorliegende Studie befasst sich damit, eine Verbesserung in der Stimmdiagnostik zu finden. Ziel dieser Arbeit war es, ein für die Stimmdiagnostik entwickeltes ambulantes Monitoringsystem an einer homogenen Gruppe stimmgesunder, männlicher Probanden zu validieren. Als ambulantes Aufnahmeverfahren wurde an der Hochschule Trier eine Stimmtagebuch Applikation für ein Handy programmiert. Zum einen sollte der Einsatz der App für ein ambulantes Monitoring getestet und anhand der Anwendungshäufigkeit die Compliance bewertet werden. Zum anderen sollten stimmdiagnostisch relevante Parameter und mögliche Schwankungen im Tagesverlauf untersucht werden. Eine geringe Abweichung zwischen Handydaten und Referenzdaten aus der Literatur würde das ambulante Monitoring als ein gültiges Instrument in der Diagnostik zur Aufzeichnung der täglichen Stimmentwicklung bestätigen. Da die Stimme im Tagesverlauf ein sehr variables Verhalten zeigt, soll in dieser Studie weiterhin gezeigt werden, wie sich die Ergebnisse bezüglich der bereits beschriebenen Referenzdaten aus der Literatur verhalten.

In den folgenden Abschnitten werden zunächst Grundlagen über die Anatomie und die Physiologie der Stimme erklärt und näher auf die derzeitig etablierte Stimmdiagnostik eingegangen.

2 Grundlagen

Im folgenden Abschnitt wird zunächst die Anatomie des Kehlkopfes und die Physiologie der Stimme erklärt. Anschließend werden die häufigsten organischen und funktionellen Stimmpathologien beschrieben und die in der Diagnostik angewendeten Bildgebungen. Abschließend wird die Bedeutung der Stimmbewertung erklärt.

2.1 Anatomie des Larynx

Der *Larynx* (Kehlkopf) ist das für die Stimmgebung relevante Organ. Er liegt auf Höhe des 3. bis 5. Halswirbels und ist aus muskulären und knorpeligen Anteilen aufgebaut. Er ist beweglich am *Os hyoideum* (Zungenbein) befestigt und bildet den Übergang von *Trachea* (Lufttröhre) und *Pharynx* (Rachen). Der Kehlkopf wird durch die *Plicae vocales* (Stimmfalten) und *Plicae vestibulares* (Taschenfalten) in die supraglottische, transglottische und subglottische Ebene gegliedert (vgl. Abbildung 1).

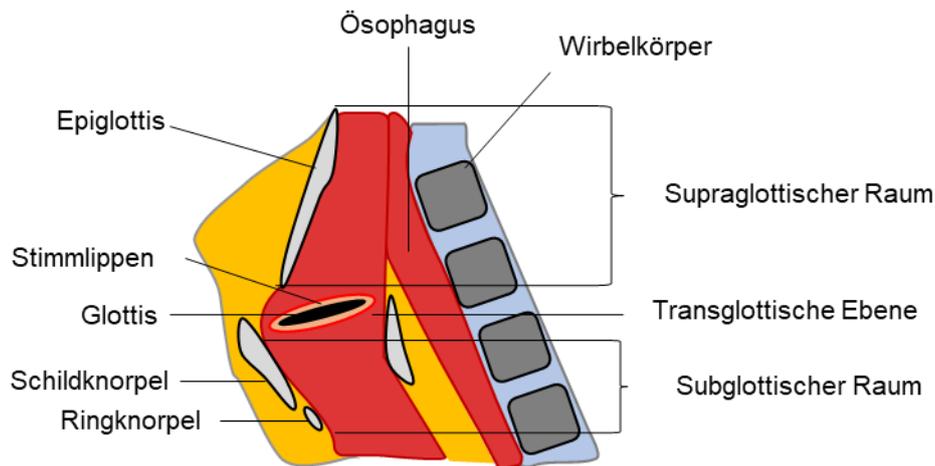


Abbildung 1: Schematische Darstellung des Kehlkopfes im Mediansagittalschnitt in Anlehnung an [7].

Größtenteils ist er mit mehrschichtigem, unverhorntem Plattenepithel und Respirationsepithel ausgekleidet. Seine drei wichtigsten Aufgaben umfassen die Atemregulation, die Stimmbildung und die Schutzfunktion beim Schlucken. Das Kehlkopfskelett umfasst vier funktionell wichtige Knorpel: *Epiglottis* (Kehlideckel), *Cartilago thyroidea* (Schildknorpel), *Cartilago cricoidea* (Ringknorpel) und *Cartilago arytenoidea* (Stellknorpel, paarig) sowie zwei Gelenkanteile *Articulatio cricothyroidea* (Ring- und Schildknorpelgelenk) und *Articulatio cricoarythenoidea* (Ring- und Stellknorpelgelenk) [116].

Elastische Bänder verbinden die Knorpel untereinander und gewähren dadurch Flexibilität und Elastizität, die vor allem für die Phonationsbildung von großer Bedeutung sind [3].

Auf die Funktion und den Aufbau der obigen Strukturen wird im Folgenden näher eingegangen. Die Epiglottis verschließt den Larynxeingang beim Schluckvorgang und verhindert so den Eintritt von Flüssigkeit oder Nahrung in die Lufttröhre [3]. Die Epiglottis ist aus elastischem

Knorpel aufgebaut und ist über das *Ligamentum thyreoepiglotticum* (Band zwischen Schildknorpel und Kehldeckel) an dem Schildknorpel verankert. Durch die elastischen Fasern im Gewebe kann sich der sonst steife Knorpel verformen. Die mukösen Drüsen des Kehldeckels befinden sich auf der perforierten Innenseite des Knorpels [116]. In diesen sind Lymphozyten, Mastzellen und eosinophile Granulozyten eingebettet, die der Immunabwehr dienen [3]. Durch das Anheben der suprahyoidalen Muskulatur beim Schlucken wird ein Fettkörper aktiviert, welcher vor der Epiglottis lokalisiert ist. Dieser ermöglicht ein Abknicken des Kehldeckels und somit den Verschluss des Larynxeinganges. Die beiden Schildknorpelplatten bestehen aus hyalinem Knorpel und bilden an ihrer Spitze die Prominentia laryngea (vgl. Abbildung 2). Der Winkel dieser beiden Platten beträgt bei Männern um die 90°, bei Frauen und Kinder etwa 120° [116]. Die *Membrana thyrohyoidea* (Membran zwischen Schildknorpel und Zungenbein) befestigt den *Larynx* (Kehlkopf) am *Os hyoideum* (Zungenbein) und enthält die Eintrittsstelle für die *Arteria laryngea superior* (obere Kehlkopfarterie) und den *Nervus laryngeus superior* (oberer Kehlkopfnerv) (vgl. Abbildung 2). Über das *Ligamentum cricothyroideum* (Band zwischen Ring- und Schildknorpel) und über das Scharniergelenk *Articulatio cricothyroidea* (Ring- und Schildknorpelgelenk) stehen Ringknorpel und Schildknorpel miteinander in Verbindung (vgl. Abbildung 2). Die Membran sorgt für eine Festigung, das Gelenk wird von einer losen Kapsel umgeben und ermöglicht dadurch eine Längenvariation im Kehlkopfdurchmesser.

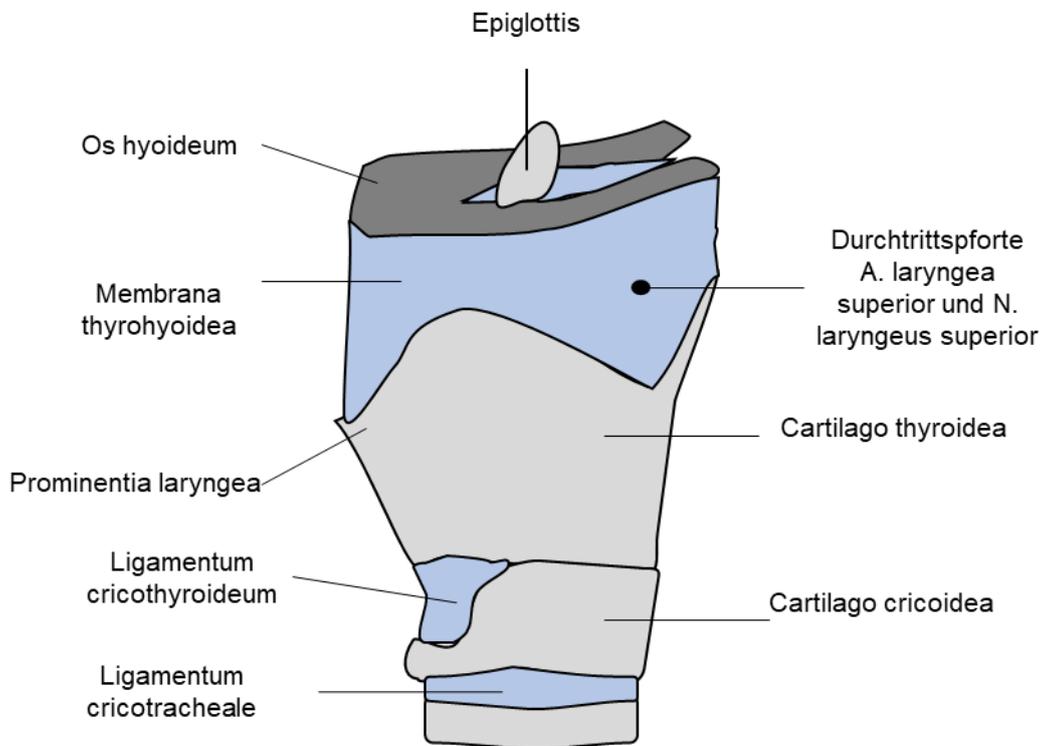


Abbildung 2: Schematische Darstellung des Kehlkopfes in der Schrägansicht von links-ventral in Anlehnung an [95].

Die Hauptfunktion des *Cartilago arytenoidea* (Stellknorpel) ist das Positionieren beider Stimmlippen. In Abbildung 3 sind die Stellknorpel eingezeichnet. Ihre dreieckige Form hat einen ventralen Anteil (*Processus vocalis*), an dem die Stimmlippen bzw. der *Musculus vocalis* (Stimmuskel) ansetzen und einen lateralen Anteil (*Processus muscularis*), an welchem die Muskeln ansetzen. Mit Hilfe einer wiederum losen Gelenkkapsel um das *Articulatio cricoarythenoidea* (Ring- und Stellknorpelgelenk), werden durch die Innervation unterschiedlicher Muskeln die Stimmlippen in unterschiedliche Positionen gebracht. Durch Kippbewegungen der Apex (Spitze der Aryknorpel) kann die Länge der Stimmlippen variiert werden. Translationsbewegungen ermöglichen das Flüstern, Rotationsbewegungen an den Stellknorpel sind für das Öffnen und Schließen der Stimmlippen verantwortlich [3].

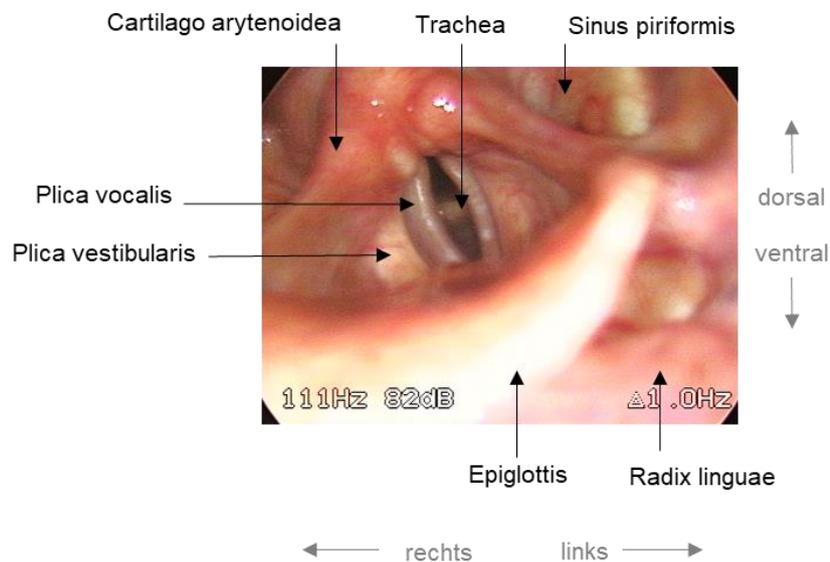


Abbildung 3: Stroboskopisches Einzelbild eines männlichen, stimmgesunden Probanden, aufgenommen während der Phonation mit einer Frequenz von 111 Hz und 82 dB.

Die für die Phonation relevanten Stimmlippen sind paarig angelegt und bilden in ihrer Mitte die *Rima glottis* (Stimmritze) [1]. Die Schleimproduktion der mukösen Drüsen versorgt die Taschenfalten mit ausreichend Feuchtigkeit. Die Länge der Stimmlippen beträgt bei Männern durchschnittlich 28 mm, bei Frauen etwa 17 mm [3]. Der Aufbau aus elastischem Plattenepithel, Bindegewebe und quergestreifter Muskulatur (*Musculus vocalis*) ist Voraussetzung für die Stimmbildung.

Das *Ligamentum vocale* wird als Stimmband bezeichnet. Die *Lamina propria* (Bindegewebsschicht unter dem Plattenepithel) ist aus 3 Schichten aufgebaut. Die Oberste bildet den sogenannten Reinke-Raum. Die eigentlichen Stimmlippen bestehen aus der mittleren und der untersten Schicht der *Lamina propria*, die mit dem *Musculus vocalis* (Stimmuskel) verbunden sind [93].

Die Stimmlippen werden in zwei Abschnitte gegliedert, die durch getrenntes oder gemeinsames Aktivieren die Stimmintensität beeinflussen können. Die *Pars*

intermembranacea (Anteil zwischen den Membranen) bildet 3/5 der Stimmbänder, die vor dem *Processus vocalis* (Vorsprung an dem der *Musculus vocalis* ansetzt) liegen. Die hinteren 2/5 reichen vom *Processus vocalis* bis zur hinteren Kommissur und werden als *Pars intercartilaginea* (Anteil zwischen den Knorpeln) bezeichnet. Das Flüstern beispielsweise entsteht durch die Aktivierung, das heißt durch das alleinige Öffnen des hinteren Anteils [3]. Das Verhältnis 2:1 zwischen männlicher und weiblicher Sprechstimmlage lässt sich durch die Längendifferenz in den mittleren Zonen der Stimmlippen erklären, welche beim Mann nahezu doppelt so lang sind [30].

Muskeln

Im Aktivierungsprozess der Stimmlippen und der Phonation spielen die Muskeln und ihre innervierenden Nerven die Hauptrolle [3].

Der einzige Außenmuskel, der *Musculus cricothyroideus* (Muskel zwischen Ring- und Schildknorpel) wird ausschließlich vom *Nervus laryngeus superior* (Oberer Kehlkopfnerve) innerviert. Dieser Muskel liegt ventral zwischen dem Schild- und Ringknorpel und ermöglicht durch die Stellungsänderung der beiden Knorpel zueinander eine Veränderung der Stimmlippenlänge und sorgt für Spannungsauf- und abbau [32].

Alle anderen Kehlkopfmuskeln sind Binnenmuskeln und werden vom *Nervus laryngeus recurrens* (zurücklaufender Kehlkopfnerve) innerviert, welcher wie der *Nervus laryngeus superior* ein Ast des *Nervus vagus* (Vagusnerv) ist [116]. Einer der wichtigsten Muskeln ist der *Musculus cricoarytenoideus posterior* (hinterer Ring- und Stellknorpelmuskel), der die Stimmlippen abduziert (vgl.

Abbildung 4). Dies ist der einzige Muskel, der die Stimmlippen öffnen kann, indem er durch einen Zug nach innen die Ary-Knorpel nach außen dreht [22]. Eine einseitige *Nervus laryngeus recurrens* Läsion führt zu Heiserkeit, da nun eine Stimmlippe dauerhaft geschlossen ist. Eine beidseitige Recurrensparese ist aus diesem Grund eine akut-lebensbedrohliche Situation, da der geschlossene Larynx einen Luftstrom in die Lunge verhindert. Der *Musculus cricoarytenoideus lateralis* (seitlicher Ring- und Stellknorpelmuskel) ist zum oben Genannten der Gegenspieler und bewirkt eine Adduktion der Stimmlippen (vgl.

Abbildung 4). Er ermöglicht daher ein Schließen der Stimmlippen. Der *Musculus thyroarytenoideus* (Schild- und Stellknorpelmuskel) ermöglicht eine Verdickung und Verkürzung der Stimmlippen und den Verschluss der Glottis. Der Schluss der Stimmlippen wird auch durch die Muskeln *Musculus arytenoideus transversus* und *Musculus arytenoideus obliquus* (querlaufender und schräglaufender Stellmuskel) ermöglicht. Der *Musculus vocalis* ist ein Bestandteil der Stimmbänder und bewirkt durch Spannungsaufbau den Klangcharakter der Stimme. Die Schutzfunktion der Epiglottis beim Schluckvorgang wird durch den *Musculus*

aryepiglotticus (Muskel zwischen Stellknorpel und Kehledeckel) verstärkt, da dieser den Larynxeingang zusätzlich verengt [3].

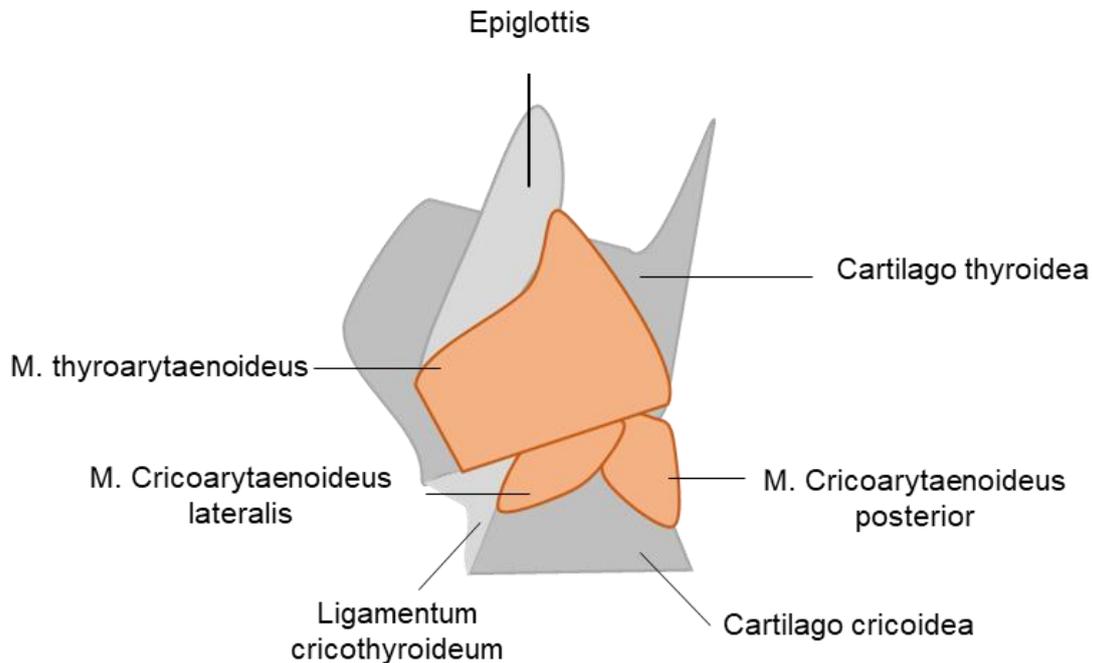


Abbildung 4: Schematische Darstellung einiger muskulärer Strukturen im Bereich des Kehlkopfes nach [95]. Ansicht von links, linke Hälfte des Schildknorpels entfernt und Sicht auf dorsal liegende Strukturen.

2.2 Grundlagen der Stimmproduktion

Im folgenden Abschnitt wird zuerst die Stimmstehung mittels der myoelastischen-aerodynamischen Theorie erklärt. Anschließend wird das Quelle-Filter-Modell beschrieben, welches eine maßgebliche Rolle in der Stimmproduktion spielt.

Physiologie der Stimmstehung

Bei der Stimmbildung handelt es sich um einen komplexen, mechanischen Vorgang, der durch Schwingungen der Stimmlippen entsteht [111]. Die an der Stimmbildung beteiligten Organe sind die Lunge, der Kehlkopf als Klangquelle und der Vokaltrakt als Resonanzraum, der eine Transfer- und Filterfunktion besitzt [27, 113].

Die menschliche Stimme entsteht, wenn Luft aus den Lungen durch die Stimmlippen hindurch strömt und die bewegungsfähigen Anteile der Stimmlippen durch unterschiedliche Atemdrucke passiv bewegt werden. Durch Modulation des Luftstroms in Folge der Stimmlippenschwingung und der Stimmlippenspannung entsteht der Primärklang im Kehlkopf [105].

Auf Stimmlippenebene findet folgendes statt: Vor der Phonation werden die Stimmlippen adduziert, dadurch sammelt sich Ausatemstrom unterhalb der Stimmlippen [116]. Der Larynx wechselt aus seinem Respirations- in den Phonationszustand. Der Prozess der Entstehung von Stimmlippenschwingungen lässt sich mittels der myoelastischen-aerodynamischen

Theorie von Van de Berg beschreiben [113]. Laut diesem Denkansatz beeinflussen drei Kräfte den Stimmlippenschluss: Der Bernoulli Effekt, die myoelastischen Rückstellkräfte und die aerodynamischen Komponenten [109]. Der Bernoulli Effekt beschreibt einen Druckabfall der durch eine höhere Strömungsgeschwindigkeit entsteht [102].

Werden nun die Stimmlippen durch den zunehmenden subglottischen Druck geöffnet (vgl. Abbildung 5 a), strömt die angestaute Luft entlang der medialen Fläche der Stimmlippen [22]. Durch die beschleunigte Luftsäule über der Glottis entsteht ein Überdruck, der die Stimmlippen weiter auseinander drängt. Aufgrund ihrer myoelastischen Eigenschaften lassen sich die Stimmlippen vom Luftstrom verformen (vgl. Abbildung 5 b). Der verringerte Querschnitt auf Stimmlippenebene im Vergleich zur *Trachea* (Luftröhre) und das gleichbleibende Luftvolumen provozieren zunächst eine steigende Strömungsgeschwindigkeit des Luftstroms (vgl. Abbildung 5 b) [102]. Im Verlauf folgt der schnell ausströmenden Luft ein Druckabfall (Bernoulli Effekt), der subglottale Druck (aerodynamische Komponente) sinkt [69] und aufgrund der myoelastischen Eigenschaften der Stimmlippen resultieren elastische Rückstellkräfte, durch welche die Stimmlippen wieder geschlossen werden [69, 77] (vgl. Abbildung 5 c). Da die Bernoulli-Kräfte zusammen mit der Rückstellkraft die Kraft des subglottalen Druckes übertreffen, schließt sich die Glottis und der Zyklus beginnt erneut (vgl. Abbildung 5 d). Bei ausbleibender Strömung herrscht kein Unterdruck mehr, Luft sammelt sich erneut unterhalb der Glottis, die Stimmlippen werden wieder geöffnet [93]. Die zyklische Aktivierung der Stimmlippen gliedert den Luftstrom und das Aufeinandertreffen der Stimmlippen erzeugt Schall [77], der Primärschall entsteht [22].

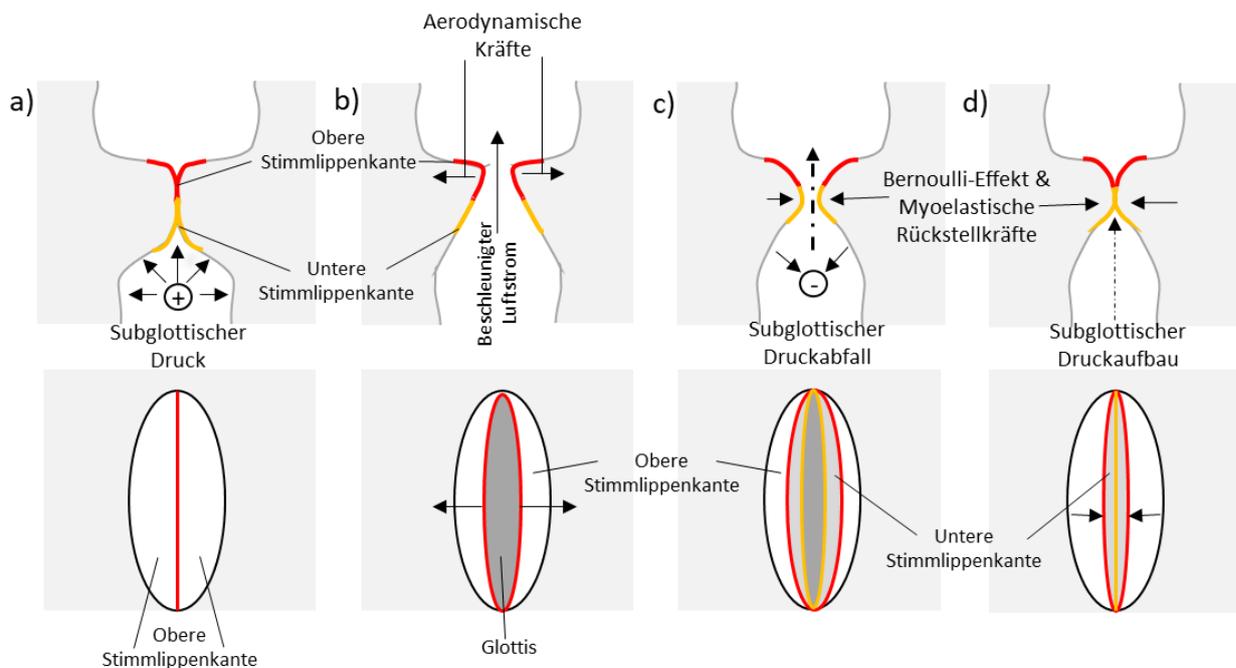


Abbildung 5: Stimmlippenebene während der Phonation im sagittalen Querschnitt (obere Bildreihe) und kranialer Aufsicht (untere Bildreihe). a) Subglottischer Druckaufbau bei noch geschlossenen Stimmlippen. b) Beginn der Phonation und Öffnen der Stimmlippen durch aerodynamische Kräfte. c) Subglottischer Druckabfall und Beginn

des Schließungsprozesses durch Bernoulli-Effekt und myoelastische Rückstellkräfte. Untere Stimmlippenkante wird sichtbar. d) Kompletter Verschluss der Stimmlippen beginnend an der unteren Stimmlippenkante.

Der Glottisverlauf über die Zeit ist in Abbildung 6 anhand von Bildausschnitten einer Hochgeschwindigkeitssequenz während der Phonation veranschaulicht.

Durch den subglottischen Druck wird zuerst die untere Stimmlippenkante und dann die obere Stimmlippenkante auseinander gedrängt (vgl. Abbildung 5 b), daraufhin werden die Stimmlippen nach lateral geöffnet [93]. Wie in Abbildung 6 gezeigt, öffnen die Stimmlippen zuerst im vorderen Drittel der Stimmlippen. Daraufhin folgt erst die Öffnung der restlichen Anteile. Die Öffnung verläuft demnach entlang der Stimmlippen von ventral nach dorsal (vgl. Abbildung 6). Die maximale Öffnung, welche im Rahmen der größten Strömungsgeschwindigkeit entsteht, wird als Amplitude bezeichnet und ist für die Lautstärke der Stimme maßgeblich [77]. Durch die myoelastischen Rückstellkräfte und den abnehmenden Druck beginnt nach maximaler Öffnung die Schließungsphase. Im Gegensatz zur Öffnungsphase beginnt diese zuerst an der unteren Stimmlippenkante und setzt sich dann bis zur oberen Stimmlippenkante fort (vgl. Abbildung 5 c und d).

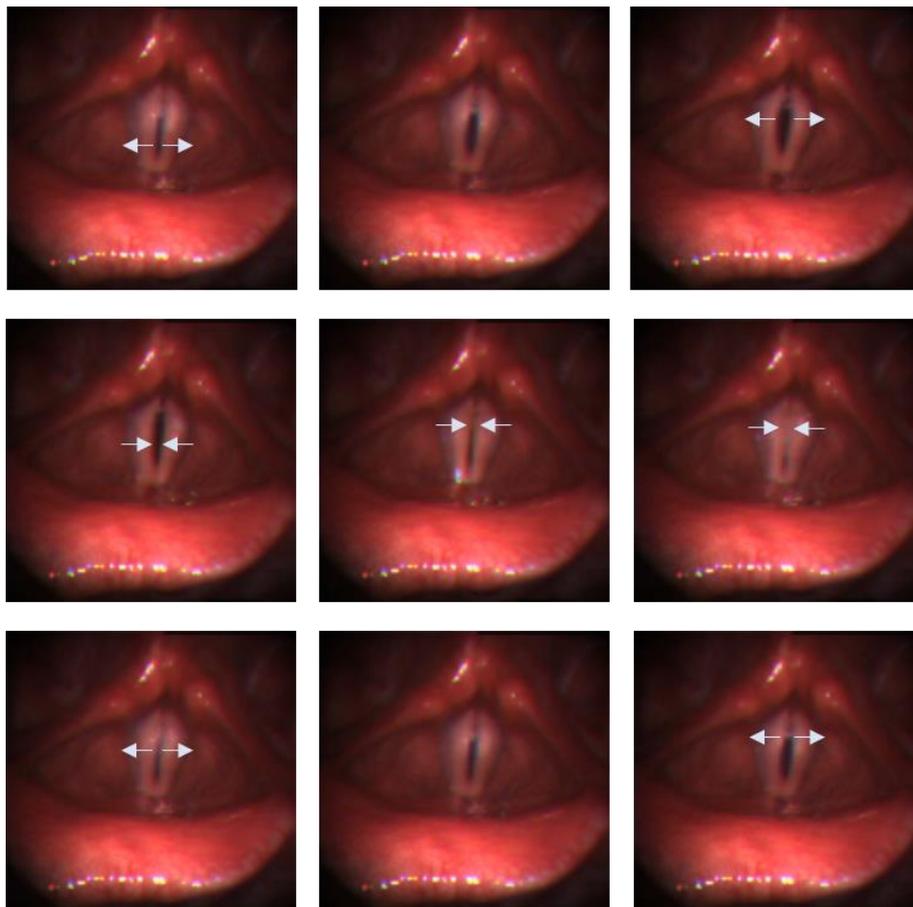


Abbildung 6: Darstellung des Glottisverlauf über die Zeit anhand von Einzelbildern aus einer Hochgeschwindigkeitssequenz, aufgenommen während der Phonation. Die weißen Pfeile geben die Bewegungsrichtung der Stimmlippen an.

Neben der Amplitude ist die Stimmfrequenz maßgeblich für den Stimmklang. Die Grundfrequenz ($F_0 = 1/T$) lässt sich aus der Periodendauer ermitteln und beschreibt die Anzahl

Schwingungszyklen pro Sekunde in Hertz ($Hz = 1/s$). Sie repräsentiert hierbei die Höhe (Pitch) des Stimmklanges. Amplitude und Periodendauer sind am Beispiel einer periodischen Sinusschwingung in Abbildung 7 schematisch erklärt. Abbildung 7 b) soll die Veränderung der Werte zeigen, wenn der Stimmklang lauter und heller wird. Stimmfrequenz und Stimmintensität beeinflussen nicht nur den Stimmklang, sondern auch die Dauer des Glottisschlusses. Eine zunehmende Frequenz hat eine steigende Anzahl von Schwingungszyklen und eine zunehmende Stimmlippenspannung zur Folge, der Glottisschluss wird immer kürzer. Anders verhält es sich bei einer zunehmenden Intensität der Stimme. Der glottale Luftstrom nimmt zu und die Amplitude und die Randkantenverschiebung sind deutlich sichtbar [116]. Durch die erhöhte Stimmlippenspannung und größeren Rückstellkräfte, schließt sich die Glottis schneller. Daraus resultiert ein längerer Glottisschluss [77].

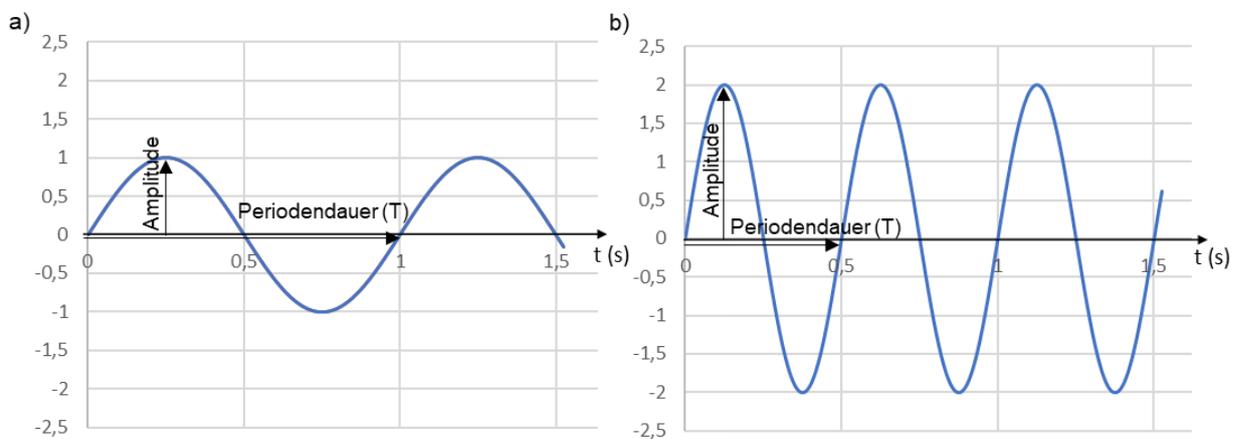


Abbildung 7: Vereinfachte Darstellung von Frequenz- und Amplitudenänderung anhand periodischer Schwingungen. Die Auslenkung der Schwingungen wird als Amplitude bezeichnet, die Periodendauer ist die Dauer eines Schwingungszyklus. a) Die Frequenz dieser Schwingung liegt bei 1 Hz. Die Amplitude liegt bei eins. b) Die Periodendauer ist aufgrund einer höheren Frequenz deutlich verkürzt, zusätzlich ist beispielhaft die Amplitude der Schwingung verdoppelt.

Quelle-Filter-Modell

Die entstandenen Schwingungen werden in der Luft als Schallwellen fortgeleitet und werden akustisch wahrnehmbar. Der primäre Kehlkopftton setzt sich aus einer Grundfrequenz und ihren Obertönen zusammen. Bei den Obertönen handelt es sich um Frequenzanteile, welche einem ganzzahligen Vielfachen einer Grundfrequenz entsprechen. Sie entstehen durch den glottalen Schluss während eines Schwingungszyklus [93] und bestimmen die Klangfarbe der Stimme [28, 91]. Die Frequenzzusammensetzung des primären Kehlkopftons ist in Abbildung 8 a) als Spektrum dargestellt [28].

Der Kehlkopfklang gelangt über Schallweiterleitung in den Vokaltrakt, der aus flexiblen und starren Anteilen aufgebaut ist. Flexibel sind der Kehlkopfeingang, der Rachen, die Mundhöhle und der Mundvorhof. Zu den starren Anteilen zählen die Nasenhaupt- und Nebenhöhlen [93]. Die Modifikation des primären Kehlkopfklang erklärte Fant mit seiner Filter Theorie [27]. Aufgrund der akustischen Eigenschaften des Vokaltrakts erfolgt eine Filterung des primären

Stimmsignals. Hierbei werden infolge der spezifischen Vokaltraktkonfiguration Frequenzanteile verstärkt bzw. unterdrückt. Dieses Verhalten wird als Resonanz bezeichnet [28, 93, 116]. In Abbildung 8 b) ist beispielhaft eine derartige Filterfunktion gezeigt. Nur die Frequenzen 300 Hz, 1300 Hz und 2300 Hz des Kehlkopftones wurden durch das Phänomen der Resonanz weitergeleitet. Entsprechend dieser Funktion wird das Spektrum des primären Stimmsignals multiplikativ gewichtet (vgl. Abbildung 8 c) und als akustisches Stimmsignal über die Lippen abgestrahlt. Die fortgeleiteten Frequenzen stellen sich in einem Frequenzspektrum als Energiemaxima dar. Diese Maxima werden als Formanten (F1, F2, F3) bezeichnet (vgl. Abbildung 8 c). Die Formanten sind anhand ihrer Amplitude, ihrer Bandbreite und ihrer Frequenzmaxima definiert [28, 116]. Diese Energie wird in Form von Druckwellen weitergeleitet und ein hörbarer Klang entsteht. Die Lautstärke des Klangs entspricht dem Schallpegel, welcher in dieser Studie in *dB* (SPL) gemessen wird.

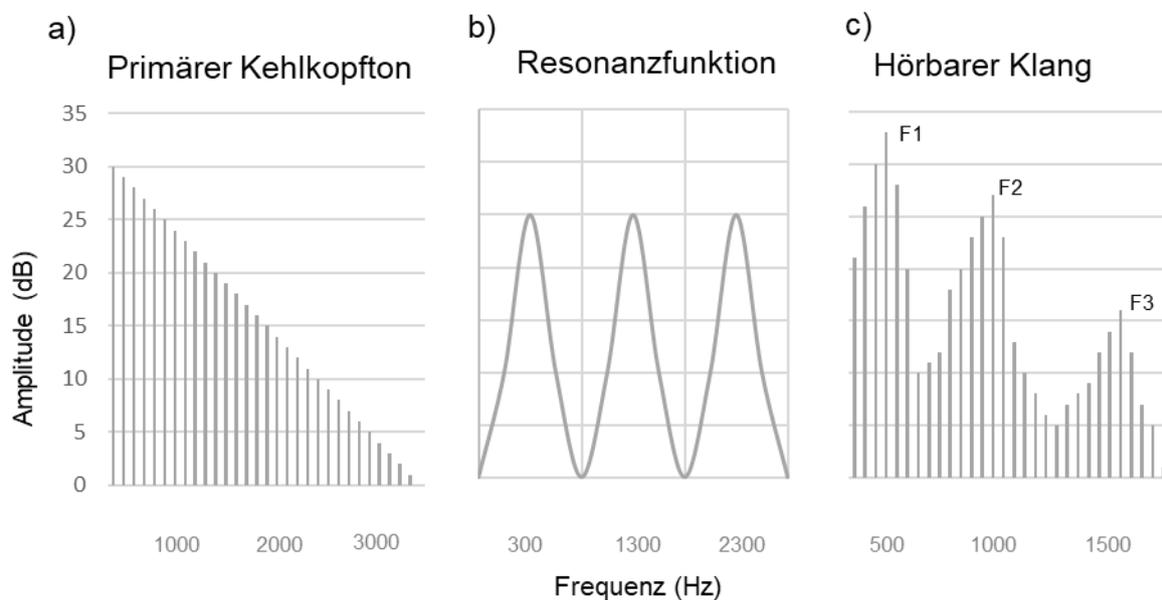


Abbildung 8: Schematische Darstellung der Entstehung des hörbaren Klangs aus dem primären Kehlkopftone. a) Primärer Kehlkopftone als Spektrum. b) Resonanzfunktion des Ansatzrohres und Verstärkung nur bestimmter Frequenzmaxima. c) Fortgeleitete Frequenzen als Energiemaxima. Entstehung von Formanten (Anlehnung an Abb 6.4. [12]).

Die Variabilität im anatomischen Aufbau des Sprachrohres ist maßgeblich für den individuellen Klang einer Stimme, da die Schallwellen unterschiedlich reflektiert werden und so unterschiedliche Formanten entstehen. Bei dem beispielsweise kürzeren weiblichen Vokaltrakt entstehen höhere Formanten und dadurch eine hellere Stimme. Dies zeigt, dass der Vokaltrakt mit seiner Resonanzfunktion eine Hauptrolle in der Variationsvielfalt des Stimmklang spielt.

2.3 Pathologien der Stimme

Stimmlippenerkrankungen werden als Dysphonien bezeichnet [116]. Die Heiserkeit ist das Leitsymptom einer Dysphonie und beschreibt eine Verschlechterung des Stimmklanges, welche beispielsweise durch unregelmäßige Stimmlippenschwingungen bedingt ist [114]. Dadurch entstehen pathologische Nebengeräusche, die zu einem rauhen oder behauchten Stimmklang führen können. Neben der Heiserkeit gelten Missempfindungen wie Kratzen und Räusperzwang oder eine verminderte Belastbarkeit der Stimme [28] als zwei weitere Hauptsymptome einer Dysphonie [114]. Aufgrund der Stellung als Leitsymptoms einer Dysphonie, sollte die Ursache und der Ursprung der Heiserkeit untersucht werden. Ursächlich für Heiserkeit können organische und funktionelle Stimmstörungen sein, die im Folgenden näher beschrieben werden [114].

Organische Dysphonien

Im Verlauf werden einige der häufigsten organischen Ursachen genannt, die Dysphonien auslösen können. Dazu zählen die akute und chronische Laryngitis, das Reinke-Ödem und benigne und maligne Veränderungen, wie der Larynxpolyp und das Larynxkarzinom.

Zu den entzündlichen Erkrankungen zählen die akute und chronische Laryngitis. Diese werden meist durch virale Infekte ausgelöst [116]. Differentialdiagnostisch sollte auch an eine durch Überanstrengung ausgelöste Hyperämie, der „Arbeitshyperämie“ [93], gedacht werden. Die Stimmlippen können gerötet und geschwollen sein. Eine Studie hat gezeigt, dass Heiserkeit bei verminderter oder aufgehobener Regelmäßigkeit und Beweglichkeit der Stimmlippenschwingung auftritt [79]. Der Stimmklang ist oft tiefer, da die Stimmfrequenz aufgrund einer Stimmlippenschwellung herabgesetzt ist.

Der Ursprung chronischer Kehlkopfentzündungen sind chronische Schleimhautprozesse, die auf dem Boden von akuten Laryngitiden entstehen können. Diese chronischen Entzündungsprozesse können durch Gase, Rauch, Staub oder auch nur physiologische Säure entstehen [116].

Ein saurer Reflux bei GERD (Gastro-ösophago-pharyngealer Reflux) kann die Larynxschleimhaut schädigen. Saurer Magensaft steigt hierbei unkontrolliert entlang der Speiseröhre nach oben, da der Ösophagusphinkter unzureichend schließt und schädigt die Schleimhaut des Larynx. Hauptsächlich äußern die Patienten Schmerzen, ein retrosternales Brennen und ein Kloßgefühl im Hals. Rauigkeit, Heiserkeit und Behauchtheit sind in unterschiedlichem Ausmaße vertreten [101].

Eine weitere Pathologie in diesem Zusammenhang ist das Reinke-Ödem, das vor allem bei Frauen zwischen 40 bis 60 Jahren auftritt [8] (vgl. Abbildung 9). Durch Gefäßschäden ist die Permeabilität der Schleimhaut der Stimmlippen erhöht und es bildet sich ein Ödem aus, welches sich aufgrund von Lymphabflussstörungen nicht zurückbilden kann [24]. Diese

Einlagerung verursacht Heiserkeit und Virilisierung der Stimme, was vor allem bei Frauen akustisch auffällt [63].

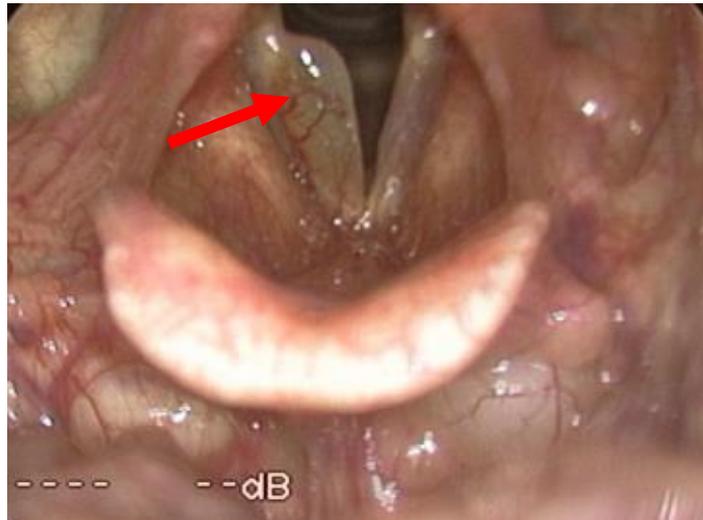


Abbildung 9: Darstellung eines Reinke-Ödems an der rechten Stimmlippe. Aufnahmen aus dem Patientenkollektiv von Dr. Bader (HNO-Klinik, Universitätsklinikum des Saarlandes).

Als postinfektiöse Stimmlippenveränderungen wird der *Sulcus vocalis* (Furchenbildung der Stimmlippen) diskutiert. Die genaue Genese ist noch nicht eindeutig geklärt, ein Überlegungsansatz ist eine vermutete Narbenbildung, die zu einem Verlust der Randkantenverschieblichkeit und des Schwingungsvermögens der Stimmlippen führt. Daraus resultiert eine heisere und rauere Stimme [93].

Dysphonien können auch durch maligne Veränderungen ausgelöst werden. In Abbildung 10 a) ist das Bild eines Larynxkarzinoms gezeigt. Die Differenzierung zwischen Larynxkarzinom und Präkanzerose erfolgt über die erhaltene Stimmlippenmobilität in der Stroboskopie. Eine infiltratives Wachstum, welches bei einem Karzinom vorliegt, verändert die Stimmlippenbeweglichkeit [93]. Karzinome im Bereich des Kehlkopfes sind am häufigsten Plattenepithelkarzinome [116]. Zwei Drittel aller Larynxkarzinome liegen in der glottalen Ebene. Länger bestehende Heiserkeit sollte daher im Rahmen einer Bildgebung abgeklärt werden [93]. Maligne Larynxläsionen dürfen jedoch nicht mit der benignen Larynxpapillomatose verwechselt werden. Diese weisen auch ein raumforderndes Wachstum auf (vgl. Abbildung 10 b). Die häufigsten Erreger sind HPV-6 und HPV-11, die ein geringes Entartungsrisiko besitzen [93].

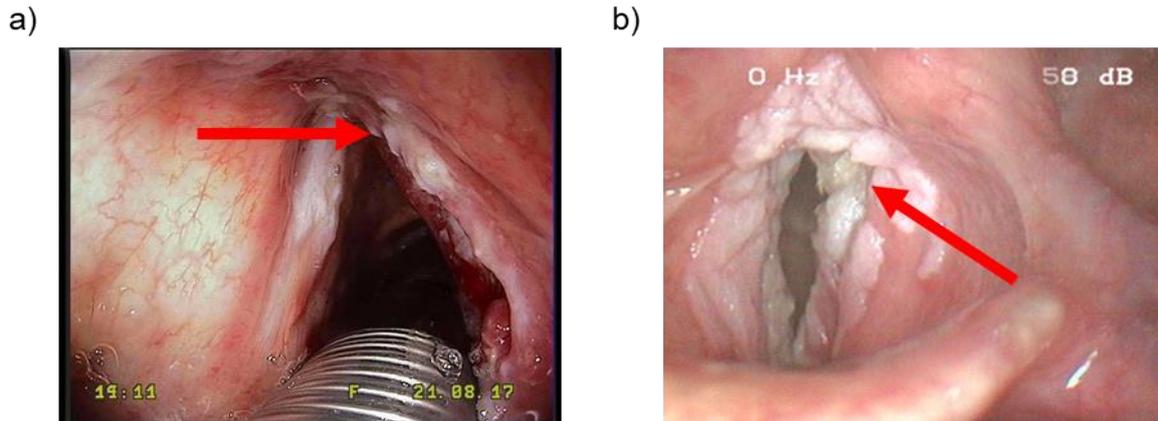


Abbildung 10: Maligne und benigne Larynxläsionen im Vergleich a) Larynxkarzinom an der rechten Stimmlippe. b) Larynxpapillomatose als Differentialdiagnose zur malignen Entartung. Aufnahmen aus dem Patientenkollektiv von Dr. Bader (HNO-Klinik, Universitätsklinikum des Saarlandes).

Gutartige Stimmveränderungen sind bei mehr als der Hälfte der Stimmpatienten Grund für Heiserkeit [8]. Zu den benignen Stimmlippenumtoren zählen die Stimmlippenzysten, die circa 14% dieser Gruppe ausmachen [8]. Diese entstehen entweder durch einen Stau in den Drüsen oder durch Überbeanspruchung [8]. Sie führen neben Heiserkeit zu einem Engegefühl beim Schlucken und ständigem Räuspern. Differentialdiagnostisch kann dies von Polypen und Knoten endoskopisch abgegrenzt werden, die in Abbildung 11 gezeigt werden. Polypen und Knoten zählen zu den häufigsten phonationsassoziierten Dysphonien und entstehen bei stimmlicher Überlastung, die oft zu vaskulären Einblutungen führen und ödematöse und entzündliche Veränderungen als Folge haben [8]. Polypen sind entzündliche Schleimhauthyperplasien. Knoten hingegen entstehen durch eine mechanische Überbelastung der Stimmlippen, sie stellen sich als paarig angelegte Gewebeverdickung dar [7] (vgl. Abbildung 11 a). Polypen können durch ihre Lokalisation von Granulomen, die eher im kartilaginären Drittel auftreten, unterschieden werden (Vgl. Abbildung 11 b) [8].

Für die Entstehung von Phonationsverdickungen durch mechanische Überbeanspruchung der Stimmlippen, werden zwei Theorien diskutiert. Einerseits ein zu kräftiger Glottisschluss aufgrund einer Hyperfunktionalität, andererseits eine strömungsbedingte Überlastung der Stimmlippen [116].

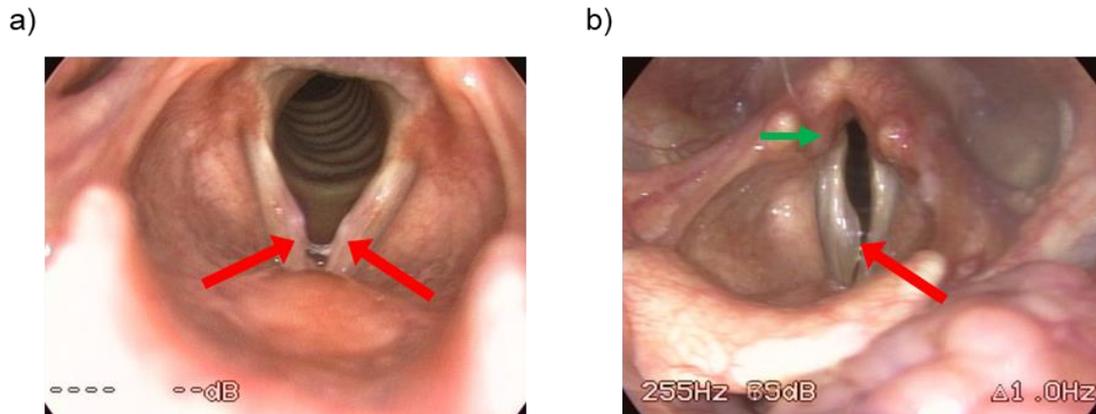


Abbildung 11: Benigne Stimmlippentumore, siehe rote Pfeile a) Stimmlippenknötchen parallel an beiden Stimmlippen. b) Stimmlippenpolyp rechts. Grüner Pfeil zeigt kartilaginären Anteil der Stimmlippen. Aufnahmen aus dem Patientenkollektiv von Dr. Bader (HNO-Klinik, Universitätsklinikum des Saarlandes).

Dysphonien können auch im Rahmen eines neuronalen Schadens entstehen. Der *Nervus laryngeus recurrens* (zurücklaufender Kehlkopfnerve) und der *Nervus laryngeus superior* (oberer Kehlkopfnerve) innervieren den Larynx. Bei Nervenschädigungen kann es entweder zu einer Parese (inkomplette Lähmung) oder zu einer Paralyse (komplette Lähmung) der betroffenen Stimmlippe kommen. Die Art der Beschwerden hängt von der Stellung der gelähmten Stimmlippen ab. Steht bei einseitiger Recurrensparese die beeinträchtigte Stimmlippe in der Medianstellung, sind Sprachbildung und Atmung noch unauffällig (vgl. Abbildung 12). Ist die Parese hingegen beidseitig, kommt es zur schweren Atemnot, da beide Stimmlippen durch die fehlende Innervation nicht angespannt werden können und die *Rima glottis* (Stimmritze) verschließen. Befindet sich die gelähmte Stimmlippe in der Lateralposition, so ist die Atmung weitgehend unbehindert, die Sprachbildung jedoch schwer gestört, da keine Spannung mehr aufgebaut werden kann. Zudem ist die Aspirationsgefahr erhöht, da die protektive Funktion des Larynx durch einen kompletten Verschluss der Stimmlippen beim Schlucken ausfällt. Eine Läsion des *Nervus laryngeus superior* (oberer Kehlkopfnerve) führt hingegen nur zu einem Verlust des Spannungsaufbaus durch den *Musculus cricothyroideus* (Muskel zwischen Ring- und Schildknorpel) und somit zu einem verminderten Tonumfang in hoher Stimmlage [93].

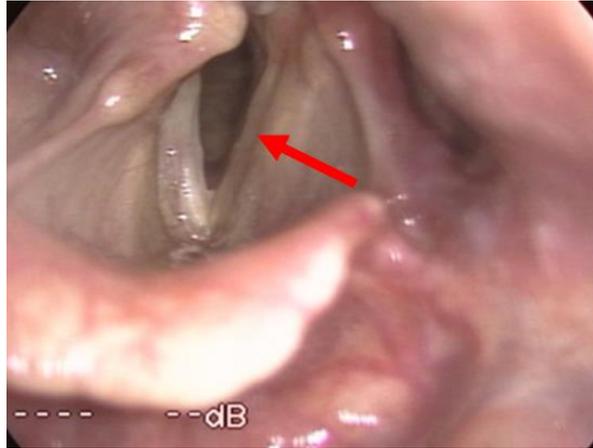


Abbildung 12: Stimmlippenparese links mit atrophischen Veränderungen. Aufgrund eines länger bestehenden Funktionsverlusts deutlich atrophierte linke Stimmlippe im Vergleich zur rechten Stimmlippe. Aufnahmen aus dem Patientenkollektiv von Dr. Bader (HNO-Klinik, Universitätsklinikum des Saarlandes).

Funktionelle Dysphonien

Funktionelle Dysphonien können durch hyper- und hypofunktionellen Störungen entstehen [93]. Bei einer hypofunktionellen Stimmstörung herrscht eine zu geringe Muskelspannung. Der Stimmcharakter ist sowohl behaucht, als auch klangarm, Rauigkeit hingegen findet sich selten. In der Stimmfeldmessung fällt oft ein verminderter Schalldruckpegel und ein reduziertes Frequenzspektrum auf. Die fehlende Muskelspannung kann sich auch ganzheitlich an der Körperhaltung bemerkbar machen [116]. Die invasive Bildgebung zeigt oft keine schwerwiegende Auffälligkeit. Die Hyperfunktion weist ebenfalls ein vermindertes Stimmfeld auf, da aufgrund des starken Muskeltonus die Bildung leiser Töne nicht möglich ist und so die unteren Schalldruckpegel nicht erreicht werden können. Die Untersuchung mit einem Stroboskop kann das Bild einer Hyperfunktion vortäuschen [93]. In dieser Situation kann die Anspannung des Patienten oder die unangenehme Fixierung der Zunge durch den Untersucher eine hyperfunktionelle Störung vortäuschen und die Ergebnisse verfälschen. Stroboskopiebefunde sind bei der Abklärung funktioneller Dysphonien deshalb oft wenig aufschlussreich. Bei der Hyperfunktion klingt die Stimme gepresst, heiser und vor allem rau. In der Anamnese klagen die Patienten über Stimmermüdung, Kratzen im Hals und vermehrtes Räuspern [93]. In der Bildgebung können sich durch eine erhöhte Muskelspannung sekundäre organische Veränderungen, wie Stimmlippenknoten zeigen [116].

2.4 Bildgebende Diagnostik des Larynx

Für die visuelle Darstellung des Larynx gibt es mehrere Möglichkeiten. Zu den Routineuntersuchungen zählt die Laryngoskopie. Hierbei werden die Stimmlippen mit einem Kehlkopfspiegel indirekt eingesehen. Eine Beurteilung der Stimmlippenbewegungen ist mit der Untersuchung nicht möglich, da die Stimmlippen in der mittleren Sprechstimmlage mit einer Frequenz von 80-400 Hz schwingen können, der Mensch jedoch nur bis zu 15 Bilder/s erkennen kann [116]. In der klinischen Diagnostik lassen sich Stimmlippenschwingungen

mittels Stroboskopie, Videokymographie, und Hochgeschwindigkeitsvideoendoskopie untersuchen.

2.4.1 Stroboskopie

Bei der Stroboskopie werden die Stimmlippen während der Phonation mittels eines Laryngoskops, das unter Fixierung der Zunge in den Oropharynx eingeführt wird, gefilmt. Zudem ist eine Beurteilung der umliegenden Strukturen, wie Kehldeckel und Taschenfalten, möglich [12]. Neben dem Endoskop besteht die Stroboskopieeinheit aus einem Mikrofon, einer Kamera und aus einer Speichermöglichkeit [93]. Die Darstellung der Stimmlippenschwingung basiert auf dem „stroboskopischen Effekt“ [85, 85]. Die Frequenz des gesungenen Tons bei der Phonation wird mit Hilfe eines Mikrofons am starren Endoskop ermittelt. Die Schwingungen werden anschließend mit einem Blitzlicht beleuchtet, dessen Frequenz sich minimal von der Grundfrequenz der Stimmlippenbewegung (meist 1-2,5 Hz) unterscheidet [116]. So entstehen einzelne „Fotos“ zu unterschiedlichen Zeitpunkten des Schwingungszyklus. Aus den einzelnen Bildern entsteht eine geschnittene, fiktive Version der realen Schwingung, die sich gut beurteilen lässt [112]. Die Bewertung der Aufnahmen orientieren sich an dem Protokoll der European Laryngological Society (ELS) zur Stimmdiagnostik [18, 29]. Bereits frühere Studien ergaben ähnliche Untersuchungsansätze [94]. Die Interpretation des Bildbefundes sollte immer in Hinblick auf Frequenz und Lautstärke erfolgen [12]. Bei der Beurteilung der Stimmlippenschwingungen werden mehrere Parameter berücksichtigt:

- a) Die maximale Amplitude: Das entspricht dem Öffnungsdurchmesser, der bei lauten Tönen größer und bei leisen Tönen kleiner wird.
- b) Die Randkantenverschieblichkeit: Diese entspricht der Beweglichkeit des Epithels an der Stimmlippenkante. Eine verminderte Verschieblichkeit beispielsweise führt zu einer erhöhten Trägheit der Stimmlippen und verursacht einen längeren Stimmlippenschluss.
- c) Der vollständige Glottisschluss: Bei einer Schlussinsuffizienz kommt es zu einem unregelmäßigen Luftstrom. Die Stimme erhält einen behauchten Klang.
- d) Die Periodizität und Regelmäßigkeit der Schwingung: Unregelmäßige und asymmetrische Bewegungen verursachen einen rauen Stimmcharakter.
- e) Die Symmetrie der Stimmlippenschwingung: Bewegen sich beide Stimmlippen asymmetrisch zueinander, wird die Stimme heiser.

Auf eine entspannte Untersuchungs-Atmosphäre sollte geachtet werden, um eine überhöhte Stimmlippenspannung und damit verbundene unregelmäßige Stimmlippenschwingungen zu vermeiden. Hier liegen die Grenzen der Stroboskopie, eine entspannte Untersuchungssituation kann in der klinischen Praxis nicht immer gewährleistet werden. Die Befunde können durch angespannte Haltung verfälscht werden. Außerdem können maximal

nur zwischen 25 und 50 Bilder/s aufgenommen werden, für eine reelle Betrachtung der Stimmlippen sind jedoch 1000-4000 Bilder/s notwendig [93]. Da die Stroboskopie eine gängiges Untersuchungsmedium im klinischen Alltag ist, wurde das Verfahren in dieser Studie angewendet.

2.4.2 Videokymographie

Ein weiteres Verfahren zur Darstellung von Stimmlippenschwingungen ist die Videokymographie. Hierbei wird eine Linie transversal über die Stimmlippenschwingung gelegt und nur dieser Bereich in Echtzeitaufnahmen dargestellt [103, 112]. Durch die Verwendung einer Hochgeschwindigkeitstechnik lassen sich bis zu 8000 Bilder/s aufnehmen [103]. Die einzelnen Aufnahmen der Linien werden zu einem Film zusammengefasst. Die Linien sind bei Erfassung der Stimmlippen hell, bei Aufnahmen der Glottis dunkel. So reihen sich helle und dunkle Bilder zu einer zweidimensionalen Bilderabfolge zusammen. Anhand der Bilder kann das Öffnen und Schließen der Stimmlippen, die Randkantenverschieblichkeit und die Öffnungsamplitude sehr gut beurteilt und gemessen werden [49, 112, 116]. Nachteil dieses Verfahrens ist die Betrachtung der Schwingung anhand nur einer Linie, der festgelegten Kymographiezeile [19, 103]. Zudem wird das longitudinale Bewegungsmuster der Stimmlippe vernachlässigt [112, 116].

2.4.3 Hochgeschwindigkeitsvideoendoskopie

Die Hochgeschwindigkeitsvideoendoskopie ermöglicht Aufnahmen von 4000-8000 Bilder/s [93]. Zur Beurteilung der Filme werden diese in Zeitlupe mit etwa 25 Bilder/s abgespielt [112]. So können feinste Irregularitäten während einer Stimmlippenschwingung detektiert werden [26, 89]. Mit dieser Untersuchung können Stimmlippenschwingungen genauer visualisiert und besser als Stroboskopieaufnahmen beurteilt werden, da es sich hierbei im Gegensatz zur Stroboskopie nicht um eine Schein- sondern um eine reale Zeitlupenaufnahme handelt [80]. Dies ermöglicht auch das Aufzeichnen irregulärer Schwingungen [112].

Die große Datenmenge, der Zeitaufwand für die Aufnahmen und Analysen und die damit verbundenen Kosten sind derzeit der limitierende Faktor der Hochgeschwindigkeitsendoskopie, um als Standarduntersuchung in der objektiven Stimm diagnostik Einzug zu finden [84, 112].

2.5 Subjektive Stimmbewertung

Ein weiterer Aspekt der Stimmbeurteilung ist neben der apparativen Diagnostik die Bewertung der Stimme durch den Patienten und durch den Untersucher. Dysphonien beeinflussen sowohl das berufliche als auch das soziale Leben [6, 62, 100]. Aufgrund einer eingeschränkten

stimmlichen Leistungsfähigkeit, müssen Betroffene beispielsweise den Job wechseln oder können sich im sozialen Umfeld nicht wie gewünscht repräsentieren oder interagieren [6, 47]. Für die Stimmbeurteilung durch den Patienten selbst werden meist mehrstufige Antwortbögen verwendet, welche gezielte Fragen zur Stimmqualität und stimmlichen Beeinträchtigungen im Alltag enthalten. Es gibt eine Vielzahl an Fragebögen für unterschiedliche Fragestellungen. Im Voice Handicap Index (VHI) und Voice Activity and Participation Profile (VAPP) beispielsweise wird das subjektive Stimmempfinden befragt. Der Voice-related Quality of Life (V-RQOL) kann zur Evaluation von Dysphonien und von Therapieergebnissen heran gezogen werden [54], [46]. Der Voice Outcome Survey (VOS) ist ein Test, der spezifisch bei einseitiger Stimmlippenlähmung Anwendung findet [83]. Da alle der folgenden Tests geeignete Instrumente zur Beurteilung der Stimmwahrnehmung sind, sollte im Hinblick auf die Fragestellung möglichst nur ein Test verwendet werden, um eine möglichst eindeutige und vergleichbare Interpretation der Ergebnisse zu gewährleisten [54].

2.5.1 Selbstbewertung der Stimme mit VHI-Test

Der VHI findet überall in Europa Anwendung [76] und wurde entwickelt, um psychosoziale Folgen, die durch Stimmstörungen hervorgerufen werden, möglichst verständlich und einfach zu messen [47]. Zu den psychosozialen Konsequenzen zählen beispielsweise Rückzug aus dem öffentlichen Leben aufgrund von Schamgefühl über die eigene Stimme oder ein Jobwechsel aufgrund von Stimmermüdung [47]. Dieser Test berücksichtigt vor allem die funktionellen, physischen und subjektiven Beeinträchtigungen der Stimme. Im deutschen Sprachraum existieren drei Versionen des VHI mit einem Fragenkatalog aus 30, 12 oder 9 Fragen. Eine höhere Anzahl von Fragen ergab keine signifikant höhere Spezifität, Sensitivität oder Reliabilität [84], so dass als Standardtest mittlerweile der VHI-9 international verwendet wird [54]. Die Fragen beziehen sich auf das subjektive Stimmempfinden des Patienten, seiner stimmlichen Interaktion mit der Umgebung und auf die psychologische Ebene beim täglichen Stimmgebrauch [29]. Jede Antwort kann mit einer Zahl von null bis vier bewertet werden. Hierbei wird angegeben wie oft eine Frage zutrifft. Es wird zwischen immer (4), oft (3), manchmal (2), selten (1) bis zu nie (0) unterschieden. Aus den Punkten wird am Ende eine Gesamtpunktzahl errechnet. Je höher die erreichte Gesamtpunktzahl, desto auffälliger ist die Stimme und desto wahrscheinlicher liegt eine Dysphonie vor. Nawka et al. (2003) haben durch eine multizentrische Studie in Deutschland die Reliabilität dieses Tests gesichert. Ist die Gesamtpunktzahl eines VHI-12 zwischen 0 bis 6 besteht kein Verdacht auf eine vorliegende Dysphonie. Der VHI findet bei Fragestellungen zu Dysphonien Anwendung und ist ein fester Bestandteil in der Erhebung der Erstdiagnostik, sowie für Verlauf- und Abschlussuntersuchung [33].

2.5.2 Fremdbewertung

Eine alleinige Berücksichtigung der subjektiven Stimmwahrnehmung durch den Patienten bei der Diagnosestellung ist jedoch nicht ausreichend ist, da die Bewertung der Stimme durch den Betroffenen und durch den Untersucher unterschiedlich sein kann. Die Diagnose sollte immer unter Berücksichtigung beider Einschätzungen erfolgen [29], da beide Stimmbewertungen relevant und nicht durch die jeweils andere austauschbar sind [38].

Die Problematik der auditiven-perzeptiven Beurteilung im Vergleich zur objektiven Bewertung, ist die Subjektivität eines jeden Untersuchers. Standardisierte Test erlauben zuverlässige und vergleichbare Ergebnisse, die von einzelnen Untersuchern erhoben wurden [57]. Eine weitere Voraussetzung ist, dass die Untersucher in der Stimmbeurteilung geübt sind. Im deutschsprachigen Raum hat sich für die Beurteilung der Stimme durch den Untersucher die RBH-Skala durchgesetzt [93]. Die drei Qualitäten, die im RBH-Test betrachtet werden, sind die „Rauigkeit“ (R), die „Behauchtheit“ (B) und der Gesamteindruck der „Heiserkeit“ (H). Dieser Test ist zur auditiven Beurteilung der Heiserkeit geeignet. Diese Merkmale lassen sich von 0-3 in 4 Schweregrade einteilen. Wobei der Wert 0 keine Symptome, der Wert 3 hingegen eine hochgradige Störung der Stimme [82] angibt.

2.6 Apparative Stimmdiagnostik

Die objektive akustische Analyse des Stimmklangs ist ein unverzichtbarer Teil der Stimmdiagnostik [18]. Die grundlegende Annahme ist, dass Abnormitäten in der Stimmlippenstruktur und funktionelle Veränderungen durch veränderte Schwingungsabläufe eine Verschlechterung der Stimmqualität hervorrufen.

Mit Hilfe von Stimmanalysen und Software Programmen, wie „LingWAVES“ werden charakteristische Eigenschaften der Stimme quantitativ ermittelt. Die in der Klinik etablierten Parameter, wie Tonhöhe, Sprechlautstärke und die Irregularitätsparameter Shimmer und Jitter werden im Weiteren vorgestellt.

Wie auch bei der subjektiven Stimmbewertung muss in der apparativen Diagnostik auf eine korrekte Datenerhebung und auf die Wahl geeigneter Parameter geachtet werden. Die Objektivität von Ergebnissen hängt von den verwendeten Softwareprogrammen ab [29], da die Verwendung unterschiedlicher Programme derzeit verschiedene Referenzwerte ergibt [66]. Die Interpretation der Werte sollte daher unter Berücksichtigung der verwendeten Geräte und der Softwarekomponente erfolgen [92, 93, 99].

2.6.1 Stimmfeldmessung

Das Aufzeichnen des Stimmfeldes ermöglicht die Erhebung einer Vielzahl an stimmrelevanten Daten und ist daher als fester Bestandteil in der apparativen Stimmdiagnostik etabliert [29, 92, 93, 96]. Hierbei kann der Frequenzumfang (Hz) und die Lautstärke (dB) von Sprech- und

Singstimme bestimmt werden. Neben dem dynamischen Stimmumfang werden Daten, wie die mittlere Sprechstimmlage und der Tonhöhenumfang, ermittelt. Diese Werte ermöglichen einen Vergleich mit beschriebenen Referenzwerten aus der Literatur und eine Einordnung und Bewertung der Stimme. Abweichungen von den Referenzbereichen können Hinweise auf eine pathologische Stimmfunktion geben. Ein reduzierter Tonhöhenumfang beispielsweise beschreibt eine verminderte Leistungsfähigkeit der Stimme. Die Grundvoraussetzungen für diese Untersuchung sind ein schallgedämpfter Raum mit einem Geräuschpegel unter 40 dB [96] und ein konstanter Mund-Mikrofonabstand von 30 cm [96]. Für die Datenerhebung wird neben dem Mikrofon ein Programm für die Verarbeitung und Analyse der Stimme verwendet. Die Aufnahmen sollten im Stehen aufgezeichnet werden, da eine schlechte Körperspannung die Stimme negativ beeinflussen kann [12, 116]. Bei der Messung werden einzelne Töne hintereinander aufgezeichnet, die vom Programm automatisch in ein Koordinatensystem eingetragen werden. In der Abszissenachse steht die Frequenz (in Hz), in der Ordinatenachse wird der Schalldruckpegel (in dB) des Tones verzeichnet. Die einzelnen Punkte werden am Ende verbunden und ergeben bestimmte Stimmprofile. Hierbei steht das Flächenmaß der Stimmprofile jedoch in keiner Relation zur Stimme [92]. Ein Problem dieser Methode ist die fehlende Normierung. Eine Vielzahl unterschiedlich verwendeter Messprogramme ergeben abweichende Parameter [18]. Um die Problematik zu verbessern, hat die European Laryngological Society (ELS) einen Leitfaden mit unterschiedlichen Komponenten erstellt, um sich einer Standardisierung anzunähern [29]. Ungeachtet dieser Grenzen ist die Stimmfeldmessung in der apparativen Diagnostik fest etabliert, da sie ein individuelles Stimmbild für jeden einzelnen Patienten ermöglicht und eine Auswertung der Stimme so schnell erfolgen kann. Bei der Analyse der Aufnahme ist das Geschlecht entscheidend und ob es sich um einen geübten oder ungeübten Sänger handelt [118]. Geübte Sänger können ihre stimmlichen Kapazitäten mehr ausschöpfen. Bei ungeübten Sängern ist zu beachten, dass es derzeit kein Unterscheidungsverfahren zwischen ungeübten und funktionell beeinträchtigten Stimmen gibt [9, 78].

2.6.1.1 Sprechstimme

Die Messung des Sprechstimmfeldes wird in 3 unterschiedlichen Lautstärken durchgeführt. Begonnen wird mit der indifferenten Sprechstimmlage, danach folgen Aufnahmen mit einer leisen, jedoch keiner flüsternden Sprechstimme und einer lauten, jedoch keiner schreienden Vortragsstimme. Es sollten nichtemotionale Worte verwendet werden, wie Monatsnamen oder Zahlenfolgen. Aus diesen drei Sprechstimmlagen lässt sich die mittlere Sprechstimmlage ($mSsl$) mit der Einheit Hz bestimmen [92]. Die Frequenz liegt bei Frauen zwischen 164 -260 Hz und bei Männern zwischen 89-175 Hz [12]. In einer anderen Studie liegt der Bereich von Frauen zwischen 200-250 Hz und bei Männern zwischen 100-150 Hz [29]. Aus der Messung

ergeben sich unter anderem der Tonumfang beim Sprechen. Dieser liegt bei Frauen zwischen 5-16 Halbtönen, bei Männern zwischen 7-21 Halbtönen. Als pathologische Grenze gilt ein Wert unter 12 Halbtönen [29, 93]. Neben der mittlere Sprechlautstärke wird die maximale Ruflautstärke ermittelt, die Spanne liegt bei Frauen zwischen 82-98 *dB* bei Männern zwischen 85-100 *dB* [12].

2.6.1.2 Singstimme

Wie bei der Sprechstimme wird die Singstimme zunächst in der indifferenten Stimmlage, danach in leiser und anschließender mit lauter, kräftiger Singstimme aufgenommen. In den drei Lautstärken wird ein Vokal, wie zum Beispiel /a:/, eine Dur-Tonleiter absteigend und dann aufsteigend gesungen. Die Tonhaldedauer sollte möglichst 2 Sekunden betragen, da viele Stimmverarbeitungsprogramme kürzere Töne nicht registrieren können [92]. Wichtig zu beachten ist, dass eine einmalig erhobenen Stimmanalyse keine absoluten Stimmwerte ergibt [92]. Die menschliche Stimme ist nicht immer gleich, sondern leicht variabel. Eine Studie hat gezeigt, dass die Lautstärke der Grundfrequenz eines gesunder Probanden an unterschiedlichen Tagen eine durchschnittliche Standardabweichung von bis zu drei *dB* beträgt [35]. Weitere Werte der Singstimme sind die minimale und maximale Lautstärke. Diese liegen bei Frauen in etwa zwischen 46-55 *dB* und 88-100 *dB*. Bei Männern ist der Bereich von 46-57 *dB* und 88-107 *dB* [12]. Die minimalen und maximalen Frequenzen sind ein Maß für den Bewegungsraum der Stimmen. Bei Frauen liegt der Rahmen für die minimale Frequenz ungefähr zwischen 123-164 *Hz*, die maximale Spanne zwischen 587-987 *Hz*. Bei Männern liegt der Rahmen etwas tiefer, die minimale Frequenz zwischen 65-97 *Hz* und die maximale Frequenz bei 391-698 *Hz* [12].

Während der Stimmumfangsprofilmessung, können aus der Tonhaldedauer eines phonierten Vokals /a:/ zusätzlich die Irregularitätsparameter Shimmer und Jitter und der Dysphonia Severity Index (DSI), der Rückschlüsse über die Stimmqualität gibt, errechnet werden. Bei dieser Untersuchung wird der Vokal so lange wie möglich phoniert. Markiert man daraus nun einen Bereich von 3 bis 5 Sekunden, kann eine entsprechende Software die Werte für diesen Zeitraum bestimmen.

2.6.2 Akustische Irregularitätsparameter

Shimmer und Jitter geben Variationen im Stimmsignal an [12]. Der Jitter beschreibt kurzzeitige Veränderungen in der Periodenlänge, die sich als zeitliche Veränderungen in der Grundfrequenz der Stimme darstellen. Bereits frühere Erstuntersuchungen haben ergeben, dass ein Kehlkopftönen Variationen in der Wellenform und der Periodizität aufweist [61]. Der Shimmer beschreibt eine kurzzeitige Amplitudenschwankung (Lautstärkenschwankung) der Stimme [115]. Ist der Shimmer ausgeprägt, wird die Stimme brummend und knisternd

empfundener [107] [12]. Ein gesungener Ton mit einem hohen Jitter erscheint sehr unregelmäßig und instabil. Der Jitter liegt bei gesunden Stimmen zwischen 0,1 und 1 %. Die Grenze für den Normwert des Shimmers liegt bei unter 2,5 % [10, 78]. Diese Grenzwerte wurden mit dem Programm „LingWAVES“ analysiert. Eine Studie aus 2008 ergab für Analysen mit Praat einen Jitter von $\leq 1,04\%$ und für den Shimmer $\leq 3,8\%$ [70]. Eine retrospektive Studie, die zur Analyse das Programm „Dr. Speech“ verwendet hat, ergab für gesunde Männer einen Median von 0,25 % für den Jitter und 2,9 % als Median für den Shimmer [117]. Die Ergebnisse müssen jeweils in Bezug auf das verwendete Softwareprogramm analysiert und interpretiert werden [66]. Ähnliche Jitter- und Shimmernormen ergaben sich aus Analysen mit „Computerized Speech Lab“ und mit dem „Multidimensional Voice Program“ [92]. Die unterschiedlichen angegebenen Normwerte zeigen, dass eine Standardisierung dieser Werte noch nicht erfolgt ist.

Jitter und Shimmer werden von unterschiedlichen Faktoren beeinflusst. Sie stehen unter dem Einfluss der Sprechlautstärke, des Geschlechts und verhalten sich je nach Vokal unterschiedlich. Der Einfluss der Grundfrequenz F_0 war entgegen der bisherigen Annahme gering [10]. Heiserkeit kann zu erheblichen Unregelmäßigkeiten in diesen beiden Komponenten führen und als Ergebnis pathologische Werte hervorbringen [20]. Außerdem sind Jitter und Shimmer anfällig bei inkorrekten Messungen und es kann zu fehlerhaften Ergebnissen kommen [11].

Studien zeigten, dass Jitter und Shimmer allein betrachtet keine zuverlässige und gültige Aussage über einen tatsächlich gesunden Stimmstatus ergeben und in ihrer klinischen Anwendung daher begrenzt sind. Eine Kombination mit bildgebenden Verfahren und subjektiven Bewertungsbögen steigert die Aussagekraft der Werte [14].

2.6.3 Dysphonia Severity Index (DSI)

Ein weiterer Parameter der apparativen Diagnostik ist der, 1997 von Wuyts et al. entwickelte Dysphonia Severity Index (DSI) [119]. Er stellt eine objektive Messmethode für Dysphonien dar, die jedoch von den verwendeten Programmen zur Stimmfeldmessung, welche von Wuyts et al. in der Studie verwendet wurden, abhängig ist. Ein Vergleich der Ergebnisse aus „LingWAVES“ und „DiVAS“ zeigte, dass eine Standardisierung der Programme zur Erhebung vergleichbarer DSI-Werte notwendig ist [2]. Der DSI wird am Ende der Stimmfeldmessung aus der maximalen Tonhaldedauer (MPT), aus dem oberen Grenzwert des Stimmfeldes ($F_{0_{high}}$), dem leisesten Schallpegel (I_{low}) des Stimmfeldes und dem berechneten Jitter ermittelt [120]. Die Berechnungsformel für den DSI lautet:

$$DSI = 0,13 \times MPT + 0,0053 \times F_{0_{high}} - 0,26 \times I_{low} - 1,18 \times \text{Jitter} + 12,4$$

Formel 1 Der DSI wird aus der maximalen Tonhaldedauer von /a:/ (MPT) in Sekunden, aus der höchstmöglichen Frequenz ($F_{0_{high}}$) in Hz, aus der geringsten Stimmintensität (I_{low}) in dB und aus dem Jitter (%) berechnet.

Die DSI-Werte liegen allgemein zwischen -5 (pathologische Stimme) und 5 (gute Stimme), diese Grenzen können weit übertroffen oder unterschritten werden. Personen ohne Stimmtraining und mit unauffälliger Stimme erreichen Werte über +1,6 [78]. Einen Wert von 1 sollten normale, gesunde Stimme ohne besondere stimmliche Anforderungen bei der Aufzeichnung erreichen [116].

Mit dem DSI werden Stimmen in unterschiedliche Dysphonie Schweregrade eingeteilt. Je negativer der DSI eines Patienten, desto dysphoner ist sein Stimmklang [120]. Zudem besteht ein linearer Zusammenhang zwischen einem absteigenden DSI und einer subjektiv zunehmenden Heiserkeit [119]. Aus diesem Grund ist der DSI ein nützlicher Wert zu Beurteilung von Therapieerfolgen bei Patienten mit Dysphonien [120].

2.6.4 Die Grundfrequenz F_0

Das Stimmsignal zeigt im zeitlichen Verlauf ein wellenförmiges Muster. Wie bereits in Abbildung 7 beschrieben, ist die Frequenz der reziproke Wert der Periodenlänge. Dieser Wert wird als Grundfrequenz F_0 bezeichnet. Die Werte für F_0 sind Geschlechter abhängig und weisen in unterschiedlichen Sprachen jeweils eigene Werte auf [39]. Die Annahme, dass F_0 bei Frauen höher ist als bei Männern hat sich auch in neueren Studien bestätigt [25]. Traunmüller hat in einer Studie die F_0 Ergebnisse mehrerer Sprachen aus unterschiedlichen Studien zusammengefasst [39]. F_0 beträgt in der deutschen Sprache bei Männern im Durchschnitt $129 \pm 2,3 \text{ Hz}$ und bei Frauen $238 \pm 1,9 \text{ Hz}$. Bei englischsprachigen Amerikanern ergab sich bei Männer eine ähnlich Grundfrequenz von $127 \pm 3,8 \text{ Hz}$, bei Frauen hingegen eine etwas verminderte von $186 \pm 5,4 \text{ Hz}$ [39]. Eine weitere Quelle gibt einen Grundfrequenzbereich von Frauen zwischen 175 - 245 Hz und bei Männern zwischen 105 – 160 Hz an [71]. Zwei neuere Studien ergaben bei Amerikanern einen F_0 Mittelwert von $119 \pm 19 \text{ Hz}$ [25] und einen MW von 121,5 Hz bei einer brasilianisch sprechenden Gruppe aus 20-30 jährigen Männern [99]. Hier wurde als Analyseprogramm Praat verwendet. Zu beachten ist, dass die Grundfrequenz von mehreren Faktoren beeinflusst werden kann. Ein erhöhter Muskeltonus und eine größere Anstrengung, können den Wert von F_0 erhöhen [86, 110]. Pathologien an den Stimmlippen verfälschen den F_0 Wert, da sie die Periodizität der Schwingung beeinflussen und eine periodische Schwingung Grundvoraussetzung für einen korrekten F_0 -Wert ist [71].

2.7 Akustische Parameter für Rauschen

Eine gesunde Stimme wird als angenehm empfunden, wenn mit wenig Anstrengung ein Ton erzeugt werden kann und die gesamte Energie des Luftstroms für die Tonbildung verwendet wird. Keine Strömungsenergie geht an Nebengeräusche verloren und es entsteht ein voller Stimmklang. Liegt hingegen beispielsweise eine Glottisschlussinsuffizienz vor, kann

ungenutzte Strömungsenergie entweichen. Diese verursacht durch Verwirbelungen in der Luft unregelmäßige Wellenformen, die als Geräusche wahrnehmbar sind und sich als Hintergrundrauschen über das Stimmsignal legen [116]. Für die Analyse von Rauschkomponenten werden die Parameter Harmonics-to-noise Ratio (HNR), Noise-to-harmonic Ratio (NHR) und Normalized-Noise Energy (NNE) verwendet. Bei der Interpretation der Normwerte muss auf das Analyseverfahren geachtet werden, mit dem die einzelnen Werte erhoben wurden, da es zu Abweichungen kommen kann [116]. Diese drei Werte wurden auch in diesem Studiendesign verwendet, um bei den Handyaufnahmen das Maß an Rauschen zu bestimmen und dieses vom Stimmsignal zu differenzieren.

2.7.1 Harmonics-to-noise ratio (HNR)

Der Harmonics-to-noise ratio (HNR) beschreibt das Verhältnis des harmonischen Stimmanteil und der Nebengeräusche [121] und wird im Spektrum ermittelt [55]. Eine Studie von Yumoto et al. zeigte, dass HNR ein objektives Maß zur Beschreibung des Geräuschsignal ist. Zudem kann der Therapieerfolg bei Behandlung von Heiserkeit mit dem HNR evaluiert werden [121]. Die Untersuchungen von Gesunden ergaben bei Männern einen Mittelwert von 12,2 *dB* und bei Frauen einen Mittelwert von 11,5 *dB*. Die Abweichung der Mittelwerte von Männern und Frauen war in dieser Studie nicht signifikant. [121]. Eine nicht signifikante Differenz des Wertes zwischen Frauen und Männern ergab sich auch in einer anderen Studie [106]. Die Spanne der ersten und dritten Quartile einer gesunden Probandengruppe in anderen Studien lag zwischen 7,0 *dB* bis 17,0 *dB* [121] und zwischen 16,7 *dB* bis 23,5 *dB* in einer Gruppe aus nur männlichen Probanden [117]. Der Mittelwert von 63 männlichen, gesunden Probanden lag bei 24 *dB*, die Ergebnisse wurden mit Praat ermittelt [106]. Eine Studie aus dem Jahr 2016 ergab bei jungen männlichen Soldaten aus der brasilianischen Armee einen MW $17,8 \pm 3,9$ *dB* [74].

Die Verteilung von 42 gesunden Probanden zeigte, dass 95 % der gesunden Testpersonen einen HNR über 7,4 *dB* aufweisen [121]. Das verwendete Programm war hier Voice Analysis Software [4]. Bei Werth et al. wurde zur Analyse das Programm Doctor Speech [117] verwendet. Die unterschiedlichen Ergebnisse der Studien zeigen, dass die fehlende Standardisierung eines Normwertes auch bei der Interpretation des HNR ein Problem ist.

Bei der Interpretation des HNR ist zu beachten, dass der Geräuschpegel und der Jitter einen starken Einfluss auf den HNR Wert haben. Steigt beispielsweise der Jitter, sinkt der HNR. Der HNR kann daher ein Maß für die Stabilität eines akustischen Signals sein [59]. Dadurch ist er in der Analyse der Stimmqualität nützlich, auch wenn durch ihn keine konkreten Aussagen über organische Veränderungen der Stimmlippen gemacht werden können [59]. Zudem gibt der HNR keine spezifische Aussage über die zyklischen Parameter wie die Amplitude, die Schwingungsform oder die Periode [107].

2.7.2 Noise-to-harmonic ratio (NHR)

Der Noise-to-harmonic ratio (NHR) ist neben dem HNR ein weiterer Wert, der den allgemeinen Geräuschanteil eines analysierten Stimmsignals in Bezug auf den klanglichen Anteil beschreibt [71]. Er dient dem generellen Einschätzen bezüglich der Präsenz von Hintergrundgeräuschen in einem analysierten Signal [23, 71]. Der NHR verhält sich reziprok zum HNR, da er das umgekehrte Verhältnis beschreibt. Im Gegensatz zu einem hohen HNR, spricht ein niedriger NHR für eine gesunde Stimme [5, 117]. Ein hoher NHR Wert entsteht beispielsweise durch unregelmäßige Strömungen entlang der Stimmlippen oder durch einen inkompletten Glottisschluss [87]. Der NHR ist neben Shimmer und Jitter ein aussagekräftiger Parameter in der postoperativen Verlaufskontrolle und im Verlauf von Stimmtherapien [23]. In der Literatur finden sich unterschiedliche Normwerte. So geben Di Nicola et al. (2006) als Mittelwert von Stimmgesunden $0,125 \text{ dB}$ an, bei Rusz et al. (2011) ist der Mittelwert $0,02 \pm 0,04 \text{ dB}$. In einer weiteren Studie liegen die NHR Werte von 15 gesunden Probanden zwischen $0,01\text{-}0,2 \text{ dB}$ [5]. Diese Analysen wurden mit Praat durchgeführt. Eine Studie, die Daten unterschiedlicher Mobiltelefontypen und unterschiedlicher Analyseprogramme miteinander verglichen hat, ergab für einen mit Praat ermittelten NHR bei den unterschiedlichen Handymodellen einen ähnlichen Wert um $0,01 \text{ dB}$ [36]. In einer Studie von Maryn 2010 ergab sich ein Mittelwert von $0,04 \pm 0,01 \text{ dB}$ [64].

2.7.3 Normalized-noise energy (NNE)

Der Wert Normalized-noise energy (NNE) ist ein Maß für die Geräuschenergie in einem Stimmsignal in Bezug auf die gesamte Energie des Signals. Beide Energien werden in der Einheit dB angegeben. Für eine korrekte Messung des NNE, müssen die Geräusche von den harmonischen Anteilen differenziert werden. Die Geräuschanteile, die zwischen den harmonischen Schwingungen liegen, lassen sich direkt aus dem Spektrum ablesen. Man nimmt an, dass innerhalb einer harmonischen Schwingung der Geräuschanteil der Mittelwert aus beiden benachbarten Minima des Spektrums ist [17]. Der optimale Frequenzbereich zur Bestimmung des NNE liegt zwischen $1\text{-}5 \text{ kHz}$ [17]. Eine ältere Studie gibt in ihrer Kontrollgruppe einen NNE Normbereich von -18 bis -11 dB an. Zudem zeigt sie, dass der NNE durch Dysphonien beeinflusst wird, die durch organische Ursachen, wie Stimmlippenknoten, Larynxkarzinomen oder durch Nervenschädigungen verursacht werden [51]. In neueren Untersuchungen ergaben sich für gesunde Probanden, dass der NNE von 50 % der Teilnehmer zwischen -12 und $-5,2 \text{ dB}$ lag [117]. Diese Studie zeigte außerdem, dass es keinen geschlechterspezifischen Unterschied für diesen Wert gibt [117]. In der Studie von Michaelis et al. (1997) wurde zudem eine starke Abhängigkeit des NNE vom Jitter beschrieben. Der Umfang des NNE in dem physiologischen Jitter Bereich von 0.01 bis 1% betrug -30 bis -5 dB .

Ungeachtet der Abhängigkeit von den Irregularitätsparametern gilt der NNE laut Wendler et al. als wichtigster Rauschparameter in der Stimmdiagnostik [116].

2.8 Spektrale Stimmqualitätsparameter

Neben den Parametern des Rauschens, die zu den spektralen Parametern der Stimmanalyse zählen, gibt es noch weitere Werte mit denen die Qualität eines Signals bewertet werden kann. Zu diesen zählen der H_1H_2 , die spektrale Neigung und die Smoothed cepstral peak prominence. Deren Vorteil ist, dass vor deren Berechnung keine Analyse der Periodizität des Signals stattfinden muss [50]. Zudem können neben Vokalphonationen auch ganze Worte und Sätze analysiert werden [50]. Dies wird durch die Analyse eines langfristigen Durchschnittsspektrums, des Long-Term Average Spectrum (LTAS), mittels Fourier Transformation ermöglicht. Hierbei wird der Zeitbereich in den Frequenzbereich umgewandelt [31]. Die Parameter, welche in dieser Arbeit Anwendung gefunden haben, werden im Folgenden besprochen.

2.8.1 H_1H_2

Die Phonation erfolgt unter anderem durch das Öffnen und Schließen der Stimmlippen. Das Verhältnis von Offen- und Schlussphase innerhalb einer Schwingungsperiode ist maßgeblich für die Stimmqualität und wird als Offenquotient bezeichnet [116]. Das normale Verhältnis einer mittellauten Stimme liegt bei $\frac{2}{3}$ zu $\frac{1}{3}$ [116]. Das bedeutet, dass die Offenphase physiologisch länger als die Schlussphase ist. Ein geeignetes akustisches Korrelat dieses Verhältnisses ist die Amplitudendifferenz der ersten (H_1) und der zweiten harmonischen Schwingung (H_2) [50]. H_1 entspricht hierbei der Grundfrequenz F_0 und das Verhältnis aus H_1H_2 wird in dB angegeben.

Um zu verstehen, wie die Differenz von H_1 und H_2 den Stimmklang beeinflusst, ist das Modell von unterschiedlichen Phonationstypen von Ladefoged 1971 hilfreich [34]. Dieser beschreibt darin, dass der Übergang von einem behauchten, zu einem knarrenden Stimmklang fließend ist. Die modale Stimme liegt in der Mitte dieses Bereichs. Für den Charakter des Stimmklangs ist maßgeblich das Verhältnis von Offen- und Schlussphase verantwortlich [34]. Dieses Verhältnis wiederum wird durch die Differenz von H_1 und H_2 dargestellt.

Die Grundlagen für diesen Denkansatz wurden von Bickley bestätigt, die in der Amplitudendifferenz eine gültige Alternative für das Maß an Behauchtheit gefunden hat [42]. Ist nun die Amplitude von H_1 leicht größer als H_2 ($H_1 > H_2$), so liegt eine längere Öffnungsphase vor. Die Differenz ergibt einen leicht positiven Wert. Dies würde für eine gesunde, normale Stimme sprechen. Ist H_2 jedoch deutlich größer als H_1 ($H_1 < H_2$), so ergibt sich ein negativer Wert. Es liegt kein normales Verhältnis von Offen- und Schlussphase vor. Die Stimme hört sich knarrend an. Ist H_1 nun wesentlich größer als H_2 , ergibt sich ein stark positiver Wert.

Diesen findet man vor allem bei behauchten Stimmen [50]. Weitere Studien ergaben, dass der H_1H_2 geschlechterspezifisch betrachtet werden muss, da Frauen eine behauchtere Stimme aufweisen als Männer [42, 53].

Die Aussagekraft des H_1H_2 über den Stimmcharakter ist komplex und abhängig von der untersuchten Person [58]. So kann der physiologische Stimmklang einer Person bei einer anderen bereits als pathologisch gelten [34]. Außerdem weist der Wert eine hohe Variabilität in unterschiedlichen Sprachen auf, so dass der Standardisierung des Wertes Grenzen gesetzt sind [52, 56]. Eine Studie mit deutschsprachigen, gesunden, männlichen Probanden im Alter von 20 – 34 Jahren, die den Einfluss nasaler Aussprache auf den H_1H_2 untersuchte, ergab für den H_1H_2 während der Phonation des Vokals /a:/ eine Spanne von -5 dB bis 0 dB und ein Median bei etwa $-2,5 \text{ dB}$. Bei zunehmender nasaler Phonation wurde der H_1H_2 negativer, der Median lag bei etwa $-4,5 \text{ dB}$ [98]. In Frankreich beispielsweise liegen Referenzwerte von männlichen Probanden bei der Phonation des Vokals /a:/ zwischen $-1,5$ bis 0 dB . Bei englischsprachigen Probanden reicht die Spanne von $-3,6$ bis 0 dB [81]. Eine Studie aus dem Jahr 2016 ergab bei jungen männlichen Soldaten aus der brasilianischen Armee einen Mittelwert von $-0,3 \pm 1,6 \text{ dB}$ [74].

2.8.2 Spektrale Neigung

Die spektrale Neigung ist neben dem H_1H_2 ein weiterer Parameter zur Beurteilung der Stimmqualität. Die gesamte Energieverteilung innerhalb eines Spektrums wird als spektrale Neigung bezeichnet. Ermittelt man das Frequenzspektrum einer stimmhaften Sequenz so sind die tieferen Frequenzen energiereicher als die hohen Frequenzen [50]. Eine behauchte Stimme beispielsweise hat einen Energieverlust in den höheren Frequenzen. Erklärt man diesen Zustand vereinfacht an nur zwei Frequenzen wäre die Frequenz f_1 deutlich größer als die Frequenz f_2 . Legt man nun eine Regressionslinie durch dieses Energiespektrum so würde die Neigung steiler werden. Knarrende Stimmen hingegen haben ein Energiedefizit in den niedrigen Frequenzen, die Neigung wird flacher, da f_1 nun kleiner wird [50]. *Tilt* und *Slope* sind geeignete Methoden, um diese spektrale Neigung zu messen und anhand dieser die Qualität der Stimme zu beschreiben.

Der *Slope* wird aus der Differenz der Energiespektren von $1\text{-}1000 \text{ Hz}$ und $1\text{-}10 \text{ kHz}$ in einem langfristigen Durchschnittsspektrum (LTAS) berechnet, er beschreibt die Neigung des gesamten LTAS [64]. Der Frequenzbereich ist nicht obligat, sondern kann variabel vergrößert oder verkleinert werden. Der *Slope* einer gesunden Stimme liegt im negativen Bereich, da die physiologische Neigung von den energiereichen tiefen Frequenzen zu den energiearmen, hohen Frequenzen abfallend ist. Maryn hat eine gesunde und einer behauchten Stimme gegenübergestellt. Der *Slope* hat bei gesunden Stimmen eine Trendlinie von $-13,3 \text{ dB}$, der *Slope* der behauchten Stimme lag bei $-8,2 \text{ dB}$ [64]. Der Normbereich liegt zwischen -15 und $-$

25 [50]. Eine anderen Quelle ergab einen Slope Mittelwert von 23 gesunden Probanden bei $-23,3 \text{ dB}$ und ein Spanne von -33 dB bis -14 dB [64].

Als Tilt wird die Differenz der Energiespektren von $1\text{-}1000 \text{ Hz}$ und $1000\text{-}10000 \text{ Hz}$ bezeichnet, welche anhand einer Regressionslinie, die durch das LTAS läuft, berechnet wird [64, 65]. Der berechnete Mittelwert von 23 gesunden Probanden liegt bei $-10,5 \text{ dB}$. Der Normbereich liegt zwischen $-12,1 \text{ dB}$ und $-8,6 \text{ dB}$ [64]. Einzelne Tilt- oder Slope-Werte sind wenig aussagekräftig, daher sollten bei der Beurteilung die Werte aus mehreren Messungen betrachtet werden [50]. Ein Vorteil des Tilts gegenüber dem HNR ist, dass der Tilt in Messungen des Geräuschanteils eines Stimmsignals unabhängig gegenüber Störungen ist, die durch Shimmer oder Jitter verursacht werden. Der HNR gibt Gesamtinformationen über die Gleichmäßigkeit eines Signals an, ist jedoch weniger spezifisch. Die Berechnung von Tilt und Slope ist zudem von der Grundfrequenz F_0 abhängig [72].

2.8.3 Smoothed cepstral peak prominence (CPPs)

Als der wichtigste spektrale Parameter zur Beurteilung der Stimmqualität gilt der Smoothed cepstral peak prominence (CPPs) [41, 64], da er im Vergleich zu anderen Messwerten eine hohe Korrelation zu behauchten Stimmen zeigte und Hinweise auf Heiserkeit in der Stimme gibt [40]. Der CPPs wird aus der inversen Fourier-Transformation des Spektrums („cepstrum“) berechnet. Das bedeutet, dass sich an die Spektralanalyse eines Signals wiederum eine Spektralanalyse des Ergebnisses anschließt [50]. An dem Punkt mit der höchsten „Rahmonische“ (cepstral peak) wird die Amplitudendifferenz aus dem smoothed cepstrum und der Tilt Regressionlinie berechnet, hieraus ergibt sich der CPPs [44, 65]. Eine Studie von gesunden Probanden ergab einen Mittelwert von $8,0 \pm 0,94 \text{ dB}$, ein Minimum von $5,9 \text{ dB}$ und ein Maximum von $10,1 \text{ dB}$ [64]. Eine Studie von Heman-Ackah et al. 2014 ergab ähnliche CPPs Werte, Mittelwert $4,77 \pm 0,97 \text{ dB}$, Minimum $1,61 \text{ dB}$ und Maximum $7,02 \text{ dB}$ [41]. Die Grundüberlegung dieser Messung ist: Je gleichmäßiger sich ein Stimmsignal ausbreitet, desto harmonischer verhält sich das Signal und desto ausgeprägter ist die Höhe des hergeleiteten cepstrums. Ein prominenter Gipfel zeigt, dass die Grundfrequenz dominanter ist als die Nachbarfrequenzen. In der Stimmauswertung spricht ein prominenter Gipfel des cepstrums daher für eine gute Stimmqualität. Bei bestehender Heiserkeit geht Energie währende des Sprechens an Nebengeräusche verloren. Da kaum ein Unterschied zwischen Grundfrequenz und den Nachbarfrequenzen besteht, ist auch das Cepstrum und damit der CPPs verringert [44].

CPPs gilt als zuverlässiger Messwert von Stimmstörungen und korreliert am besten mit akustisch wahrnehmbaren Dysphonien [41, 64, 88]. Mit dem CPPs können dysphone und behauchte Stimmanteile in pathologischen Stimmaufnahmen besser detektiert werden, als beispielsweise mit dem NHR [40, 64]. Zudem kann ein eingeschränkter CPPs Hinweise auf

Kraftlosigkeit der Stimme in Patientengruppen mit motorischer Stimmeinschränkung geben, beispielsweise bei Parkinson Erkrankten [48].

Zusammenfassend sind die spektralen Parameter weniger fehlerabhängig als andere Werte. Eine praktische Fehlerfreiheit ist jedoch auch hier nicht möglich. Diese Problematik setzt eine gewisse Erfahrung in der Stimmbeurteilung voraus [50].

3 Material und Methoden

Im Folgenden wird die in dieser Arbeit durchgeführte Studie im Detail beschrieben. Die vier wichtigsten Kriterien für einen Einschluss in die Studie waren ein unauffälliger Anamnesebogen, ein physiologischer Stroboskopie-Befund, ein unauffälliger VHI-12 Test und ein normwertiges Stimmprofil. Zudem wird anschließend die Einstellung und Benutzung der App erklärt.

3.1 Ethikantrag

Bei dieser Studie handelt es sich um eine prospektive Longitudinalstudie, in welcher die Wirksamkeit und der Erfolg von Therapiekontrollen bei Stimmrehabilitationsmaßnahmen objektiviert werden soll. Zur Abbildung einer gesunden Stimmvariabilität mittels telemedizinischen Monitoring Verfahren wird in diesem Studienteil eine Kontrollgruppe aus gesunden, männlichen Probanden betrachtet. Laut Ethikvotum bestehen bei der Durchführung der Studie weder berufsrechtliche noch ethische Bedenken.

3.2 Probanden

Gesuchte Teilnehmer für diese Studie waren Männer, mit einem Mindestalter von 18 Jahren. Die Studie erfolgte auf freiwilliger und unentgeltlicher Basis am Universitätsklinikum des Saarlandes. Ein Proband galt als gesund, wenn der Anamnesebogen unauffällig war und das bildgebende Verfahren und Selbst- und Fremdbewertung der Stimme ohne Befund waren. Die genauen Anforderungen werden im Folgenden erklärt.

3.3 Einschluss in die Studie

Für einheitliche und valide Ergebnisse in den klinischen Untersuchungen, wurde die Datenerhebung entsprechend des Protokolls der European Laryngological Society (ELS) [18] durchgeführt. Dieses Modell umfasst eine bildgebende Untersuchung der Stimmlippen und eine Messung akustischer und aerodynamischer Komponenten der Stimme. Zur Datenerhebung wurden hierfür die Stroboskopie und die Stimmfeldmessung hinzugezogen. Außerdem wurden das subjektive Stimmbefinden des Patienten und die Beurteilung der Stimme durch einen erfahrenen klinischen Untersucher ermittelt. Eine Studie aus dem Jahr 2005 hat ergeben, dass dieses Untersuchungsschema gültige Daten ergeben hat [29].

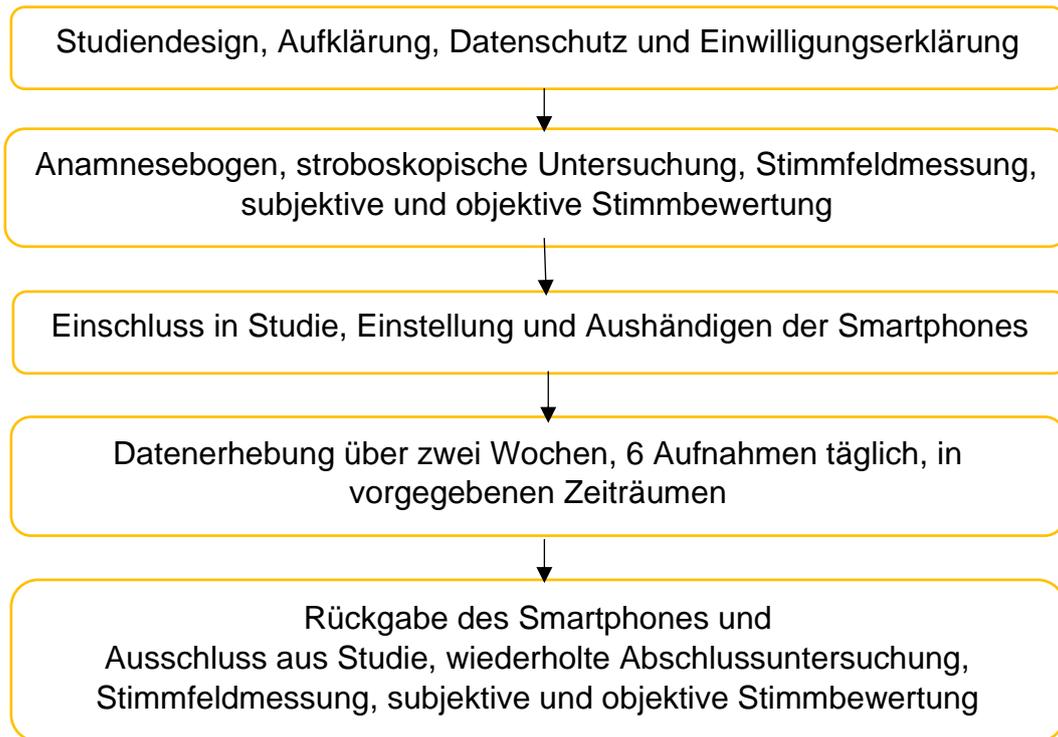


Abbildung 13: Schematische Auflistung der einzelnen Untersuchungsschritte.

3.3.1 Einverständniserklärung und Anamnese

Zu Beginn wurden die Probanden in der Einverständniserklärung ausführlich über den Ablauf der Studie informiert. Neben der Motivation und Zielsetzung waren auch die Risiken der Studie aufgelistet. Die Probanden wurden über den Datenschutz und über ihr Rücktrittsrecht informiert. Durch die Unterschrift gestatteten die Probanden die Verwendung der Daten, die im Zusammenhang mit der Studie erhoben wurden. Anschließend wurde den Probanden ein standardisierter Anamnesebogen ausgehändigt, mit welchem sie über ihre Stimmbelastung, ihr Heiserkeitsprofil und über ihren Beruf befragt wurden. Zudem wurden vorliegende Stimmerkrankungen, und eine vorangegangene logopädische Behandlung ausgeschlossen. Weiterhin sollten Raucher, sowie Probanden mit Asthmaerkrankungen, welches medikamentös behandelt werden muss, nicht in die Studie aufgenommen werden.

3.3.2 VHI-12 Test

Im Anschluss wurde jeder Proband gebeten den „Voice Handicap Index“ – Test auszufüllen (vgl. Abbildung 61). In diesem Test wurde das subjektive Sprechempfinden durch den Probanden evaluiert. Hierzu erfolgte zunächst eine Befragung des Probanden über den Gebrauch seiner Sprech- und Singstimme. Ein Ausschlusskriterium in dieser Studie war der berufliche Einsatz der Singstimme. In diesem Test sollte der Proband seine Gesprächigkeit von eins bis zehn einschätzen, bevor er zwölf Fragen bewerten sollte. In den Fragen wurden vier Faktoren in jeweils drei Punkten abgefragt: „Negative Stimmerfahrung, Selbstunsicherheit, Mangelnde Tragfähigkeit der Stimme und Negative Emotionalität“ [104]. Die Fragen konnten

jeweils auf einer Skala von 0 bis 4 bewertet werden. „0“ sollte angekreuzt werden, wenn nie eine Stimmstörung vorlag, „1“, wenn selten und „2“ bei Gelegenheit. „3“ und „4“ sagen aus, dass die Aussagen immer zu treffen. Zum Schluss wurde konkret nach dem Stimmbefinden des Probanden vor der Untersuchung gefragt. Als stimmgesund, gemäß dem VHI-12, galten alle Probanden, die in der Befragung einen Gesamtwert von maximal 6 Punkten aufwiesen [33].

3.3.3 Stroboskopie

Die klinische Untersuchung zum Ausschluss von funktionellen und organischen Pathologien der Stimmlippen erfolgte in dieser Studie mit dem Stroboskop. Für die Untersuchung wurde ein starres Endoskop mit 90°-Optik verwendet, das aufgrund guter Lichtqualität eine optimale Beleuchtung des Kehlkopfes gewährleistet [93]. Auf ein Schleimhautoberflächenanästhetikum (z.B. Xylocain-Spray) wurde im Rahmen der Studie verzichtet, da es sich zum einen um eine nicht schmerzhaft Untersuchung handelte, zum anderen sollte ein möglichst realitätsgetreues Bild der Stimmlippenschwingungen abgebildet werden. Der Frequenzunterschied der Blitzlichtfrequenz Δf und der Stimmfrequenz in dieser Studie wurde automatisch bestimmt und lag zwischen 1 und 2,5 Hz. Die Untersuchung wurde bei gehaltener Phonation des Vokals /a:/ durchgeführt. Hierzu wurden die Probanden aufgefordert, diesen Vokal in angenehmer Singstimmlage und mittlerer Lautstärke zu singen. Für einen aussagekräftigen Befund sollte die Dauer der Aufnahme mehr als 3 sec betragen. Der Befund wurde anschließend mit Hilfe eines vorgefertigten Untersuchungsbogens bewertet. Die Aufnahmen wurden auf eine normale Randkantenverschiebung, gleichmäßige Amplituden und Phasensymmetrie untersucht (vgl. Abbildung 62). Da eine genaue Beurteilung der Stimmlippen eine gute Einsicht auf die Stimmlippen erfordert, wurde die Verdeckung der Stimmlippen durch den *Petiolus* (Spitze des Schildknorpels) berücksichtigt. Waren diese Voraussetzungen erfüllt und keine organischen oder funktionellen Auffälligkeiten auf Stimmlippenebene erkennbar, so galten die Probanden als klinisch unauffällig. Neben der subjektiven Bewertung der Stimme durch den Probanden selbst, erfolgte auch eine auditive Bewertung der Stimme im Anschluss an die Stimmfeldmessung. Alle Untersuchungen wurden vom selben Untersucher erhoben, um eine möglichst hohe Vergleichbarkeit der Werte zu gewährleisten.

3.3.4 Stimmanalyse

Die Ermittlung der akustischen und aerodynamischen Komponenten der Stimme erfolgte mittels einer Stimmfeldmessung. Hierzu wurde die Sprech- und Singstimme mit einem Aufnahmemikrophon unter Zuhilfenahme des Programmes „LingWAVES“ aufgezeichnet. Die Aufnahmen erfolgten im Stehen und ein Mund-Mikrophonabstand von 30 cm wurde eingehalten. Zunächst erfolgte die Bestimmung des Sprechstimmfeld. Zur Darstellung des

Sprechstimmprofilen wurden die mittleren Sprechstimmlage, der Dynamik- und Tonumfang und die maximalen Rufstimme erhoben. Zum Vergleich wurden die Mittelwerte und die Standardabweichung dieser Werte von allen Probanden zu Beginn und am Ende der Studie berechnet. Die Probanden zählten hierbei zum einen mit sehr leiser, aber noch nicht geflüsterter Stimme, zum anderen mit ihrer Indifferenzstimme und mit ihrer Vortragsstimme von 21 bis 40. Die Rufstimme wurde durch einmaliges lautes Rufen des Beispielsatzes „Hey Paul, komm doch mal rüber“ bestimmt. Aus diesen vier Frequenzen wurde mittels „LingWAVES“ die mittlere Sprechstimmlage (mSsl) berechnet. In dieser Frequenz spricht der Proband am häufigsten.

Des Weiteren wurde das Singstimmfeld vermessen. Hierbei wurden Maxima und Minima von Lautstärke und Frequenz, sowie der Dynamikbereich der Singstimme und der Jitter der Probanden zu Beginn und am Ende der Studie ermittelt. In der Auswertung wurden die Werte der einzelnen Teilnehmer und des Gesamtkollektivs betrachtet. Die Probanden wurden hierbei aufgefordert, zunächst mit möglichst leiser Stimme ein mindestens zwei Sekunden lang phoniertes /la:/ in möglichst kleinen Tonleiterschritten, ausgehend von ihrer mSsl, nach unten zu singen. An ihrer physiologischen tiefen Stimmgrenze angekommen, sollten die Probanden mit leiser Singstimme, in kleinen Tonleiterschritten aufsteigend und so hoch wie möglich singen. Dieser Vorgang wurde anschließend mit lauter und kraftvoller Stimme wiederholt. Ein mehrfaches Wiederholen ergab auch bei unmusikalischen und ungeübten Probanden ein aussagekräftiges Messresultat.

Im Anschluss an die Stimmfeldmessung wurde die Tonhaldedauer ermittelt. Der Proband phonierte möglichst lange ein gleichmäßig gehaltenes /a:/. Jitter und Shimmer waren relevante Parameter, die sich bei dieser Untersuchung ergeben. Zudem wurde in diesem Untersuchungsschritt der DSI erhoben, für dessen Berechnung die maximale Frequenz, die minimale Lautstärke der Singstimme und der Jitter notwendig sind.

Der letzte Teil der Untersuchung befasste sich mit der objektiven Bewertung der Stimme. Mit dem RBH Test kann die Stimme von einem erfahrenen Kliniker klassifiziert werden. Zur Bestimmung der RBH Klassifikation wurde der Standardtext „Der Nordwind und die Sonne“ [93, 114] verwendet. RBH beschreibt die Rauigkeit, die Behauchtheit und die Heiserkeit einer Stimme. Diese können in vier Schweregraden ausgeprägt sein 0 = „nicht gestört“, 1 = „geringgradig gestört“, 2 = „mittelgradige“ und 3 = „hochgradige“ Beeinträchtigung. Heiserkeit ist sowohl bei Rauigkeit, als auch bei Behauchtheit oder bei beiden vorliegenden Pathologien nachweisbar. Voraussetzung für den Einschluss in die Studie war ein R0B0H0 bei den Probanden.

3.3.5 Stimmmonitoring und Aufbau der Handyapp Voicediary

Im Rahmen dieser Studie erfolgte das Stimmmonitoring mit Hilfe der an der Hochschule Trier entwickelten telemedizinischen Smartphone-Applikation „Voice Diary“. Um Einflüsse unterschiedlicher Mikrofonttechnologien auszuschließen wurden alle Monitoringsysteme auf baugleichen Smartphones der Marke HUAWEI installiert. Jedem Probanden wurde nach einer unauffälligen Erstuntersuchung ein Handy ausgehändigt, nachdem er in der App mit einer individuellen Kennung registriert wurde (vgl. Abbildung 14). Neben der Registrierung wurde die ermittelte mittlere Sprechstimmlage als orientierende Frequenz für die Stimmaufnahmen eingestellt. Die erste Probestimmaufnahme erfolgte zusammen mit dem Untersucher am Untersuchungstag. Hierbei wurde kontrolliert, ob die Aufnahme im Anschluss im integrierten Stimmtagebuch der App verzeichnet wurde. Zudem wurde der Fragebogen, der jeder Stimmaufnahme folgte, zusammen mit dem Probanden besprochen. Der Fragebogen wurde von der Hochschule Trier in der „Voice Diary“ Applikation integriert. Wurde der Fragebogen nach einer Aufnahme nicht ausgefüllt, so wurde die Aufnahme nicht gespeichert. Zu Beginn wurde der Gesundheitszustand des Patienten und der Raucherstatus befragt, ob derzeit eine Erkrankung der Atemwege vorliege und ob es sich bei dem Probanden um einen Raucher oder Nicht-Raucher handle. Die Fragen des Bewertungsbogen befragten die Hauptsymptome und fakultative Nebensymptome von Dysphonien [28]. Es wurde nach Missempfindungen bei der Stimmaufnahme und deren Art gefragt, sowie nach der Art und dem Grad der Stimmbelastung. Die Bewertung des Grades war von 1 „keine Stimmbelastung“ bis 6 „sehr starke Stimmbelastung“ möglich. Eine wichtige Anmerkung für die weitere Auswertung ist, dass bei Probanden, die keine Stimmbelastung angegeben haben (Grad 1) die Frage nach der Art der Belastung übersprungen wurde. Außerdem wurde der subjektive Ermüdungsgrad der Stimme und die Stimmqualität bei jeder Aufnahme befragt. Eine Bewertung war hier von 1 „keine Ermüdung“ bis 6 „sehr starke Ermüdung“ möglich. Die Bewertung der Stimmqualität war von 1 „sehr gute Stimmqualität“ bis 6 „sehr schlechte Qualität“ möglich. Dieser Fragebogen ermöglichte eine individuelle, tagesaktuelle Bewertung jeder einzelnen Aufnahme. Das Studienmodell war so aufgebaut, dass die Probanden mindestens 6 Aufnahmen pro Tag ungefähr zur gleichen Zeit über 14 Tage durchführen sollten. Die sechs Zeitrahmen wurden den Probanden zu Beginn der Studie mitgeteilt. Diese waren von 6-9 Uhr, 9-11 Uhr, 11-14 Uhr, 14-17 Uhr, 17-20 Uhr und 20-24 Uhr.

3.3.6 Einstellung und Bedienung der Handyapplikation

Waren alle vorangegangenen Untersuchungen unauffällig, konnte der Proband als Stimmgesunder an der Studie teilnehmen. Im nächsten Schritt wurde der freiwillige Teilnehmer auf einem, im Rahmen der Studie, zur Verfügung gestellten Handy in der App „Voice Diary“ registriert. Jeder Teilnehmer erhielt bei der Registrierung eine eigene ID. So konnten auch

mehrere Probanden ein Handy hintereinander benutzen ohne dass die Gefahr des Datenverlustes bestand. Nach der Registrierung wurde die mittlere Sprechstimmlage des jeweiligen Probanden mit der App erfasst (vgl. Abbildung 14). Der Aufnahmeprozess findet sich unter „Einstellung“ und wurde von der App selbst angeleitet. Die mittlere Sprechstimmlage, die mittels Stimmfeldmessung ermittelt wurde, wurde zusammen in der Klinik eingestellt und die Probanden darauf hingewiesen, dass sie keine Veränderungen in den Einstellungen vornehmen sollten, da die Aufnahmen der gehaltenen Phonation immer im Rahmen der eingestellten mittleren Sprechstimmlage erfolgten.

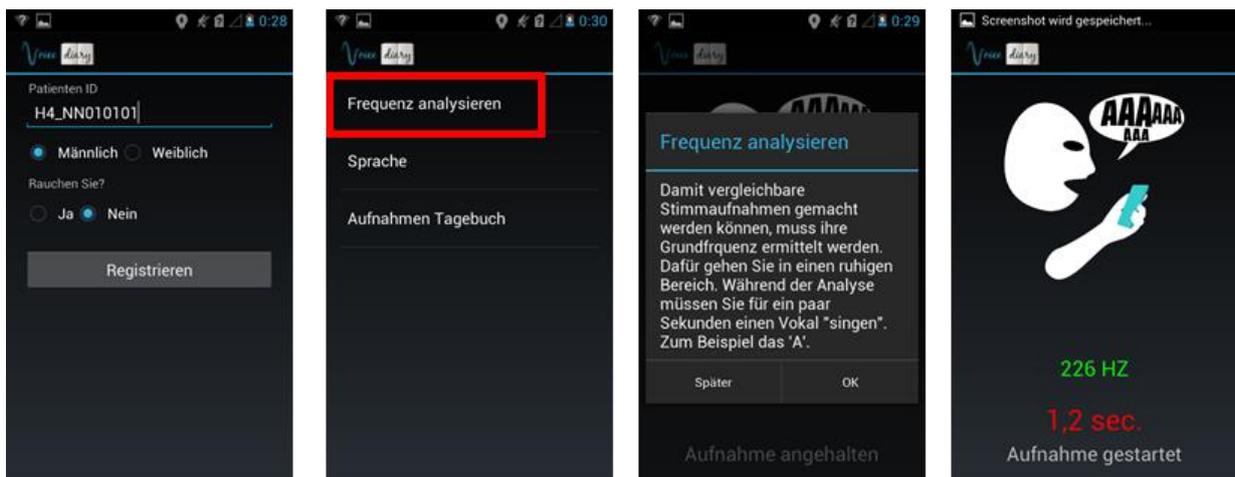


Abbildung 14: Registrierung und Frequenzeinstellung der Probanden auf die individuelle mittlere Sprechstimmlage. Die erste Aufnahme führte der Proband eigenständig durch, jedoch im Beisein des Untersuchers, um mögliche Frage zu klären. Wichtig bei den Aufnahmen war eine möglichst leise Umgebung und ein Handy-Mund-Abstand von etwa 10cm, um normal laute Aufnahmen zu erhalten. Eine Aufnahme wurde erfolgreich aufgezeichnet, wenn der phonierte Ton /a:/ 3 Sekunden gleichmäßig gehalten wurde und die Tonhöhe innerhalb eines vorgegebenen Frequenzbereichs um die mSsl lag. Dieser wurde bei der Aufnahme als grüner Bereich angezeigt (vgl. Abbildung 15). Aufnahmen, die um mehr als das doppelte von der Standardabweichung der mittleren Sprechstimmlage eines Probanden abwichen, wurden als nicht auswertbar definiert und nicht verwendet.

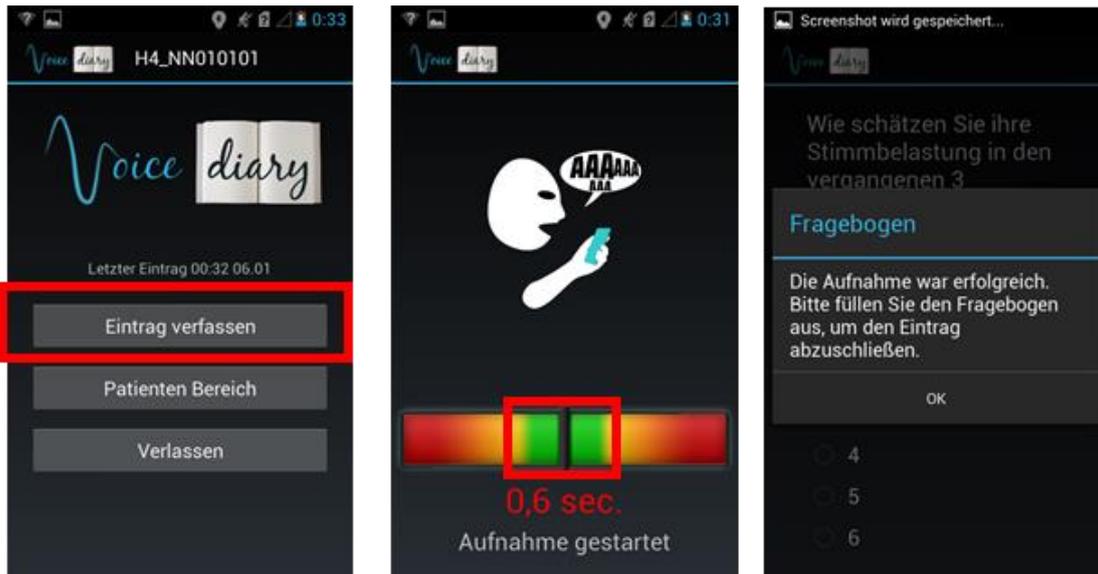


Abbildung 15: Durchführung einer Aufnahme. Vorgegebener grüner Frequenzbereich während der Aufnahme im mittleren Bild rot markiert.

Nach jeder Aufnahme wurde automatisch von der App der kurze Fragebogen abgerufen. Hier sollte der Proband seine eigene Stimme bewerten und Auskunft über die aktuelle Stimmbelastung geben und somit seinen subjektiven Eindruck dokumentieren. Die zwei ersten Fragen des Katalogs haben eine akut bestehende Krankheit und den Raucher Status abgefragt. Da die Probanden der Studie konstant Nichtraucher waren und bei keinem eine akute Krankheit vorlag, wurden die Ergebnisse graphisch nicht dargestellt, da keine Ausreißer vorlagen.

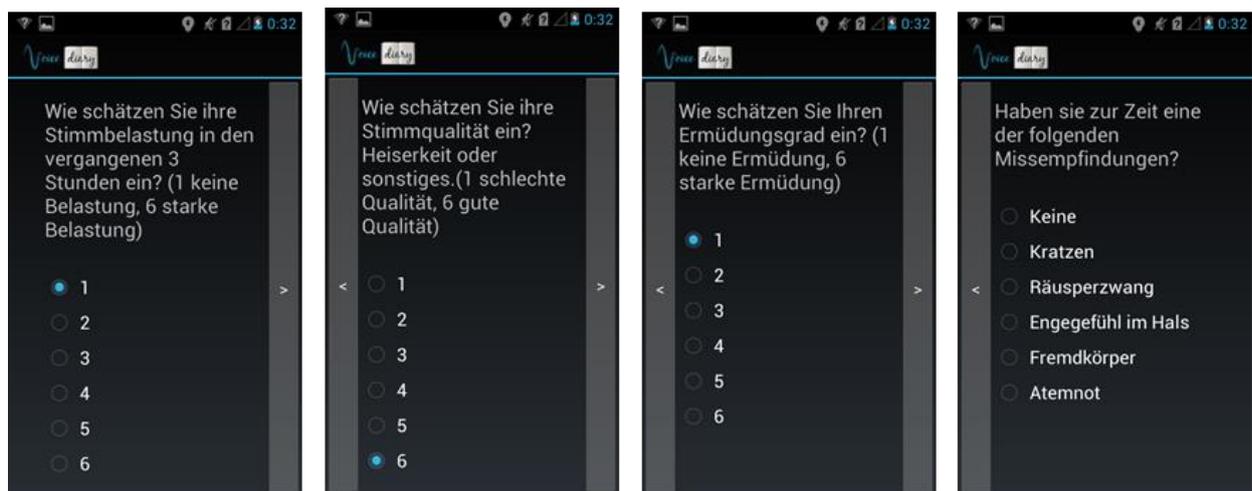


Abbildung 16: Bewertung der Stimme nach Stimmbelastung, Stimmqualität und Stimmermüdung, sowie die Qualität von Missempfindungen.

War die erste Aufnahme des Probanden erfolgreich so wurde kontrolliert, ob die Aufnahme im Aufnahmetagebuch der App gespeichert wurde und jeweils Datum und Uhrzeit der Aufnahme richtig waren (vgl. Abbildung 17). Im Aufnahmetagebuch konnten die Freiwilligen auch ihre eigenen Aufnahmen kontrollieren und so über die Aufnahmen einen guten Überblick erhalten.

Waren keine weiteren Fragen mehr offen, bekam der Proband das Handy für einen Zeitraum von zwei Wochen ausgehändigt. In diesem Zeitraum sollte er zu möglichst gleichen Zeitpunkten 6 Aufnahmen pro Tag machen. Um eine Vergleichbarkeit zwischen den Probanden zu gewährleisten, sollten während der zweiwöchigen Studienteilnahme Aufnahmepausen vermieden werden. Nach diesem Zeitraum wurde der Proband für eine klinische Abschlussuntersuchung wieder in die Klinik einbestellt.

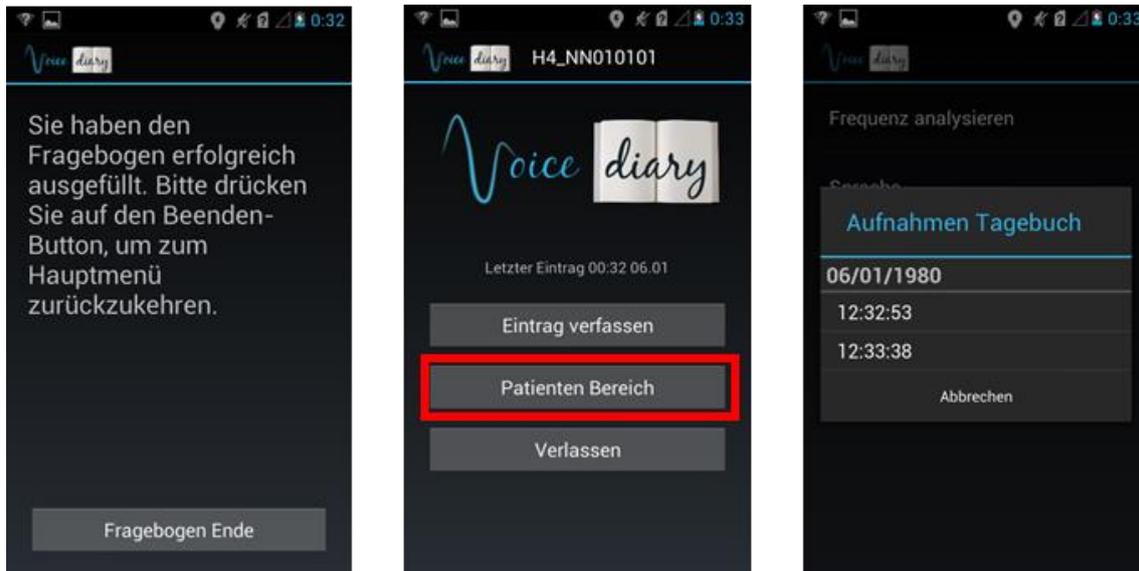


Abbildung 17: Abschluss des Fragebogens und Kontrolle der erfolgten Aufnahme im Stimmtagebuch über den Patienten Bereich.

3.3.7 Abschlussuntersuchung der Studie

Abgesehen von der bildgebenden Untersuchung mit dem Stroboskop, war die Vorgehensweise der Abschlussuntersuchung identisch zur Voruntersuchung. Die Stimmfeldmessung und die objektive und subjektive Stimmeinschätzung wurden erneut erhoben. Ziel der Abschlussuntersuchung war sicherzustellen, dass die Probanden auch nach Abschluss der Studie über eine unauffällige Stimmanamnese verfügen.

3.4 Statistische Auswertung

In dieser Studie handelt es sich um eine rein deskriptive Analyse der erhobenen Normdaten. Die Darstellung der einzelnen Daten erfolgte durch Mittelwerte und Standardabweichungen. Zudem wurde die Verteilung der Werte in Boxplots und in Balkendiagrammen dargestellt. Die Ermittlung der Werte erfolgte unter Verwendung von Excel und MATLAB R2017a. Um einen Überblick über die Vielfalt der Daten und ihrer Verteilung aus dem telemedizinischen Monitoring zu erhalten, erfolgte eine dimensionsreduzierende Visualisierung der Daten mittels „t-distributed stochastic neighbor embedding“ (t-SNE). Mit diesem Verfahren kann die Beziehung hochdimensionaler Daten in einem niederdimensionalen Koordinatensystem visualisiert werden. Durch die Reduktion der Dimensionen ist jeder Parameter in einem

bestimmten Bereich auf der Map verteilt [60]. Mit dem t-SNE können so nicht nur Gemeinsamkeiten der einzelnen Werte, sondern auch mögliche Differenzen einiger Parameter dargestellt werden, die sich im Vergleich zum Gesamtkollektiv anders verteilen. In dieser Studie wurden 9 Parameter (Shimmer und Jitter, HNR, NHR, NNE, H1H2, Tilt und Slope und von CPPs) aus dem telemedizinischen Monitoring in einer Grafik mit Hilfe des t-SNE Verfahren zusammengefasst. Die einzelnen Probanden wurden mit unterschiedlichen Farben kodiert.

4 Ergebnisse

In diesem Kapitel werden die im Rahmen der Studie erzielten Ergebnisse aus der klinischen Ein- und Abschlussuntersuchung präsentiert. Hierbei wird näher auf die Anamnesedaten, auf die Stroboskopieaufnahmen und auf die Stimmanalyse eingegangen. Weiterhin erfolgt die Darstellung der Ergebnisse aus dem „VoiceDiary“ Fragebogen. Als letztes werden die Ergebnisse aus dem telemedizinischen Monitoring vorgestellt. Hierbei werden die Irregularitätsparameter, die Rauschwerte und die spektralen Parameter betrachtet und mit Referenzwerten aus der Literatur verglichen.

4.1 Probanden

In die Studie wurden 20 männliche Probanden eingeschlossen. Hierbei handelte es sich um deutschsprachige Studenten im Alter von 22 bis 33 Jahren, die nach eigener Angabe alle Nichtraucher und gesund waren. Es wurden die Ergebnisse der Anamnese, der VHI-12 Auswertung, der stroboskopische Untersuchung und der Stimmanalyse betrachtet. Von den 20 gesammelten Datensätzen konnten nur 17 für die Auswertung verwendet werden, da von 3 Probanden die Datenerhebung nicht kontinuierlich durchgeführt wurde. Diese Probanden wurden aufgrund der Unvollständigkeit der Datensätze nach Abschluss der Studie nicht in die Analyse mit einbezogen.

4.2 Zusammenfassung der Anamnesedaten

Die Auswertung der Anamnesedaten ergab, dass der Stimmgebrauch der Probandengruppe vornehmlich in der Freizeit stattfindet und sie unter keinerlei stimmlichen Einschränkungen leiden. Ausgebildete Sänger befinden sich nicht darunter. Kein Proband gab an Raucher zu sein oder an einer asthmatischen Erkrankung zu leiden. Eine logopädische Behandlung wurde bei keinem der Probanden je durchgeführt. Alle Teilnehmer verneinten eine häufig auftretende Heiserkeit, mit Ausnahme bei gewöhnlichen Infekten der oberen Atemwege. Das Auftreten eines Refluxes wurde mit niemals oder selten bewertet und in dieser Studie als unauffällig bewertet. Die Anamnese ergab bei den schnarchenden Probanden keinen erheblichen Unterschied in der Stimmbelastung im Vergleich zu den Nicht-Schnarchern. Da eine andere Studie keine signifikanten Unterschiede zwischen Schnarchern und Nicht-Schnarchern fand und das Schnarchen von nur 3 Probanden in der Gruppe angegeben wurde [37], wurde dieser Punkt im Verlauf der Studie nicht weiter berücksichtigt.

Das Auftreten von Stimmstörungen bei intensiver Verwendung der Stimme wurde im Anamnesebogen von 11 Probanden verneint. Der Rest empfand beim Singen, bei Stress, bei intensivem Vortragen oder bei anderer stimmlicher Belastung ein vermehrtes Auftreten von Stimmstörungen. Die Regenerationsfähigkeit der Stimme wurde anhand der Dauer einer

belastungsbedingten, vorübergehenden Heiserkeit abgeschätzt. 13 Probanden gaben eine Regeneration der Stimme bereits nach 8 Stunden an. Nur zwei Probanden gaben eine Regenerationszeit von über 12 Stunden an.

4.3 VHI-12 Auswertung

Die subjektive Stimmbeurteilung der Probanden erfolgte zu Beginn und am Ende der Studie. Bei allen 17 Probanden ergab die VHI-12 Auswertung zu beiden Zeitpunkten jeweils immer eine Punktzahl unter 6 [33]. Abbildung 18 zeigt die Ergebnisse jedes einzelnen Probanden. Bei Proband 8,10 und 20 sind keine Balken zu sehen, da die Gesamtpunktzahl der Probanden 0 betrug.

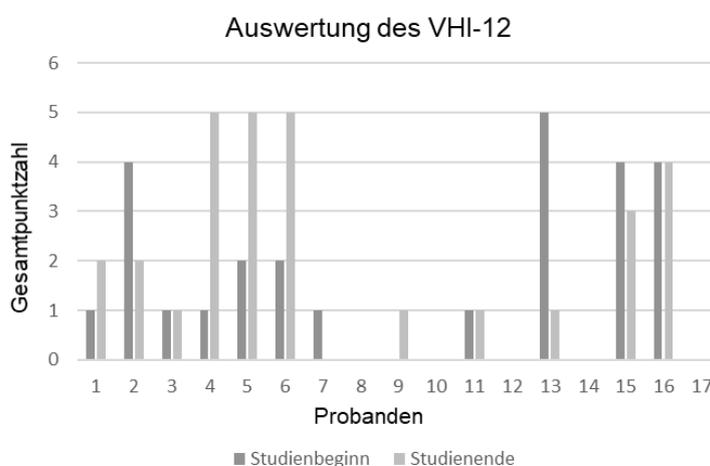


Abbildung 18: Verteilung der VHI-Werte der 17 Probanden bei Beginn und am Ende auf der x-Achse. Die Gesamtpunktzahl des VHI-12 Tests auf der y-Achse der Grafik, zeigt, dass kein Proband die 6 Punktegrenze überschreitet. Eine Punktzahl ab 8 Punkte würden für ein geringgradiges Missempfinden sprechen. Die Gesamtpunktzahl 0, die zum Beispiel bei Proband 8 angegeben wurde, entspricht einer gesund empfundenen Stimme.

4.4 Stroboskopiebefunde

Die maximale Verdeckung in den Stroboskopieaufnahmen lag bei 5 Probanden bei 25% der Stimmlippen. Alle Aufnahmen waren zur Beurteilung geeignet. Die Randkantenverschieblichkeit und die Amplitude waren bei allen Probanden unauffällig. Hinweise auf mögliche irreguläre Schließungsformen oder auf Insuffizienzen ergaben sich bei keinem der Probanden. Vereinzelt traten bei einzelnen Probanden während der Untersuchung Phasendifferenzen der Stimmlippenschwingung auf.

Bei keinem der Probanden fanden sich pathologische Veränderungen oder Entzündungen an den Stimmlippen. Die Patientenhistorie in Bezug auf bekannte Veränderungen oder zurückliegende Entzündungen, war bei allen Probanden unauffällig. In einigen Bildern finden sich weißliche Schleimauflagerungen auf den Stimmlippen, die keinen pathologischen Wert besitzen (vgl. Abbildung 19 a).

a) Proband (VG300393), m, 26 Jahre



b) Proband (GF110694), m, 25 Jahre



Abbildung 19: Öffnungsphase zweier Probanden während der Phonation. Im Bild a) beträgt die Frequenz des Probanden (VG300392) 150 Hz und die Lautstärke 82 dB. In b) phoniert der zweite Proband (GF110694) mit 138 Hz und 76 dB. Die Blitzlichtfrequenz Δf zur Grundfrequenz f_0 ist in beiden Bildern ± 1 Hz. Auf den Stimmlippen in Bild a) lassen sich weiße, nicht pathologische Schleimauflagerungen erkennen.

4.5 Ergebnisse der Stimmfeldanalyse

In der Stimmfeldanalyse wurden charakteristische Stimmparameter (mittlere Sprechstimmlage, Frequenz- und Lautstärke, Jitter, Shimmer und DSI), die Belastbarkeit und der Umfang der Stimme gemessen. Die Daten der einzelnen Probanden wurden zu Beginn und am Ende der Studie erhoben und zunächst mit sich selbst und anschließend dann mit dem Probandenkollektiv verglichen. Der durchschnittliche Zeitraum zwischen Anfangs- und Enduntersuchung betrug in dieser Studie 34 Tage. In der Auswertung der Stimmanalyse werden die Datensätze der 17 Probanden dargestellt, die die Einschlusskriterien erfüllten.

4.5.1 Sprechstimmprofil

Für die mittlere Sprechstimmlage der 17 Teilnehmer ergab sich für die Eingangsuntersuchung eine Frequenz von $106 \text{ Hz} \pm 11 \text{ Hz}$ und $109 \text{ Hz} \pm 10 \text{ Hz}$ für die Ausgangsuntersuchung. Die mittlere Sprechlautstärke zu Beginn liegt bei $62 \text{ dB} \pm 3 \text{ dB}$ und am Ende der Studie bei $66 \text{ dB} \pm 4 \text{ dB}$. Es gibt keine großen Verteilungsdifferenzen zu den beiden Zeitpunkten. Abbildung 20 a) zeigt, dass die Verteilung der mittleren Sprechstimmlage zu Beginn und Ende der Studie kaum voneinander abweichen. Der Median liegt zu Beginn bei 104 Hz und bei 105 Hz am Ende der Studie und ist nahezu identisch. Zudem liegen 50 % der Werte beider Zeitpunkte in einem Frequenzbereich von $99 - 110 \text{ Hz}$, Ausreißer liegen keine vor. Abbildung 20 b) stellt die für den jeweiligen Probanden individuell erhobene mittlere Sprechstimmlage zu Beginn und am Ende der Studie gegenüber. Die Frequenzmessung der mittleren Sprechstimmlage unterscheidet sich zu den zwei unterschiedlichen Zeitpunkten bei den einzelnen Probanden nur geringfügig.

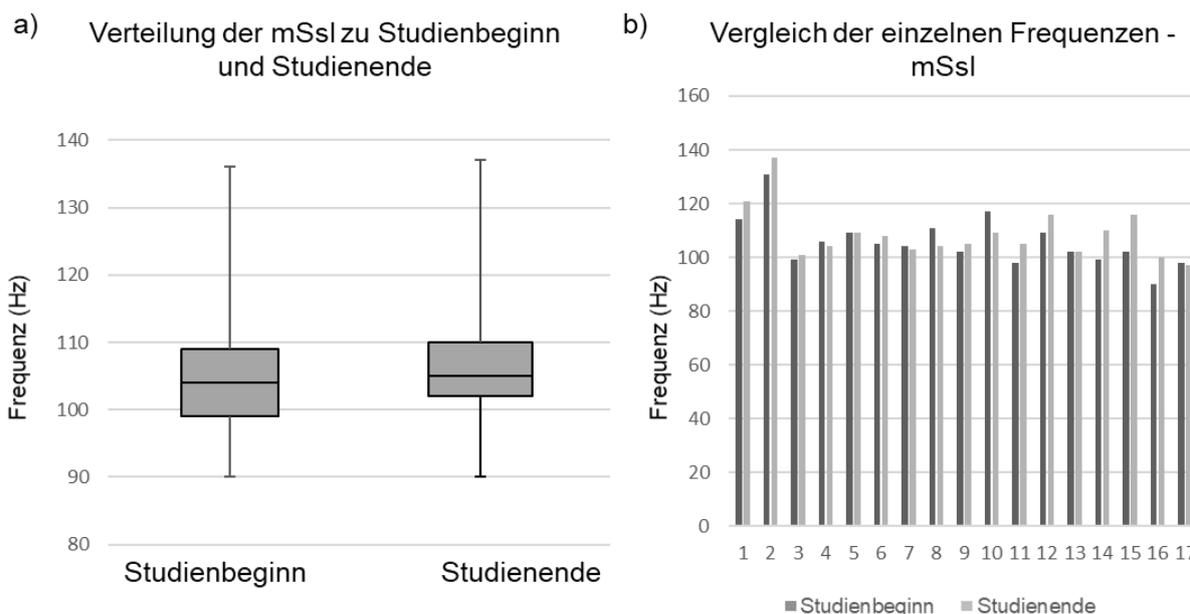


Abbildung 20: a) Die Grafik zeigt anhand zweier Boxplots die Verteilung der mittleren Sprechstimmlage zu Studienbeginn und -ende. Beide Boxplots zeigen, dass 50 % der mSsl aller Probanden zu beiden Zeitpunkten in einem Frequenzbereich von 99 – 110 Hz liegen. Die Verteilung zu beiden Zeitpunkten ist sehr ähnlich. b) Das Balkendiagramm zeigt jeweils nur eine sehr geringe Abweichung der mSsl zu Beginn und am Ende der Studie bei den jeweiligen Probanden.

Für die maximale Ruflautstärke ergaben sich zu Beginn eine Lautstärke von $104 \text{ dB} \pm 5 \text{ dB}$ und $106 \text{ dB} \pm 6 \text{ dB}$ am Ende. Die Frequenz der Rufstimme wird aufgrund ihrer physiologischen weiten Spanne zur Vervollständigung der Daten aufgeführt. Der Mittelwert zu Beginn der Studie liegt bei $316 \text{ Hz} \pm 56 \text{ Hz}$ und am Ende bei $301 \text{ Hz} \pm 59 \text{ Hz}$. Die Berechnung des Mittelwertes des Tonumfangs zu beiden Zeitpunkten konnte bei fehlenden Werten aufgrund einer Speicherproblematik bei „LingWAVES“ nicht erfolgen. Der durchschnittliche Dynamikbereich von 13 Probanden liegt in der Eingangsuntersuchung bei $36 \text{ dB} \pm 5 \text{ dB}$ und bei $39 \text{ dB} \pm 8 \text{ dB}$ in der Ausgangsuntersuchung. Die Abbildungen zu Lautstärke der mittleren Sprechstimmlage, Dynamikbereich und Frequenz und Lautstärke der Rufstimme sind im Anhang in Abbildung 63 zu finden, da sie keinen direkten Einfluss auf den weiteren Verlauf der Studie hatten. Zu beachten ist, dass aufgrund technischer Schwierigkeiten bei der Datenspeicherung nicht bei allen Probanden ein vollständiges Stimmprofil gespeichert wurde. Da jedoch nur einzelne Werte fehlen, die nicht zu den Ein- oder Ausschlusskriterien der Studie zählten und die keine Relevanz für die Einstellung der Handy-App und für das Stimmmonitoring hatten, besteht kein negativer Einfluss auf die Studie.

Eine Stimmaufnahme des Textes „Der Nordwind und die Sonne“ [93] wurde für jeden Probanden mit dem RBH-Test bewertet. Hier ergab sich für alle Probanden das klinisch unauffällige Ergebnis R0B0H0.

4.5.2 Singstimmprofil

In der Singstimmprofilmessung wurden Maxima und Minima von Lautstärke und Frequenz, sowie der Dynamikbereich der Singstimme, der Jitter und der DSI der 17 teilnehmenden Probanden zu Beginn und am Ende der Studie ermittelt.

Bei Studieneinschluss liegt der Mittelwert der maximalen Frequenz der Singstimme aller Probanden bei $597 \text{ Hz} \pm 163 \text{ Hz}$, bei Studienausschluss bei $598 \text{ Hz} \pm 146 \text{ Hz}$. Die Abbildung 21 a) zeigt den Mittelwert und die Standardabweichung der maximalen Frequenz aller 17 Probanden zu Beginn und am Ende der Studie, die Verteilung ist zu beiden Zeitpunkten ähnlich. Bei der Erhebung der maximalen Frequenz, war einigen Probanden das Singen der hohen Töne unangenehm und eine maximale Ausschöpfung der Frequenz wurde nicht immer erreicht. Das Bild b) der Abbildung 21 zeigt die Mittelwerte und die SD des leisesten Schallpegel des Probandenkollektivs bei der Ein- und Ausschlussuntersuchung. Die Werte sind identisch und liegen bei $47 \text{ dB} \pm 4 \text{ dB}$.

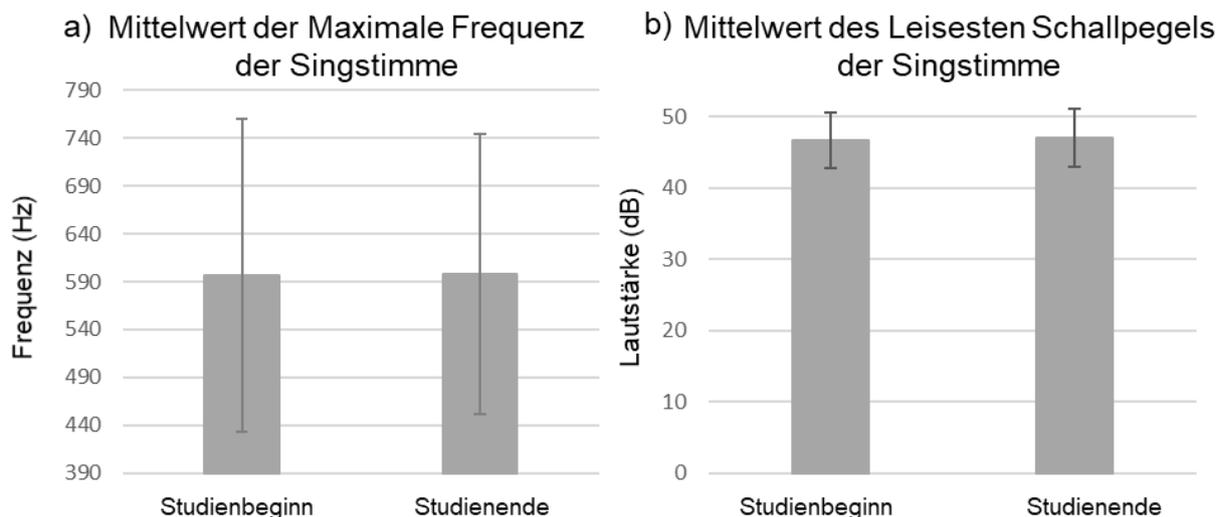


Abbildung 21: a) Die beiden errechneten Mittelwerte der maximalen Frequenz der 17 Teilnehmer sind zu Studienbeginn ($597 \text{ Hz} \pm 163 \text{ Hz}$) und Studienende ($598 \text{ Hz} \pm 146 \text{ Hz}$) nahezu identisch und ähnlich verteilt. b) Der Mittelwert und die Abweichung des leisesten Schallpegels ist zu beiden Zeitpunkten identisch und liegt bei $47 \text{ dB} \pm 4 \text{ dB}$.

Im Singstimmprofil wurden zudem die minimale Frequenz, der maximale Schallpegel und der Dynamikumfang ermittelt. Die Grafiken der Werte sind im Anhang in Abbildung 64 zu finden. Aufgrund der Speicherproblematik waren nicht mehr alle Daten von allen Probanden vorhanden, die jedoch keine weitere Relevanz für die Studie hatten. Der Mittelwert der minimalen Frequenz von 13 Probanden war bei Studieneinschluss $85 \text{ Hz} \pm 8 \text{ Hz}$, bei Ausschluss $87 \text{ Hz} \pm 5 \text{ Hz}$. Die Mittelwerte der durchschnittlich maximalen Lautstärke sind nahezu identisch, $91 \text{ dB} \pm 6 \text{ dB}$ und $91 \text{ dB} \pm 7 \text{ dB}$. Das laute, kräftige Singen wurde von einigen Probanden als unangenehm empfunden und eine maximale Ausschöpfung der Stimmleistung war nicht bei allen gegeben. Der Mittelwert der Tonhöhendynamik von 16 Probanden liegt jeweils bei $44 \text{ dB} \pm 7 \text{ dB}$ und $45 \text{ dB} \pm 8 \text{ dB}$.

Ergebnisse

Am Ende der Singstimmfeldmessung wurde der Jitter aus der gehaltenen Phonation eines Vokals berechnet. In der Einschlussuntersuchung lag der Mittelwert des Jitters von den 17 Teilnehmern bei $0,12\% \pm 0,047$, in der Ausschlussuntersuchung bei $0,11\% \pm 0,035$. Die Abbildung 22 zeigt, dass sich die Mittelwerte und ihre Standardabweichungen kaum unterscheiden. Außerdem zeigt die Abbildung, dass alle erhobenen Werte unter der pathologischen 1% Grenze liegen.

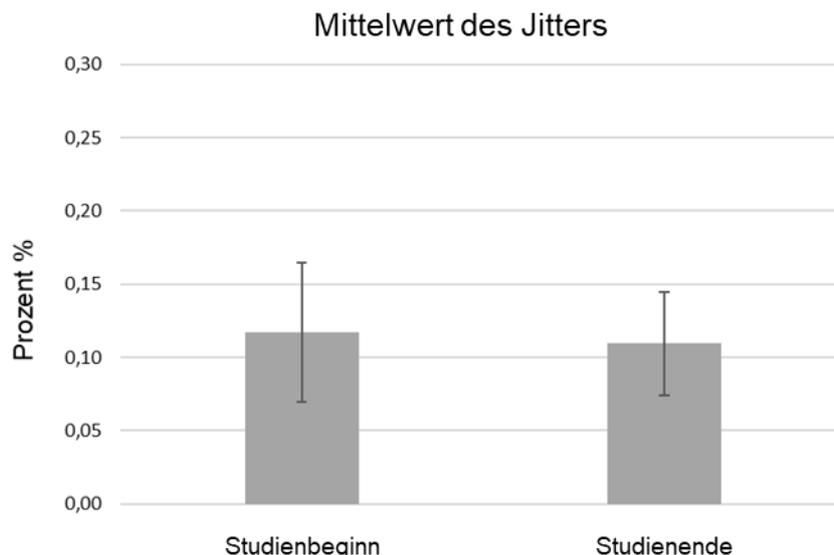


Abbildung 22: Darstellung der berechneten Mittelwerte des Jitters der 17 Studienteilnehmer zu Beginn und am Ende der Studie. Die Werte sind zu beiden Zeitpunkten sehr ähnlich ($0,12\% \pm 0,047$ und $0,11\% \pm 0,035$). Zudem liegt die Verteilung aller Jitter Werte in einem sehr kleinen Bereich und im physiologischen Bereich $< 1\%$.

Die in der Singstimmfeldmessung ermittelten Werten ermöglichen nur die Berechnung des Dysphonia Severity Index (DSI). Die Berechnung erfolgt aus der Tonhaltedauer, der oberen Stimmfrequenz, dem Jitter und dem leisesten Schallpegel des Singstimmfeldes. Der Mittelwert der einzelnen Parameter aus der Studie, sowie der durchschnittliche DSI sind in Tabelle 1 dargestellt. Ein angestrebter $DSI > 4,2$, welcher für eine physiologische Stimme spricht [78], wurde bei fast jedem Probanden bei Studienbeginn und -ende erreicht. Der errechnete Mittelwert liegt mit $6,4 \pm 1,8$ bei Beginn und $6,3 \pm 1,5$ am Ende der Studie über dem erstrebten Grenzwert $>4,2$.

Tabelle 1: Berechnung des DSI aus den Mittelwerten der Tonhaltedauer, der oberen Grenzfrequenz, dem niedrigsten Schallpegel und des Jitters der 17 Studienteilnehmer, die zu Studienbeginn und Studienende in der Singstimmfeldmessung bestimmt wurden. Anschließend Vergleich der erhobenen Werte mit vergleichbaren Referenzmittelwerten aus der Literatur [12, 78, 93].

	<i>Einheit</i>	Studienbeginn	Studienende	Referenzwerte
Tonhaltedauer	<i>Sek</i>	25	23	>15
Obere Grenzfrequenz	<i>Hz</i>	597	598	523
Niedrigster Schallpegel	<i>dB</i>	47	47	51
Jitter	<i>%</i>	0,12	0,11	<1
DSI		6,4	6,3	$>4,2$

Die Abbildung 23 zeigt die Verteilung um den Median des mit „LingWAVES“ berechneten DSI zu Beginn 6,3 und am Ende der Studie 6,6. Mittelwert und Median des DSI sind fast identisch. Die Verteilung um den Median ist gleichmäßig, es gibt keine Ausreißer.

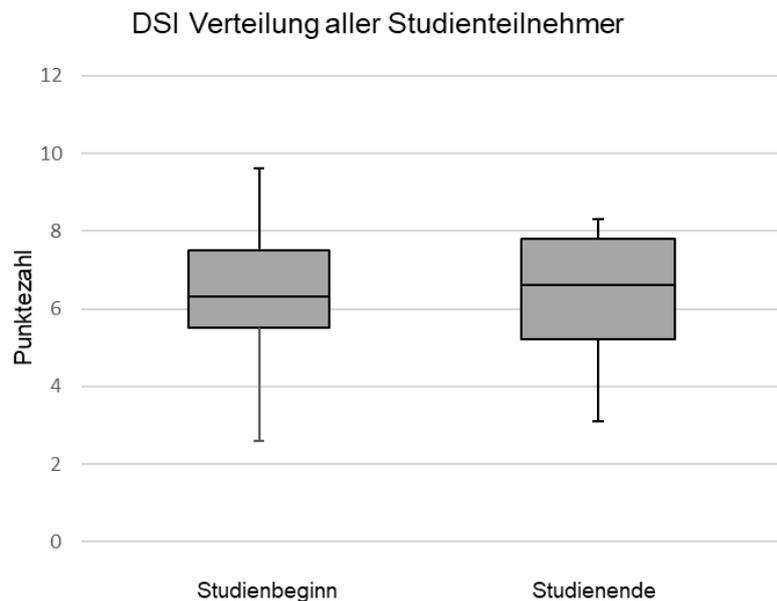


Abbildung 23: Die Boxplots stellen die Verteilung der DSI Werte aller Probanden zu Studienbeginn und- ende gegenüber. Der Mediane ist zu Beginn 6,3 und am Ende der Studie 6,6, Die Verteilung ist sehr regelmäßig und unterscheidet sich zu besagten Zeitpunkten wenig. Es gibt keine Ausreißer und mehr als 75% der Werte liegen über der gewünschten Norm von 4,2 [78].

Die individuelle Betrachtung zeigt, dass bei Proband 1 beide DSI-Werte (3 und 3,1) unterhalb der gewünschten 4,2 sind, bei Proband 8 nur der Anfangswert, der bei 2,6 lag.

4.6 Ergebnisse von „Voice Diary“

Aus der App „Voice Diary“ wurde die Anzahl der Aufnahmen ermittelt und der Fragebogen ausgewertet und zusammengefasst. Für die Auswertung des Stimmtagebuchs waren wiederum die Datensätze von 17 Probanden vergleichbar. Insgesamt wurden 1386 Aufnahmen in dieser Studie analysiert, das entspricht im Mittel etwa 81 Aufnahmen je Probanden. Die ursprünglich 6 festgelegten Zeitintervalle, in welchen die Probanden je eine Aufnahme machen sollten, wurde in Anbetracht der Verteilung der Aufnahmen und der Auswertung auf 5 Zeitintervalle gekürzt, welche von 5-10 Uhr, 10-14 Uhr, 14-18 Uhr, 18-23 Uhr und 23-5 Uhr waren. Der Boxplot in Abbildung 24 a) zeigt die allgemeine Verteilung aller Aufnahmen von allen Studienteilnehmern zu den unterschiedlichen Zeiten an einem Tag. Die meisten Aufnahmen erfolgten von 14-18 Uhr und die wenigsten in den Morgenstunden von 23-5 Uhr. Im Zeitraum von 14-18 Uhr haben 50 % der Probanden zwischen 17 und 25 Aufnahmen gemacht. Die Mediane liegen angefangen beim ersten Intervall (23-5 Uhr) bei 9, im zweiten Intervall (5-10 Uhr) bei 12, im dritten Intervall (10-14 Uhr) bei 21, im vierten Intervall (14-18 Uhr) bei 19 und (18-23 Uhr) bei 18 Aufnahmen. Die Mittelwerte der Aufnahmeanzahl

liegen angefangen beim ersten Intervall bei 10 ± 4 , 12 ± 3 , 21 ± 5 , 19 ± 5 und 18 ± 4 Aufnahmen. Median und Mittelwert sind in den jeweiligen Intervallen fast identisch.

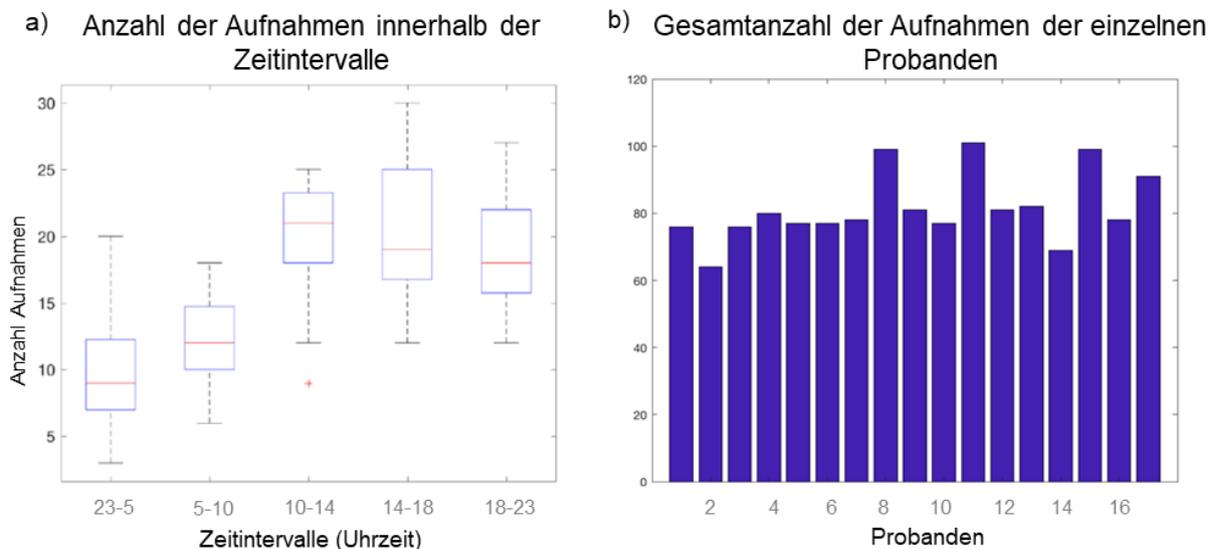


Abbildung 24: a) Der Boxplot zeigt die Verteilung aller Aufnahmen von allen 17 Probanden in den einzelnen Zeitintervallen, welche von 5-10 Uhr, 10-14 Uhr, 14-18 Uhr, 18-23 Uhr und 23-5 Uhr waren. Ein Großteil der Aufnahmen wurde zwischen 10 und 23 Uhr aufgenommen. b) Das Säulendiagramm zeigt die Gesamtanzahl der Aufnahmen von jedem einzelnen Probanden innerhalb der vorgegebenen zwei Wochen. Bei allen Probanden stehen mehr als 60 Aufnahmen zur Auswertung zur Verfügung.

Abbildung 24 b) zeigt, wie viele Aufnahmen die Probanden im Einzelnen über den 14-tägigen Zeitraum gemacht haben. Die Betrachtung der Gesamtanzahl der Aufnahmen ergab ein Minimum von etwa 62 und ein Maximum von etwa 103 Aufnahmen. Die Probanden 8, 12 und 15 haben die App häufig verwendet und mehr als 100 Aufnahmen aufgezeichnet.

4.6.1 Stimmbelastung

Im Fragenkatalog folgten zunächst die Fragen zu den Belastungsgraden der Stimme und nach der Art der Stimmbelastung, der die Stimme vor der Aufnahme ausgesetzt war. Die Verteilung der Ergebnisse wird im folgenden Boxplot dargestellt (vgl. Abbildung 25). 50% der Probanden, die keine Belastung beim Stimmgebrauch angaben, gaben dies zwischen 10 und 18 Uhr an. Die späteste Uhrzeit für eine Belastung des Grad 1 war bei 22 Uhr, die Früheste lag bei 6 Uhr morgens. Ähnlich dazu verhält sich die Verteilung eines leichten Belastungsgrad der Stimme (Grad 2). Bei den Aufnahmen, die am Nachmittag oder am Abend aufgezeichnet wurden, wurde eine Höhergradige Stimmbelastungen wie Grad 3 und 4 angegeben. 50 % der Probanden, die Grad 3 gewählt haben, gaben den Belastungsgrad bei Aufnahmen zwischen 15 bis 19 Uhr an, die Grad 4 gewählt haben, zwischen 14 bis 21 Uhr. Im Gegensatz dazu beschränken sich die zeitlichen Angaben einer sehr starken Stimmbelastung (Grad 6) bei 50 % der betroffenen Probanden auf 13-15 Uhr.

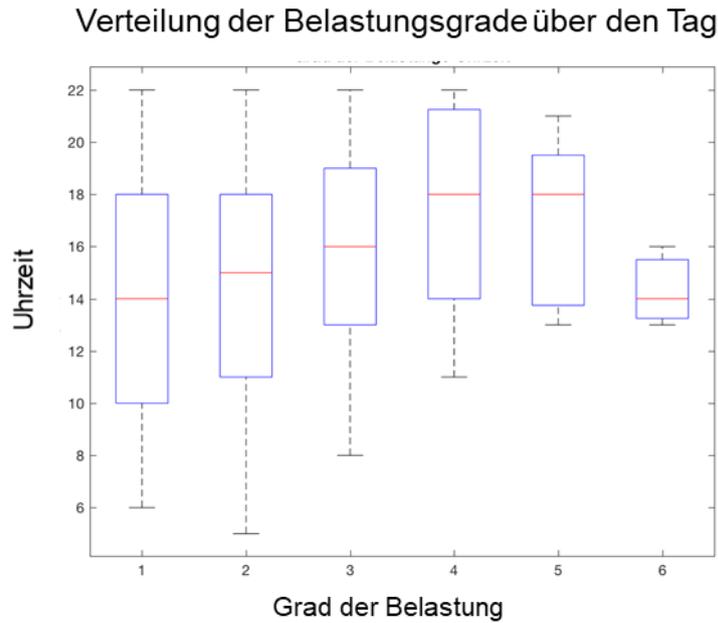


Abbildung 25: Der Boxplot zeigt die Verteilung der sechs unterschiedlichen Belastungsgrade der Stimme im Tagesverlauf, welche nach jeder Stimmaufnahme abgefragt wurde. Höhergradige Stimmbelastungen treten vorwiegend im späteren Verlauf des Tages auf.

Da sich die Informationen des Boxplots nur auf den Grad der Belastung in Bezug zur Uhrzeit beziehen und die absolute und relative Häufigkeit der gewählten Stimmbelastung vernachlässigt werden, soll Abbildung 26 einen besseren Überblick über diese Verteilung geben. Abbildung 26 a) zeigt die absolute Häufigkeit der unterschiedlichen Belastungsgrade, die Verteilung ist im Tagesverlauf ungleich. Abbildung 26 b), in welcher die relative Häufigkeit der einzelnen Belastungsgrade visualisiert ist, zeigt, dass der Belastungsgrad der Stimme nach der Nachtruhe niedriger war als zu anderen Uhrzeiten. Von 10 bis 22 Uhr gibt jeweils nur noch etwa die Hälfte der Probanden keine Stimmbelastung an, der Rest gibt mit absteigender Häufigkeit Stimmbelastungen bis Grad 4 an. Während sich die relative Häufigkeit von Belastungsgrad 2 und 3 kaum ändern, steigt die Häufigkeit der Grad 4 Belastung bis 22 Uhr an. Stimmaufnahmen, die nachts aufgenommen wurden, wurden häufiger mit einem Belastungsgrad 3 bewertet als in den Stunden zuvor. Weiterhin erwähnenswert ist, dass in dieser Studie eine Stimmbelastung von Grad 5 oder 6 fast nie empfunden wurde.

Stimmbelastung

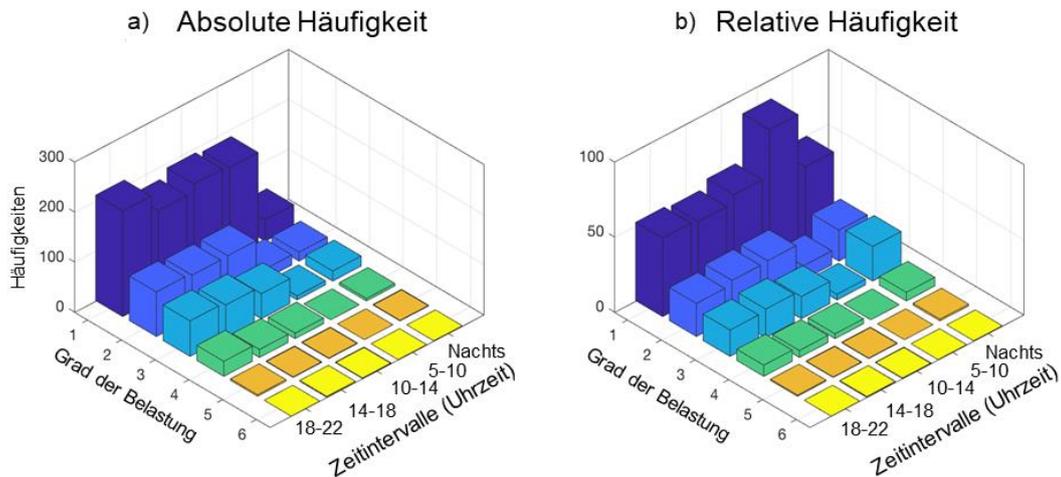


Abbildung 26: a) Die Verteilung der absoluten Häufigkeit der Belastungsgrade zeigt in den unterschiedlichen Intervallen eine ungleiche Verteilung. b) Hier wird die relative Häufigkeit der 6 Belastungsgrade gezeigt. Die Verteilung zeigt, dass die Stimme ab 10 Uhr einer zunehmenden Stimmbelastung ausgesetzt war. Zudem steigt die relative Häufigkeit der Grad 4 Belastung von 10 bis 22 Uhr kontinuierlich an, wohingegen sich die Häufigkeiten von Belastungsgrad 2 und 3 in diesem Zeitraum kaum ändern.

Abbildung 27 stellt die relative Häufigkeit der subjektiven Stimmbelastung aufgeschlüsselt auf die einzelnen 17 Probanden dar. Die Stimmbelastung wurde von den Probanden sehr unterschiedlich wahrgenommen. Proband 1 und 5 haben selten oder nie eine verstärkte Belastung empfunden. Proband 10, der seine allgemeine Gesprächigkeit als normal einschätzte, hingegen gab Stimmbelastungen bis Grad 5 an.

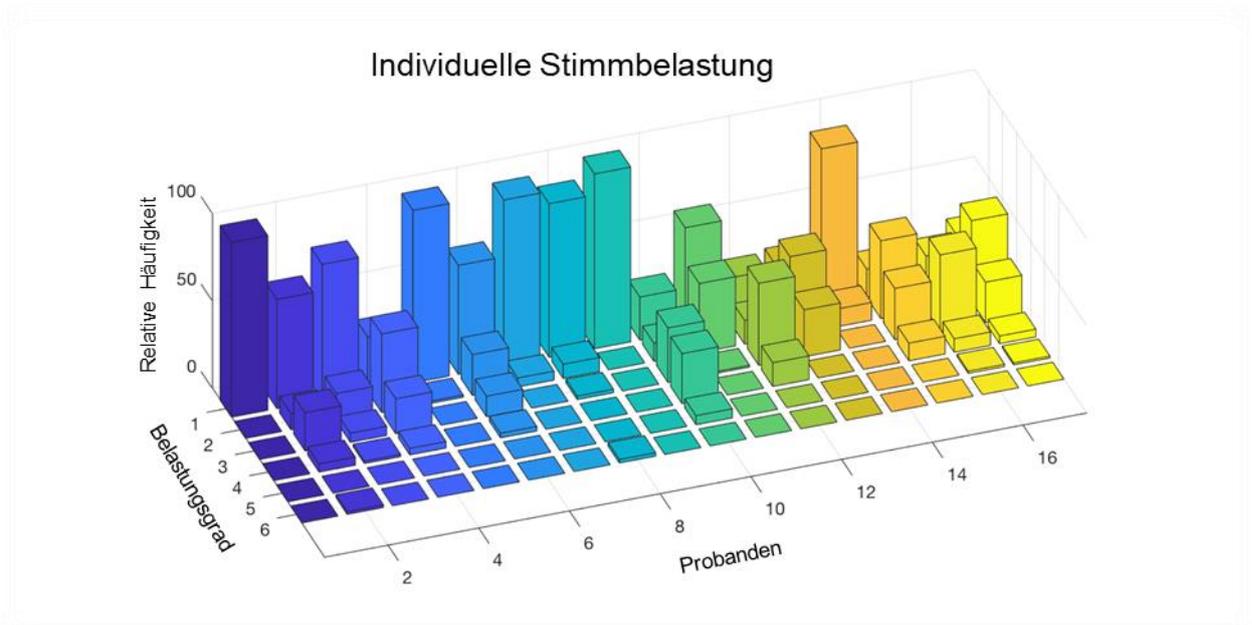


Abbildung 27: Relative Häufigkeit der subjektive Stimmbelastung aufgeschlüsselt auf die einzelnen Probanden. Die subjektive Belastung der Stimme wurde von den Probanden sehr unterschiedlich empfunden. Proband 1 und 5 haben beispielsweise nie oder nur selten eine Belastung der Stimme angegeben, bei Proband 10 hingegen lang fast immer ein Belastungsgrad vor.

Neben dem Grad wurde auch nach der Art der Belastung gefragt. Auszuwählende Optionen waren hier „Sprechen“, „Singen“ oder „Rufen“. Die Abbildung 28 beschreibt in a) die relative Häufigkeit der Art der Stimmbelastung in den unterschiedlichen Zeitintervallen. Hier zeigt sich, dass vor allem „Sprechen“ die häufigste Belastungsart im Studienzeitraum war. Außerdem fällt auf, dass die relative Aufnahme Häufigkeit von 5-10 Uhr sowie nachts geringer ist, als die Aufnahmen von 10–22 Uhr. Der Hauptnutzungszeitraum der App liegt also zwischen 10-22 Uhr. Dies zeigt auch die Verteilung der Aufnahmeanzahl in Anbetracht der Zeitintervalle (vgl. Abbildung 24). Die Abbildung 28 b) zeigt, wie häufig jeder Proband die drei Arten der Stimmbelastung in Bezug auf all seine Aufnahmen gewählt hat. Fast jeder Proband gab als häufigste stimmliche Belastung das Sprechen an. Gelegentlich wurden auch Rufen und Singen als Belastung angegeben. Anders verhält es sich bei Proband 1, 5 und 14. Bei Proband 1 wird keine relative Häufigkeit angegeben. Dies liegt daran, dass er während des Studienzeitraums seine Stimme nicht als belastet gewertet hat (vgl. Abbildung 27). Die Frage nach der Art der Belastung wurde im Fragebogen daher übersprungen. Proband 5 und 14 gaben eine geringe individuelle Stimmbelastung an (vgl. Abbildung 27). Trat nun eine Stimmbelastung auf, zeigt Abbildung 28 b), dass bei Proband 5 diese hauptsächlich durch „Rufen“ verursacht wurde. Proband 14 gab eine Stimmbelastung vor allem nach dem Singen an.

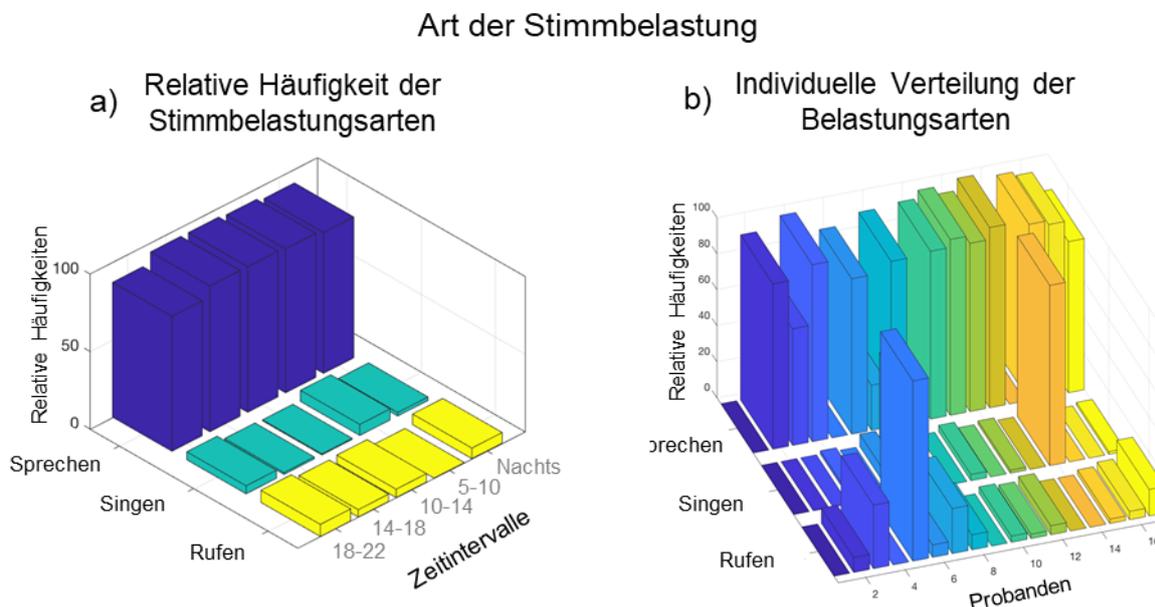


Abbildung 28: a) Relative Häufigkeit der Stimmbelastungsart in Bezug auf alle Aufnahmen. Die Auswertung zeigt, dass Sprechen die häufigste Belastung für die Stimme der Teilnehmer war. Abbildung 28 b) zeigt die relative Häufigkeit der individuellen Verteilung der unterschiedlichen Belastungsarten bei jedem einzelnen Probanden. Proband 5 und 14 empfanden eine Stimmbelastung hauptsächlich beim Rufen und Singen.

4.6.2 Stimmqualität

Anschließend wurden die Probanden nach jeder Aufnahme gebeten Ihre Stimme einzuschätzen. Diese konnte von 1 „sehr schlechte Stimmqualität“ bis 6 „sehr gute Stimmqualität“ bewertet werden (vgl. Abbildung 16). Einige Probanden gaben nach der Studie an, dass sie eine gute Stimme analog zum deutschen Notensystem mit einer 1 anstatt einer 6 bewertet haben. Die umgekehrte Skalierung der Bewertungsgrade bei dieser Frage hat laut einigen Probanden für Missverständnisse gesorgt. Die Verteilung der relativen Häufigkeit in Abbildung 29 a) zeigt, dass nach kaum einer Aufnahme die Stimmqualität als gut oder sehr gut bewertet wurde und ein Großteil der Aufnahmen mit „sehr schlecht“ (Grad 1) oder „schlecht“ (Grad 2) bewertet wurden und dass die Stimmqualität in über der Hälfte der Fälle zu allen Uhrzeiten als sehr schlecht bewertet wurde. Eine kleine Abweichung findet man bei den Aufnahmen von 5 bis 10 Uhr. Hier werden im Vergleich zu anderen Uhrzeiten weniger Aufnahmen mit „sehr schlecht“ bewertet. Die relative Häufigkeit von Aufnahmen mit der Stimmqualität Grad 4 ist in allen Zeitabschnitten gleichmäßig verteilt.

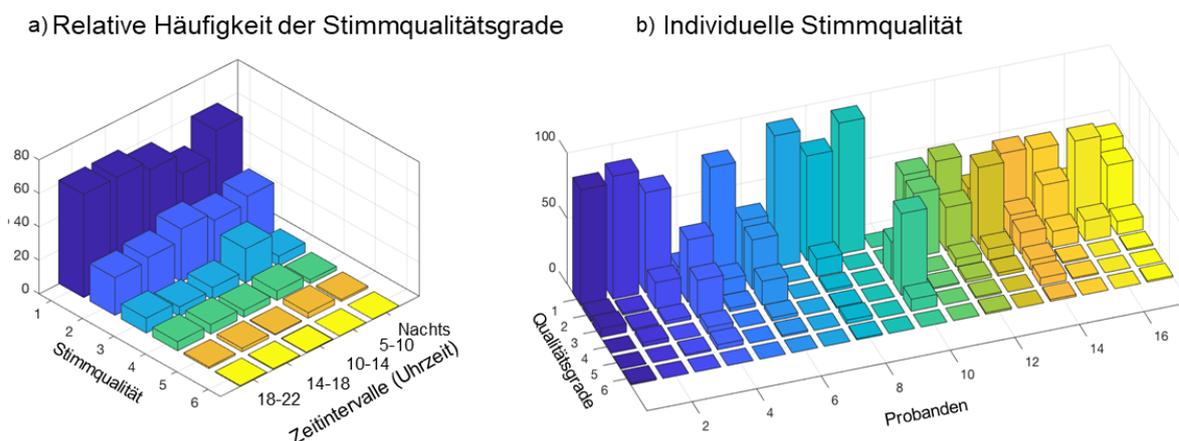


Abbildung 29: a) Relative Verteilung der Stimmqualitätsgrade von „sehr schlecht“ (Grad 1) bis „sehr gut“ (Grad 6) in den jeweiligen Zeitintervallen. Über die Hälfte der Aufnahmen wurden mit „schlecht“ oder „sehr schlecht“ bewertet. Die relative Häufigkeit der Verteilung von 5-10 Uhr zeigt, dass die Aufnahmen im Vergleich etwas besser bewertet wurden. Grad 3 und 4 treten häufiger als zu anderen Zeiten auf. b) Subjektive Einschätzung der Stimmqualität durch die Probanden. Ein Großteil der Probanden bewertet die Aufnahmen mit einer „schlechten“ oder „sehr schlechten“ Stimmqualität.

In Abbildung 29 b) wird die Verteilung der 6 unterschiedlichen Bewertungsgrade der Stimmqualität für die einzelnen Probanden im Hinblick auf die gesamte Anzahl ihrer Aufnahmen veranschaulicht. Die subjektive Einschätzung kann man anhand der Verteilungsmuster in Abbildung 29 b) in drei Gruppen gliedern. Die erste Gruppe besteht aus Proband 1-9 mit Ausnahme von Proband 4 und 6. Diese Gruppe gab in über der Hälfte der Aufnahmen eine sehr schlechte Stimmqualität an. Die zweite Gruppe, die Proband 11 bis 17 umfasst, bewertet die Qualität der Stimme vor allem mit Grad 1 und 2. Proband 4, 6 und 10 zählen zu den „Ausreißern“. Die Stimmen wurden eher als schlecht oder nur mittelmäßig

bewertet. Auffallend ist vor allem Proband 10, der seine Stimmqualität zum größten Teil eher als ausreichend (Grad 4) wahrnimmt. Vergleicht man nun die Ergebnisse aus Abbildung 29 a) und b) miteinander, wird deutlich, dass das Auftreten des Stimmqualitätsgrades 4 vor allem durch Proband 10 verursacht wurde.

Die Verteilung der subjektiven Bewertung im Tagesverlauf zeigt keinen bestimmten Trend in der Qualitätsentwicklung (vgl. Boxplot Abbildung 30). Die Grade 1 bis 5 wurden durchgehend über den Tag verteilt angegeben. 50 % der Bewertungen mit „sehr schlechter“ (Grad 6) Stimmqualität wurden im Zeitraum von 15 bis 21 Uhr abgegeben.

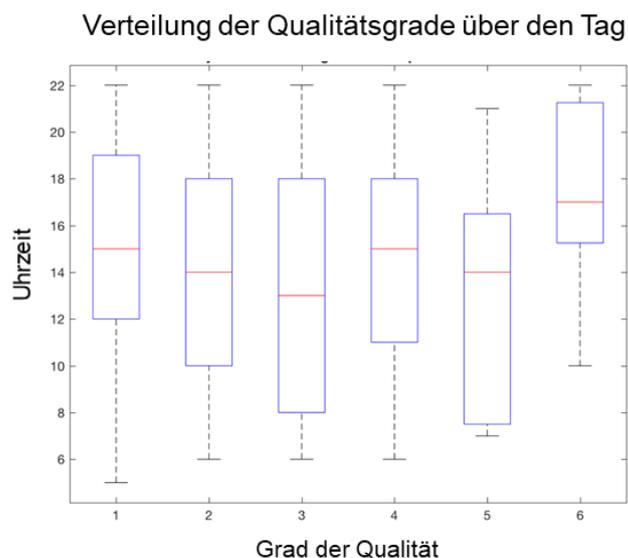


Abbildung 30: Der Boxplot zur zeitlichen Verteilung der unterschiedlichen Stimmqualitäten aller Aufnahmen zeigt keine bestimmte Verteilung der unterschiedlichen Qualitätsgrade. 50 % der Stimmaufnahmen mit „sehr guter“ (Grad 6) Qualität wurden zwischen 15 und 21 Uhr aufgenommen.

4.6.3 Missempfindungen

Mit Hilfe der Frage über vorliegende Missempfindungen beim Stimmgebrauch wurden Qualitäten wie „Kratzen“, „Räuspern“, ein „Enge- oder Fremdkörpergefühl“ oder „Atemnot“ abgefragt. In Abbildung 31 a) wird die relative Häufigkeit des Auftretens unterschiedlicher Missempfindungsqualitäten in den zeitlichen Intervallen dargestellt. Die Verteilung zeigt, dass bei einem Großteil der Aufnahmen keine Missempfindungen vorlagen. Bei der ersten Aufnahme des Tages (5-10 Uhr) wurden von einzelnen Probanden ein Kratzen, Räuspern und seltener ein Engegefühl angegeben. Im Tagesverlauf zeigt sich, dass die Tendenz dieser Empfindung rückläufig ist (vgl. Abbildung 31 a). Da die Ergebnisse der relativen und absoluten Wahrscheinlichkeit stark übereinstimmen, wurde in der Grafik nur die relative Häufigkeitsverteilung angegeben. Abbildung 31 b) ermöglicht einen genauen Einblick über die individuelle Bewertung von Missempfindungen der einzelnen Probanden. Die Grafik zeigt mit welcher relativen Häufigkeit ein Proband bestimmte Qualitäten von Missempfindungen in Bezug auf die Gesamtanzahl seiner Aufnahmen angegeben hat. Es zeigt sich, dass nur

wenige Probanden bei den Aufnahmen ein Kratzen oder ein Räuspern angegeben haben. Die Probanden 6, 12 und 14 haben in einigen Aufnahmen häufiger ein Kratzen empfunden. Die Probanden 4, 6, 8 und 14 haben etwas häufiger ein Räuspern empfunden. Angaben über ein öfter auftretendes Engegefühl wurden nur von Proband 14 gelegentlich beschrieben. Ein Fremdkörpergefühl oder Atemnot wurde von keinem Probanden im Verlauf der Studie angegeben. Der Vergleich der Abbildung 27 und Abbildung 31 b) ergibt, dass auch bei einer stärker empfundenen Stimmbelastung Missempfindungen nicht vermehrt aufgetreten sind. Die Probanden 15 bis 17 beispielsweise geben im Tagesverlauf Belastung bis Grad 4 an und eher selten keine Belastung an. Die relative Häufigkeit an Missempfindungen ist trotz der höheren Belastung sehr gering.

Art der Missempfindungen

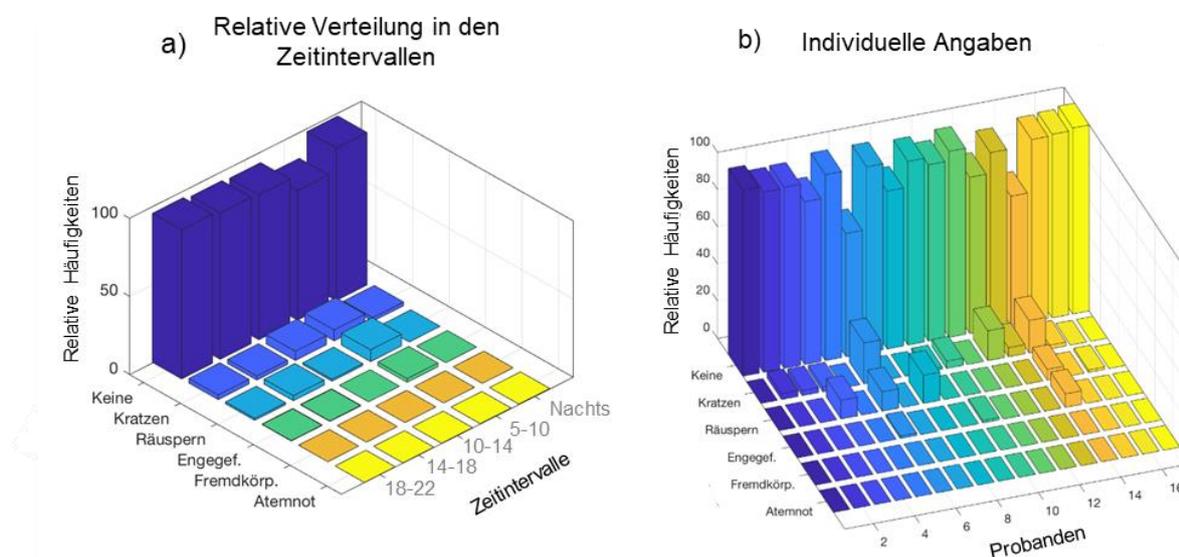


Abbildung 31: a) Verteilung der relativen Häufigkeit von unterschiedlichen Missempfindungen zu unterschiedlichen Zeitpunkten. Die Ergebnisse der Analyse zeigen, dass nur in einzelnen Fällen Räuspern oder Kratzen im Hinblick auf alle Aufnahmen angegeben wurden und ein Missempfinden häufiger morgens auftritt. b) Geringe Relative Häufigkeit bestimmter Qualitäten von Missempfindungen in Bezug auf die Gesamtzahl der einzelnen Aufnahmen der einzelnen Probanden.

4.6.4 Ermüdungsgrade

In der letzten Frage sollten die Probanden den Ermüdungsgrad ihrer Stimme von 1 „keine Ermüdung“ bis 6 „sehr starke“ Ermüdung bewerten. Zu beachten ist, dass einige Probanden am Ende der Studie angegeben haben, dass die Fragen nach dem Ermüdungsgrad etwas unpräzise sei. Daher haben einzelne Probanden die Frage nach dem Ermüdungsgrad auf ihren Allgemeinzustand bezogen. In Abbildung 32 wird die relative Häufigkeit der Ermüdungsgrade in Bezug auf die 5 Zeitintervalle und das individuelle Empfinden der Probanden in Bezug auf die Ermüdung gezeigt. Die relative Häufigkeit in Abbildung 32 a) zeigt, dass bei mehr als der Hälfte aller Aufnahmen in den jeweiligen Zeitabschnitten keine erhöhte

Müdigkeit (Grad 1) angegeben wurde. Die relativen Häufigkeiten höherer Ermüdungsgrade nehmen in allen Zeitintervallen ab. Im Tagesverlauf zeigt sich eine steigende Tendenz der Stimmermüdung. In den Intervallen von 18-22 Uhr und nachts haben weniger Probanden Grad 1 angegeben und die Stimmermüdungsgrade von Grad 4 und Grad 5 wurden vermehrt empfunden. In Abbildung 32 b) wird gezeigt, wie häufig jeder einzelne Proband in Anbetracht all seiner Aufnahmen, die jeweils 6 unterschiedlichen Ermüdungsgrade gewählt hat. Während einige nie eine Ermüdung ihrer Stimmen empfanden (Proband 1 und 5), gaben andere, beispielsweise die Probanden 4, 10 und 12 häufiger eine leichte bis mittelschwere Stimmermüdung an.

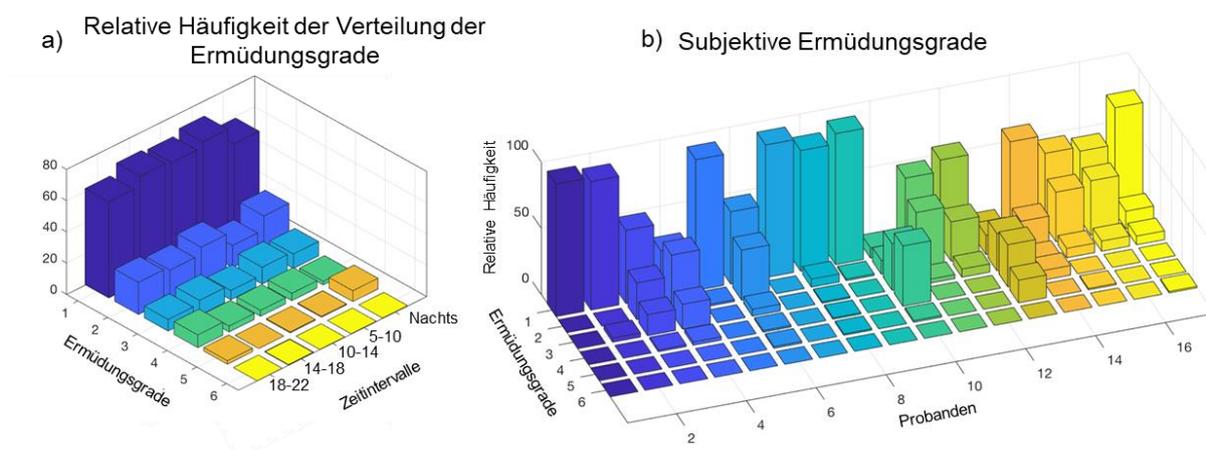


Abbildung 32: a) Relative Häufigkeiten der Ermüdungsgrade in den unterschiedlichen Zeitintervallen. Bei über der Hälfte aller Aufnahmen wurde keine Stimmermüdung angegeben, die relative Häufigkeit der Ermüdungsgrade war in jedem Zeitintervall abnehmend. Zwischen 18-5 Uhr morgens traten im Vergleich höhere Ermüdungsgrade (Grad 4 und 5) auf. b) Subjektive Einschätzung der Ermüdungsgrade der einzelnen Probanden. Hierbei liegen große Unterschiede vor, während Proband 1 und 5 nie eine Belastung empfinden, geben andere, beispielsweise Proband 11 häufiger mittelschwere Ermüdungsgrade an.

In Abbildung 33 wird die Verteilung der Ermüdungsgrade aller Aufnahmen im Hinblick auf den Tagesverlauf gezeigt. Wie bereits in Abbildung 32, nimmt in Abbildung 33 die Stimmermüdung zu späteren Uhrzeiten zu. Der Boxplot zeigt, dass sich die Aufnahmen, die jeweils mit Ermüdungsgrad 1 bis 3 bewertet wurden, in einem ähnlichen Zeitfenster (10-18 Uhr) verteilen. Höhergradige Stimmermüdung treten zu späteren Zeiten auf. 50 % der Grad 4 Bewertungen wurden zwischen 13 bis 20 Uhr, 50% der Grad 5 Bewertungen zwischen 16 bis 22 Uhr angegeben. Eine Stimmermüdung 6 Grades kam bei einer Aufnahme um 17 Uhr vor.

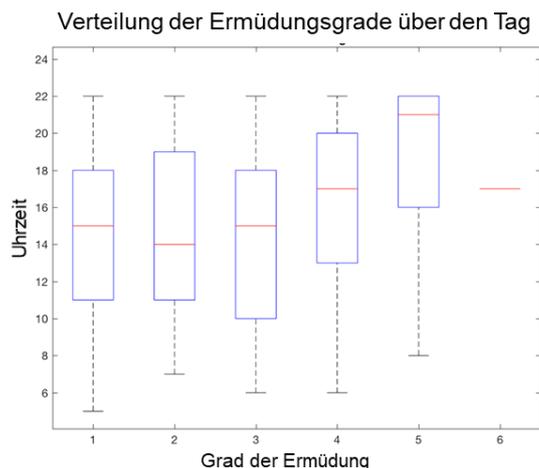


Abbildung 33: Verteilung der unterschiedlichen Bewertungsgrade zur Stimmermüdung aller Aufnahmen im Hinblick auf den Tagesverlauf. Die Ermüdungsgrade 1 bis 3 verteilen sich hauptsächlich zwischen 10 bis 18 Uhr. Das Auftreten von höhergradigen Stimmermüdungen wurde in den Abendstunden und nachts angegeben. Bei einer Aufnahme in der gesamten Studie wurde ein „sehr starker“ Ermüdungsgrad angegeben.

4.7 Ergebnisse des telemedizinischen Monitorings

Die Analyse der via „Voice Diary“ aufgezeichneten Sprachaufnahmen erfolgte anhand mehrerer Parameter, die bereits im Grundlagenteil erklärt wurden. Bei der Auswertung wurden die Grundfrequenz F_0 , die Irregularitätsparameter Jitter und Shimmer, die akustischen Parameter für Rauschen und die spektralen Parameter aus der Stimmdiagnostik betrachtet. Die Auswertung erfolgte mit dem Signalverarbeitungsprogramm Praat. Bei der Auswertung wurden die 1386 Aufnahmen der 17 Probanden im 14-tägige Untersuchungszeitraum betrachtet. Die Ergebnisse wurden in unterschiedlichen Grafiken dargestellt. Anhand von ausgewählten Beispielen werden aufgetretene Unterschiede erklärt.

4.7.1 Grundfrequenz F_0

Als erster Auswertungspunkt wird die Grundfrequenz betrachtet. Das Histogramm in Abbildung 34 a) zeigt die allgemeine Häufigkeitsverteilung aller F_0 Werte der Aufnahmen. Die Analyse ergab einen Mittelwert von $110 \pm 9,9 \text{ Hz}$. Die Verteilung ist in Anbetracht des Mittelwertes leicht schief und fällt im Bereich der niedrigen Frequenzen steiler ab, als im Bereich der hohen Frequenzen. Zudem fällt die Kurve im Bereich vor 95 Hz und nach 118 Hz steil ab. Das absolute F_0 Minimum liegt bei etwa 78 Hz , das absolute Maximum bei 144 Hz . Der Boxplot in Abbildung 34 b) zeigt die Entwicklung von F_0 im Tagesverlauf. Hier wurde der Untersuchungszeitraum von 14 Tagen betrachtet. Bis auf einzelne Ausreißer (rote Kreuze) liegen 50 % der Aufnahmen in allen 5 Zeitintervallen in einem ähnlichen Frequenzbereich, der zwischen $100\text{-}180 \text{ Hz}$ liegt. Der Median ist in allen Intervallen bei circa 110 Hz zu finden. Die Zeitintervalle von 5-10 Uhr und 10-14 Uhr weichen ab. Hier liegt der Median mit etwa 108 Hz etwas tiefer.

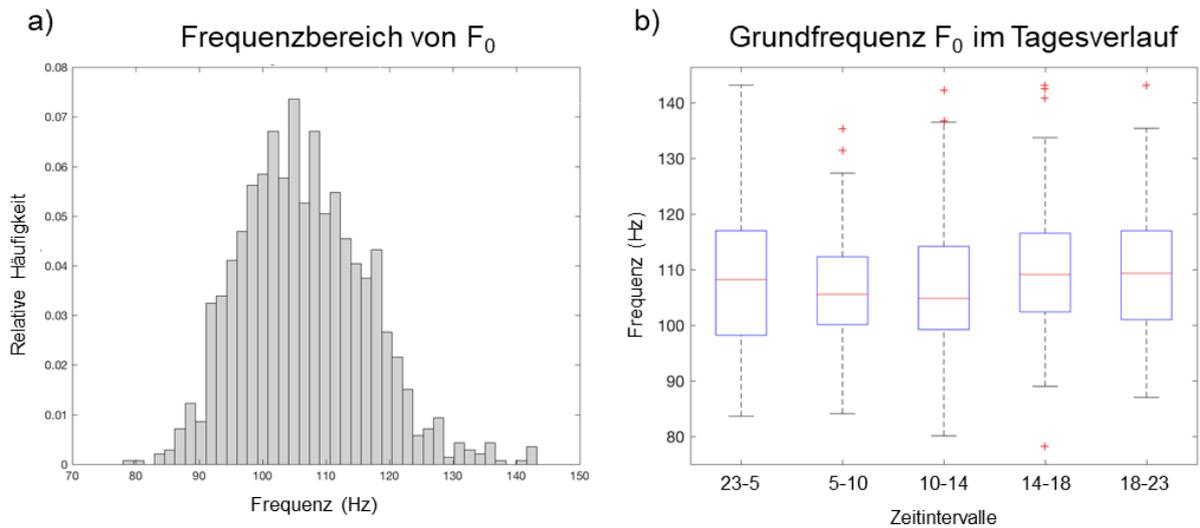


Abbildung 34: a) Das Histogramm zeigt die allgemeine Häufigkeitsverteilung von F_0 . Der Mittelwert liegt bei $110 \pm 9,9 \text{ Hz}$. Die Verteilung ist im Vergleich zum Mittelwert leicht schief, jedoch regelmäßig verteilt. b) Der Boxplot zeigt eine regelmäßige Verteilung der Grundfrequenz im Tagesverlauf. Von 5-14 Uhr ist der Median etwas niedriger als in den anderen Zeitintervallen.

Beide Grafiken geben Hinweis auf eine stabile und konstante Verteilung der Grundfrequenz im Tagesverlauf und im Verlauf der Studie. In Abbildung 35 a) wird die Differenz zwischen Grundfrequenz F_0 und mittlerer Sprechstimmlage (mSsl) gezeigt. Hier zeigt sich, dass in der Nacht und in den Morgenstunden $F_0 < \text{mSsl}$ ist. Ab 10 Uhr bis 23 Uhr ist $F_0 > \text{mSsl}$, die Differenz wird positiv. Die Differenz der beiden Werte ist bei 50 % der 1386 Aufnahmen < 10 , die beiden Werte sind bei den 17 Probanden im gesamten Studienverlauf sehr ähnlich. Am Beispiel des Probanden 13 wird in Abbildung 35 b) die Entwicklung der Grundfrequenz im Studienzeitraum gezeigt. Die rote Linie bei 102 Hz stellt die mittlere Sprechstimmlage des Probanden dar. Die Grundfrequenz ist gleichmäßig um den Mittelwert verteilt. Einzelne Ausreißer weichen maximal um 10 Hz ab. Da sich das Verhältnis der beiden Werte bei allen Probanden ähnlich verhält, wurde nur Proband 13 exemplarisch dargestellt.

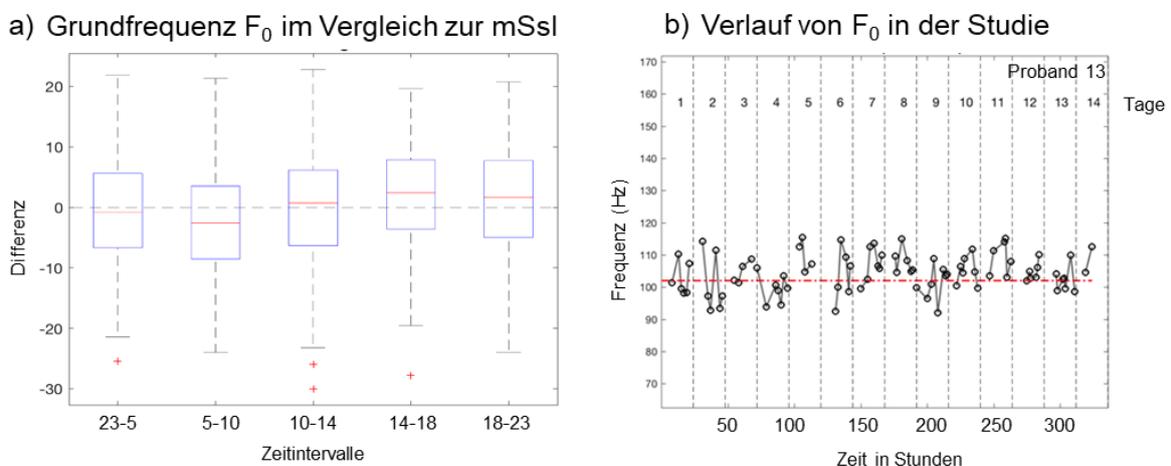


Abbildung 35: a) Boxplot, der das Differenzverhältnis zwischen F_0 und mSsl im Verlauf des Tages darstellt. In den Morgenstunden besteht ein leichtes negatives Verhältnis, das ab 10 Uhr positiv wird. b) Entwicklung von F_0 im

Ergebnisse

Verlauf der gesamten Studie am Beispiel von Proband 13. Es zeigt sich eine gleichmäßige Verteilung der Grundfrequenz um die mittlere Sprechstimmlage (rote Linie).

Ein weiterer Betrachtungspunkt ist, wie sich F_0 der einzelnen Probanden gegenüber dem errechneten Mittelwert und Standardabweichung für das gesamte Kollektiv in den unterschiedlichen Zeitintervallen verhält. In Abbildung 36 sind exemplarisch Proband 2 und 7 miteinander verglichen worden. Die Verteilung der Grundfrequenz von Proband 7 in Abbildung 36 a) liegt in dem Rahmen des Mittelwertes $110 \pm 9,9 \text{ Hz}$ der für das Gesamtkollektiv berechnet wurde (grauer Bereich). Die anderen Probanden liegen ähnlich verteilt. Große Abweichungen hingegen zeigt Proband 2. F_0 liegt hier weit über dem ermittelten Bereich. Auffallend ist, dass bei beiden Probanden die Grundfrequenz morgens tiefer ist als im Tagesverlauf. Dieser Trend zeigte sich bereits in den Ergebnissen aus Abbildung 34 b).

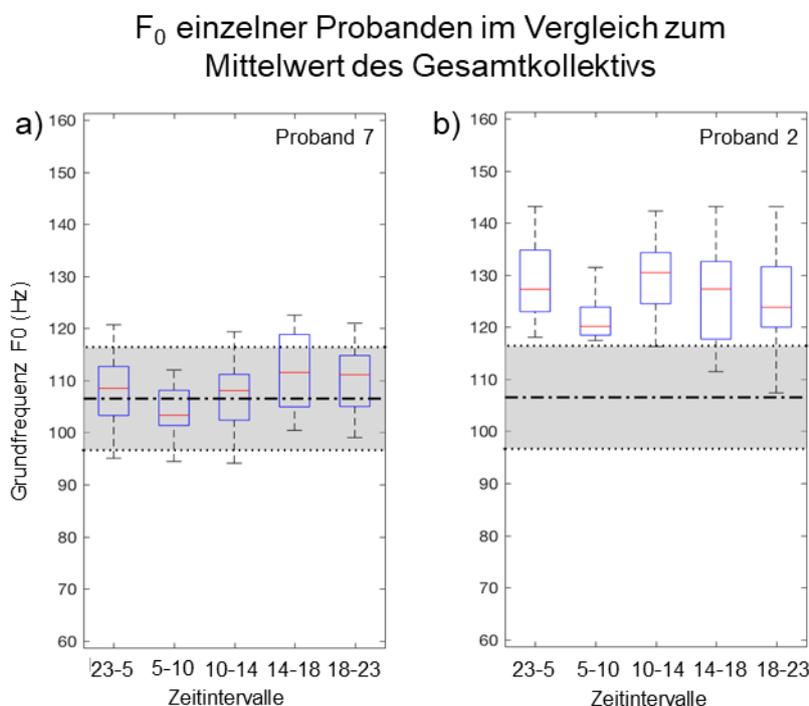


Abbildung 36: Vergleich der Probanden 2 und 7 miteinander und Vergleich mit dem F_0 Wert des Gesamtkollektivs. Die Grundfrequenz des Probanden 2 weicht stark von der mittleren Verteilung von F_0 in der Studie ab, die an dem Beispiel von Proband 7 dargestellt ist.

4.7.2 Irregularitätsparameter: Jitter & Shimmer

Anhand der Parameter Jitter und Shimmer werden die Aufnahmen auf Irregularitäten untersucht. In Abbildung 37 a) ist die Verteilung der relativen Häufigkeit der Jitter Werte in Anbetracht aller Aufnahmen dargestellt. Der Mittelwert aller Aufnahmen liegt bei $0,52 \pm 0,22 \%$. Die Abbildung zeigt, dass sich die Verteilung auf einen sehr kleinen Bereich beschränkt. Die Spanne des Jitters liegt zwischen 0,2 - 1 %. Nur einzelne Aufnahmen weisen einen Jitter $>1\%$ auf. Bild b) in Abbildung 37 zeigt einen beständigen Median bei 0,05% über allen Zeitintervallen. Eine Entwicklung im Tagesverlauf ist nicht zu erkennen. 50 % aller Aufnahmen

des Studienzeitraums ergeben einen Jitter zwischen etwa 0,45% und 0,62%. Die Streuung der oberen Whisker liegt fast immer unter 1%. Die roten Kreuze zeigen einzelne Ausreißer in der Studie.

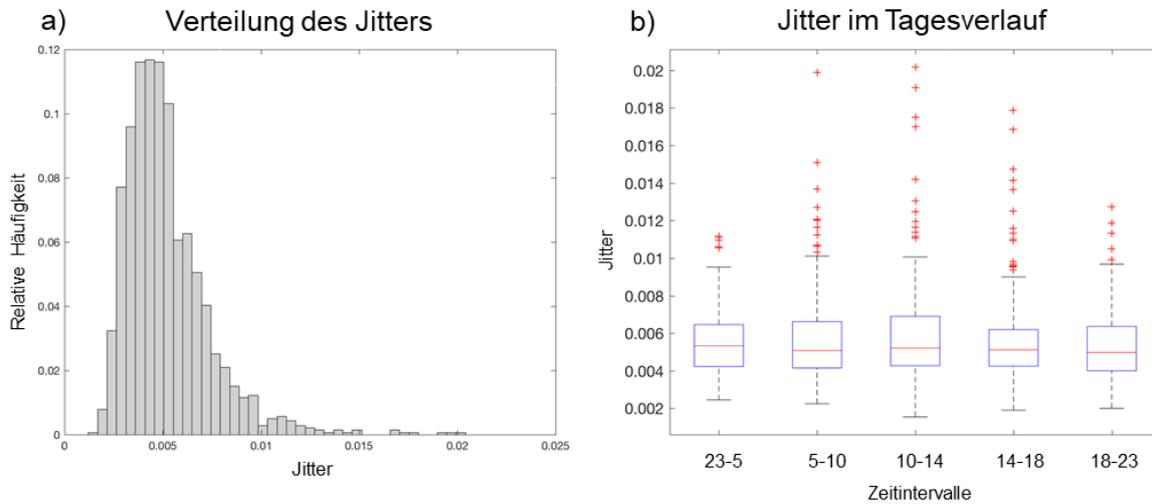


Abbildung 37: a) Regelmäßige und kompakte Verteilung des Jitters um den Mittelwert 0,5%. Es finden sich kaum Werte über 1%. b) Verteilung des Jitters im Tagesverlauf um einen Median bei 0,5%. Nur einzelne Ausreißer (rote Kreuze) liegen über 1%.

Aus Abbildung 37 wird ersichtlich, dass in der gesamten Studie Aufnahmen selten einen Jitter $>1\%$ erreichten. Abbildung 38 vergleicht den Verlauf des Jitters über die Zeit von Proband 4 und Proband 12. Der grau markierte Bereich stellt Mittelwert und Standardabweichung des Jitters dar. Proband 4 wurde als Beispiel gewählt, weil er die stärksten Abweichungen des Jitters im Bereich der Standardabweichung in der Studie aufweist. Im Zeitraum von 5-10 Uhr weist der Jitter dieses Probanden eine hohe Streuung auf. Mehr als 50 % der Aufnahmen liegen in diesem Zeitraum außerhalb der Standardabweichung. Gerade morgens weist dieser Proband erhöhte Frequenzschwankungen in der Stimme auf. Trotz dieser starken Abweichungen liegt dennoch der Großteil seiner Aufnahmen unter einem Jitter von 1 %. Die Abbildung 38 Grafik zeigt exemplarisch Proband 12, dessen Jitter Werte im errechneten Bereich liegen. Der Median des Jitters ist hier etwas niedriger als der Mittelwert.

Ergebnisse

Jitter einzelner Probanden im Vergleich zum Mittelwert des Gesamtkollektivs

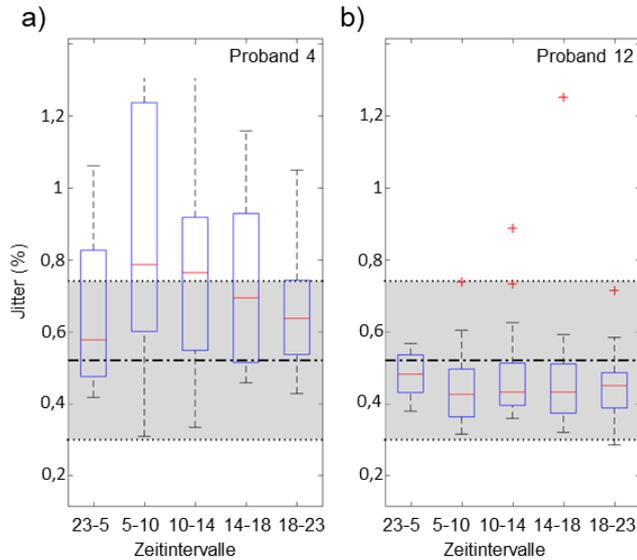


Abbildung 38: Vergleich des Jitters von Proband 4 und 12 in Bezug zum errechneten Mittelwert des Gesamtkollektivs. Trotz stärkerer Abweichungen vom durchschnittlichen Bereich, haben die Aufnahmen von Proband 4 selten einen Jitter über 1%. Der Jitter des Probanden 12 liegt etwas unterhalb des mittleren Jitters des Gesamtkollektivs.

Neben der Frequenzschwankung muss auch die Amplitudenschwankung auf Irregularitäten untersucht werden. Dies wird durch die Bestimmung des Shimmers ermöglicht. Ein Großteil der Shimmer Werte in dieser Studie liegt im Bereich von 1% bis 8% (vgl. Abbildung 39 a). Der Mittelwert des Shimmers, errechnet aus allen Aufnahmen des Gesamtkollektivs, liegt bei $4,6 \pm 2,2$ %. Der Median des Shimmers im Tagesverlauf liegt bei 5% und verhält sich ähnlich wie der beschriebene Jitter Wert. Die unteren Quartile liegen etwa bei 3,9%, die oberen etwa bei 6,2%. Bei etwa Dreiviertel aller Aufnahmen liegt der Shimmer unter 6,1%. Die roten Kreuze in Abbildung 39 b) zeigen einzelne Ausreißer.

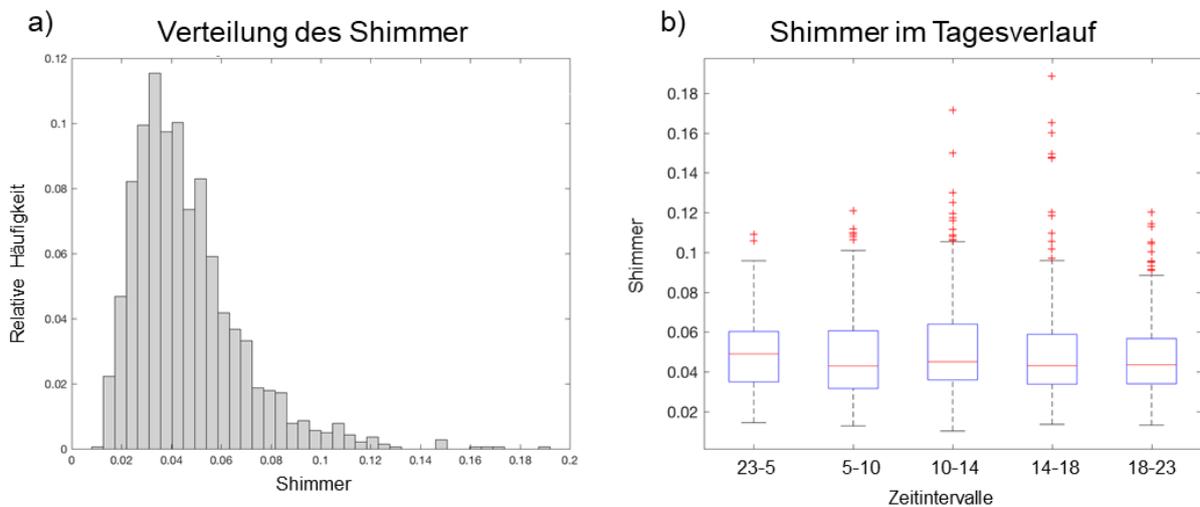


Abbildung 39: a) Absteigende und kompakte Verteilung um den Mittelwert 4,6%. b) Verteilung des Shimmers im Tagesverlauf um einen Median bei 0,5%. Bei 75% der Aufnahmen liegt der Shimmer $\leq 6,1$ %.

In Abbildung 40 ist die exemplarische Gegenüberstellung von Proband 4 und Proband 12 dargestellt. Einige Aufnahmen von Proband 4 liegen außerhalb der Standardabweichung, bei Proband 12 hingegen verteilen sich die Shimmer Werte in allen Zeitintervallen im Bereich der Standardabweichung. Auffallend ist, dass bei Proband 12 die nächtlichen Shimmer Werte deutlich höher sind als in den anderen Zeitintervallen und somit die Amplitudenschwankung der Stimme mehr ausgeprägt war. Vergleicht man nun diese Abbildung mit Abbildung 38, wird ersichtlich, dass sich der Shimmer der beiden Probanden ähnlich zum Jitter verhält. In Anbetracht aller Probanden, liegen die meisten Aufnahmen in den einzelnen Zeitintervallen im Bereich des Mittelwertes und der Standardabweichung. Einzelne Probanden liegen über diesem Bereich, wie zum Beispiel Proband 4 und 9 (vgl. Abbildung 65 im Anhang).

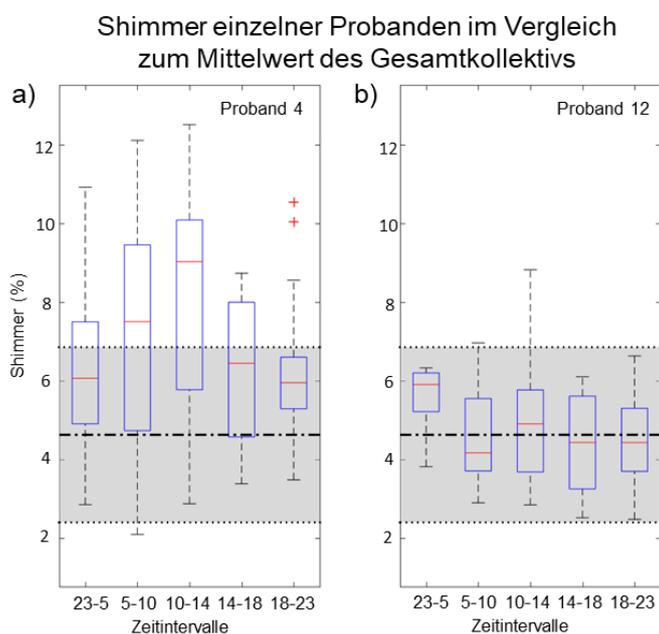


Abbildung 40: Vergleich des Shimmers von Proband 4 und 12 in Bezug zum errechneten Mittelwert $4,6 \pm 2,2$ % des Gesamtkollektivs. Proband 4 zeigt in einigen Aufnahmen eine starke Abweichung vom durchschnittlichen Bereich des Shimmers. Der Shimmer des Probanden 12 ist nachts etwas höher als im Tagesverlauf, liegt jedoch in allen Zeitintervallen im errechneten Durchschnitt.

4.7.3 HNR

Neben den Irregularitäten sind auch die Rauschanteile ein wichtiges Kriterium, um die Aufnahmen via Handy zu beurteilen. Ein geeigneter Parameter zur Darstellung des Rauschens in einer Aufnahme ist die Harmonics-to-noise ratio (HNR). Wäre der Rauschanteil der Aufnahmen sehr groß, so wäre die Anwendung eines Handys für ein ambulantes Stimmmonitoring nicht brauchbar. Abbildung 41 a) zeigt die relative Verteilung aller HNR-Werte in dieser Studie. Der errechnete Mittelwert in dieser Studie liegt bei $15 \pm 3,2$ dB. Die Verteilung um diesen Mittelwert ist gleichmäßig und die Streuung ist gering. Grafik b) zeigt anhand von Boxplots die Entwicklung des HNR in den 5 Zeitintervallen im untersuchten Zeitraum. Der Median liegt im Durchschnitt bei 14 dB und ist etwas niedriger als der Mittelwert.

Ergebnisse

Im Tagesverlauf bleibt der Median des HNR gleich. Der Bereich, in dem 50 % der Aufnahmen liegen, bleibt annähernd konstant zwischen 12,5-15,5 dB (vgl. Abbildung 41 b). Einzelne Ausreißer sind auch bei der HNR Auswertung vorhanden, jedoch ist die Verteilung insgesamt sehr gleichmäßig und ohne großen Streuungsrahmen.

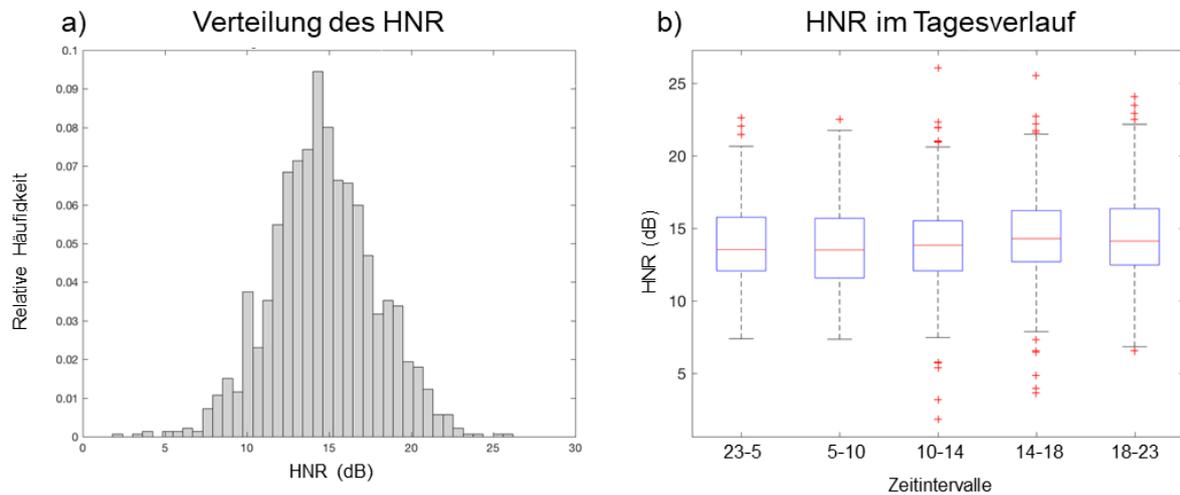


Abbildung 41: a) Darstellung der Verteilung der relativen Häufigkeiten des HNR. Der Mittelwert ist $15 \pm 3,2$ dB. Zu beiden Seiten des MW ist die Verteilungskurve stark abfallend. b) Der Median liegt in allen Zeitintervallen ungefähr bei 14 dB. Ebenso weichen die 1. und 3. Quartile in den 5 Bereich nicht stark voneinander ab.

Niedrige HNR Werte sprechen für einen erhöhten Rauschanteil. Der Verlauf des HNR bei Proband 9 soll hier als Beispiel dienen, da er die kleinsten HNR Werte im Kollektiv aufweist (vgl. Abbildung 42 a). Die Mediane in den Zeitintervallen liegen bei diesem Probanden unterhalb der Standardabweichung des Mittelwertes. Der HNR zwischen 10-14 Uhr zeigt im Boxplot die größte Streuung. Bei Proband 9 liegt der Median abends, nachts und morgens zwar unterhalb des MW, dieser ist jedoch für den Probanden annähernd konstant und besitzt eine geringe Standardabweichung. Zwischen 10-18 Uhr treten vermehrt Aufnahmen mit einem niedrigen HNR auf. Zum Vergleich ist der Proband 5 gegenübergestellt, der genau im Bereich des Mittelwertes und der SD liegt. Die Probanden 6, 10 und 17 zeigten höhere HNR Werte (vgl. Abbildung 66).

HNR einzelner Probanden im Vergleich zum Mittelwert des Gesamtkollektivs

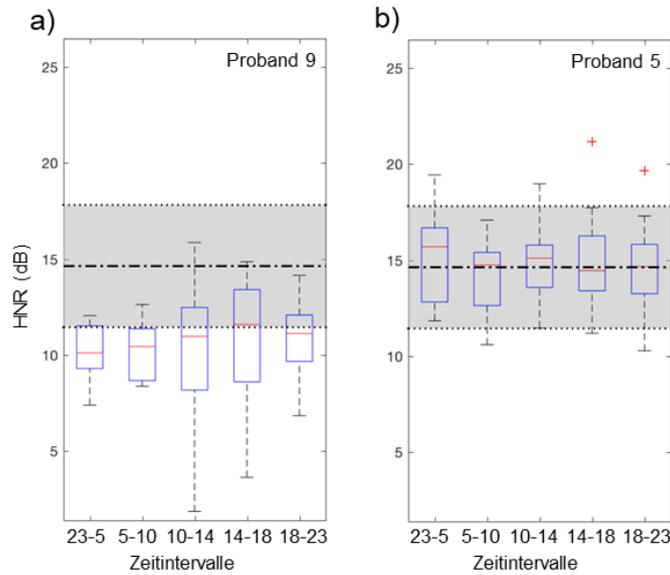


Abbildung 42: a) Der HNR des Probanden 9 liegt zum größten Teil außerhalb des grauen Bereichs ($MW \pm SD$). Der Median des Probanden liegt im zeitlichen Verlauf zwischen 10-12 dB . Auffallend sind niedrige HNR Werte zwischen 10 und 18 Uhr, die für einen vermehrten Geräuschanteil bei einzelnen Aufnahmen sprechen. b) Proband 5 als Beispiel für eine regelmäßige Verteilung des HNR in Anbetracht des Gesamtkollektiv.

4.7.4 NHR

Der reziproke Wert zum HNR ist der NHR. Er beschreibt, wie der HNR, den Geräuschanteil in einem Stimmsignal. Der Unterschied ist, dass ein sehr niedriger NHR für gute Aufnahmen mit wenig Störeffekten spricht. Der Mittelwert des NHR in dieser Studie ist $0,062 \pm 0,05 \text{ dB}$. Die Standardabweichung ist sehr gering. Die relativen Häufigkeiten in Abbildung 43 a) verteilen sich größtenteils innerhalb eines sehr kleinen Bereichs zwischen 0 - 0,02 dB . 25 % aller Aufnahmen haben einen NHR zwischen etwa 0,035-0,05 dB . Die Abbildung 43 b) zeigt wie sich der NHR zu unterschiedlichen Tageszeiten verteilt. Der Median liegt wie der Mittelwert im Bereich zwischen 0,06-0,07 dB . Bei 50 % der Aufnahmen liegt der NHR Wert zwischen 0,05 - 0,09 dB . Der NHR zeigt keine Trendentwicklung in den unterschiedlichen Zeitrahmen. Die roten Kreuze markieren einzelne Ausreißer.

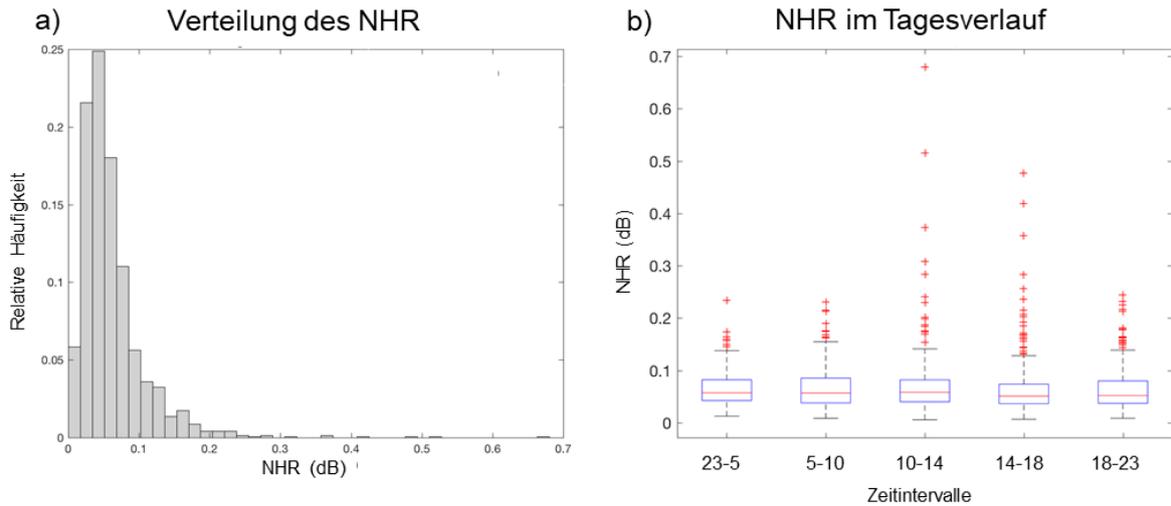


Abbildung 43: a) Der NHR aller Aufnahmen verteilt sich größtenteils zwischen 0 - 0,2 dB. Der Mittelwert liegt bei $0,062 \pm 0,05$ dB. Die Standardabweichung zeigt eine sehr geringe Streuung der Werte. b) Der Median ist dem Mittelwert sehr ähnlich und liegt in allen 5 Zeitbereichen zwischen etwa 0,06 - 0,07 dB.

Obwohl der NHR wenig streut, heben sich zwei Probanden von den anderen 15 ab, da der Großteil ihrer Aufnahmen nicht im Rahmen des Mittelwertes und der Standardabweichung liegen. Ihre NHR Werte sind in Abbildung 44 dargestellt. Bei Proband 4 treten im Vergleich zu allen Aufnahmen am Morgen erhöhte Werte auf. Fast 75% der Aufnahmen zwischen 5-10 Uhr liegen über 0,1 dB. Im Tagesverlauf nimmt dieser Trend ab und nähert sich dem Mittelwert des Kollektivs an. Zwischen 23 - 5 Uhr steigt der NHR wieder leicht an. Dies spricht dafür, dass die Stimme des Probanden 4 nachts und in den Morgenstunden von einem erhöhten Rauschanteil geprägt ist (vgl. Abbildung 44). Bei Proband 9 liegen fast 50% aller Aufnahmen über 0,1 dB. Der leicht erhöhte NHR dieses Probanden bestätigt die Ergebnisse des HNR von Proband 9 (vgl. Abbildung 42 a).

NHR einzelner Probanden im Vergleich zum Mittelwert des Gesamtkollektivs

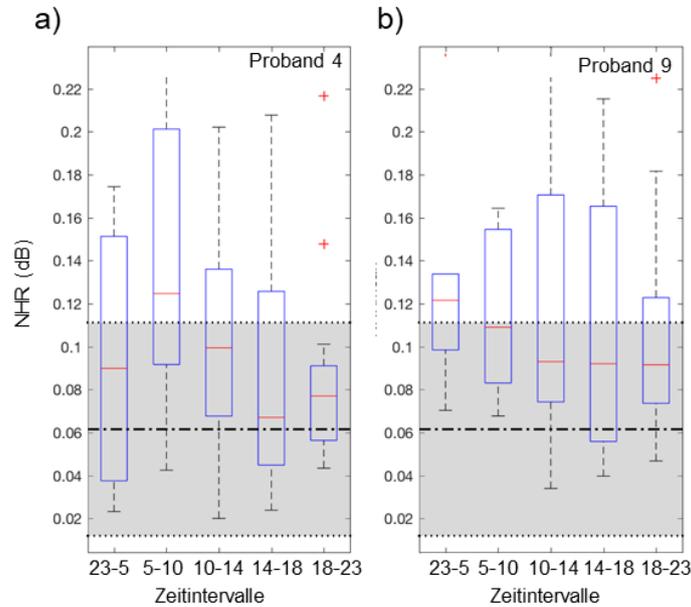


Abbildung 44: Vergleich von Proband 4 und 9, die beide etwas vom Mittelwert des Gesamtkollektivs abweichen. a) Bei Proband 4 weichen überwiegend die Aufnahmen nachts und in den Morgenstunden ab. Zwischen 10 – 23 Uhr liegt der NHR im Bereich der Standardabweichung des Gesamtkollektivs. b) Bei Proband 9 liegt der NHR durchschnittlich etwas höher als der Mittelwert und die Standardabweichung des Gesamtkollektivs.

4.7.5 NNE

Mit dem NNE wird der Geräuschanteil in einem Signal anhand dessen Geräuschenergie bestimmt (vgl. 2.7.2). In Abbildung 45 a) ist die Verteilung der relativen Häufigkeit aller NNE Werte aus der Studie dargestellt. Der errechnete Mittelwert liegt bei $-3,4 \pm 3,2 \text{ dB}$. In Abbildung 45 a) fallen neben der normalen Verteilung viele unterschiedliche kleine NNE Werte auf. Diese Ausreißer, die in Abbildung 45 b) durch rote Markierungen gekennzeichnet sind, liegen im Bereich von $-17,5 \text{ dB}$ bis -5 dB . Diese sind in allen Zeitintervallen ähnlich verteilt. Die Mediane in den unterschiedlichen Zeitintervallen liegen zwischen -2 und -3 dB . 95% aller Aufnahmen haben einen NNE zwischen 0 und -6 dB .

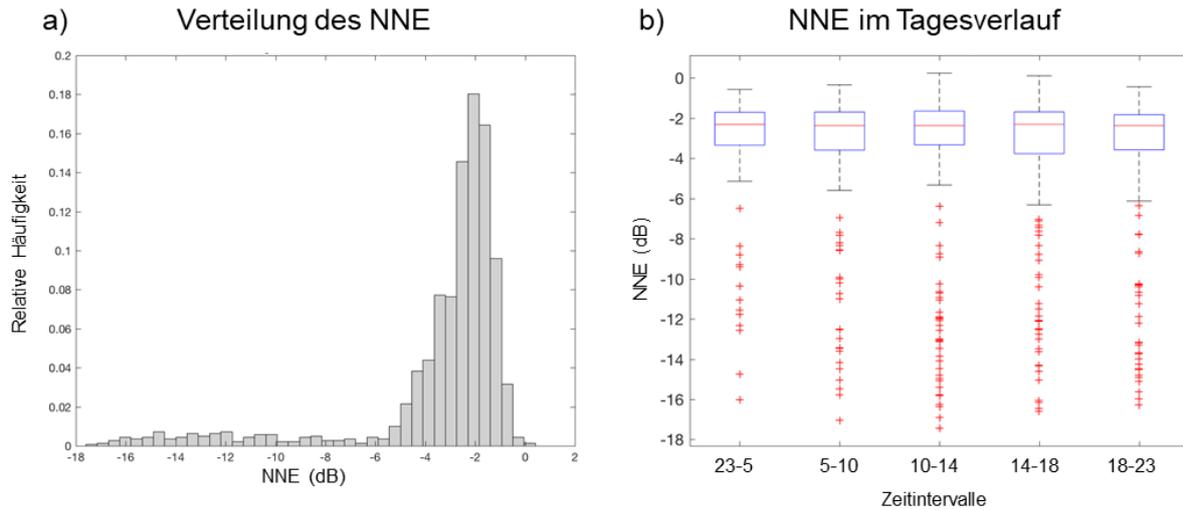


Abbildung 45: a) Der NNE Mittelwert aller Aufnahmen ist bei $-3,4 \pm 3,2 \text{ dB}$. Die meisten NNE Werte sind zwischen 0 und -6 dB verteilt. Auffallend sind viele kleine Ausreißer. Diese sind mit roten Kreuzen in b) gekennzeichnet. Der Median liegt im Bereich des Mittelwertes und über die Zeitintervalle ist der NNE näherungsweise konsistent verteilt.

Um die Ausreißer besser darstellen zu können, wurde anstatt der Boxplots, die die Verteilung der Aufnahmen in den 5 Zeitintervallen darstellen, die zeitliche Entwicklung des NNE für jeden Probanden im Verlauf des 14-tägigen Studienzeitraums begutachtet. Im Gegensatz zu der Verteilung des HNR und des NHR, zeigen sich bei vielen Probanden im zeitlichen Verlauf starke Schwankungen im NNE. In Abbildung 46 wurde exemplarisch die NNE Entwicklung von Proband 3 und 15 innerhalb der 14 Tage dargestellt, da diese die größten NNE Schwankungen zeigten. Im Vergleich zum Mittelwert weichen viele NNE Werte bei beiden Probanden stark ab und eine klare Tendenz ist bei keinem Probanden vorhanden. Die Rauschparameter HNR und NHR hingegen zeigen sich bei Proband 3 und 15 gleichmäßig verteilt (vgl. Abbildung 67 im Anhang). Die breite Streuung des NNE zeigt sich auch bei vielen anderen Probanden, die im zeitlichen Verlauf einen gleichmäßig HNR und NHR haben. Der NNE aller Aufnahmen ist breiter verteilt, als in der anfänglichen Betrachtung der relativen Häufigkeit auffällt.

Entwicklung des NNE im Studienverlauf

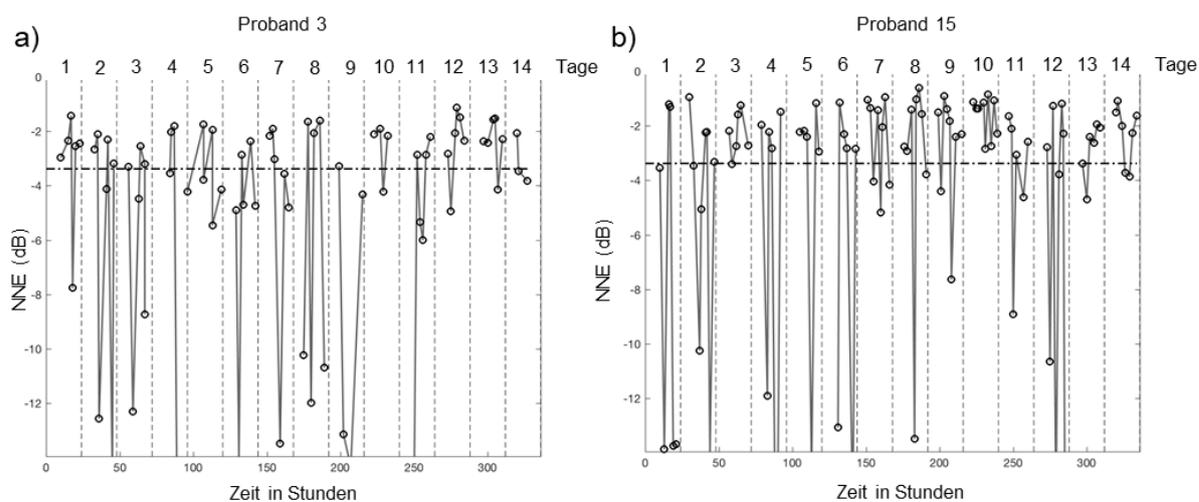


Abbildung 46: a und b) zeigen den Verlauf der NNE Ergebnisse exemplarisch von Proband 3 und 5 im Zeitraum von 14 Tagen. Bei beiden Probanden ist eine sehr unregelmäßige Verteilung der Werte erkennbar. Diese breite Streuung der Wert findet man bei vielen weiteren Probanden in der Studie.

4.7.6 H_1H_2

Das Verhältnis von Offen- und Schlussphase, das maßgeblich für die Stimmqualität ist, wurde in der Studie anhand des Parameters H_1H_2 betrachtet. Die Amplitudendifferenz von H_1 und H_2 gilt als ein Maß für die Behauchtheit [42]. In Abbildung 47 a) wird die relative Verteilung der H_1H_2 Ergebnisse aus der gesamten Studie gezeigt. Der errechnete Mittelwert liegt bei $-0,68 \pm 4,2 \text{ dB}$, die Standardabweichung ist bei diesem Parameter groß. Die Verteilung um den MW ist zu beiden Seiten gleichmäßig abfallend und ein Großteil aller Werte liegt zwischen -10 und 10 dB . Abbildung 47 b) stellt anhand von Boxplots die Verteilung des H_1H_2 innerhalb der unterschiedlichen Zeitintervalle dar. Der Median liegt in den unterschiedlichen Zeitintervallen zwischen 0 und -1 dB und ist dem Mittelwert sehr ähnlich. 50% aller Aufnahmen zu allen Zeitpunkten liegen zwischen -3 dB bis 3 dB . Die Verteilung des H_1H_2 ist näherungsweise konstant im Tagesverlauf. Ausreißer sind auch bei diesem Parameter vereinzelt vorhanden und mit roten Markierungen in Abbildung 47 b) gekennzeichnet.

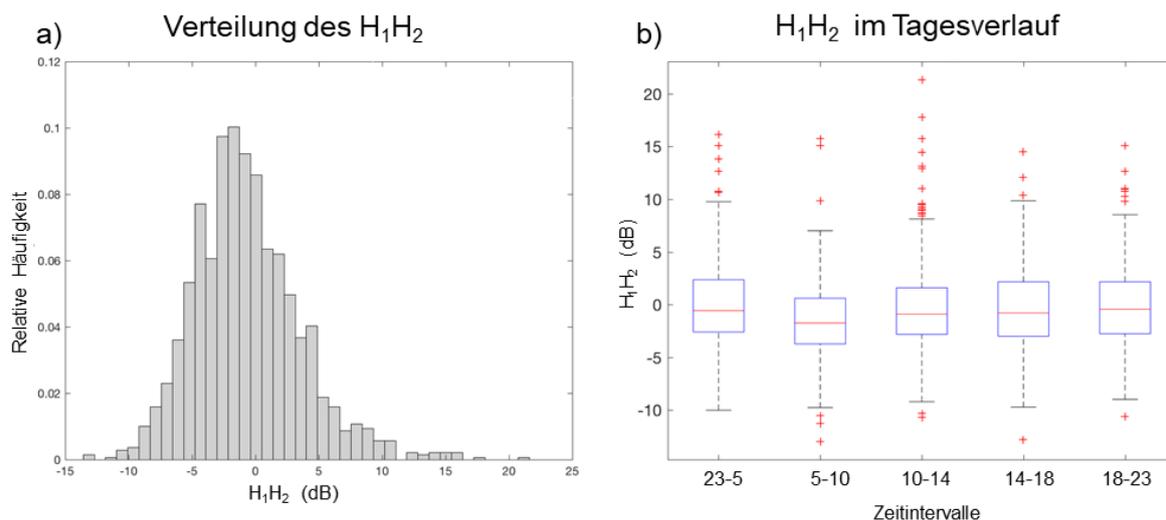


Abbildung 47: a) Die Werte verteilen sich zu beiden Seiten des MW $-0,68 \pm 4,2 \text{ dB}$ gleichmäßig abfallend. Ausreißer, die sich in der Größe ihrer Anzahl in der relativen Häufigkeit widerspiegeln sind nur in geringer Anzahl vorhanden. b) Der Median schwankt im Tagesverlauf im Bereich rund um 0 und -1 dB verteilt. Eine große Übereinstimmung zwischen Median und Mittelwert liegt vor.

Da sich der H₁H₂ nicht ähnlich gleichmäßig bei den einzelnen Probanden verteilt, wie die bisher besprochen Parameter, und es einige Abweichungen gibt werden exemplarisch einige Probanden miteinander verglichen, um die Unterschiede darzustellen. In Abbildung 48 wird zunächst die Verteilung des H₁H₂ von 2 Probanden in den 5 Zeitfenstern verglichen. Der graue Bereich kennzeichnet den Bereich der einfachen Standardabweichung, die schwarze Linie zeigt den Mittelwert des Gesamtkollektivs. Abbildung 48 a) zeigt, dass sich der Median bei Proband 1 zwischen 18 - 5 Uhr zwischen $4-6 \text{ dB}$ befindet und deutlich über dem Mittelwert liegt. In den anderen Zeitintervallen liegt er zwischen $0-2 \text{ dB}$ im grauen Bereich, der Anteil an Behauchtheit ist kleiner und der Stimmklang entspricht dem durchschnittlichen Klang des Gesamtkollektivs. Ein ähnlich unregelmäßiges Verhalten liegt auch bei Proband 8 vor. Die H₁H₂ Werte von Proband 17 liegen mit einem individuellen Median bei etwa -5 dB etwas tiefer als der Mittelwert, dennoch sind die Werte in den Zeitintervallen etwa gleichmäßig verteilt. Zwischen 5-10 Uhr haben die Aufnahmen tendenziell einen höheren H₁H₂.

H_1H_2 einzelner Probanden im Vergleich zum Mittelwert des Gesamtkollektivs

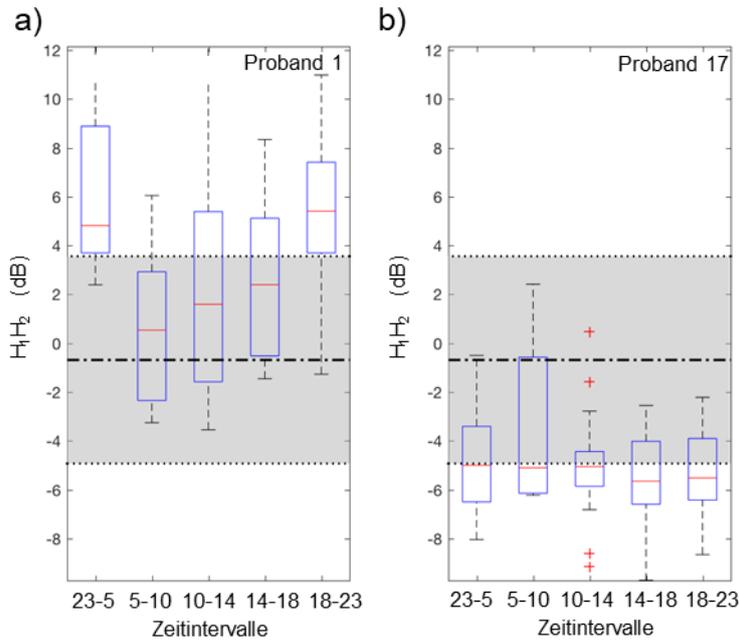


Abbildung 48: Vergleich zweier Probanden mit dem Gesamtkollektiv. a) Die H_1H_2 Werte zwischen 18-5 Uhr weichen von der Medianverteilung im Lauf des Tages ab und sind wesentlich größer als der Mittelwert. b) Die Werte von Proband 17 liegen außerhalb der unteren SD, der Median in den einzelnen Zeitintervallen ist gleich homogen verteilt und liegt bei ca. -5 dB.

Da der H_1H_2 des Probanden 8 in den zeitlichen Intervallen ebenfalls unterschiedlich streut (vgl. Abbildung 68), soll der genaue zeitliche Verlauf des Wertes über den gesamten Zeitraum der Studie betrachtet werden. Abbildung 49 a) zeigt, dass die H_1H_2 Werte der einzelnen Aufnahmen von Proband 8 oberhalb des Mittelwertes liegen und sehr unregelmäßig verteilt sind. Einzelne Ausreißer liegen bei 15 dB und eine Tendenz um einen individuellen Mittelwert ist nicht zu erahnen. Im Vergleich dazu ist der Parameter H_1H_2 bei Proband 2 sehr regelmäßig und ohne große Streuung um den Mittelwert verteilt.

Entwicklung des H_1H_2 im Studienverlauf

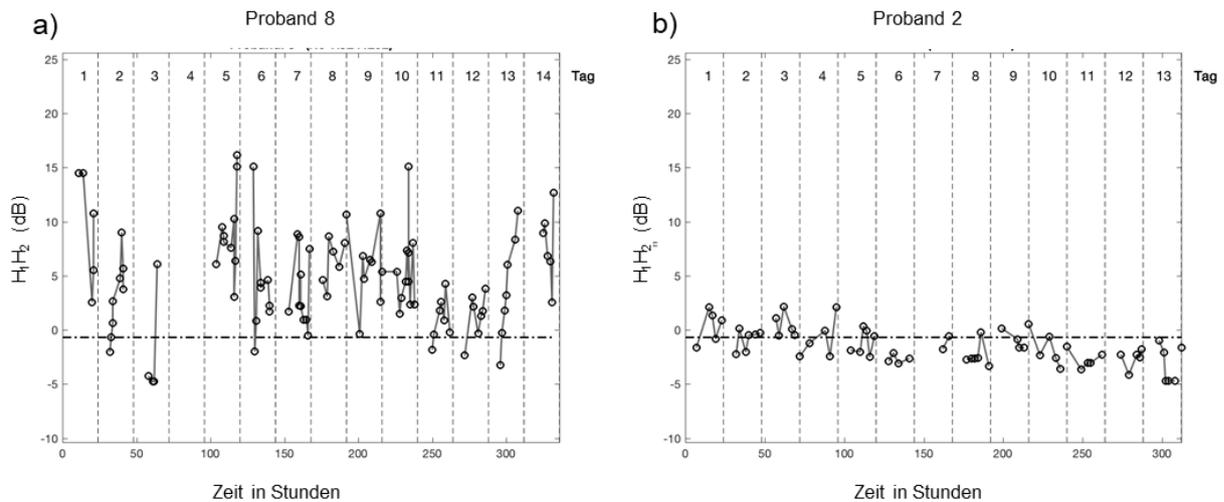


Abbildung 49: a) Breite Streuung der einzelnen H_1H_2 Werte bei Proband 8. Ein Trend zu einem bestimmten Mittelwert liegt nicht vor. Selten wird der Mittelwert des Gesamtkollektivs erreicht. b) Im Vergleich regelmäßige Verteilung der H_1H_2 Ergebnissen um den errechneten Mittelwert bei Proband 2.

4.7.7 Spektrale Neigung: Tilt & Slope

Die Spektrale Neigung beschreibt die Verteilung der niedrigen und hohen Frequenzen im Spektrum (vgl. 2.8.2). Die Neigung ist bei unauffälligen Stimmen negativ, da die tieferen Frequenzen in der Regel energiereicher sind als die hohen Frequenzen und die Regressionslinie durch das Frequenzspektrum daher eine negative Neigung besitzt. Die spektrale Neigung wird anhand von Slope und Tilt gemessen. In Abbildung 50 werden die relativen Häufigkeiten der einzelnen Slope Werte und die Verteilung des Slopes in den einzelnen Zeitintervallen gezeigt. Der Mittelwert des Slope liegt bei $-17 \pm 3,5 \text{ dB}$. Die Verteilung ist zu beiden Seiten des Mittelwertes abfallenden, zu den niedrigeren Slope Werten ist der Abfall etwas flacher als zu den hohen Slope Werten. Etwa 90% der Aufnahmen haben einen Wert zwischen -25 bis -10 dB (vgl. Abbildung 50 a). Diese Spanne bleibt in den 5 betrachteten Zeitintervallen etwa gleich (vgl. Abbildung 50 b). Der Median in Abbildung 50 b) hat zwischen 5-10 Uhr einen etwas kleineren Wert als zu den anderen Uhrzeiten und liegt bei circa $-17,5 \text{ dB}$. Der Median zu den anderen Uhrzeiten liegt zwischen -16 und -17 dB . Der Slope von etwa 50 % der Aufnahmen liegt zwischen -20 und -15 dB . Bis auf einzelne Ausreißer, zeigt Abbildung 50 b, dass die Slope Werte in einem kleinen Wertebereich ähnlich verteilt liegen. Mittelwert und Median sind sich sehr ähnlich. In anderen Stimmanalysen wurden Slope Bereiche zwischen -25 dB bis -15 dB [50] und -33 dB bis -14 dB [64] angegeben. Die Berechnung des Mittelwertes einer Gruppe aus 23 stimmgesunden Männern und Frauen mit durchschnittlichem Alter von 24 Jahren ergab $-23 \pm 5 \text{ dB}$ [64].

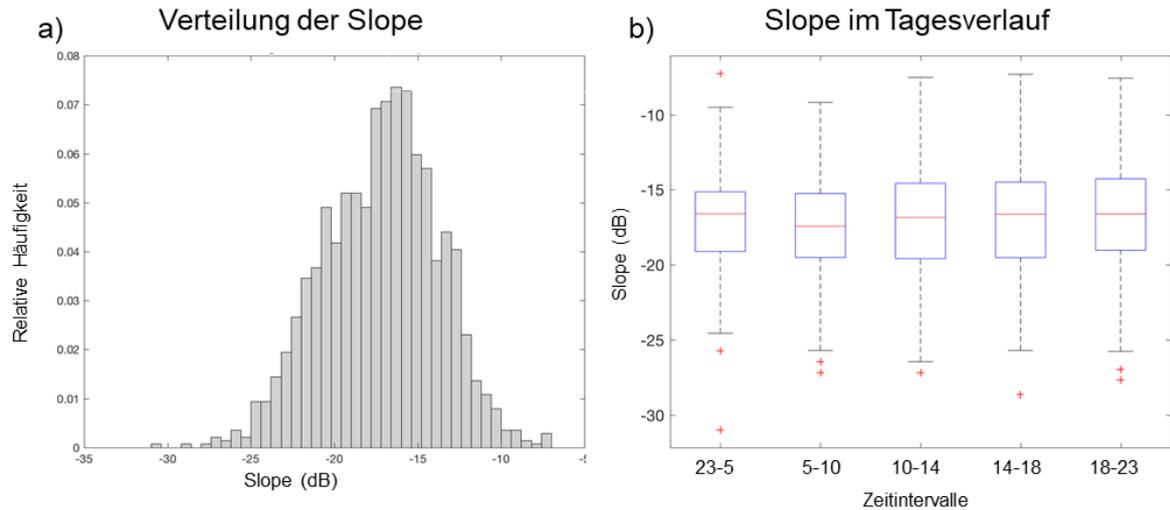


Abbildung 50: a) Der Mittelwert des Slope liegt bei $-17 \pm 3,5 \text{ dB}$. Die relative Häufigkeit der anderen Werte nimmt zu beiden Seiten des MW ab. Zu den niedrigeren Werten ist das Gefälle etwas flacher, zu den höheren Werten steiler. b) Im Tagesverlauf ist der Median des Slope zwischen -16 und -17 dB . Zwischen 5-10 Uhr ist der Median bei etwa $17,5 \text{ dB}$ etwas vermindert.

Betrachtet man die Verteilung des Slope bei den einzelnen Probanden im Zeitverlauf so fällt auf, dass die Slope Werte eine größere Variabilität in ihrer Verteilung aufweisen, als die anderen Rauschparameter. Die Verteilung der Slope Werte zu den unterschiedlichen Zeiten gestaltet sich bei allen Probanden individuell. Dennoch liegt ein Großteil der Werte bei fast allen Probanden trotz der Variabilität im Bereich der Standardabweichung. Exemplarisch ist in Abbildung 51 a) der Verlauf der einzelnen Slope Werte von Proband 7 im Studienzeitraum dargestellt. Auffallend ist, dass die Werte von Proband 7 durchschnittlich oberhalb des Mittelwertes $-17 \pm 3,5 \text{ dB}$ liegen. Die Verteilung ist jedoch in einem Bereich zwischen -11 dB und -16 dB regelmäßig. Einzelne Ausreißer liegen vor. In Abbildung 51 b) wird am Beispiel von Proband 6 die Verteilung des Slope innerhalb der 5 Zeitintervalle betrachtet. Die Werte sind bei diesem Probanden im Bereich der Standardabweichung verteilt. Bei einigen Probanden liegt der Slope zwischen 5-10 Uhr im Vergleich zu den anderen Zeitintervallen etwas ober- oder unterhalb des Probandendurchschnittes. Die Probanden wurden aufgefordert direkt nach dem Aufstehen die Aufnahmen zu beginnen, die Stimmen waren demnach rauher oder behauchter als im Tagesverlauf.

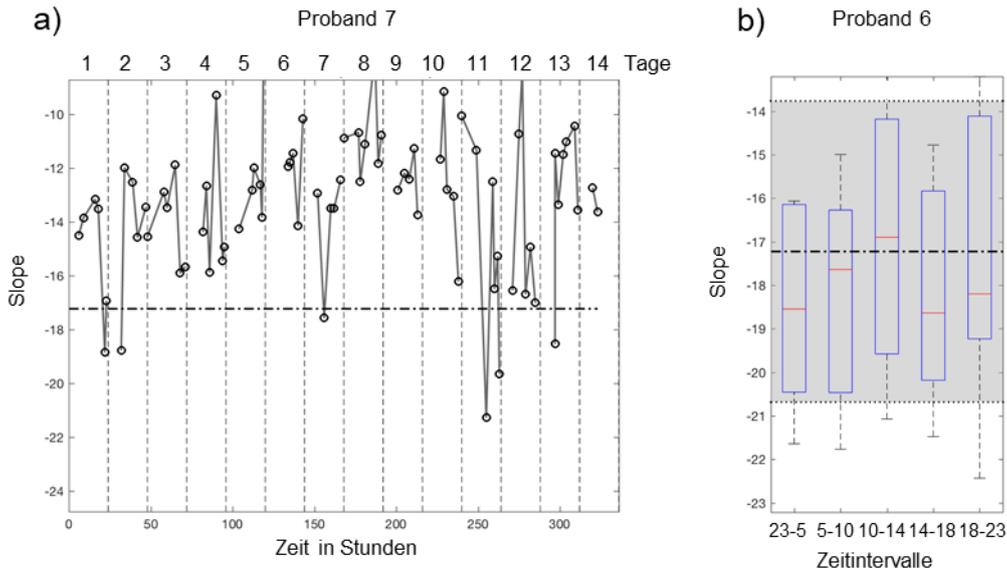


Abbildung 51: a) Die Verteilung des Slope bei Proband 7 soll die individuelle Verteilung bei den einzelnen Probanden zeigen. Ein Großteil der Werte von Proband 7 liegt oberhalb des errechneten Mittelwertbereichs. Die Verteilung ist bis auf einzelne Ausreißer jedoch regelmäßig b) Proband 6 zeigt in den betrachteten Zeitintervallen eine regelmäßige Verteilung des Slopes innerhalb der Standardabweichung. Die Verteilung gestaltet sich bei den anderen Probanden ähnlich.

Neben dem Slope eignet sich auch der Tilt dazu, die Neigung eines Spektrums zu beschreiben. Abbildung 52 zeigt in a) die Verteilung der einzelnen Tilt Werte und ihre relative Häufigkeit. Der Mittelwert aller Tilt Ergebnisse dieser Studie liegt bei $-8 \pm 1,5 \text{ dB}$. Um diesen Wert verteilen sich die Werte regelmäßig und die Häufigkeit ist zu beiden Seiten hin abnehmend. Zu den höheren Tilt Werten ist das Gefälle etwas flacher als zu den kleineren Tilt Werten. Abbildung 52 b) zeigt die Streuung des Tilt Wertes in den unterschiedlichen Tagesintervallen. Der Median liegt zwischen -8 bis -9 dB . Eine große Übereinstimmung zwischen Mittelwert und Median ist gegeben. Nur morgens zwischen 5-10 Uhr ist der Median größer -8 dB . 50% der Aufnahmen liegen in etwa zwischen -7 bis $-11,5 \text{ dB}$. Nur zwischen 5-10 Uhr gibt es eine Abweichung. Etwa 50 % der Aufnahmen liegen zwischen $-8,5 \text{ dB}$ und $-6,5 \text{ dB}$. Im Vergleich zum Slope ist der Tilt in einem wesentlich kleineren Bereich verteilt.

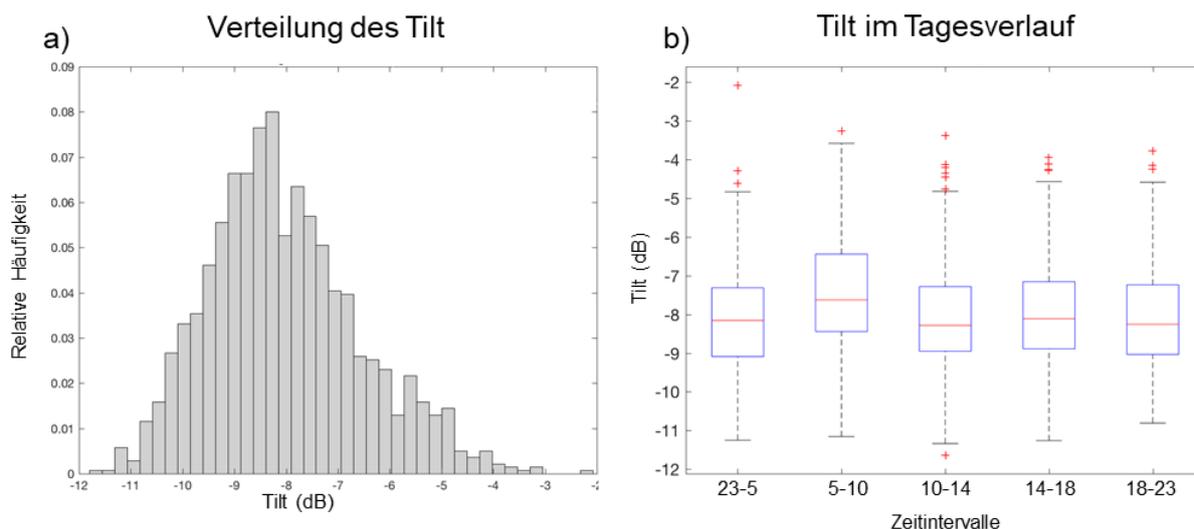


Abbildung 52: a) Verteilung aller Tilt Werte aus der Studie um den Mittelwert $-8 \pm 1,5$ dB. Die Verteilung nimmt zu den größeren Werten hin flacher ab als zu kleineren Tilt Werten. b) Die Verteilung um den Median bei circa -8 dB ist gleichmäßig. 50 % der Aufnahmen liegen zwischen -8,5 und -6,5 dB. Zwischen 5 bis 10 Uhr ist der Median des Tilts etwas größer als in den anderen Intervallen.

Der zeitliche Verlauf des Tilts zeigt bei jedem einzelnen Probanden leichte Streuungen, die aber im Vergleich zum Slope viel geringer ausfallen. Wie auch beim Slope zeigt sich ein individuelles Tilt Profil in den unterschiedlichen Zeitintervallen. Wie beim Slope sind auch die Werte des Tilts zwischen 5-10 Uhr anders verteilt als in den anderen Zeitintervallen.

4.7.8 Smoothed cepstral peak prominence (CPPs)

Die Ergebnisse des Smoothed cepstral peak prominence (CPPs) sind in Abbildung 53 dargestellt. Für die Beurteilung der Stimmqualität gilt er als der wichtigste Parameter, da durch in Stimmstörungen besonders gut detektiert werden können (vgl. 2.8.3). Der errechnete Mittelwert aus dieser Studie beträgt $12 \pm 2,1$ dB und die Verteilung der einzelnen Werte ist in Abbildung 53 a) dargestellt. Die relative Häufigkeit der Werte nimmt zu beiden Seiten des Mittelwerts gleichmäßig ab. Abbildung 53 b) zeigt, dass der Median in allen 5 Zeitintervallen etwa zwischen 12 und 13 dB liegt. Auch für diesen Wert zeigt sich eine große Übereinstimmung zwischen Mittelwert und Median. Etwa 50 % der Aufnahmen ergeben einen CPPs zwischen ca. 10 dB und 13,5 dB. 97% der CPPs Werte liegen in etwa zwischen 5 dB – 17,5 dB.

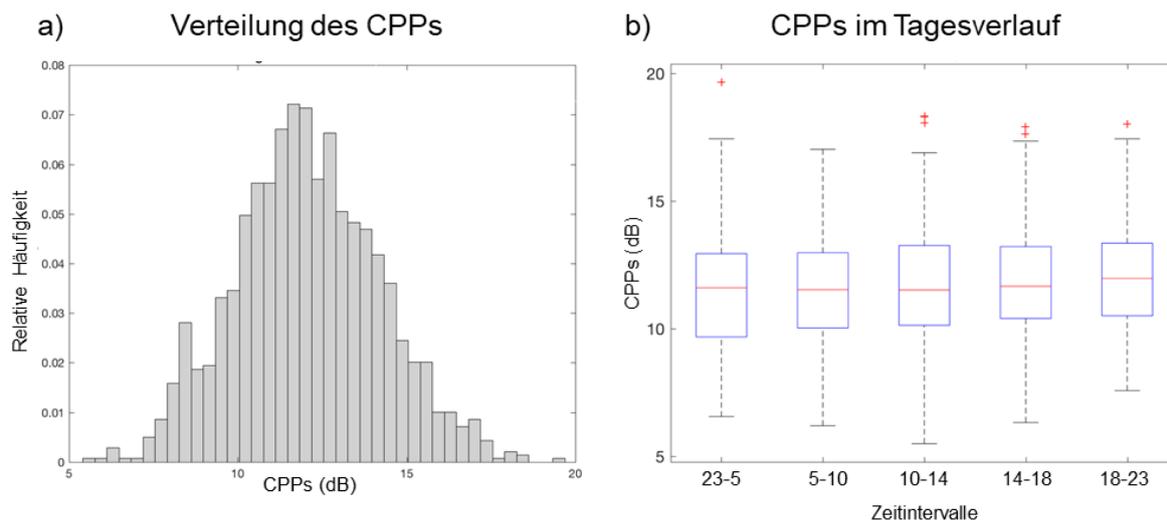


Abbildung 53: a) Der Mittelwert der Smoothed cepstral peak prominence ist bei $12 \pm 2,1$ dB. Die Werte verteilen sich gleichmäßig abfallend zu beiden Seiten. b) Der Median ist dem Mittelwert in allen Zeitintervallen sehr ähnlich. Die Aufnahmen liegen in allen 5 Zeitintervallen in einem ähnlichen Bereich.

Zwischen den CPPs der einzelnen Probanden zeigen sich keine wesentlichen Unterschiede. Weder in Anbetracht des zeitlichen Verlaufs, noch innerhalb der 5 Zeitintervalle. Die Streuung des CPPs Wertes ist im Studienzeitraum insgesamt etwas größer als beispielsweise beim Jitter, dennoch ist die Verteilung bei den einzelnen Probanden in ihrem individuellen Bereich regelmäßig. Diese Verteilung ist in Abbildung 54 a) gezeigt, der CPPs des Probanden 1 liegt etwas unterhalb des Mittelwertes, der Median ist im Tagesverlauf etwas ansteigend, liegt jedoch im Bereich von 10 dB. Die Streuung des CPPs von 50 % der Aufnahmen von Proband 1 ist zudem sehr gering. Die geringe Streubreite zeigt sich auch bei Proband 2 in Abbildung 54 b). Bei diesem Probanden liegen fast alle CPPs Werte im Bereich der Standardabweichung. Im Gegensatz zum Slope und Tilt ist die Streuung dieses spektralen Parameters bei allen Probanden geringer.

CPPs einzelner Probanden im Vergleich zum Mittelwert des Gesamtkollektivs

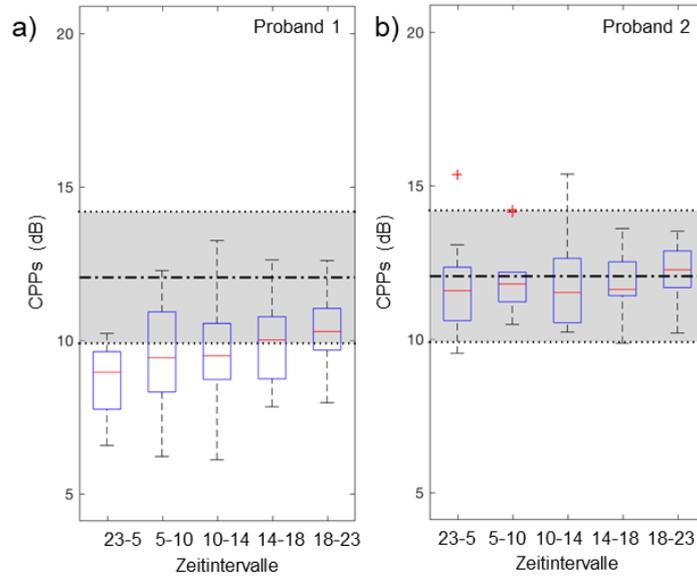


Abbildung 54: Verteilung des CPPs in den unterschiedlichen Zeitintervallen am Beispiel zweier Probanden. a) Bei Proband 1 ist der Median im Tagesverlauf ansteigend und liegt etwas unterhalb des berechneten Mittelwertes $12 \pm 2,1$ dB. Die Verteilung des Wertes ist jedoch regelmäßig und ohne große Streuung. b) Der CPPs bei Proband 2 liegt fast immer im Bereich der Standardabweichung.

4.7.9 t-SNE

Das Ergebnis in Abbildung 55 zeigt, dass die Daten sich nach Dimensionsreduktion innerhalb eines großen und eines kleinen Clusters verteilen.

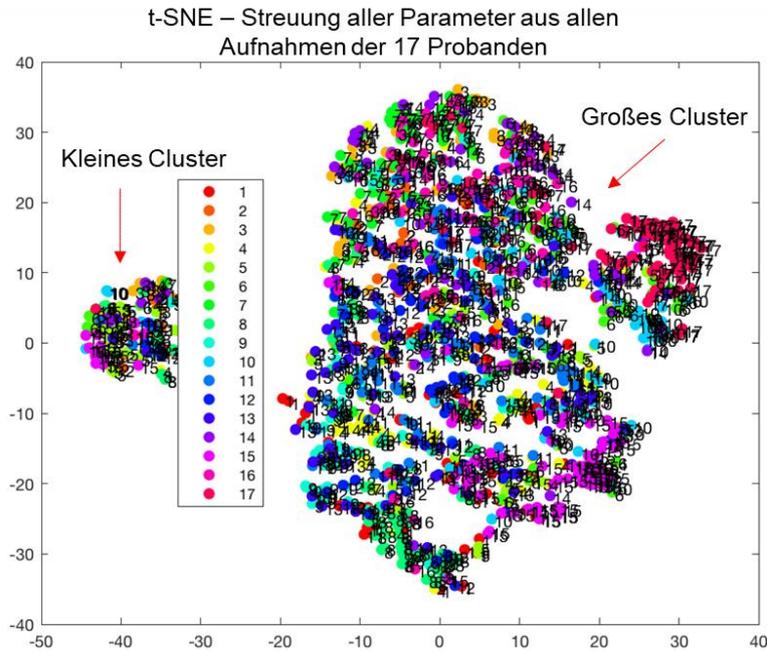


Abbildung 55: Verteilung aller 9 Parameter aus dem telemedizinischen Monitoring nach Dimensionsreduktion mittels t-SNE Verfahren. Berücksichtigung aller Aufnahmen der 17 Probanden. Entstehung von zwei unterschiedlich großen Clustern.

Um zu untersuchen aus welchem Grund sich diese Clusterung ergibt, wurden die 9 Parameter der Aufnahmen, die sich dem kleinen und großen Cluster zuordnen lassen, identifiziert und gegenübergestellt. Um die Werte des kleinen Clusters zu finden, wurde die Verteilung von jedem Wert aus dem telemedizinischen Monitoring, bis auf die Grundfrequenz, im großen und kleinen Cluster in einem Boxplot dargestellt und die beiden Abbildungen miteinander verglichen. Beispielhaft wurde dies mit den Parametern Jitter und NNE in Abbildung 56 gezeigt, da sich, bis auf den NNE, alle anderen Werte ähnlich wie der Jitter verhalten (vgl. Abbildung 69). Während der Jitter zwischen den beiden Clustern ähnlich ist, zeigen sich nur beim NNE große Unterschiede zwischen den beiden Clustern. Im Gegensatz zum Jitter und allen anderen Parametern ist der NNE in beiden Clustern sehr unterschiedlich verteilt (vgl. Abbildung 56). Die Verteilung des NNE zeigt weder übereinstimmende Mediane noch Quartile. Die Streuung des kleinen und großen Clusters ist sehr unterschiedlich.

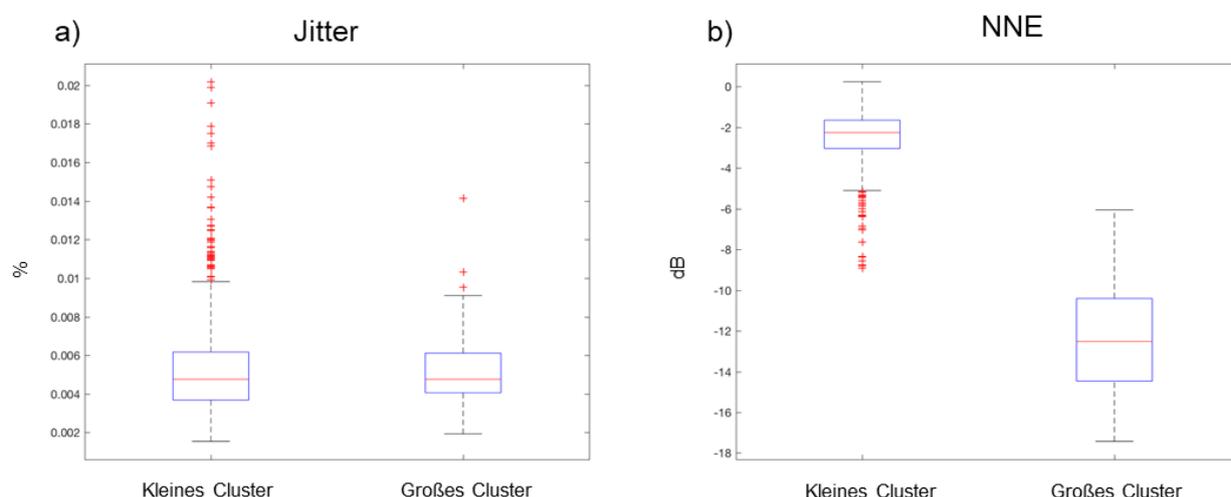


Abbildung 56: Vergleich der beiden Cluster von Jitter und NNE. a) Gleichmäßig und fast deckungsgleiche Verteilung des Jitters in beiden Clustern. Die roten Markierungen zeigen die einzelnen Ausreißer. b) Keine Übereinstimmung der NNE Werte in beiden Clustern.

Da der NNE im Vergleich zu den anderen Werten sehr unregelmäßig in seiner Verteilung ist, wurden die Werte ohne den NNE erneut mit dem t-SNE Verfahren in einer Grafik dargestellt. Das Ergebnis in Abbildung 57 zeigt, dass das kleine Cluster verschwunden ist. Das kleine Cluster bestand überwiegend aus NNE Werten. Im neuen Cluster liegen die Werte von Shimmer und Jitter, HNR, NHR, H_1H_2 , Tilt und Slope und von CPPs verteilt.

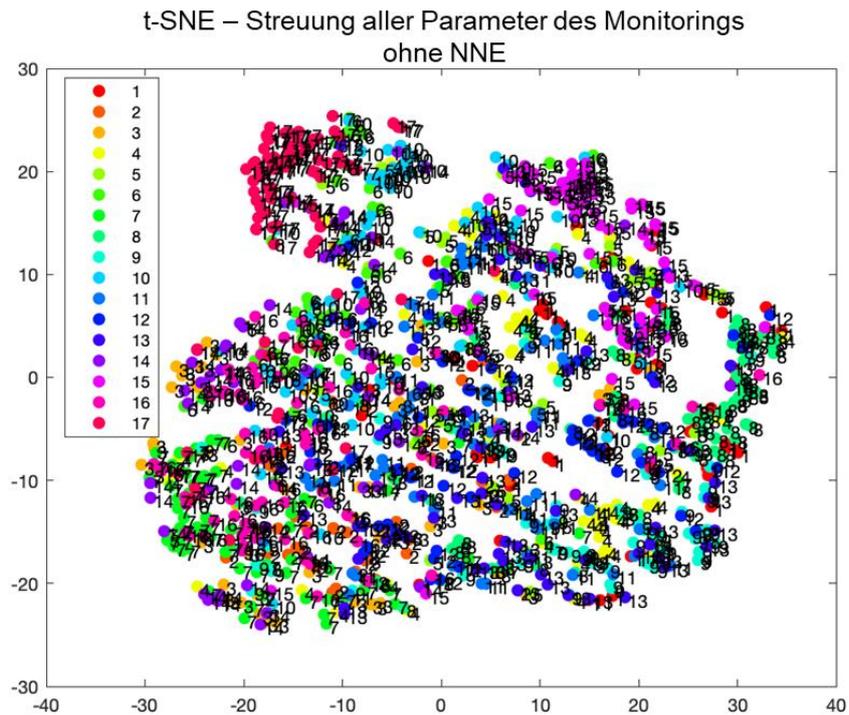


Abbildung 57: Verteilung der Parameter aus dem telemedizinischen Monitoring ohne NNE. Das kleine Cluster ist nach einer Auswertung ohne NNE Werte verschwunden.

Anschließend wird nun das Cluster nach den jeweiligen Probanden aufgeschlüsselt und die Verteilung aller Werte für jede einzelne Testperson betrachtet. Als Beispiel wurden die Cluster der ersten 4 Probanden betrachtet, die in Abbildung 58 dargestellt sind. Die Ergebnisse der einzelnen Probanden häufen sich in individuellen Bereichen. Bei Proband 1 beispielsweise liegen die Werte in der rechten Bildhälfte des Clusters, die Werte von Proband 2 sind vor allem im linken unteren Quadranten verteilt. Die Werte von Proband 3 sind ebenfalls im linken unteren Quadranten verteilt. Bei Proband 4 konzentrieren sich die Werte im oberen und unteren rechten Quadranten.

Verteilung der Werte bei den einzelnen Probanden

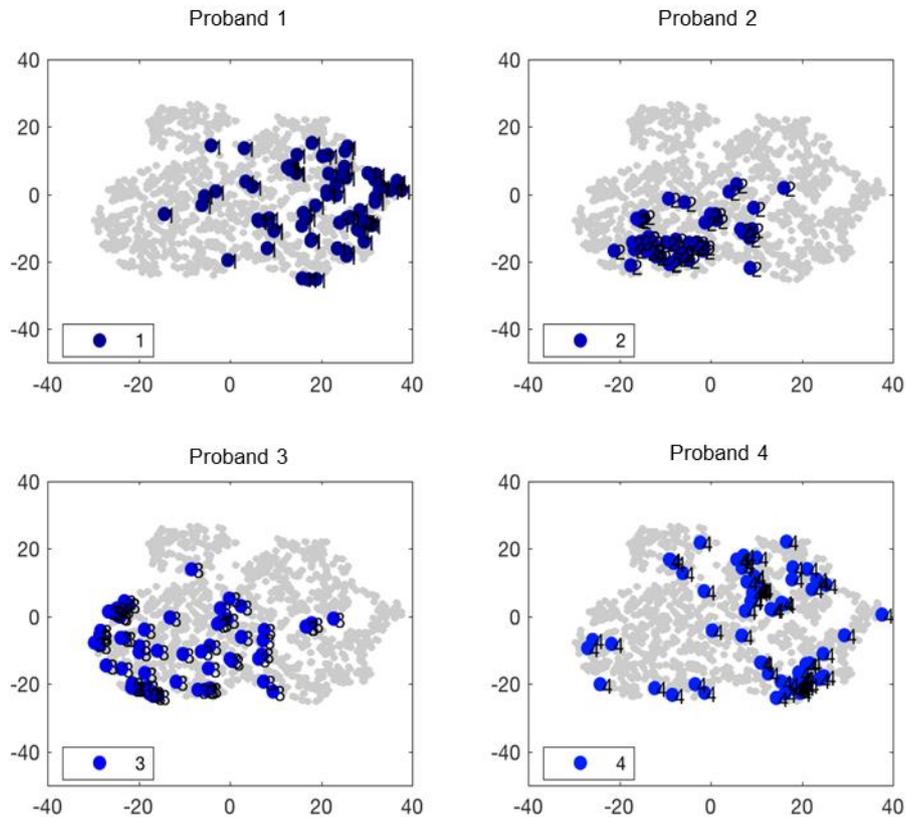


Abbildung 58: Betrachtung der Verteilung der erhobenen Parameter bei einzelnen Probanden. Die Probanden 1-4 zeigen, dass sich die Werte aus dem telemedizinischen Monitoring in einem bestimmten Bereich (Proband 3) oder um einen bestimmten Punkt (Proband 2) im Cluster ansammeln.

Nachdem die Betrachtung der einzelnen Probanden gezeigt hat, dass jeder Proband ein bestimmtes Gebiet im Cluster in Anspruch nimmt, soll nun die Verteilung der einzelnen Probanden Daten in den unterschiedlichen Zeitintervallen betrachtet werden. Dazu wurden wieder alle erhobenen Werte ausschließlich des NNE betrachtet. Aufgrund des ähnlichen Streuverhaltens unter den einzelnen Probanden werden beispielhaft zwei unterschiedliche Probanden besprochen. Die t-SNE Verteilung des Proband 13, aufgeschlüsselt nach den Zeitintervallen, ist in Abbildung 59 gezeigt. Hierbei zeigt sich, dass unabhängig von der Uhrzeit die Werte in einem ähnlichen Bereich liegen und ähnlich verteilt sind.

t- SNE - Verteilung in den 5 Zeitintervallen
ohne NNE

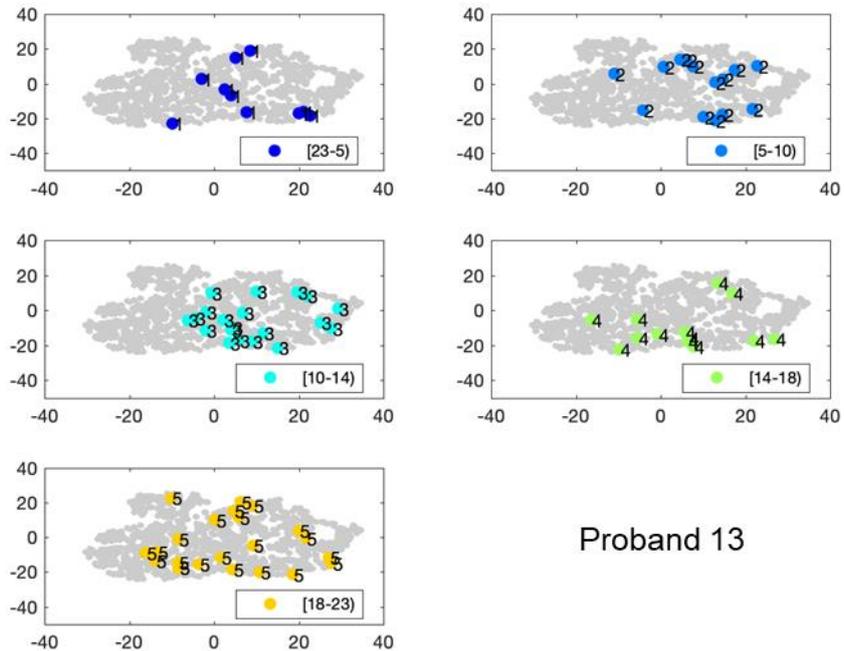


Abbildung 59: Verteilung der Parameter des telemedizinischen Monitorings mit Ausnahme des NNE in den 5 unterschiedlichen Zeitintervallen am Beispiel von Proband 13. Die Werte verteilen sich unabhängig von der Uhrzeit immer in einem ähnlichen Bereich. In diesem Bereich liegt eine gewisse Variabilität vor.

Um ein weiteres Beispiel zu zeigen, wurde die Verteilung der Werte ausschließlich NNE in den unterschiedlichen Zeitintervallen noch bei Proband 17 in Abbildung 60 betrachtet. Bei diesem Probanden ist die Verteilung der Werte zu den unterschiedlichen Zeiten fast identisch und liegt im oberen linken Quadranten. Die Streuung der Werte ist bei diesem Probanden noch geringer, als bei Proband 13. Die einzelnen Aufnahmen bei Proband 17 sind in sich sehr ähnlich.

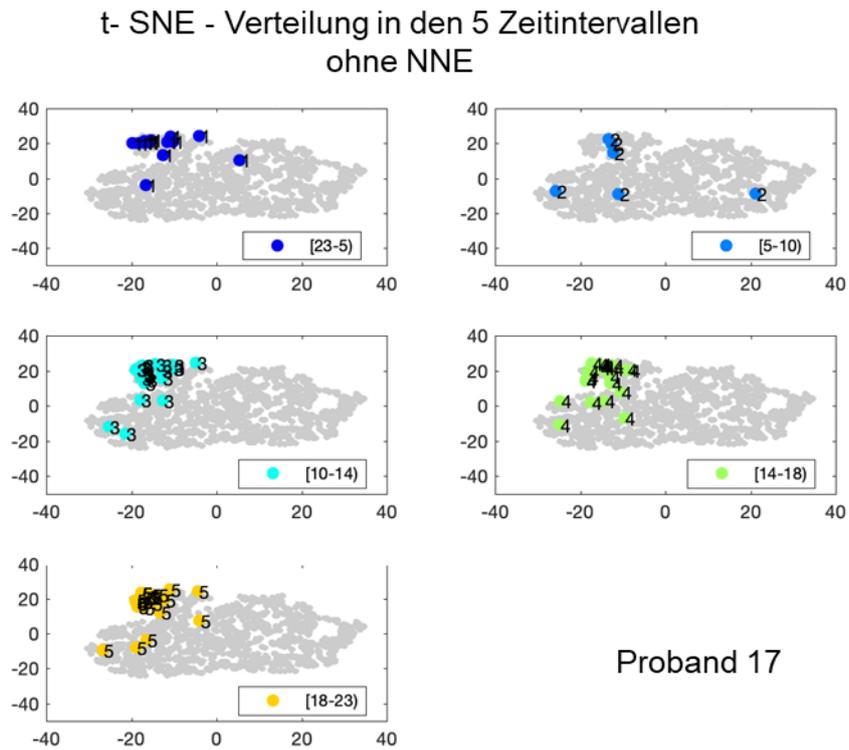


Abbildung 60: Verteilung der Parameter des telemedizinischen Monitorings mit Ausnahme des NNE in den unterschiedlichen Zeitintervallen am Beispiel von Proband 17. Die Verteilung weist eine deutlich geringere Variabilität auf als bei Proband 13, deren Verteilung in allen 5 Bildern nahezu identisch ist.

5 Diskussion

Die Einschlusskriterien dieser Studie zeigten, dass es sich bei den Probanden um 20 gesunde Freiwillige gehalten hat. In der Stimmfeldmessung traten keine großen Abweichungen zu vergleichbaren Referenzwerten aus der Literatur auf. Im integrierten Fragebogen der „VoiceDiary“-App zeigte sich bei fast allen Probanden eine normale Stimmbeanspruchung und eine gute Stimmregeneration. Die Frage der Stimmqualität hingegen erbrachte abweichende Ergebnisse. Zudem konnte in der Studie zwischen Werten, welche im telemedizinischen Monitoring zur Stimmanalyse geeignet sind und welche in diesem Verfahren anfällig für Fehler sind, differenziert werden. Die Übereinstimmung der Ergebnisse aus dem Studieneinschluss und dem telemedizinischen Monitoring mit Referenzwerten aus der Literatur war ebenfalls sehr groß. Daher kann angenommen werden, dass das Aufzeichnungsverfahren mit Hilfe der Handy App in der Diagnostik verwendet werden kann.

5.1 Einschluss in die Studie

Die Bedingungen für eine Teilnahme an der Studie waren stimmgesunde, männliche Probanden. Ein Einschluss in die Studie war möglich, wenn in den Voruntersuchungen der Anamnesebogen unauffällig war, die Gesamtpunktzahl im VHI-12 Test > 5 Punkte nicht überschritten wurde [33] und eine unauffällige stroboskopische Untersuchung vorlag. Die VHI-12 Auswertung zeigt, dass die subjektive Beurteilung der Stimme bei den Probanden unauffällig war. Vereinzelt war während der stroboskopischen Untersuchung eine Phasendifferenz zu erkennen, die sich bei Stimmgesunden vor allem in hohen Stimmlagen feststellen lässt [93]. Zum einen haben ungeübte Personen Schwierigkeiten in hohen Frequenzen gleichmäßig zu phonieren, zum anderen kann die Phasendifferenz durch Irritationen während der Stroboskopie erklärt werden. Hierbei sind die Schwingungsabläufe der Schwingungen nicht synchron.

Die Voruntersuchungen zeigten, dass die 20 untersuchten Probanden in Zusammenschau der Daten als stimmgesund eingestuft wurden und somit als Kandidaten für die Erhebung von Normdaten bei Stimmgesunden in dieser Studie geeignet waren. Da die Handyapp von 3 Probanden täglich nicht hinreichend oft genutzt wurde und kein zusammenhängender Zeitraum von 14 Tagen gegeben war, wurden lediglich die erhobenen Daten von 17 Probanden in die Studie ausgewertet. Im Rahmen der Voruntersuchung wurden die Sprech- und Singstimme, sowie der DSI betrachtet. Zu beachten ist, dass der exakt 14-tägiger Zeitraum zwischen Ein- und Ausschluss der Studie in der Realität nicht verwirklicht war.

Die Ergebnisse der mittleren Sprechstimmlage der Studie, $106 \text{ Hz} \pm 11 \text{ Hz}$ und $62 \text{ dB} \pm 3 \text{ dB}$ zu Studienbeginn und $109 \text{ Hz} \pm 10 \text{ Hz}$ und $66 \text{ dB} \pm 4 \text{ dB}$ am Ende der Studie, sind zu Referenzwerten der mittleren Sprechstimmlage aus der Literatur (112 Hz und 62 dB) ähnlich

und befinden sich im Bereich der Referenzwertspanne 89 - 175 Hz und 57 – 68 dB [12]. Friedrich et al. (2000) gab ebenfalls Referenzwerte in einem Frequenzspektrum von 98 – 131 Hz an [28]. Die erhobenen Werte der mittleren Sprechstimmlage weichen zu Beginn und am Ende für die einzelnen Probanden nur geringfügig voneinander ab. Diese Übereinstimmung und das Fehlen von Ausreißern in der Datenerhebung, sprechen für gesunde Probanden mit einer stabilen mittleren Sprechstimmlage. Dieser Parameter hat für die Studie eine wichtige Bedeutung, da die sich die Aufnahmen der Handy App nach der ermittelten und eingestellten mittleren Sprechstimmlage der jeweiligen Probanden richtet. Wird beispielsweise eine zu hohe mittlere Sprechstimmlage in der Stimmfeldmessung ermittelt und als Grundfrequenz der Handyapp eingestellt, kann es durch die unnatürliche Tonhöhe zu stimmlicher Fehlbelastung bei dem Probanden kommen, was wiederum die Ergebnisse des telemedizinischen Monitorings verfälschen würde. Da die mittlere Sprechstimmlage der einzelnen Probanden zu Beginn und am Ende fast identisch ist und von keinem Probanden das Aufzeichnen der Aufnahmen als unangenehm empfunden wurde, kann angenommen werden, dass ein Fehler dieser Art in der Studie nicht vorgekommen ist.

Der berechnete Mittelwert der maximalen Ruflautstärke war mit $104 \text{ dB} \pm 5 \text{ dB}$ und $106 \text{ dB} \pm 6 \text{ dB}$ am Ende gegenüber einem Referenzwert von 92 dB erhöht [12]. Das Alter des Kollektivs von diesem Referenzwert war jedoch zwischen 18 und 67 Jahren und somit im Mittel deutlich älter als das Probandenkollektiv in dieser Studie. Andere vergleichbare Daten ergaben Werte zwischen 100 - 110 dB [28]. Ein guter Vergleich ist mit den Daten aus Schneider-Stickler et al. 2013 [93] möglich, da es sich bei dieser Probandengruppe ebenfalls um ausschließlich junge Männer um die 30 Jahre handelt und die Daten mit „LingWAVES“ erhoben wurden. Der errechnete Mittelwert in dieser Gruppe liegt bei $101 \text{ dB} \pm 8 \text{ dB}$ und stimmt mit dem Mittelwert aus dieser Studie nahezu überein. Die geringen Unterschiede zu anderen Referenzwerten aus der Literatur können durch unterschiedliche Anweisung bei der Messung entstanden sein. Aufgrund der geringen Differenz können diese jedoch vernachlässigt werden. Entweder wird der Abstand zum Mikrofon nicht korrekt eingehalten oder die Probanden rufen zu stark bei den Aufnahmen. Alle Probanden dieser Studie haben eine Ruflautstärke $> 90 \text{ dB}$ erreicht. Das Erreichen dieser Lautstärke beim Rufen spricht in der Literatur für eine normale Leistungsfähigkeit einer gesunden Stimme [93].

Die Messungen der maximalen Frequenz der Singstimme in der Studie lagen bei $597 \text{ Hz} \pm 163 \text{ Hz}$ zu Beginn und bei $598 \text{ Hz} \pm 146 \text{ Hz}$ am Ende der Studie. Die Verteilung zu beiden Zeitpunkten ist sehr ähnlich. Zu Referenzwerten aus der Literatur (Mittelwert 523 Hz und Wertespanne zwischen 391-698 Hz) besteht ebenfalls eine große Übereinstimmung [12]. Die größere Wertespanne der Referenzwerte aus der Literatur im Vergleich zu dieser Studie lassen sich anhand des im Mittel älteren Patientenkollektivs (18-67Jahre) erklären. Die Standardabweichung der maximalen Frequenz in dieser Studie ist auffallend groß, da die

Probanden ihre Stimme bei den Aufnahmen unterschiedlich ausgeschöpft haben. Das hohe Singen wurde von einigen Probanden als unangenehm empfunden und die maximale Singkraft wurde dadurch nicht ausgeschöpft.

Die Mittelwerte der minimalen Frequenz der Singstimme (zu Beginn $85 \text{ Hz} \pm 8 \text{ Hz}$, bei Ausschluss $87 \text{ Hz} \pm 5 \text{ Hz}$) stimmen mit einem Mittelwert von 82 Hz aus der Literatur überein [12]. Die Mittelwerte des Schalldruckpegels werden in der Literatur teilweise mit 98 dB etwas höher angegeben [12], in dieser Studie lagen die Mittelwerte bei Studieneinschluss und Abschlussuntersuchung bei $91 \text{ dB} \pm 6 \text{ dB}$ und $91 \text{ dB} \pm 7 \text{ dB}$. Andere Quellen fordern bei physiologischen Stimmen Werte $> 90 \text{ dB}$ [93]. Die leicht erniedrigten Werte in dieser Studie lassen sich durch die Singerfahrung und durch die Compliance der Probanden erklären. Bei der Untersuchung fiel auf, dass den Probanden teilweise das laute und kräftige Singen aufgrund ihrer wenigen Singerfahrung unangenehm war und sie so nicht richtig mitarbeiteten. Da die Probanden klinisch gesund sind und auch die bisher erhobenen Stimmanalyseparameter einen physiologischen Zustand der Probanden bekräftigen, werden diese Unterschiede auf den Patienten und seine Mitarbeit zurückgeführt.

Die Spanne für den leisesten Schallpegels liegen bei stimmgesunden Männern zwischen 18-67 Jahren zwischen $46 - 57 \text{ dB}$, der Mittelwert bei 51 dB [12]. Die Mittelwerte zu Beginn und am Ende der Studie liegen jeweils bei $47 \text{ dB} \pm 4 \text{ dB}$. Der Dynamikbereich der Singstimme in dieser Studie ($44 \text{ dB} \pm 7 \text{ dB}$ und $45 \text{ dB} \pm 8 \text{ dB}$) ist nahezu identisch mit einer vergleichbaren Probandengruppe mit einem Mittelwert von 47 dB und einer Wertespanne von $35 - 58 \text{ dB}$ [12]. Die Ergebnisse der Stimmdynamik und der Vergleich mit bisher bestehenden Werten von gesunden Probanden zeigen, dass die Probanden in der Variation der Lautstärke nicht beeinträchtigt sind, was wiederum Zeichen einer gesunden Stimme ist. Der Mittelwert des Jitters in dieser Studie ist $0,12\% \pm 0,047$ und liegt bei allen Probanden unterhalb der pathologischen 1% Grenze. Dies spricht für Frequenzregelmäßigkeit und Stimmstabilität während der Singstimmfeldmessung [78]. Die nahezu identischen Mittelwerte und Standardabweichungen des Singstimmprofil zeigen, dass die Stimme der Probanden sich im Verlauf des Studienzeitraums wenig verändert haben und die Stimmfeldmessung ordentlich durchgeführt worden ist.

Der letzte erhobene Wert bei Einschluss der Studie war der DSI. Ein angestrebter DSI $> 4,2$ [78], welcher für eine physiologische Stimme spricht, wurde bei 15 von 17 Probanden erreicht. Der errechnete Mittelwert liegt mit $6,4 \pm 1,8$ zu Beginn und mit $6,3 \pm 1,5$ am Ende der Studie weit über dem erstrebten Grenzwert. Bei Proband 1 lagen die DSI Werte zu Beginn bei 3 und am Ende der Studie bei 3,1. Bei Proband 8 war nur der Anfangswert mit 2,6 erniedrigt. Der erniedrigte DSI-Werte von 2,6 bei Proband 8 kann aufgrund der normalen Messergebnisse bei Studienende durch fehlende Mitarbeit während der Stimmfeldmessung erklärt werden, da sich beim Ausschöpfen der gesamten Stimmkraft ein normwertiger DSI ergeben hat. Bei

Proband 1 liegen beide DSI-Werte unter 4. Dies kann im Rahmen einer fehlenden stimmlichen Begabung erklärt werden, denn auch bei gesunden Probanden ohne Dysphonien kann der DSI unter 4 liegen, ein Normwert für Stimmunbegabte wurde bei $> 1,6$ ermittelt [78, 116]. So spricht ein erniedrigter DSI nicht immer automatisch für eine kranke Stimme und sollte daher stets unter Berücksichtigung aller anderen erhobenen Daten, beispielsweise der bildgebenden Diagnostik sowie der subjektiven Stimmbewertung, betrachtet werden. Abschließend ergibt sich, dass auch das Singstimmprofil in der Ein- und Abschlussuntersuchung vergleichbare Ergebnisse hervorgebracht hat und es zu bereits erhobenen Literaturdaten kaum Abweichungen gibt.

Insgesamt zeigte die gesamte Stimmfelduntersuchung der erhobenen Parameter nur geringfügige Unterschiede zu Studienbeginn und -ende, sowie eine hohe Übereinstimmung mit Referenzwerten aus der Literatur. Daraus lässt sich schlussfolgern, dass die Messungen richtig durchgeführt wurden und dass die Ergebnisse aus „LingWAVES“ und aus dem telemedizinischen Monitoring in Anbetracht aller Voruntersuchungen von gesunden, männlichen Probanden stammen.

5.2 Auswertung des Fragebogens in „Voice Diary“

Die Auswertung des Fragebogens ergab, dass die Probanden die Handy-App ähnlich oft und vor allem zwischen 10 bis 23 Uhr verwendet haben. Die Probanden beurteilten jede Aufnahme individuell hinsichtlich Haupt- und Nebensymptomen von Stimmstörungen. Die Auswertung zeigte, dass kein Proband eine höhergradige Stimmstörung empfand. Ziel des Fragebogens war es, die subjektive Einschätzung der eigenen Stimme durch gesunde Probanden abzuschätzen und diese mit den errechneten Werten aus dem telemedizinischen Monitoring zu vergleichen. Starke Ausreißer der errechneten Werte aus dem telemedizinischen Monitoring beispielsweise können mithilfe der Aufnahmebewertung erklärt werden.

Die Auswertung der Frage nach der Stimmbelastung zeigt, dass morgens die Stimmbelastung geringer war als im Tagesverlauf. Dies spricht für eine gute Regenerationszeit der Stimme (vgl. Abbildung 26). Stimmbelastungen traten bei den meisten Probanden während des Sprechens auf. Proband 5 hingegen gibt eine hohe Belastung vor allem beim Rufen an. Bei Proband 14 tritt eine Stimmbelastung lediglich beim Singen auf. Dies lässt sich am besten anhand der Freizeitgestaltung der jungen, nicht berufstätigen Probandengruppe erklären. Da die Probanden allesamt Studenten und keine professionellen Sänger sind, ist dieses Ergebnis in sich schlüssig. In Abbildung 26 wird gezeigt, dass im Tagesverlauf die Stimmbelastungen des Gesamtkollektivs zunahm, da gegen Nachmittag und Abend die Belastung als höhergradig eingestuft wurde. Die individuelle Bewertung der Belastungsgrade zeigte Variationen, wie anhand eines Vergleichs von Proband 1 und 10 deutlich wird (vgl. Abbildung 27). Die Belastung sollte im Hinblick auf erhöhte Stimmermüdung und auf Auftreten von

Missempfindungen während der Sprachproduktion betrachtet werden. Bei fehlenden Symptomen kann die Stimme dann dennoch als gesund bewertet werden [28].

Die Daten der Stimmqualität aus dieser Studie sind nicht vergleichbar mit den anderen Daten. Die Bewertungsgrade für die Qualität 1 „schlecht“ bis 6 „gut“ sind im Vergleich zu den anderen Fragen genau umgekehrt. So hat die gesunde Probandengruppe in dieser Studie laut Fragebogenauswertung eine überdurchschnittliche schlechte Stimmqualität, was für eine pathologische Stimme sprechen würde [28]. Dies scheint nicht mit der Realität übereinzustimmen, da in einer nachträglichen Befragung die Probanden ihre Stimme zum Großteil als „gut“ empfunden haben. Einige Probanden haben am Ende der Studie angegeben, dass sie die Veränderung in den Bewertungsgraden bei dieser Frage nicht bemerkt haben (vgl. Abbildung 16). Eine gültige Interpretation der Stimmqualität ist im Nachhinein schwierig, da nicht mehr nachvollzogen werden kann, welcher der Probanden nun seine Stimmqualität tatsächlich oder aufgrund der Inversion der Bewertungsgrad mit „sehr schlecht“ oder „sehr gut“ bewertet hat. Die Befragung der Stimmqualität wird daher nachträglich aus der Auswertung ausgeschlossen. Bei erneutem Durchführen einer Studie mit ähnlichem Befragungsbogen sollte auf eine einheitliche Nummerierung des Bewertungssystems bei allen Fragen beachtet werden.

Die Ergebnisse der relativen Häufigkeit von aufgetretenen Missempfindungen haben gezeigt, dass nur in einzelnen Fällen Missempfindungen, wie Räuspern oder Kratzen im Gesamtbild aller Aufnahmen angegeben wurden. Aufgrund dessen ist die relative Häufigkeit der Missempfindungen insgesamt gering. Da Probanden, die wie Proband 10 und 12, trotz vermehrter Stimmbelastung, keine höhergradigen Missempfindungen empfunden haben, spricht dies für gesunde und regenerationsfähige Stimmen in dieser Studie [116].

Die Auswertung der Frage nach Stimmermüdung zeigt, dass bei über der Hälfte der Aufnahme keine Stimmermüdung empfunden wurde (vgl. Abbildung 32). Der Vergleich der Ermüdungsgrade (vgl. Abbildung 32) und der Stimmbelastung (vgl. Abbildung 28) zeigt einen Zusammenhang. Proband 16 und 17 beispielsweise empfinden im Vergleich zu den Probanden 5 und 9 eine höhere Stimmbelastung und auch eine höhere Stimmermüdung. Zu beachten ist, dass jedoch auch bei höhergradiger Belastung nur eine leichte bis mittlere Stimmermüdung bei Proband 16 und 17 beschrieben wurde. Da auch bei schwerer Stimmbelastung die Stimmermüdung von den Probanden meist nur als leicht empfunden wurde, kann die Stimme als gesund bewertet werden [28]. Auffallend in dieser Studie ist, dass schwere Ermüdungsgrade in den Abendstunden und vor allem auch nachts (23-5 Uhr) angegeben wurden. Diese Aufnahmen sind in Anbetracht des jungen Probandenkollektivs aus Studenten daher am ehestem im Rahmen der abendlichen Freizeitgestaltung, wie auf Feiern und auf Partys, entstanden. Da in dieser Studie ein gesundes Probandenkollektiv mit einem regulären Tagesablauf mit Einhalten der Nachtruhe betrachtet werden sollte, können diese

späteren Aufnahmen mit hohen Ermüdungsgraden als möglicher systemischer Fehler gewertet werden und nachträglich aus der Studie ausgeschlossen werden. An dieser Stelle ist zudem anzumerken, dass einzelne Probanden, wie beispielsweise Proband 6, die Frage nach dem Ermüdungsgrad auf ihren Allgemeinzustand bezogen. Ein Verbesserungsvorschlag hier ist, in der Formulierung der Frage den „Ermüdungsgrad der Stimme“ explizit noch einmal zu erwähnen und bei Einschluss der Studie und während der Erklärung der „Handy-App“ mögliche Unklarheiten mehrmals abzufragen.

Proband 10 sollte gesondert betrachtet werden, da er eine vermehrte Stimmbelastung, eine anders verteilte Stimmqualität und einen höheren Ermüdungsgrad als die anderen Probanden angegeben hat. Da die Parameter aus dem telemedizinischen Monitoring jedoch keine Abweichungen vom Gesamtkollektiv zeigen, ist die Auswertung aus Voice Diary bei diesem Probanden, trotz der Abweichung, nicht als pathologisch zu werten.

Zusammenfassend ist die subjektive Auswertung der Stimmaufnahmen durch die Probanden sinnvoll, wenn die Frage nach der Stimmqualität ausgeschlossen wird. Da es sich um ein gesundes Probandenkollektiv handelt und auch die subjektive Einschätzung der Stimme gut ist, können die aus der Stimmanalyse gewonnen Parameter als Referenzwerte einer gesunden Stimme betrachtet werden. Wie sich die Parameter tatsächlich im Vergleich zu in der Literatur genannten Werten verhalten, wird im letzten Punkt diskutiert.

Eine Optimierung der Fragen auf der „Handy-App“ sollte zudem in Betracht gezogen werden, um Missverständnisse vorzubeugen. Die Skalierung der Fragen sollte einheitlich sein und die Fragen einfach und klar formuliert werden.

Eine maßgebliche Rolle für auswertbare Ergebnisse der Studie spielt die Compliance der Probanden. Die Compliance während der zweiwöchigen Datenerhebung war gut, nur 3 von 20 Probanden mussten aufgrund von unregelmäßigen Zeitabständen zwischen den Aufnahmen aus der Studie ausgeschlossen werden. Dennoch sollte beachtet werden, dass es trotz guter Compliance der Probanden vereinzelt zu Fehlern während der Aufzeichnung der Stimme oder während der Datenerhebung kommen kann. Zum Beispiel, wenn Aufnahmen in einem lauten Umfeld aufgezeichnet wurden oder der Mund-Mikrofonabstand nicht eingehalten wurde. Einzelne Ausreißer in der Studie können dadurch erklärt werden.

5.3 Parameter aus dem telemedizinischen Monitoring

Die meisten analysierten Werte aus dem Monitoring, stimmen mit Werten aus früheren Studien oder aus der Literatur überein. Die Grundfrequenz, Shimmer und Jitter haben eine hohe Übereinstimmung. Bei den spektralen Parametern zeigen die Einzelwerte teils eine große Variabilität. In Hinblick auf alle Daten zeigt sich aber auch eine regelmäßige und gleichmäßige Verteilung. Einzelne Ausreißer in den Graphiken ergeben sich bei fast allen Parametern. Da die Aufnahmen von den Probanden nicht nur im häuslichen, sondern auch im freien Umfeld

oder bei nicht optimalen Messbedingungen aufgenommen wurden, lassen sich die abweichenden Werte durch teilweise nicht adäquate Messsituationen erklären. Der einzige Wert, der in dieser Studie nicht sinnvoll erscheint, ist der NNE. Dieser wurde bei der Betrachtung aller Werte durch das t-SNE Verfahren ausgeschlossen. Auch diese Studie zeigt, dass bei der Interpretation der Ergebnisse das Analyseprogramm berücksichtigt werden sollte, da die Verwendung von unterschiedlichen Analyseprogrammen teilweise abweichende Werte ergeben können. Eine neue Studie aus dem Jahr 2016 zeigt, dass neben dem Analyseprogramm auch die Wahl des Smartphone Modells eine Rolle spielt [36]. Die Daten dieser Studie, die mit einem HUAWEI Smartphone aufgezeichnet wurden, reihen sich problemlos neben den Werten, die mit anderen Smartphone Modellen aufgezeichnet wurden, ein [36].

5.3.1 F_0 und Irregularitätsparameter

Die Verteilung der F_0 Werte gestaltet sich in der Studie sehr regelmäßig. Die Grundfrequenzen von einem Großteil aller Aufnahmen lagen im Bereich des Mittelwertes $110 \pm 9,9 \text{ Hz}$. Zudem waren die Grundfrequenzen der einzelnen Aufnahmen aller Probanden um deren jeweilige mittlere Sprechstimmlage verteilt, die in der Einschlussuntersuchung bestimmt wurde. Die Grundfrequenz von Proband 2 liegt in der Studie weit über dem ermittelten Bereich. Da sich seine Grundfrequenz und seine ermittelte mSsl bei 131 Hz ähnlich verteilen, handelt es sich nicht um einen Messfehler. Das Verhalten von F_0 gegenüber der mittleren Sprechstimmlage ist interessant, da bei genauer und korrekter Bestimmung der Grundfrequenz F_0 und die mittlere Sprechstimmlage übereinstimmende Werte zeigen [75]. Dies ist in dieser Studie gegeben. Im Vergleich zu anderen Quellen zeigen sich leichte Unterschiede. Eine Studie mit männlichen, deutschsprachigen Probanden von Rappaport 1958 ergab eine mittlere F_0 von $129 \pm 2,3 \text{ Hz}$ [39]. Eine Gruppe englischsprachiger, stimmgesunder Probanden in einer ähnlichen Altersklasse ergab eine mittlere Grundfrequenz von $119 \pm 19 \text{ Hz}$ [25]. Ähnlich verteilte sich die Grundfrequenz einer gleichaltrigen, portugiesisch sprechenden Probandengruppe, Mittelwert $121 \pm 19 \text{ Hz}$ [99]. Der Mittelwert $110 \pm 9,9 \text{ Hz}$ dieser Studie ist etwas kleiner, aber weiterhin im von Mukhopadhyay beschriebenen Grundfrequenzrahmen von $105\text{-}160 \text{ Hz}$ [71]. Anhand der unterschiedlichen Ergebnisse kann eine valide Bestimmung der Grundfrequenz im telemedizinischen Monitoring angenommen werden.

Der Bereich eines normalen Jitter liegt $<1\%$ [78]. Aus Abbildung 37 wird ersichtlich, dass in der gesamten Studie Aufnahmen selten einen Jitter $>1\%$ erreichten. Proband 12 beispielsweise weist einen niedrigeren Jitter Wert und eine geringere Streubreite des Wertes im Vergleich zum Gesamtkollektiv auf. Dies zeigt, dass die Aufnahmen von Proband 12 wenig Irregularitäten aufweisen und somit wenig Frequenzschwankungen vorliegen. Auch anderen Studien, die zur Analyse die Software Praat verwendet haben, geben einen Jitter $\leq 1,02\%$ bei

Stimmgesunden an [70]. Die Ähnlichkeit der Werte aus der Literatur und dieser Studie spricht dafür, dass der Jitter in der Studie Werten von Stimmgesunden entspricht. Da es sich in dieser Studie um gesunde Probanden handelt, kann angenommen werden, dass mit der Aufnahmetechnik via Handy die Stimme ohne große Frequenzschwankungen aufgezeichnet wurde. Daraus resultiert, dass die Qualität des Handymikrofons für ein telemedizinisches Monitoring ausreichend ist und kein hochwertigeres Mikrofon oder andere spezielle Aufnahmetechniken benötigt werden.

Die Mediane des Shimmers in den 5 Zeitintervallen sind ebenso gleichmäßig verteilt wie die des Jitters. Die Werte des Shimmers aller 17 Probanden sind in der Studie ähnlich verteilt, innerhalb der Standardabweichung gleichwertig und somit plausibel, dennoch gibt es eine Abweichung zu den in der Literatur angegebenen Referenzwerten. Der Mittelwert des Shimmers in dieser Studie ($4,6 \pm 2,2$ %) liegt über der von beschriebenen Referenzwerten, die bei $< 2,5$ % [78] oder bei $\leq 3,8$ % angegeben werden [70]. Zum Vergleich eignet sich vor allem der Referenzwert $\leq 3,8$ % von Minnema 2008, da als Software ebenfalls Praat verwendet wurde. Der durchschnittlich leicht erhöhte Shimmer in dieser Studie kann auf die abweichende Aufnahmetechnik via Smartphone zurückgeführt werden.

Diese Vermutung wird durch eine Studie von Grillo et al. 2016 gestützt, die das Aufnahmeverfahren mit unterschiedlichen Handymodellen unter anderem an männlichen Probanden untersucht hat und unter Verwendung der Software Praat vergleichbare Shimmerwerte ergab. Hierbei wurden für das Handymodell iPhone 5 Shimmerwerte von $5,87 \pm 3,35$ % erhoben, für das Handymodell Samsung Galaxy 5 $4,67 \pm 3,05$ % [36]. Die vergleichbaren Ergebnisse zeigen, dass im telemedizinischen Monitoring die Shimmer Werte höher ausfallen, als die, die in einer klinischen Stimmfeldmessung erhoben wurden.

5.3.2 Akustische Parameter für Rauschen

Die akustischen Parameter fürs Rauschen HNR und NHR sind bis auf den NNE alle gleichmäßig um einen Mittelwert in der Studie verteilt und besitzen keine große Standardabweichung. Eine Übereinstimmung mit beschriebenen Werten aus anderen Studien liegt für die Parameter HNR und NHR vor.

Innerhalb der Studie war der HNR der Probanden regelmäßig um den Mittelwert $15 \pm 3,2$ dB verteilt. Die Hintergrundgeräusche haben bei fast allen Aufnahmen der Studie einen ähnlichen Anteil. Die Probanden 6, 10 und 17 zeigten im Vergleich zu den anderen Probanden höhere HNR Werte. Da ein größerer HNR Wert aber für eine gute Stimmqualität und geringen Rauschanteil spricht, werden diese Probanden nicht weiter diskutiert. Da ein Großteil der Aufnahmen von Proband 9 außerhalb der unteren Standardabweichung liegen, wird dieser gesondert betrachtet. Der Median liegt bei diesem Probanden etwa zwischen 10 - 12 dB. Der niedrige HNR im Vergleich zu den anderen Probanden spricht für einen größeren Rauschanteil

während der Aufnahmen. Eine mögliche Erklärung für dieses Geräuschphänomen könnte eine laute Umgebung bei den Aufnahmen sein. Da jedoch das Verteilungsmuster des HNR bei diesem Probanden gleichmäßig niedrig verteilt ist, kann auch angenommen werden, dass dies der individuelle HNR des Probanden ist.

Der HNR anderer Studien reicht von >11 dB [73] bis hin zu Mittelwerten bei $11,9 \pm 2,32$ dB [121], $15 \pm 1,26$ dB [4] und $24 \pm 4,3$ dB [106]. Der Mittelwert $15 \pm 3,2$ dB aus dieser Studie reiht sich in die Ergebnisse ein. Eine brasilianische Studie mit einem vergleichbar jungen Probandenkollektiv, bei dessen Analyse Praat verwendet wurde, ergab einen MW von $17,8 \pm 3,9$ dB [74]. Eine etwas größere Probandengruppe in der Studie von Teixeira et al. 2014, die ebenfalls Praat verwendet haben ergab einen Mittelwert von $24 \pm 4,3$ dB. Die Ähnlichkeit zu vergleichbaren Daten aus der Literatur spricht dafür, dass der Anteil an Hintergrundgeräuschen in dieser Studie ähnlich groß war, obwohl die Datenerhebung nicht in der Klinik, sondern mittels telemedizinischem Monitoring erhoben wurde. Die Schlussfolgerung daraus ist, dass es im Rahmen des Monitorings mit Handy zu keinem erhöhten Auftreten von Hintergrundgeräuschen kommt. Aufgrund seines indirekten Verhältnisses zum Jitter, kann der HNR auch Aussage über die Stabilität eines Signales treffen. Da die Analyse des Jitters keine hohen Frequenzschwankungen zeigte und der HNR Wert vergleichbar mit anderen Studien war, kann angenommen werden, dass die Signalübertragung der Handyaufnahmen in dieser Studie eine gute Stabilität aufweist. Da der HNR ein objektives Maß zur Einstufung von Heiserkeit ist und in dieser Studie als unauffällig betrachtet werden kann, kann das Vorliegen von Heiserkeit in einem Großteil der Aufnahmen in dieser Studie ausgeschlossen werden [121].

Der reziproke Wert des HNR, der NHR, wurde ebenfalls in dieser Studie herangezogen, um das Rauschen in den Aufnahmen zu beurteilen. Der Mittelwert des NHR in dieser Studie liegt bei $0,062 \pm 0,05$ dB. Die Standardabweichung ist klein und zeigt, dass ein Großteil der Aufnahmen einen niedrigen NHR Wert besitzt und daher auch das Hintergrundrauschen der Aufnahmen gering ist. Lediglich Proband 4 und 9 zeigen im Durchschnitt höher NHR Werte als das Gesamtkollektiv. Dies spricht dafür, dass die Stimme der Probanden von einem erhöhten Rauschanteil geprägt ist. In der Literatur wurde für den NHR, der bei gesunden männlichen Probanden mit „Praat“ ermittelt wurde, ein Mittelwert von $0,04 \pm 0,01$ dB [64] und $0,02 \pm 0,04$ dB [87] angegeben. In einem ähnlichen Rahmen verteilten sich NHR Werte eines Probandenkollektiv im Alter von 40 – 65 Jahren, das Minimum für den NHR lag bei $0,01$ dB, das Maximum bei $0,2$ dB [5]. Ein Vergleich der Ergebnisse des NHR aus der Studie mit anderen NHR Werten aus der Literatur zeigt große Übereinstimmungen. Die NHR Werte dieser Studie sind zu vergleichbaren Werten aus der Literatur nur geringfügig erhöht. Dies spricht wiederum für eine gute Aufnahmequalität, trotz Verwendung von Smartphone Mikrofonen. Der Vergleich des NHR mit den Werten aus der Studie von Grillo et al. 2016 ist

interessant, da hier die Werte ebenfalls durch telemedizinisches Monitoring erhoben wurden. In dieser Studie werden verschiedene Signalverarbeitungsparameter verglichen, die durch die Verwendung unterschiedlicher Smartphones und Bearbeitungsprogramme ermittelt wurden. Hierbei ergaben sich unter Verwendung von „Praat“ NHR Werte von $0,1 \pm 0,12 \text{ dB}$ für das iPhone 5 und $0,1 \pm 0,24 \text{ dB}$ für das Samsung Galaxy 5 [36]. Es zeigte sich, dass unter Verwendung von Smartphones der NHR etwas höher ist, als erhobene NHR Werte in der Klinik. Dadurch lassen sich abschließend auch die erhöhten NHR Werte von Proband 4 und 9 erklären. Die NHR Werte sind im Vergleich zum Kollektiv etwas erhöht und somit folglich auch das Hintergrundrauschen, jedoch lässt der leicht erhöhte Rauschanteil nicht sofort auf eine schlechte Qualität bei den Aufnahmen schließen. Der Mittelwert des NHR der in dieser Studie mit einem HUAWEI Handy aufgezeichnet wurde, liegt unterhalb der Mittelwerte, die mit anderen Smartphones ermittelt wurden. Dies spricht dafür, dass der Rauschanteil bei diesem Handy niedriger und die Qualität der Handyaufnahmen etwas besser ist als bei anderen Smartphone Modellen.

Die Verteilung des NNE hingegen ist in dieser Studie deutlich unregelmäßiger verteilt als der HNR und NHR. Zunächst fällt in Anbetracht des Mittelwertes und der Standardabweichung $-3,4 \pm 3,2 \text{ dB}$ eine große Streuung auf. Im Vergleich zu den anderen Werten ist hier die Standardabweichung sehr groß und beträgt fast 100% des Mittelwertes. Außerdem gibt es deutlich mehr Ausreißer in der NNE Auswertung (vgl. Abbildung 45 b). Ein Vergleich mit Werten aus der Literatur zeigt, dass der durchschnittliche NNE in dieser Studie größer ist, als bereits in anderen Studien beschrieben. Kasuya et al. 1986 gab für gesunde Probanden einen NNE Wert von etwa -18 bis -10 dB an [51]. Eine weitere Studie mit etwas älteren männlichen Probanden, deren durchschnittliches Alter bei 51 Jahren lag, ergab für den NNE einen Median von $-8,3 \text{ dB}$ [117]. Der NNE von 50 % der Probanden lag bei Werth et al. 2010 zwischen $-5,2 \text{ dB}$ und $-12,3 \text{ dB}$. Der Median in dieser Studie lag etwa zwischen -2 dB und -3 dB . Der NNE ist deutlicher höher als in der bereits bestehenden Literatur. Aufgrund dieser Unterschiede ist die Bestimmung des NNE im telemedizinischen Monitoring zu diskutieren. Da der NNE in dieser Studie zudem eine große Streuung aufweist und daher der Informationsgehalt über den Rauschanteil in einer Aufnahme sehr unterschiedlich ist, sind in dieser Studie zur Beurteilung des Hintergrundrauschen der HNR und NHR deutlich aussagekräftiger als der NNE. Die Stellung des NNE als wichtigster Rauschparameter in der bisherigen Stimmdiagnostik konnte in dieser Studie nicht auf das ambulante Stimmmonitoring übertragen werden [116].

5.3.3 Spektrale Parameter

Insgesamt zeigen die spektralen Parameter eine größere Streuung in ihren Einzelwerten. Die Verteilung der gesamten Daten ist jedoch regelmäßig. Eine Übereinstimmung mit bereits existierenden Studienwerten liegt vor.

Der H_1H_2 ist, in Anbetracht aller Aufnahmen, um den Mittelwert $-0,68 \pm 4,2 \text{ dB}$ verteilt. Auffallend bei diesem Parameter ist die große Wertespanne von -10 dB bis 10 dB , in der sich 98% der Aufnahmen verteilen. Die Verteilung von 50 % der Aufnahmen liegt hingegen in einem deutlich kleineren Bereich etwa zwischen -3 dB und 3 dB . Die Verteilung des H_1H_2 bei einzelnen Probanden über den gesamten Untersuchungszeitraum ist sehr variabel. Der erniedrigte H_1H_2 bei Proband 14 kann durch eine erhöhte nasale Phonation erklärt werden. In einer Studie konnte eine Veränderung des H_1H_2 bei nasaler Phonation gezeigt werden [98]. Die erhöhten H_1H_2 Werte von Proband 1 und Proband 8 würden in Anbetracht des deutlich kleineren Mittelwerts für eine vermehrte Behauchtheit in den Aufnahmen sprechen. Der RBH – Test war jedoch zu Beginn und am Ende der Studie bei beiden Probanden unauffällig. Die hohen H_1H_2 lassen sich daher am ehesten durch das Aufnahmeverfahren erklären. Für diese Theorie spricht der Median der Studie (etwa -1 dB), der dem Median eines gleichaltrigen, Probandenkollektivs aus der Studie von Simpson (Median $-2,5 \text{ dB}$) ähnlich war und die vergleichsweise sehr große Wertespanne von -10 dB bis 10 dB , die bei Simpson nur zwischen -5 dB bis 0 dB lag [98]. Die Schlussfolgerung hieraus wäre, dass der H_1H_2 im telemedizinischen Monitoring anfälliger gegenüber Schwankungen ist als die Referenzwerte die in einem stillen Raum aufgenommen wurden. Die große Streuung der erhobenen H_1H_2 Werte im telemedizinischen Monitoring zeigt sich auch in der großen Standardabweichung des Mittelwerts dieser Studie ($-0,68 \pm 4,2 \text{ dB}$). Die Standardabweichung ist in einem gleichaltrigen, brasilianischem Probandenkollektiv $-0,3 \pm 1,6 \text{ dB}$ deutlich kleiner [74]. Der Mittelwert hingegen ist nahezu identisch.

Zudem weist der H_1H_2 eine weitere Problematik auf. Feste Normdaten können für diesen Wert nicht festgelegt werden, da jede untersuchte Person individuelle Werte aufweist [34] und der Wert zudem vom Sprecher leicht variiert werden kann [58]. Anhand dieser Überlegungen lassen sich ebenfalls die abweichenden H_1H_2 Werte erklären. Proband 8 beispielsweise hat grundsätzlich einen höheren H_1H_2 als das Probandenkollektiv, welcher hier jedoch für eine gute Stimmqualität spricht. Hätte hingegen Proband 14 einen vergleichbaren hohen H_1H_2 , müsste dies als pathologische Qualität gewertet werden, da der individuelle H_1H_2 von Proband 14 in der Studie deutlich unterhalb des Mittelwertes des Gesamtkollektivs liegt. Der H_1H_2 ist kein fixer Wert, sondern befindet sich in einem bestimmten Normbereich [50]. Da sich der H_1H_2 in dieser Studie trotz großer Wertespanne und Standardabweichung in einem vergleichbaren Bereich von Referenzwerten aus anderen Studien befindet, kann durch ihn eine allgemeine Bewertung der Stimmqualität im telemedizinischen Monitoring erfolgen [56].

Die Beurteilung der Stimmqualität anhand der spektralen Neigung erfolgt durch die Bestimmung von Slope und Tilt. Der Mittelwert des Slopes lag in dieser Studie bei $-17 \pm 3,5 \text{ dB}$, der Mittelwert des Tilts bei $-8 \pm 1,5 \text{ dB}$. Beide Parameter wiesen in dieser Studie eine geringe Streuung um den Mittelwert auf. Betrachtet man die Verteilung einzelner Probanden

innerhalb des Untersuchungszeitraumes, so zeigen Tilt und Slope bei jedem Probanden einen individuellen Verteilungstrend mit mäßiger Streuung. Dieser Verteilungstrend weicht bei einzelnen Probanden vom Mittelwert ab.

Erwähnenswert ist zudem die Abweichung der Werte im Zeitintervall zwischen 5-10 Uhr, hierbei sind vor allem die Slope Werte, die oberhalb des Medians liegen, zu betrachten, da diese für eine behauchte Stimme und somit schlechte Stimmqualität sprechen würden. Retrospektiv ist dies am ehesten der Aufnahmesituation in der Früh geschuldet, da die Probanden gebeten wurden direkt nach dem Aufstehen die ersten Aufnahmen aufzuzeichnen und die Stimme zu diesem Zeitpunkt noch nicht ganz einsatzfähig war. Die regelmäßige Verteilung um den Median zu den anderen Uhrzeiten spricht ebenfalls für diese Erklärung. Bereits beschriebene Referenzwerte aus anderen Studien zeigen ähnlich Werte. Eine Studie mit 23 gesunden Männern und Frauen im durchschnittlichen Alter von 24 Jahren ergab für den Slope $-23 \pm 5 \text{ dB}$ und einen Tilt von $-10,5 \pm 0,7 \text{ dB}$. Die Variabilität der Streuung der Tilt Werte ist in dieser Studie wesentlich geringer als die der Slope Werte.

Zusammenfassend sprechen die Ergebnisse dafür, dass mit Tilt und Slope eine gültige Aussage über die Stimmqualität mittels Stimmmonitoring getroffen werden kann und die Qualität der Aufnahmen in dieser Studie als gut bewertet werden kann. Zu berücksichtigen ist, dass beide Werte eine gewisse Variabilität aufweisen und anhand von Einzelwerten keine valide Schlussfolgerung erfolgen darf, sondern nur in Anbetracht aller Daten [50].

Der Mittelwert des CPPs in dieser Studie liegt bei $12 \pm 2,1 \text{ dB}$. Die Standardabweichung ist sehr klein und spricht für eine geringe Streuung des Parameters. Auch in Anbetracht der Verteilung der CPPs Werte bei allen Probanden zeigt sich ebenfalls eine geringere Variabilität in der Verteilung, als bei den bereits diskutierten spektralen Parametern Tilt, Slope und H_1H_2 . Zudem sind Median und Mittelwert sehr ähnlich. Die Zahl der Ausreißer bei diesem Parameter ist zudem sehr gering. Vergleichbare Studien ergaben etwas kleinere CPPs Mittelwerte. In einer Studie mit 23 gesunden Männern und Frauen ergab sich ein Mittelwert von $8 \pm 0,9 \text{ dB}$ [64]. Eine etwas größeren Studie mit über 50 gesunden Freiwilligen ergab einen CPPs Mittelwert von $4,7 \pm 0,9 \text{ dB}$ [41]. Dies zeigt, dass die CPPs Werte aus dieser Studie mit Werten aus anderen Studien vergleichbar sind. Die Wertespanne in der Studie ist im Vergleich zur Literatur etwas umfangreicher und erreicht höhere Werte. Da höhere Werte per definitionem für eine gute Stimmqualität sprechen und erst niedrige CPPs Werte für eine Beeinträchtigung in der Qualität sprechen würden, wurde der höhere CPPs in dieser Studie akzeptiert. Da eine objektive Beurteilung der Stimme mit dem CPPs möglich ist [21] und bereits Einzelmessung Hinweise auf Pathologien geben können [88], ist der CPPs im telemedizinischen Monitoring ein sehr guter Parameter um die Qualität der Aufnahmen zu beurteilen. Aufgrund seiner geringen Streuung in dieser Studie sollte er im telemedizinischen Monitoring gegenüber anderen spektralen Analyseparametern bevorzugt verwendet werden.

5.3.4 Verteilung aller Werte in t-SNE Grafik

Mit Hilfe des t-SNE Verfahrens kann man anhand von Stimmclustern einen schnellen Überblick über die Verteilung von Werten erhalten und anhand von unregelmäßigen Verteilungsmustern Unregelmäßigkeiten in Stimmen innerhalb eines Gesamtkollektivs erkennen. Bis auf den NNE sind die restlichen Parameter im großen und kleinen Cluster sehr ähnlich verteilt, wie beispielhaft am Jitter gezeigt wurde. Die Verteilung der Werte im kleinen Cluster repräsentiert daher ihre Verteilung im großen Cluster. Nach Dimensionsreduktion und Ausschluss des NNE verteilen sich alle restlichen Parameter in einem großen Cluster. Daher scheint die gewisse Streuung, die bei einzelnen Parametern im Ergebnissteil ersichtlich war, im Gesamtzusammenhang keinen relevanten Einfluss zu haben. Zudem zeigt sich, dass alle Parameter der einzelnen Probanden jeweils in einem individuellen Bereich verteilt sind. Daraus kann geschlossen werden, dass die einzelnen Parameter der jeweiligen Probanden in sich ähnlich sind und in einer gewissen Beziehung zu einander stehen. Eine diskrete Streuung der Werte innerhalb des individuellen Bereichs, kann durch die Variabilität einer gesunden Stimme erklärt werden. Der NNE scheint anfällig gegenüber Störfaktoren zu sein, da die Ergebnisse dieses Wertes aus dem telemedizinischen Monitoring in ihrer Verteilung untereinander nicht übereinstimmen und auch von den restlichen Werten abweichen. Daher sollte der NNE aus dieser Studie ausgeschlossen werden. Obwohl er als wichtigster Diagnostikparameter für Rauschen gilt [116], scheint er für die Beurteilung der Stimmqualität im ambulanten Monitoring nicht geeignet.

In Anbetracht aller Ergebnisse sind die Parameter HNR, NHR, Jitter & Shimmer, sowie H_1H_2 , Tilt&Slope und CPPs geeignete Parameter im telemedizinischen Monitoring zur Beurteilung von Stimmaufnahmen von Stimmgesunden. Des Weiteren waren die Werte der Parameter mit bereits beschriebenen Referenzwerten aus der bisherigen Literatur vergleichbar und stimmten überein. Damit scheint das telemedizinische Monitoring einen ähnlichen Informationsgehalt über die Stimme und über ihre Qualität zu ergeben, wie herkömmliche klinische, nicht mobile Untersuchungsmaßnahmen. Dies spricht dafür, dass das telemedizinische Monitoring ein geeignetes mobiles Aufnahmesystem der Stimme ist und es eine valide Beurteilung der Stimmqualität ermöglicht.

6 Anhang

VHI-12 (vom Patient auszufüllen)

Ich brauche meine <i>Sprechstimme</i> vorwiegend für	<input type="checkbox"/> meinen Beruf <input type="checkbox"/> Freizeitaktivitäten <input type="checkbox"/> normale tägliche Unterhaltungen
Ich brauche meine <i>Singstimme</i> vorwiegend für	<input type="checkbox"/> meinen Beruf <input type="checkbox"/> Freizeitaktivitäten

Ich schätze meine Gesprächigkeit so ein (bitte ankreuzen)									
1	2	3	4	5	6	7	8	9	10
stiller Zuhörer			normaler Sprecher				äußerst gesprächig		

Bitte kreuzen Sie an, was für Sie zutrifft: 0 = nie, 1 = selten, 2 = manchmal, 3 = oft, 4 = immer

1	Bevor ich spreche, weiß ich nicht, wie meine Stimme klingen wird.	0	1	2	3	4
2	Abends ist meine Stimme schlechter.	0	1	2	3	4
3	Ich habe das Gefühl, dass ich mich anstrengen muss, wenn ich meine Stimme benutze.	0	1	2	3	4
4	Wegen meines Stimmproblems bin ich weniger kontaktfreudig.	0	1	2	3	4
5	Ich meide größere Gruppen wegen meiner Stimme.	0	1	2	3	4
6	Ich fühle mich bei Unterhaltungen wegen meiner Stimme ausgeschlossen	0	1	2	3	4
7	Anderen fällt es schwer, mich in einer lauten Umgebung zu verstehen	0	1	2	3	4
8	Meine Familie hat Schwierigkeiten, mich zu hören, wenn ich zuhause nach ihnen rufe.	0	1	2	3	4
9	Man hört mich wegen meiner Stimme schlecht.	0	1	2	3	4
10	Es ist mir peinlich, wenn man mich bittet, etwas zu wiederholen	0	1	2	3	4
11	Ich ärgere mich, wenn man mich bittet, etwas zu wiederholen.	0	1	2	3	4
12	Ich schäme mich wegen meines Stimmproblems.	0	1	2	3	4

Wie schätzen Sie Ihre Stimme heute ein?	0	1	2	3
0=normal, 1=leicht gestört, 2=mittelgradig gestört, 3=hochgradig gestört				

Abbildung 61: Verwendeter VHI-12 Test aus dem Studienprotokoll.

Stroboskopische Untersuchung (vom Untersucher auszufüllen)

Allgemeine Informationen

Endoskop	<input type="checkbox"/> 70°	<input type="checkbox"/> 90°	<input type="checkbox"/> Anderer Winkel:	_____
Position des Endoskops bezüglich der Zunge	<input type="checkbox"/> links vorbei	<input type="checkbox"/> rechts vorbei	<input type="checkbox"/> darüber	<input type="checkbox"/> unbekannt
Stimmgrundfrequenz	_____ Hz			

Subjektive Bewertung der stroboskopischen Aufnahme

Randkantenverschiebung	<input type="checkbox"/> keine	<input type="checkbox"/> reduziert	<input type="checkbox"/> normal	<input type="checkbox"/> erhöht		
Amplitude links	<input type="checkbox"/> keine	<input type="checkbox"/> reduziert	<input type="checkbox"/> normal	<input type="checkbox"/> erhöht		
Amplitude rechts	<input type="checkbox"/> keine	<input type="checkbox"/> reduziert	<input type="checkbox"/> normal	<input type="checkbox"/> erhöht		
Phasen-Symmetrie	<input type="checkbox"/> keine	<input type="checkbox"/> vorhanden				
Glottale Schlussinsuff.	<input type="checkbox"/> keine	<input type="checkbox"/> anterior	<input type="checkbox"/> oval	<input type="checkbox"/> posterior	<input type="checkbox"/> hour-glass	
Glottisschluss (gem. ELS)	<input type="checkbox"/> longitudinal	<input type="checkbox"/> dorsal	<input type="checkbox"/> ventral	<input type="checkbox"/> oval	<input type="checkbox"/> hour-glass	<input type="checkbox"/> irregulär
Petiolus - Anteil der ver-	<input type="checkbox"/> 0%	<input type="checkbox"/> ≤25%	<input type="checkbox"/> ≤50%	<input type="checkbox"/> ≤75%	<input type="checkbox"/> ≥75%	
	(Keine Verdeckung)					
Taschenfalteneinsatz						
Links	<input type="checkbox"/> kein	<input type="checkbox"/> mittel	<input type="checkbox"/> viel	<input type="checkbox"/> Taschenfaltenstimme		
A.P - Verkürzung	<input type="checkbox"/> ja	<input type="checkbox"/> nein				

Abbildung 62: Untersuchungsprotokoll zur Beurteilung der Stroboskopie Aufnahmen aus dem Studienprotokoll.

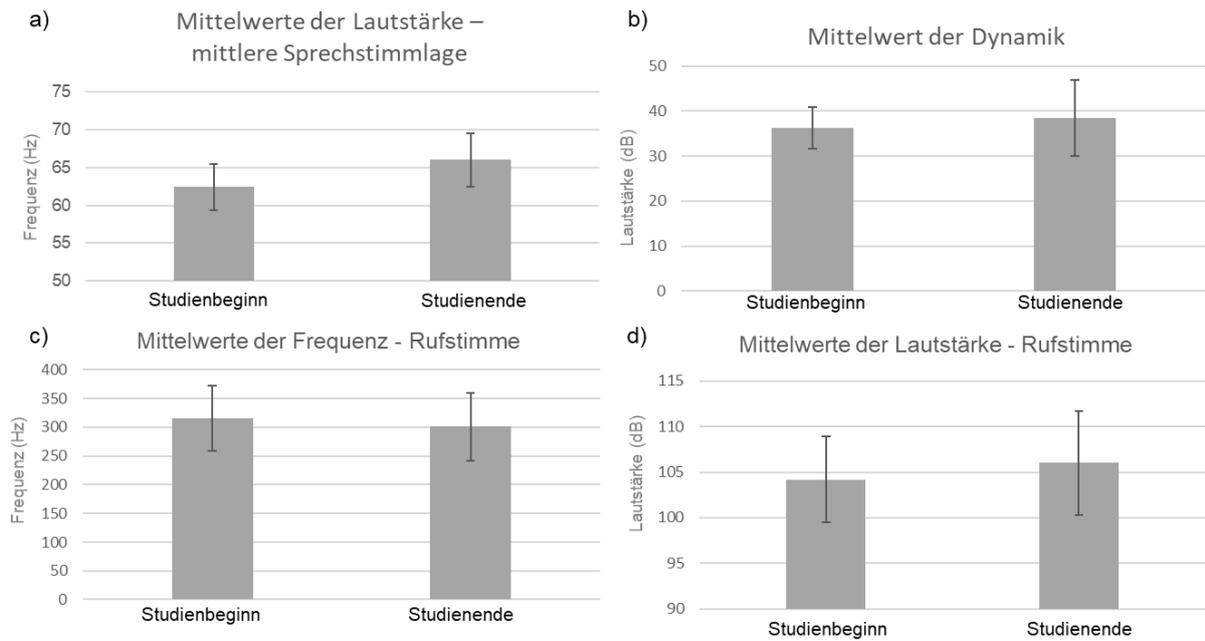


Abbildung 63: Vervollständigung der Daten des Sprechstimmfeldes a) Mittlere Lautstärke der mittleren Sprechstimmlage von 17 Probanden. Da der Frequenzbereich klein gewählt ist, erscheint die SD und die Abweichung der Mittelwerte zu Beginn und am Ende der Studie groß. b) Die Mittelwerte der Dynamik zu den beiden Zeitpunkten sind sehr ähnlich. Bei Ausschluss aus der Studie ist die SD etwas höher. c) Die beiden Mittelwerte der Rufstimmfrequenz sind zu beiden Zeitpunkten ähnlich und die SD ist ähnlich verteilt. d) Ebenso verhält sich die Lautstärke der Rufstimme. Die Mittelwerte sind minimal unterschiedlich und die SD am Ende etwas größer.

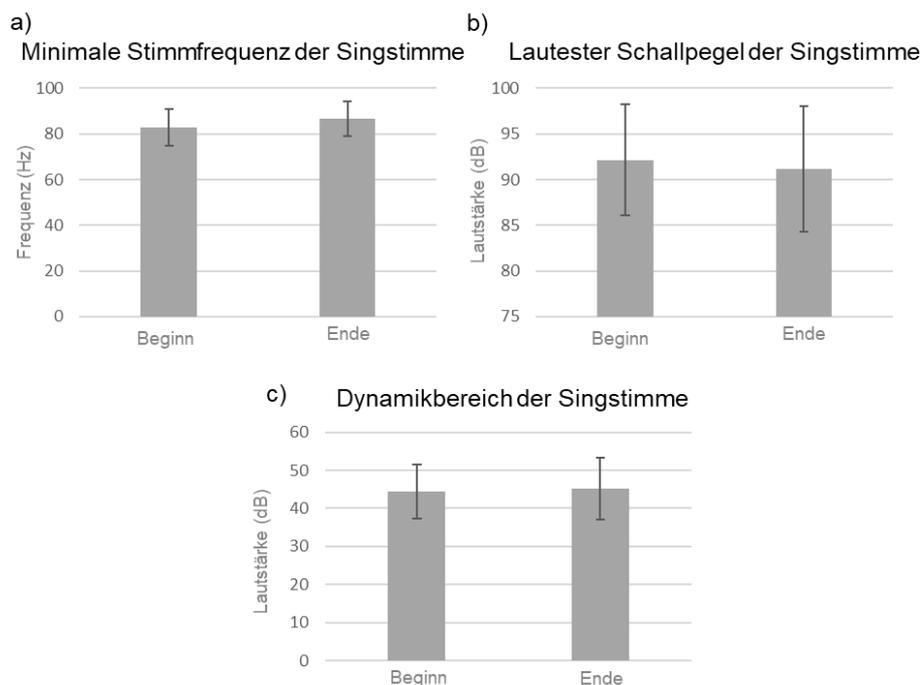


Abbildung 64: Vervollständigung der Daten des Singstimmfeldes a) Minimale Stimmfrequenz der Singstimme im Vergleich zu Beginn und am Ende der Studie. Die Daten aller Probanden liegen in einem sehr ähnlichen Bereich und die SD des Wertes ist sehr gering. b) Die Verteilung des lautesten Schallpegels der Singstimme ist in dieser Studie relativ weit. c) Die Verteilung des Dynamikbereichs der Singstimme ist nahezu identisch zu Beginn und am Ende der Studie.

Shimmer von Proband 9

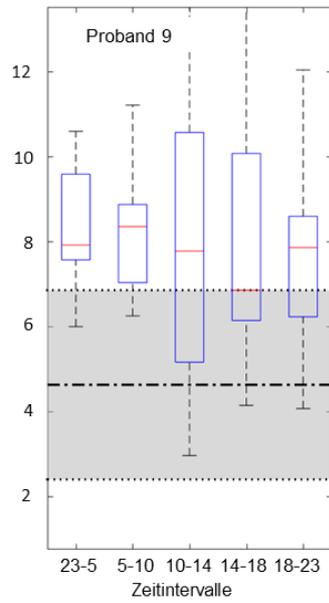


Abbildung 65: Verteilung des Shimmers von Proband 9 im Vergleich zum Mittelwert und Standardabweichung des Gesamtkollektivs ($4,6 \pm 2,2 \%$).

HNR einzelner Probanden im Vergleich zum Mittelwert des Gesamtkollektivs

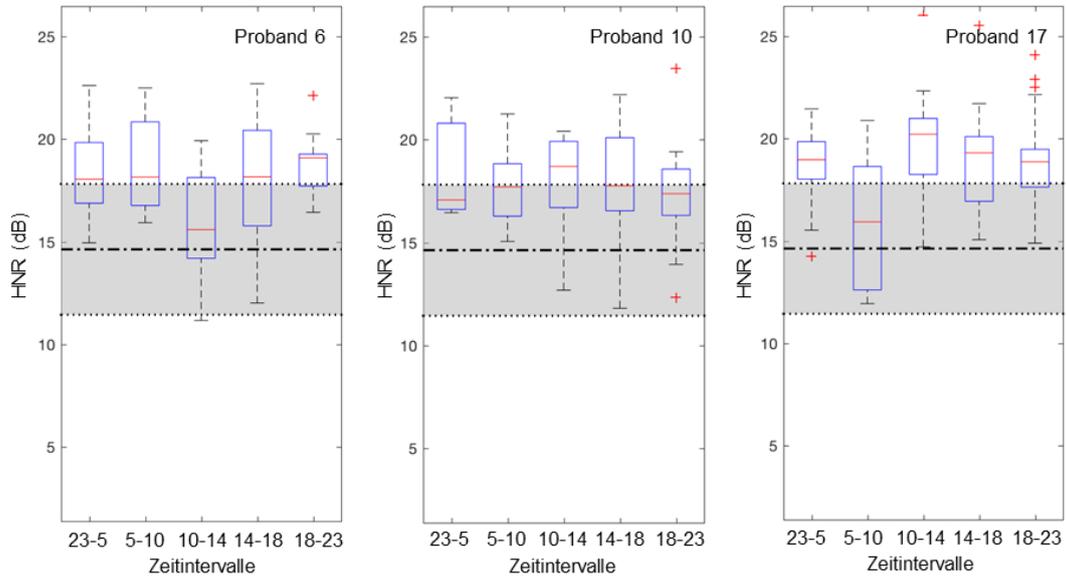
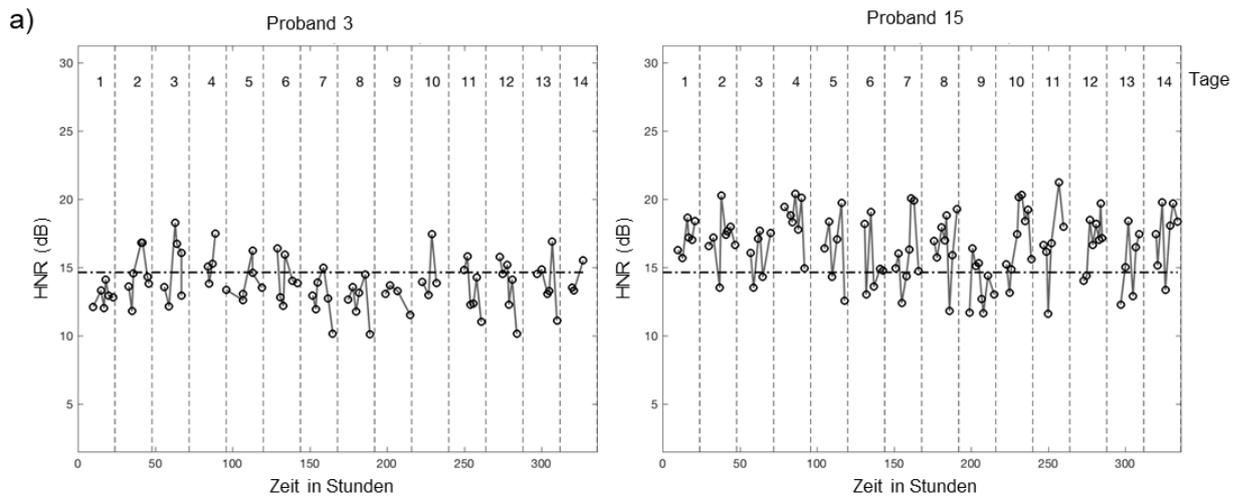


Abbildung 66: Hochnormale HNR-Werte einzelner Probanden im Vergleich zum Mittelwert des Gesamtkollektivs ($15 \pm 3,2 \text{ dB}$).

Entwicklung des HNR im Studienverlauf



Entwicklung des NHR im Studienverlauf

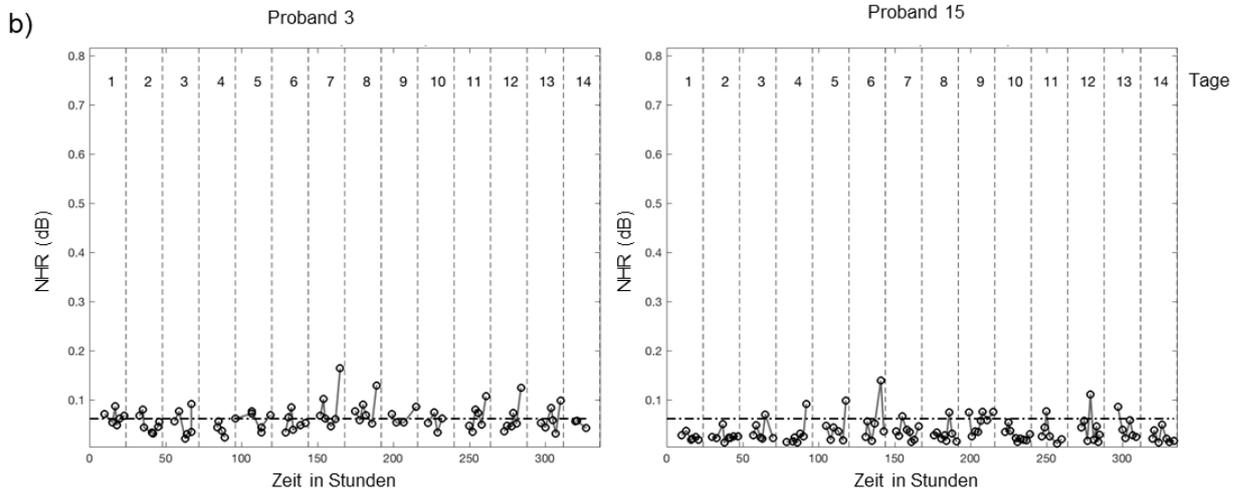


Abbildung 67: a) Entwicklung des HNR im Studienverlauf von Probant 3 und Probant 15. b) Entwicklung des NHR im Studienverlauf von Probant 3 und Probant 15.

H1H2 von Proband 8 im Vergleich zum
Mittelwert des Gesamtkollektivs

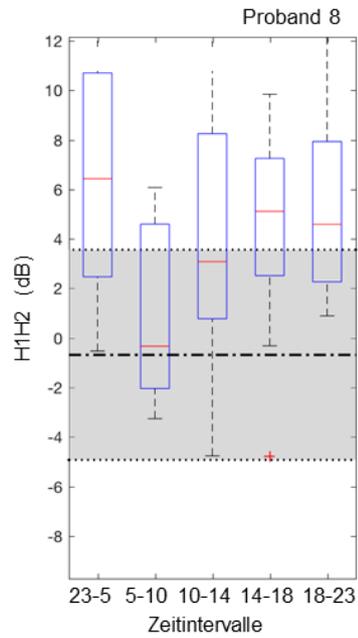
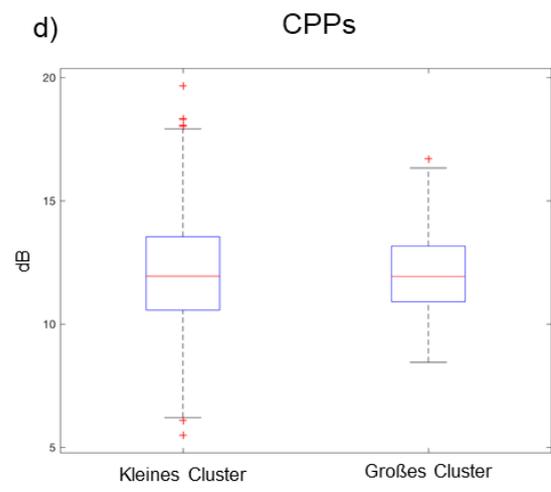
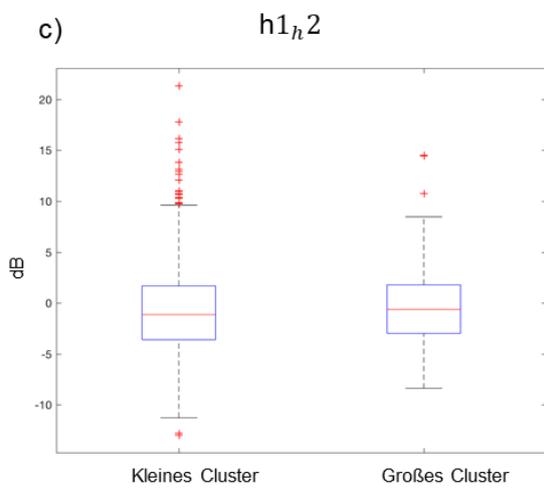
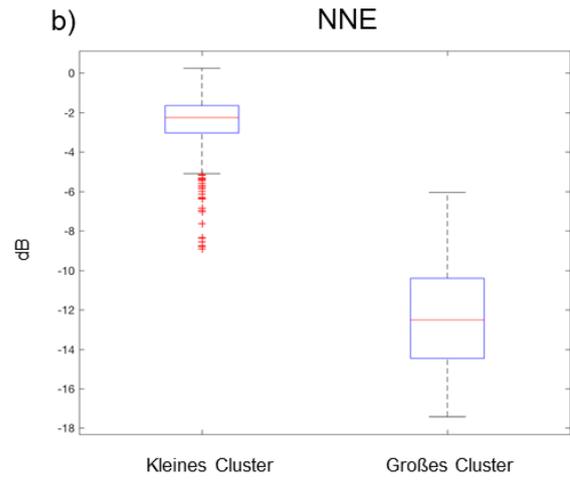
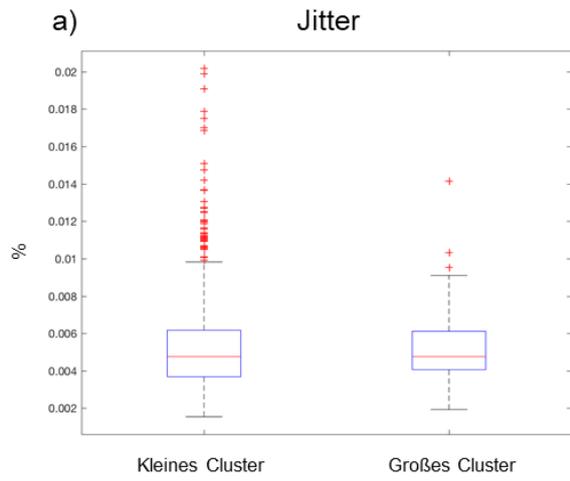


Abbildung 68: Abweichende Verteilung des H1H2 von Proband 8 im Vergleich zum Mittelwert und der Standardabweichung des Gesamtkollektivs ($-0,68 \pm 4,2$ dB).



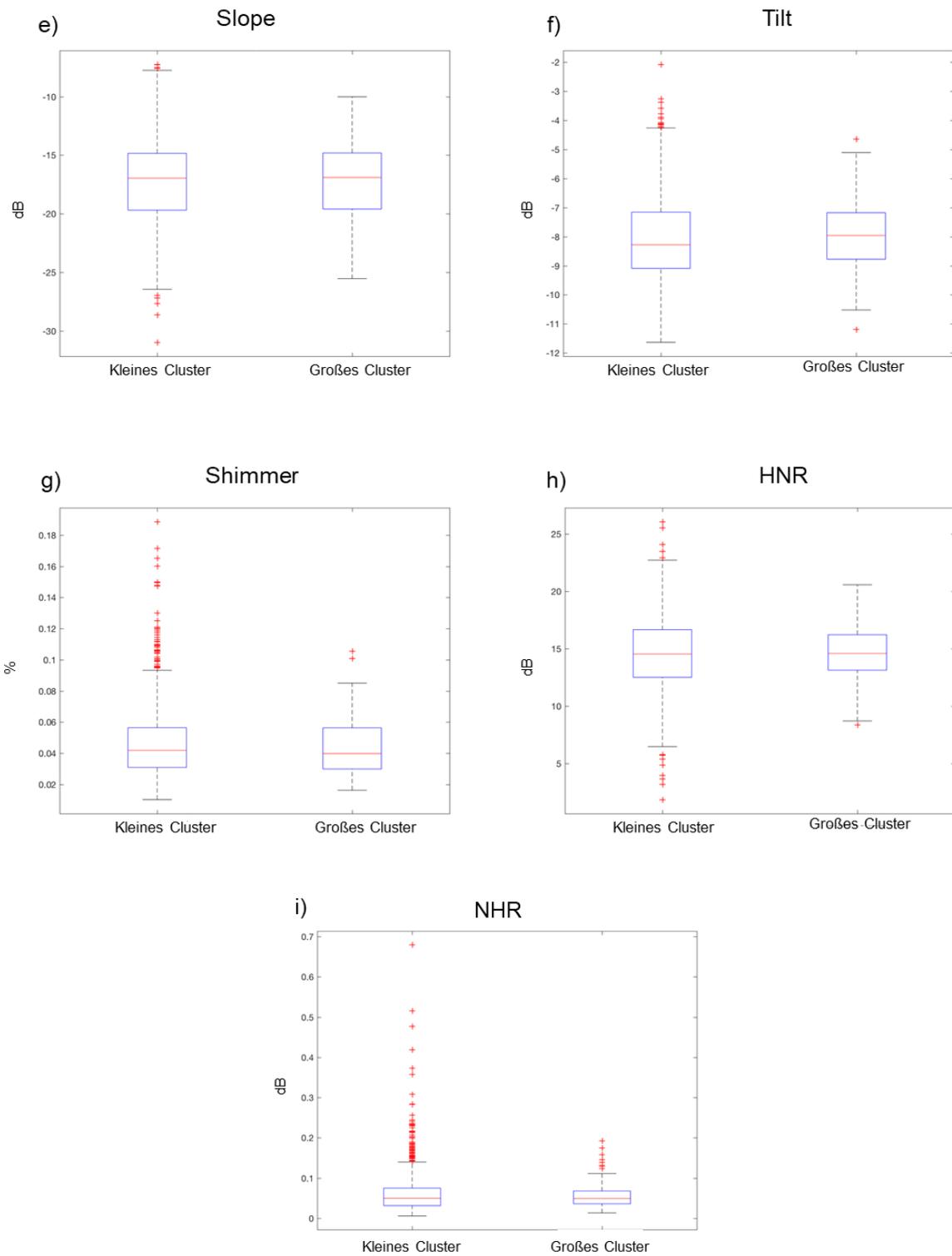


Abbildung 69: Verteilung der Parameter aus dem telemedizinischen Monitoring im kleinen und großen Cluster.

7 Literaturverzeichnis

- [1] Abdelli-Beruh Nassima, Thomas Drugman, R. H. Red Owl. 2015. Occurrence Frequencies of Acoustic Patterns of Vocal Fry in American English Speakers.
- [2] Aichinger, P., Feichter, F., Aichstill, B., Bigenzahn, W., and Schneider-Stickler, B. 2012. Inter-device reliability of DSI measurement. *Logopedics, phoniatrics, vocology* 37, 4, 167–173.
- [3] Aumüller, G. 2014. *Anatomie*. Duale Reihe. Thieme, Stuttgart.
- [4] Awan, S. N. and Frenkel, M. L. 1994. Improvements in estimating the harmonics-to-noise ratio of the voice. *Journal of Voice* 8, 3, 255–262.
- [5] Batra, Bhasin, Singh. 2015. Acoustic Analysis of voice samples to differentiate Healthy and Asthmatic persons Acoustic Analysis of voice samples to differentiate Healthy and Asthmatic persons. *International Journal Of Engineering And Computer Science*, Volume 4, 13161–13164.
- [6] Baylor, C. R., Yorkston, K. M., and Eadie, T. L. 2005. The consequences of spasmodic dysphonia on communication-related quality of life. A qualitative study of the insider's experiences. *Journal of communication disorders* 38, 5, 395–419.
- [7] Boenninghaus, H.-G. and Lenarz, T. 2007. *Hals-Nasen-Ohren-Heilkunde*. Springer-Lehrbuch. Springer Medizin Verlag Heidelberg, Berlin, Heidelberg.
- [8] Bohlender. 2013. Diagnostic and therapeutic pitfalls in benign vocal fold diseases. *GMS Current Topics in Otorhinolaryngology - Head and Neck Surgery* 2013, Vol.12.
- [9] Brockmann, M., Decoster, W., and Riedmüller, S. 2010. *Relevanz des Stimmfeldes für Diagnostik und Therapieevaluation*. Schulz-Kirchner.
- [10] Brockmann, M., Drinnan, M. J., Storck, C., and Carding, P. N. 2011. Reliable jitter and shimmer measurements in voice clinics. The relevance of vowel, gender, vocal intensity, and fundamental frequency effects in a typical clinical task. *Journal of voice : official journal of the Voice Foundation* 25, 1, 44–53.
- [11] Brockmann-Bauser, M. and Bohlender, J. 2016. Wie aussagekräftig sind die instrumentellen akustischen Irregularitätsparameter Jitter und Shimmer in der Stimmdiagnostik? *Sprache Stimme Gehör* 40, 04, 179–182.
- [12] Brockmann-Bauser, M. and Bohlender, J. E. 2014. *Praktische Stimmdiagnostik. Theoretischer und praktischer Leitfaden*. Georg Thieme Verlag KG, s.l.
- [13] Cantor Cutiva, C., Vogel, I., and Burdorf, A. 2013. Voice disorders in teachers and their associations with work-related factors. A systematic review. *Journal of communication disorders* 46, 2, 143–155.
- [14] Carding, P. N., Wilson, J. A., MacKenzie, K., and Deary, I. J. 2009. Measuring voice outcomes. State of the science review. *The Journal of laryngology and otology* 123, 8, 823–829.
- [15] Claudia Wüstenhagen. 2013. Der Klang der Seele. *Stimmforschung. Zeit Wissen*, 5.
- [16] Cohen, S. M. 2010. Self-reported impact of dysphonia in a primary care population. An epidemiological study. *The Laryngoscope* 120, 10, 2022–2032.
- [17] D. Michaelis, T. Gramss, and H.W. Strube. 1996. Glottal-to-Noise Excitation Ratio – a New Measure for Describing Pathological Voices. *Acustica* Vol.83.
- [18] Dejonckere, P. H., Bradley, P., Clemente, P., Cornut, G., Crevier-Buchman, L., Friedrich, G., van de Heyning, P., Remacle, M., and Woisard, V. 2001. A basic protocol for functional assessment of voice pathology, especially for investigating the efficacy of (phonosurgical) treatments and evaluating new assessment techniques. Guideline elaborated by the Committee on Phoniatrics of the European Laryngological Society (ELS). *European archives of oto-rhino-laryngology : official journal of the European Federation of Oto-Rhino-Laryngological Societies (EUFOS) : affiliated with the German Society for Oto-Rhino-Laryngology - Head and Neck Surgery* 258, 2, 77–82.

- [19] Dejonckere, P. H. and Peters, H. F. M. 1999. *Communication and its disorders. A science in progress : proceedings : 24th Congress International Association of Logopedics and Phoniatrics : Amsterdam, August 23-27, 1998 : The Netherlands*. Nijmegen University Press, Nijmegen.
- [20] Dejonckere, P. H., Remacle, M., Fresnel-Elbaz, E., Woisard, V., Crevier-Buchman, L., and Millet, B. 1996. Differentiated perceptual evaluation of pathological voice quality. Reliability and correlations with acoustic measurements. *Revue de laryngologie - otologie - rhinologie* 117, 3, 219–224.
- [21] Delgado-Hernández, J., León-Gómez, N. M., Izquierdo-Arteaga, L. M., and Llanos-Fumero, Y. 2018. Análisis cepstral de la voz normal y patológica en adultos españoles. Medida de la prominencia del pico cepstral suavizado en vocales sostenidas versus habla conectada. *Acta otorrinolaringologica española* 69, 3, 134–140.
- [22] Denk, D. M., Swoboda, H., and Steiner, E. 1998. Physiologie des Larynx. *Der Radiologe* 38, 2, 63–70.
- [23] Di Nicola, V., Fiorella, M. L., Spinelli, D. A., and Fiorella, R. 2006. Acoustic analysis of voice in patients treated by reconstructive subtotal laryngectomy. Evaluation and critical review. *Acta otorhinolaryngologica Italica : organo ufficiale della Società italiana di otorinolaringologia e chirurgia cervico-facciale* 26, 2, 59–68.
- [24] Elő, J., Sebök, J., and Szilágyi, G. 1980. Adatok a Reinke-ödéma keletkezésének patomechanizmusához. *Morphologiai es igazságügyi orvosi szemle* 20, 4, 269–274.
- [25] Erwan Pépiot. 2014. Male and female speech: a study of mean f₀, f₀ range, phonation type and speech rate in Parisian French and American English speakers. *Speech Prosody* 7, Dublin, Ireland (May. 2014), pp.305-309.
- [26] Eysholdt, U., Rosanowski, F., and Hoppe, U. 2003. Vocal fold vibration irregularities caused by different types of laryngeal asymmetry. *European archives of oto-rhino-laryngology : official journal of the European Federation of Oto-Rhino-Laryngological Societies (EUFOS) : affiliated with the German Society for Oto-Rhino-Laryngology - Head and Neck Surgery* 260, 8, 412–417.
- [27] Fant, G. 1981. The source filter concept in voice production. *STL-QPSR*, 22, 21–37.
- [28] Friedrich, G., Bigenzahn, W., Zorowka, P., and Brunner, E. 2000. *Phoniatrie und Pädaudiologie. Einführung in die medizinischen, psychologischen und linguistischen Grundlagen von Stimme, Sprache und Gehör*. Gesundheitsberufe Stimm-, Sprech- und Sprachheilkunde. Huber, Bern.
- [29] Friedrich, G. and Dejonckere, P. H. 2005. Das Stimm diagnostik-Protokoll der European Laryngological Society (ELS) -- erste Erfahrungen im Rahmen einer Multizenterstudie. *Laryngo- rhino- otologie* 84, 10, 744–752.
- [30] G. Friedrich, J Kainz, W. Freidl. 1993. Zur funktionellen Struktur der menschlichen Stimmlippe. *Laryngo-Rhino-Otol.* 72.
- [31] Garrett, R. 2013. *Cepstral- and Spectral-Based Acoustic Measures of Normal Voices*.
- [32] Glunz, M., Reuß, C., Schmitz, E., and Stappert, H. 2011. *Laryngektomie. Von der Stimmlosigkeit zur Stimme*. Praxiswissen Logopädie 0. Springer-Verlag Berlin Heidelberg, Berlin, Heidelberg.
- [33] Gonnermann and Nawka. 2007. Klassifikation der Werte des VHI-12 nach Schweregraden. 24. Wissenschaftliche Jahrestagung der Deutschen Gesellschaft für Phoniatrie und Pädaudiologie.
- [34] Gordon, M. and Ladefoged, P. 2001. Phonation types. A cross-linguistic overview. *Journal of Phonetics* 29, 4, 383–406.
- [35] Gramming, P., Sundberg, J., and Akerlund, L. 1991. Variability of phonetograms. *Folia phoniatrica* 43, 2, 79–92.

- [36] Grillo, E. U., Brosious, J. N., Sorrell, S. L., and Anand, S. 2016. Influence of Smartphones and Software on Acoustic Voice Measures. *International journal of telerehabilitation* 8, 2, 9–14.
- [37] Hamdan, A. L., Al-Barazi, R., Kanaan, A., Al-Tamimi, W., Sinno, S., and Husari, A. 2012. The effect of snoring on voice. A controlled study of 30 subjects. *Ear, nose, & throat journal* 91, 1, 28–33.
- [38] Hanschmann, H., Lohmann, A., and Berger, R. 2011. Comparison of subjective assessment of voice disorders and objective voice measurement. *Folia phoniatica et logopaedica : official organ of the International Association of Logopedics and Phoniatics (IALP)* 63, 2, 83–87.
- [39] Hartmut Traunmüller, A. E. 1995. The frequency range of the voice fundamental in the speech of male and female adults. *Unpublished manuscript*.
- [40] Heman-Ackah, Y. D., Michael, D. D., and Goding, G. S. 2002. The relationship between cepstral peak prominence and selected parameters of dysphonia. *Journal of Voice* 16, 1, 20–27.
- [41] Heman-Ackah, Y. D., Sataloff, R. T., Laureyns, G., Lurie, D., Michael, D. D., Heuer, R., Rubin, A., Eller, R., Chandran, S., Abaza, M., Lyons, K., Divi, V., Lott, J., Johnson, J., and Hillenbrand, J. 2014. Quantifying the cepstral peak prominence, a measure of dysphonia. *Journal of voice : official journal of the Voice Foundation* 28, 6, 783–788.
- [42] Henton, C. G. and Bladon, R.A.W. 1985. Breathiness in normal female speech. Inefficiency versus desirability. *Language & Communication* 5, 3, 221–227.
- [43] Herold, R., van den Berg, N., Dörr, M., and Hoffmann, W. 2018. Telemedical Care and Monitoring for Patients with Chronic Heart Failure Has a Positive Effect on Survival. *Health services research* 53, 1, 532–555.
- [44] Hillenbrand, J. and Houde, R. A. 1996. Acoustic Correlates of Breathless Vocal Quality. Dysphonic Voices and Continuous Speech. *J Speech Lang Hear Res* 39, 2, 311.
- [45] Hillman, R. E., Holmberg, E. B., Perkell, J. S., Walsh, M., and Vaughan, C. 1990. Phonatory function associated with hyperfunctionally related vocal fold lesions. *Journal of Voice* 4, 1, 52–63.
- [46] Hogikyan, N. D. and Sethuraman, G. 1999. Validation of an instrument to measure voice-related quality of life (V-RQOL). *Journal of Voice* 13, 4, 557–569.
- [47] Jacobson, Johnson, Grywalski, Silbergleit, Benninger, and Newman. 1997. The Voice Handicap Index (VHI): Development and Validation. *American Journal of Speech-Language Pathology*, Vol.6.
- [48] Jannetts, S. and Lowit, A. 2014. Cepstral analysis of hypokinetic and ataxic voices. Correlations with perceptual and other acoustic measures. *Journal of voice : official journal of the Voice Foundation* 28, 6, 673–680.
- [49] Jitka Vydrova, Jan G. Svec, and František Sram. 2015. Videokymography (VKG) in Laryngologic Practice. *The Journal of MacroTrends in Health and Medicine*, Volume 3.
- [50] Jörg Mayer. 2017. *Phonetische Analysen mit Praat. Ein Handbuch für Ein- und Umsteiger*. Vorlesungsskript, Universität Stuttgart.
- [51] Kasuya, H., Ogawa, S., Mashima, K., and Ebihara, S. 1986. Normalized noise energy as an acoustic measure to evaluate pathologic voice. *The Journal of the Acoustical Society of America* 80, 5, 1329–1334.
- [52] Keating and Esposito. 2006. Linguistic Voice Quality. *UCLA Working Papers in Phonetics*, 105, 85–91.
- [53] Klatt, D. H. and Klatt, L. C. 1990. Analysis, synthesis, and perception of voice quality variations among female and male talkers. *The Journal of the Acoustical Society of America* 87, 2, 820–857.
- [54] Ko, S.-R. 2016. Die Selbsteinschätzung der Stimmfunktion. *Sprache Stimme Gehör* 40, 04, 164–167.

- [55] Kojima, H., Gould, W. J., And, A. L., and Isshiki, N. 2009. Computer Analysis of Hoarseness. *Acta Oto-Laryngologica* 89, 3-6, 547–554.
- [56] Kreiman, J. and Gerratt, B. R. 2010. Perceptual sensitivity to first harmonic amplitude in the voice source. *The Journal of the Acoustical Society of America* 128, 4, 2085–2089.
- [57] Kreiman, J., Gerratt, B. R., and Ito, M. 2007. When and why listeners disagree in voice quality assessment tasks. *The Journal of the Acoustical Society of America* 122, 4, 2354–2364.
- [58] Kreiman, J., Shue, Y.-L., Chen, G., Iseli, M., Gerratt, B. R., Neubauer, J., and Alwan, A. 2012. Variability in the relationships among voice quality, harmonic amplitudes, open quotient, and glottal area waveform shape in sustained phonation. *The Journal of the Acoustical Society of America* 132, 4, 2625–2632.
- [59] Krom, G. de. 1993. A cepstrum-based technique for determining a harmonics-to-noise ratio in speech signals. *Journal of speech and hearing research* 36, 2, 254–266.
- [60] L.J.P. van der Maaten and G.E. Hinton. 2008. Visualizing High-Dimensional Data Using t-SNE. *Journal of Machine Learning Research*, 2579–2605.
- [61] Lieberman, P. 1963. Some Acoustic Measures of the Fundamental Periodicity of Normal and Pathologic Larynges. *The Journal of the Acoustical Society of America* 35, 3, 344–353.
- [62] Ma, E. P.-M. and Yiu, E. M.-L. 2001. Voice Activity and Participation Profile. *J Speech Lang Hear Res* 44, 3, 511.
- [63] Martins, R. H. G., Tavares, E. L. M., and Pessin, A. B. B. 2017. Are Vocal Alterations Caused by Smoking in Reinke's Edema in Women Entirely Reversible After Microsurgery and Smoking Cessation? *Journal of voice : official journal of the Voice Foundation* 31, 3, 380.e11-380.e14.
- [64] Maryn, Y. 2010. *Acoustic measurement of overall voice quality in sustained vowels and continuous speech. Thesis submitted in partial fulfilment of the requirements for the degree of Doctor in Health Sciences, Gent, Universiteit Gent.*
- [65] Maryn, Y., Bodt, M. de, Barsties, B., and Roy, N. 2014. The value of the acoustic voice quality index as a measure of dysphonia severity in subjects speaking different languages. *European archives of oto-rhino-laryngology : official journal of the European Federation of Oto-Rhino-Laryngological Societies (EUFOS) : affiliated with the German Society for Oto-Rhino-Laryngology - Head and Neck Surgery* 271, 6, 1609–1619.
- [66] Maryn, Y., Corthals, P., Bodt, M. de, van Cauwenberge, P., and Deliyiski, D. 2009. Perturbation measures of voice. A comparative study between Multi-Dimensional Voice Program and Praat. *Folia phoniatrica et logopaedica : official organ of the International Association of Logopedics and Phoniatrics (IALP)* 61, 4, 217–226.
- [67] Mehta, D. D., van Stan, J. H., Zañartu, M., Ghassemi, M., Guttag, J. V., Espinoza, V. M., Cortés, J. P., Cheyne, H. A., and Hillman, R. E. 2015. Using Ambulatory Voice Monitoring to Investigate Common Voice Disorders. Research Update. *Frontiers in bioengineering and biotechnology* 3, 155.
- [68] Mehta, D. D., Zañartu, M., Feng, S. W., Cheyne, H. A., and Hillman, R. E. 2012. Mobile voice health monitoring using a wearable accelerometer sensor and a smartphone platform. *IEEE transactions on bio-medical engineering* 59, 11, 3090–3096.
- [69] Merati, A. L. and Bielamowicz, S. A. op. 2007. *Textbook of laryngology*. Plural Publishing, San Diego (CA).
- [70] Minnema, S. 2008. Objektive computergestützte Stimmanalyse mit "Praat". *Forum Logopädie*, 4(22), 24–29.
- [71] Mukhopadhyay, S. C., Ed. 2010. *Advances in biomedical sensing, measurements, instrumentation and systems*. Lecture notes in electrical engineering Vol. 55. Springer, Berlin, Heidelberg.

- [72] Murphy, P. J. 2001. "Spectral tilt as a perturbation-free measurement of noise levels in voice signals." *INTERSPEECH*.
- [73] Murry, T., Brown, W. S., and Rothman, H. 1987. Judgments of voice quality and preference. Acoustic interpretations. *Journal of Voice* 1, 3, 252–257.
- [74] Nascimento, C. L., Constantini, A. C., and Mourão, L. F. 2016. Vocal Effects in Military Students Submitted to an Intense Recruit Training. A Pilot Study. *Journal of voice : official journal of the Voice Foundation* 30, 1, 61–69.
- [75] Nawka, T., Belogradski, D., Caffier, P., Gross, M., Martin, A., Reinhardt, A., and Anders, L.-C., Eds. 2011. *Automatische Bestimmung der mittleren Sprechstimmlage beim Vorlesen des Textes „Nordwind und Sonne“*. German Medical Science GMS Publishing House.
- [76] Nawka, T., Verdonck-de Leeuw, I. M., Bodt, M. de, Guimaraes, I., Holmberg, E. B., Rosen, C. A., Schindler, A., Woisard, V., Whurr, R., and Konerding, U. 2009. Item reduction of the voice handicap index based on the original version and on European translations. *Folia phoniatica et logopaedica : official organ of the International Association of Logopedics and Phoniatrics (IALP)* 61, 1, 37–48.
- [77] Nawka, T., Wirth, G., and Anders, L. C. 2008. *Stimmstörungen. Für Ärzte, Logopäden, Sprachheilpädagogen und Sprechwissenschaftler ; mit 30 Tabellen*. Dt. Ärzte-Verl., Köln.
- [78] Nawka, Franke, Galkin. 2006. Objektive Messverfahren in der Stimmdiagnostik. *Logopädie*, 4.
- [79] Ng, M. L., Gilbert, H. R., and Lerman, J. W. 1997. Some aerodynamic and acoustic characteristics of acute laryngitis. *Journal of Voice* 11, 3, 356–363.
- [80] Olthoff, A., Woywod, C., and Kruse, E. 2007. Stroboscopy versus high-speed glottography. A comparative study. *The Laryngoscope* 117, 6, 1123–1126.
- [81] Pépiot Erwan. 2012. Voice, speech and gender: male-female acoustic differences and cross-language variation in English and French speakers. *XVèmes Rencontres Jeunes Chercheurs de l'ED* 268.
- [82] Ptok, M., Schwemmler, C., Iven, C., Jessen, M., and Nawka, T. 2006. Zur auditiven Bewertung der Stimmqualität. *HNO* 54, 10, 793–802.
- [83] RICHARD, E. G., ROBERT, M. G., and WILLIAM, W. M. 2016. Validation of a Voice Outcome Survey for Unilateral Vocal Cord Paralysis. *Otolaryngol Head Neck Surg* 120, 2, 153–158.
- [84] Richter, B. 2016. Bildgebende Diagnostik. Video-Laryngo-Stroboskopie, Hochgeschwindigkeitsglottografie, dynamische Kernspintomografie. *Sprache Stimme Gehör* 40, 04, 168–172.
- [85] Roget, P. M. 1815. Explanation of an Optical Deception in the Appearance of the Spokes of a Wheel Seen through Vertical Apertures. *Proceedings of the Royal Society of London* 2, 0, 230–231.
- [86] Rostolland, D. 1982. Acoustic Features of Shouted Voice. *Acta Acustica united with Acustica, Volume 5*, Number 2, 118–125.
- [87] Rusz, J., Cmejla, R., Ruzickova, H., and Ruzicka, E. 2011. Quantitative acoustic measurements for characterization of speech and voice disorders in early untreated Parkinson's disease. *The Journal of the Acoustical Society of America* 129, 1, 350–367.
- [88] Sauder, C., Bretl, M., and Eadie, T. 2017. Predicting Voice Disorder Status From Smoothed Measures of Cepstral Peak Prominence Using Praat and Analysis of Dysphonia in Speech and Voice (ADSV). *Journal of voice : official journal of the Voice Foundation* 31, 5, 557–566.
- [89] Schade, G. and Müller, F. 2005. Hochgeschwindigkeitsglottographische Diagnostik in der Laryngologie. *HNO* 53, 12, 1085-6, 1088-91.

- [90] Schloneger, M. J. and Hunter, E. J. 2017. Assessments of Voice Use and Voice Quality Among College/University Singing Students Ages 18-24 Through Ambulatory Monitoring With a Full Accelerometer Signal. *Journal of voice : official journal of the Voice Foundation* 31, 1, 124.e21-124.e30.
- [91] Schneider, B. and Bigenzahn, W. 2003. Influence of glottal closure configuration on vocal efficacy in young normal-speaking women. *Journal of Voice* 17, 4, 468–480.
- [92] Schneider-Stickler, B. 2016. Stimmumfangsprofilmessung. *Sprache Stimme Gehör* 40, 04, 173–178.
- [93] Schneider-Stickler, B. and Bigenzahn, W. 2013. *Stimmdiagnostik. Ein Leitfaden für die Praxis*. Springer, Wien.
- [94] Schultz-Coulon, H.-J. 1980. Die Diagnostik der gestrten Stimmfunktion. *Arch Otorhinolaryngol* 227, 1-2, 1–169.
- [95] Schünke, M. 2009. *Prometheus Lernatlas - Kopf, Hals und Neuroanatomie*. Thieme, Stuttgart.
- [96] Schutte, H. K. and Seidner, W. 1983. Recommendation by the Union of European Phoniaticians (UEP). Standardizing Voice Area Measurement/Phonetography. *Folia phoniatica et logopaedica : official organ of the International Association of Logopedics and Phoniatics (IALP)* 35, 6, 286–288.
- [97] Sendlmeier, W. 2012. Die psychologische Wirkung von Stimme und Sprechweise. Geschlecht, Alter, Persönlichkeit, Emotion und audiovisuelle Interaktion.
- [98] Simpson, A. P. 2012. The first and second harmonics should not be used to measure breathiness in male and female voices. *Journal of Phonetics* 40, 3, 477–490.
- [99] Spazzapan, E. A., Cardoso, V. M., Fabron, E. M. G., Berti, L. C., Brasolotto, A. G., and Marino, V. C. d. C. 2018. Características acústicas de vozes saudáveis de adultos. Da idade jovem à meia-idade. *CoDAS* 30, 5, e20170225.
- [100] Speyer, R., Wieneke, G. H., and Dejonckere, P. H. 2004. Self-assessment of voice therapy for chronic dysphonia. *Clinical otolaryngology and allied sciences* 29, 1, 66–74.
- [101] Stein, D. J. and Noordzij, J. P. 2013. Incidence of chronic laryngitis. *The Annals of otology, rhinology, and laryngology* 122, 12, 771–774.
- [102] Stöcker, H., Ed. 1998. *Taschenbuch der Physik. Formeln, Tabellen, Übersichten*. Deutsch, Thun.
- [103] Svec, J. G. and Schutte, H. K. 1996. Videokymography. High-speed line scanning of vocal fold vibration. *Journal of Voice* 10, 2, 201–205.
- [104] T.Nawka, U.Wiesmann, and U.Gonnermann. 2003. Validierung des Voice Handicap Index (VHI) in der deutschen Fassung. *HNO*, 51, 921–929.
- [105] Tadeus Nawka, Ute Gonnermann, and Ulrich Wiesmann. Stimmstörungsindex.
- [106] Teixeira, J. P. and Fernandes, P. O. 2014. Jitter, Shimmer and HNR Classification within Gender, Tones and Vowels in Healthy Voices. *Procedia Technology* 16, 1228–1237.
- [107] Titze. 1994. Workshop on Acoustic Voice Analysis.
- [108] Titze, I., Ed. 1994. *NCVS Status and Progress Report. Effect of Voice Disorders on Quality of Life Volume 7*.
- [109] Titze, I. R. 1980. Comments on the Myoelastic - Aerodynamic Theory of Phonation. *J Speech Lang Hear Res* 23, 3, 495.
- [110] Titze, I. R. 1989. On the relation between subglottal pressure and fundamental frequency in phonation. *The Journal of the Acoustical Society of America* 85, 2, 901–906.
- [111] Titze, I. R. 2000. *Principles of voice production*. National Center for Voice and Speech, Iowa City.
- [112] Ulrich Eysholdt and Jörg Lohscheller. 2007. Diagnostik bei unklarer Heiserkeit – Bildgebung von Stimmlippenschwingungen. *Deutsches Ärzteblatt* 104, 51-52.

- [113] van den Berg, J. 1958. Myoelastic-Aerodynamic Theory of Voice Production. *Journal of speech and hearing research* 1, 3, 227–244.
- [114] Voigt-Zimmermann, S., Lampe, K., and Arens, C. 2014. Differenzialdiagnostik der Heiserkeit. *Laryngo- rhino- otologie* 93, 4, 263-84; quiz 285-6.
- [115] Wendahl, R. W. 1966. Laryngeal Analog Synthesis of Jitter and Shimmer Auditory Parameters of Harshness. *Folia phoniatica et logopaedica : official organ of the International Association of Logopedics and Phoniatrics (IALP)* 18, 2, 98–108.
- [116] Wendler, J., Seidner, W., and Eysholdt, U. 2015. *Lehrbuch der Phoniatrie und Pädaudiologie. 35 Tabellen*. Thieme, Stuttgart.
- [117] Werth, K., Voigt, D., Döllinger, M., Eysholdt, U., and Lohscheller, J. 2010. Clinical value of acoustic voice measures. A retrospective study. *European archives of oto-rhino-laryngology : official journal of the European Federation of Oto-Rhino-Laryngological Societies (EUFOS) : affiliated with the German Society for Oto-Rhino-Laryngology - Head and Neck Surgery* 267, 8, 1261–1271.
- [118] Wolfram Seidner, Ingeburg Walch, Ingeborg Kuchler, and Jörg Sandmann. 2003. Stimmumfangsprofile im Normbereich. 20. Wissenschaftliche Jahrestagung der DGPP. Rostock, 12.-14.09.2003 Düsseldorf. Deutsche Gesellschaft für Phoniatrie und Pädaudiologie, Köln: German Medical Science.
- [119] Wuyts, Van de Heyning, De Bodi, Heylen, and Remacle. 1997. Dysphonia Severity Index: A Multiparameter approach for the assessment of Voice Quality. *Larynx*, 73–76.
- [120] Wuyts, F. L., Bodt, M. S. D., Molenberghs, G., Remacle, M., Heylen, L., Millet, B., van Lierde, K., Raes, J., and van de Heyning, P. H. 2000. The Dysphonia Severity Index. *J Speech Lang Hear Res* 43, 3, 796.
- [121] Yumoto, E., Gould, W. J., and Baer, T. 1982. Harmonics-to-noise ratio as an index of the degree of hoarseness. *The Journal of the Acoustical Society of America* 71, 6, 1544–1549.

8 Abbildungsverzeichnis

Abbildung 1: Schematische Darstellung des Kehlkopfes im Mediansagittalschnitt in Anlehnung an [7].	4
Abbildung 2: Schematische Darstellung des Kehlkopfes in der Schrägansicht von links-ventral in Anlehnung an [95].	5
Abbildung 3: Stroboskopisches Einzelbild eines männlichen, stimmgesunden Probanden, aufgenommen während der Phonation mit einer Frequenz von 111 Hz und 82 dB.....	6
Abbildung 4: Schematische Darstellung einiger muskulärer Strukturen im Bereich des Kehlkopfes nach [95]. Ansicht von links, linke Hälfte des Schildknorpels entfernt und Sicht auf dorsal liegende Strukturen.....	8
Abbildung 5: Stimmlippenebene während der Phonation im sagittalen Querschnitt (obere Bildreihe) und kranialer Aufsicht (untere Bildreihe). a) Subglottischer Druckaufbau bei noch geschlossenen Stimmlippen. b) Beginn der Phonation und Öffnen der Stimmlippen durch aerodynamische Kräfte. c) Subglottischer Druckabfall und Beginn des Schließungsprozesses durch Bernoulli-Effekt und myoelastische Rückstellkräfte. Untere Stimmlippenkante wird sichtbar. d) Kompletter Verschluss der Stimmlippen beginnend an der unteren Stimmlippenkante.....	9
Abbildung 6: Darstellung des Glottisverlauf über die Zeit anhand von Einzelbildern aus einer Hochgeschwindigkeitssequenz, aufgenommen während der Phonation. Die weißen Pfeile geben die Bewegungsrichtung der Stimmlippen an.	10
Abbildung 7: Vereinfachte Darstellung von Frequenz- und Amplitudenänderung anhand periodischer Schwingungen. Die Auslenkung der Schwingungen wird als Amplitude bezeichnet, die Periodendauer ist die Dauer eines Schwingungszyklus. a) Bei dieser Schwingung ist die Frequenz circa 1 Hz. Die Amplitude liegt bei eins. b) Die Amplitude liegt bei 2 und die Periodendauer ist aufgrund einer höheren Frequenz deutlich verkürzt und beträgt circa 2 Hz.	11
Abbildung 8: Schematische Darstellung der Entstehung des hörbaren Klanges aus dem primären Kehlkopftön. a) Primärer Kehlkopftön als Spektrum. b) Resonanzfunktion des Ansatzrohres und Verstärkung nur bestimmter Frequenzmaxima. c) Fortgeleitete Frequenzen als Energiemaxima. Entstehung von Formanten (Anlehnung an Abb 6.4. [12]).	12
Abbildung 9: Darstellung eines Reinke-Ödems an der rechten Stimmlippe. Aufnahmen aus dem Patientenkollektiv von Dr. Bader (HNO-Klinik, Universitätsklinikum des Saarlandes). ..	14
Abbildung 10: Maligne und benigne Larynxläsionen im Vergleich a) Larynxkarzinom an der rechten Stimmlippe. b) Larynxpapillomatose als Differentialdiagnose zur malignen Entartung. Aufnahmen aus dem Patientenkollektiv von Dr. Bader (HNO-Klinik, Universitätsklinikum des Saarlandes).	15
Abbildung 11: Benigne Stimmlippentumore, siehe rote Pfeile a) Stimmlippenknoten parallel an beiden Stimmlippen. b) Stimmlippenpolyp rechts. Grüner Pfeil zeigt kartilaginären Anteil der Stimmlippen. Aufnahmen aus dem Patientenkollektiv von Dr. Bader (HNO-Klinik, Universitätsklinikum des Saarlandes).	16
Abbildung 12: Stimmlippenparese links mit atrophischen Veränderungen. Aufgrund eines länger bestehenden Funktionsverlusts deutlich atrophisierte linke Stimmlippe im Vergleich zur rechten Stimmlippe. Aufnahmen aus dem Patientenkollektiv von Dr. Bader (HNO-Klinik, Universitätsklinikum des Saarlandes).	17
Abbildung 13: Schematische Auflistung der einzelnen Untersuchungsschritte.	33
Abbildung 14: Registrierung und Frequenzeinstellung der Probanden auf die individuelle mittlere Sprechstimmlage.	37
Abbildung 15: Durchführung einer Aufnahme. Vorgegebener grüner Frequenzbereich während der Aufnahme im mittleren Bild rot markiert.	38

Abbildung 16: Bewertung der Stimme nach Stimmbelastung, Stimmqualität und Stimmermüdung, sowie die Qualität von Missempfindungen.....	38
Abbildung 17: Abschluss des Fragebogens und Kontrolle der erfolgten Aufnahme im Stimmtagebuch über den Patienten Bereich.....	39
Abbildung 18: Verteilung der VHI-Werte der 17 Probanden bei Beginn und am Ende auf der x-Achse. Die Gesamtpunktzahl des VHI-12 Tests auf der y-Achse der Grafik, zeigt, dass kein Proband die 6 Punktegrenze überschreitet. Eine Punktzahl ab 8 Punkte würden für ein geringgradiges Missempfinden sprechen. Die Gesamtpunktzahl 0, die zum Beispiel bei Proband 8 angegeben wurde, entspricht einer gesund empfundenen Stimme.	42
Abbildung 19: Öffnungsphase zweier Probanden während der Phonation. Im Bild a) beträgt die Frequenz des Probanden (VG300392) 150 Hz und die Lautstärke 82 dB. In b) phoniert der zweite Proband (GF110694) mit 138 Hz und 76 dB. Die Blitzlichtfrequenz Δf zur Grundfrequenz f_0 ist in beiden Bildern ± 1 Hz. Auf den Stimmlippen in Bild a) lassen sich weiße, nicht pathologische Schleimauflagerungen erkennen.....	43
Abbildung 20: a) Die Grafik zeigt anhand zweier Boxplots die Verteilung der mittleren Sprechstimmlage zu Studienbeginn und -ende. Beide Boxplots zeigen, dass 50 % der mSsl aller Probanden zu beiden Zeitpunkten in einem Frequenzbereich von 99 – 110 Hz liegen. Die Verteilung zu beiden Zeitpunkten ist sehr ähnlich. b) Das Balkendiagramm zeigt jeweils nur eine sehr geringe Abweichung der mSsl zu Beginn und am Ende der Studie bei den jeweiligen Probanden.	44
Abbildung 21: a) Die beiden errechneten Mittelwerte der maximalen Frequenz der 17 Teilnehmer sind zu Studienbeginn (597 Hz \pm 163 Hz) und Studienende (598 Hz \pm 146 Hz) nahezu identisch und ähnlich verteilt. b) Der Mittelwert und die Abweichung des leisesten Schallpegels ist zu beiden Zeitpunkten identisch und liegt bei 47 dB \pm 4 dB.	45
Abbildung 22: Darstellung der berechneten Mittelwerte des Jitters der 17 Studienteilnehmer zu Beginn und am Ende der Studie. Die Werte sind zu beiden Zeitpunkten sehr ähnlich (0,12% \pm 0,047 und 0,11% \pm 0,035). Zudem liegt die Verteilung aller Jitter Werte in einem sehr kleinen Bereich und im physiologischen Bereich < 1%.	46
Abbildung 23: Die Boxplots stellen die Verteilung der DSI Werte aller Probanden zu Studienbeginn und- ende gegenüber. Der Mediane ist zu Beginn 6,3 und am Ende der Studie 6,6, Die Verteilung ist sehr regelmäßig und unterscheidet sich zu besagten Zeitpunkten wenig. Es gibt keine Ausreißer und mehr als 75% der Werte liegen über der gewünschten Norm von 4,2 [78].	47
Abbildung 24: a) Der Boxplot zeigt die Verteilung aller Aufnahmen von allen 17 Probanden in den einzelnen Zeitintervallen, welche von 5-10 Uhr, 10-14 Uhr, 14-18 Uhr, 18-23 Uhr und 23-5 Uhr waren. Ein Großteil der Aufnahmen wurde zwischen 10 und 23 Uhr aufgenommen. b) Das Säulendiagramm zeigt die Gesamtanzahl der Aufnahmen von jedem einzelnen Probanden innerhalb der vorgegebenen zwei Wochen. Bei allen Probanden stehen mehr als 60 Aufnahmen zur Auswertung zur Verfügung.	48
Abbildung 25: Der Boxplot zeigt die Verteilung der sechs unterschiedlichen Belastungsgrade der Stimme im Tagesverlauf, welche nach jeder Stimmaufnahme abgefragt wurde. Höhergradige Stimmbelastungen treten vorwiegend im späteren Verlauf des Tages auf.	49
Abbildung 26: a) Die Verteilung der absoluten Häufigkeit der Belastungsgrade zeigt in den unterschiedlichen Intervallen eine ungleiche Verteilung. b) Hier wird die relative Häufigkeit der 6 Belastungsgrade gezeigt. Die Verteilung zeigt, dass die Stimme ab 10 Uhr einer zunehmenden Stimmbelastung ausgesetzt war. Zudem steigt die relative Häufigkeit der Grad 4 Belastung von 10 bis 22 Uhr kontinuierlich an, wohingegen sich die Häufigkeiten von Belastungsgrad 2 und 3 in diesem Zeitraum kaum ändern.	50
Abbildung 27: Relative Häufigkeit der subjektive Stimmbelastung aufgeschlüsselt auf die einzelnen Probanden. Die subjektive Belastung der Stimme wurde von den Probanden sehr unterschiedlich empfunden. Proband 1 und 5 haben beispielsweise nie oder nur selten eine	

Belastung der Stimme angegeben, bei Proband 10 hingegen lang fast immer ein Belastungsgrad vor.	50
Abbildung 28: a) Relative Häufigkeit der Stimmbelastungsart in Bezug auf alle Aufnahmen. Die Auswertung zeigt, dass Sprechen die häufigste Belastung für die Stimme der Teilnehmer war. Abbildung 28 b) zeigt die relative Häufigkeit der individuellen Verteilung der unterschiedlichen Belastungsarten bei jedem einzelnen Probanden. Proband 5 und 14 empfanden eine Stimmbelastung hauptsächlich beim Rufen und Singen.	51
Abbildung 29: a) Relative Verteilung der Stimmqualitätsgrade von „sehr schlecht“ (Grad 1) bis „sehr gut“ (Grad 6) in den jeweiligen Zeitintervallen. Über die Hälfte der Aufnahmen wurden mit „schlecht“ oder „sehr schlecht“ bewertet. Die relative Häufigkeit der Verteilung von 5-10 Uhr zeigt, dass die Aufnahmen im Vergleich etwas besser bewertet wurden. Grad 3 und 4 treten häufiger als zu anderen Zeiten auf. b) Subjektive Einschätzung der Stimmqualität durch die Probanden. Ein Großteil der Probanden bewertet die Aufnahmen mit einer „schlechten“ oder „sehr schlechten“ Stimmqualität.	52
Abbildung 30: Der Boxplot zur zeitlichen Verteilung der unterschiedlichen Stimmqualitäten aller Aufnahmen zeigt keine bestimmte Verteilung der unterschiedlichen Qualitätsgrade. 50 % der Stimmaufnahmen mit „sehr guter“ (Grad 6) Qualität wurden zwischen 15 und 21 Uhr aufgenommen.	53
Abbildung 31: a) Verteilung der relativen Häufigkeit von unterschiedlichen Missempfindungen zu unterschiedlichen Zeitpunkten. Die Ergebnisse der Analyse zeigen, dass nur in einzelnen Fällen Räuspern oder Kratzen im Hinblick auf alle Aufnahmen angegeben wurden und ein Missempfinden häufiger morgens auftritt. b) Geringe Relative Häufigkeit bestimmter Qualitäten von Missempfindungen in Bezug auf die Gesamtzahl der einzelnen Aufnahmen der einzelnen Probanden.	54
Abbildung 32: a) Relative Häufigkeiten der Ermüdungsgrade in den unterschiedlichen Zeitintervallen. Bei über der Hälfte aller Aufnahmen wurde keine Stimmermüdung angegeben, die relative Häufigkeit der Ermüdungsgrade war in jedem Zeitintervall abnehmend. Zwischen 18-5 Uhr morgens traten im Vergleich höhere Ermüdungsgrade (Grad 4 und 5) auf. b) Subjektive Einschätzung der Ermüdungsgrade der einzelnen Probanden. Hierbei liegen große Unterschiede vor, während Proband 1 und 5 nie eine Belastung empfinden, geben andere, beispielsweise Proband 11 häufiger mittelschwere Ermüdungsgrade an.	55
Abbildung 33: Verteilung der unterschiedlichen Bewertungsgrade zur Stimmermüdung aller Aufnahmen im Hinblick auf den Tagesverlauf. Die Ermüdungsgrade 1 bis 3 verteilen sich hauptsächlich zwischen 10 bis 18 Uhr. Das Auftreten von höhergradigen Stimmermüdungen wurde in den Abendstunden und nachts angegeben. Bei einer Aufnahme in der gesamten Studie wurde ein „sehr starker“ Ermüdungsgrad angegeben.	56
Abbildung 34: a) Das Histogramm zeigt die allgemeine Häufigkeitsverteilung von F_0 . Der Mittelwert liegt bei $110 \pm 9,9 \text{ Hz}$. Die Verteilung ist im Vergleich zum Mittelwert leicht schief, jedoch regelmäßig verteilt. b) Der Boxplot zeigt eine regelmäßige Verteilung der Grundfrequenz im Tagesverlauf. Von 5-14 Uhr ist der Median etwas niedriger als in den anderen Zeitintervallen.	57
Abbildung 35: a) Boxplot, der das Differenzverhältnis zwischen F_0 und $mSs1$ im Verlauf des Tages darstellt. In den Morgenstunden besteht ein leichtes negatives Verhältnis, das ab 10 Uhr positiv wird. b) Entwicklung von F_0 im Verlauf der gesamten Studie am Beispiel von Proband 13. Es zeigt sich eine gleichmäßige Verteilung der Grundfrequenz um die mittlere Sprechstimmlage (rote Linie).	57
Abbildung 36: Vergleich der Probanden 2 und 7 miteinander und Vergleich mit dem F_0 Wert des Gesamtkollektivs. Die Grundfrequenz des Probanden 2 weicht stark von der mittleren Verteilung von F_0 in der Studie ab, die an dem Beispiel von Proband 7 dargestellt ist.	58

Abbildung 37: a) Regelmäßige und kompakte Verteilung des Jitters um den Mittelwert 0,5%. Es finden sich kaum Werte über 1%. b) Verteilung des Jitters im Tagesverlauf um einen Median bei 0,5%. Nur einzelne Ausreißer (rote Kreuze) liegen über 1%.....59

Abbildung 38: Vergleich des Jitter von Proband 4 und 12 in Bezug zum errechneten Mittelwert des Gesamtkollektivs. Trotz stärkerer Abweichungen vom durchschnittlichen Bereich, haben die Aufnahmen von Proband 4 selten einen Jitter über 1%. Der Jitter des Probanden 12 liegt etwas unterhalb des mittleren Jitters des Gesamtkollektivs.....60

Abbildung 39: a) Absteigende und kompakte Verteilung um den Mittelwert 4,6%. b) Verteilung des Shimmers im Tagesverlauf um einen Median bei 0,5%. Bei 75 % der Aufnahmen liegt der Shimmer $\leq 6,1\%$60

Abbildung 40: Vergleich des Shimmers von Proband 4 und 12 in Bezug zum errechneten Mittelwert $4,6 \pm 2,2 \%$ des Gesamtkollektivs. Proband 4 zeigt in einigen Aufnahmen eine starke Abweichung vom durchschnittlichen Bereich des Shimmers. Der Shimmer des Probanden 12 ist nachts etwas höher als im Tagesverlauf, liegt jedoch in allen Zeitintervallen im errechneten Durchschnitt.....61

Abbildung 41: a) Darstellung der Verteilung der relativen Häufigkeiten des HNR. Der Mittelwert ist $15 \pm 3,2 \text{ dB}$. Zu beiden Seiten des MW ist die Verteilungskurve stark abfallend. b) Der Median liegt in allen Zeitintervallen ungefähr bei 14 dB . Ebenso weichen die 1. und 3. Quartile in den 5 Bereich nicht stark voneinander ab.....62

Abbildung 42: a) Der HNR des Probanden 9 liegt zum größten Teil außerhalb des grauen Bereichs ($MW \pm SD$). Der Median des Probanden liegt im zeitlichen Verlauf zwischen $10-12 \text{ dB}$. Auffallend sind niedrige HNR Werte zwischen 10 und 18 Uhr, die für einen vermehrten Geräuschanteil bei einzelnen Aufnahmen sprechen. b) Proband 5 als Beispiel für eine regelmäßige Verteilung des HNR in Anbetracht des Gesamtkollektiv.....63

Abbildung 43: a) Der NHR aller Aufnahmen verteilt sich größtenteils zwischen $0 - 0,2 \text{ dB}$. Der Mittelwert liegt bei $0,062 \pm 0,05 \text{ dB}$. Die Standardabweichung zeigt eine sehr geringe Streuung der Werte. b) Der Median ist dem Mittelwert sehr ähnlich und liegt in allen 5 Zeitbereichen zwischen etwa $0,06 - 0,07 \text{ dB}$64

Abbildung 44: Vergleich von Proband 4 und 9, die beide etwas vom Mittelwert des Gesamtkollektivs abweichen. a) Bei Proband 4 weichen überwiegend die Aufnahmen nachts und in den Morgenstunden ab. Zwischen $10 - 23 \text{ Uhr}$ liegt der NHR im Bereich der Standardabweichung des Gesamtkollektivs. b) Bei Proband 9 liegt der NHR durchschnittlich etwas höher als der Mittelwert und die Standardabweichung des Gesamtkollektivs.65

Abbildung 45: a) Der NNE Mittelwert aller Aufnahmen ist bei $-3,4 \pm 3,2 \text{ dB}$. Die meisten NNE Werte sind zwischen 0 und -6 dB verteilt. Auffallend sind viele kleine Ausreißer. Diese sind mit roten Kreuzen in b) gekennzeichnet. Der Median liegt im Bereich des Mittelwertes und über die Zeitintervalle ist der NNE näherungsweise konsistent verteilt.66

Abbildung 46: a und b) zeigen den Verlauf der NNE Ergebnisse exemplarisch von Proband 3 und 5 im Zeitraum von 14 Tagen. Bei beiden Probanden ist eine sehr unregelmäßige Verteilung der Werte erkennbar. Diese breite Streuung der Wert findet man bei vielen weiteren Probanden in der Studie.....67

Abbildung 47: a) Die Werte verteilen sich zu beiden Seiten des MW $-0,68 \pm 4,2 \text{ dB}$ gleichmäßig abfallend. Ausreißer, die sich in der Größe ihrer Anzahl in der relativen Häufigkeit widerspiegeln sind nur in geringer Anzahl vorhanden. b) Der Median schwankt im Tagesverlauf im Bereich rund um 0 und -1 dB verteilt. Eine große Übereinstimmung zwischen Median und Mittelwert liegt vor.....68

Abbildung 48: Vergleich zweier Probanden mit dem Gesamtkollektiv. a) Die H_1H_2 Werte zwischen $18-5 \text{ Uhr}$ weichen von der Medianverteilung im Lauf des Tages ab und sind wesentlich größer als der Mittelwert. b) Die Werte von Proband 17 liegen außerhalb der unteren SD, der Median in den einzelnen Zeitintervallen ist gleich homogen verteilt und liegt bei ca. -5 dB69

Abbildung 49: a) Breite Streuung der einzelnen H_1H_2 Werte bei Proband 8. Ein Trend zu einem bestimmten Mittelwert liegt nicht vor. Selten wird der Mittelwert des Gesamtkollektivs erreicht. b) Im Vergleich regelmäßige Verteilung der H_1H_2 Ergebnissen um den errechneten Mittelwert bei Proband 2.70

Abbildung 50: a) Der Mittelwert des Slope liegt bei $-17 \pm 3,5 \text{ dB}$. Die relative Häufigkeit der anderen Werte nimmt zu beiden Seiten des MW ab. Zu den niedrigeren Werten ist das Gefälle etwas flacher, zu den höheren Werten steiler. b) Im Tagesverlauf ist der Median des Slope zwischen -16 und -17 dB . Zwischem 5-10 Uhr ist der Medan bei etwa $17,5 \text{ dB}$ etwas vermindert.71

Abbildung 51: a) Die Verteilung des Slope bei Proband 7 soll die individuelle Verteilung bei den einzelnen Probanden zeigen. Ein Großteil der Werte von Proband 7 liegt oberhalb des errechneten Mittelwertbereichs. Die Verteilung ist bist auf einzelne Ausreißer jedoch regelmäßig b) Proband 6 zeigt in den betrachteten Zeitintervallen eine regelmäßige Verteilung des Slopes innerhalb der Standardabweichung. Die Verteilung gestaltet sich bei den anderen Probanden ähnlich.72

Abbildung 52: a) Verteilung aller Tilt Werte aus der Studie um den Mittelwert $-8 \pm 1,5 \text{ dB}$. Die Verteilung nimmt zu den größeren Werten hin flacher ab als zu kleineren Tilt Werten. b) Die Verteilung um den Median bei circa -8 dB ist gleichmäßig. 50 % der Aufnahmen liegen zwischen $-8,5$ und $-6,5 \text{ dB}$. Zwischen 5 bis 10 Uhr ist der Median des Tilts etwas größer als in den anderen Intervallen.73

Abbildung 53: a) Der Mittelwert der Smoothed cepstral peak prominence ist bei $12 \pm 2,1 \text{ dB}$. Die Werte verteilen sich gleichmäßig abfallend zu beiden Seiten. b) Der Median ist dem Mittelwert in allen Zeitintervallen sehr ähnlich. Die Aufnahmen liegen in allen 5 Zeitintervallen in einem ähnlichen Bereich.74

Abbildung 54: Verteilung des CPPs in den unterschiedlichen Zeitintervallen am Beispiel zweier Probanden. a) Bei Proband 1 ist der Median im Tagesverlauf ansteigend und liegt etwas unterhalb des berechneten Mittelwertes $12 \pm 2,1 \text{ dB}$. Die Verteilung des Wertes ist jedoch regelmäßig und ohne große Streuung. b) Der CPPs bei Proband 2 liegt fast immer im Bereich der Standardabweichung.75

Abbildung 55: Verteilung aller 9 Parameter aus dem telemedizinischen Monitoring nach Dimensionsreduktion mittels t-SNE Verfahren. Berücksichtigung aller Aufnahmen der 17 Probanden. Entstehung von zwei unterschiedlich großen Clustern.75

Abbildung 56: Vergleich der beiden Cluster von Jitter und NNE. a) Gleichmäßig und fast deckungsgleiche Verteilung des Jitters in beiden Clustern. Die roten Markierungen zeigen die einzelnen Ausreißer. b) Keine Übereinstimmung der NNE Werte in beiden Clustern.76

Abbildung 57: Verteilung der Parameter aus dem telemedizinischen Monitoring ohne NNE. Das kleine Cluster ist nach einer Auswertung ohne NNE Werte verschwunden.77

Abbildung 58: Betrachtung der Verteilung der erhobenen Parameter bei einzelnen Probanden. Die Probanden 1-4 zeigen, dass sich die Werte aus dem telemedizinischen Monitoring in einem bestimmten Bereich (Proband 3) oder um einen bestimmten Punkt (Proband 2) im Cluster ansammeln.78

Abbildung 59: Verteilung der Parameter des telemedizinischen Monitorings mit Ausnahme des NNE in den 5 unterschiedlichen Zeitintervallen am Beipsiel von Proband 13. Die Werte verteilen sich unabhängig von der Uhrzeit immer in einem ähnlichen Bereich. In diesem Bereich liegt ein gewisse Variabilität vor.79

Abbildung 60: Verteilung der Parameter des telemedizinischen Monitorings mit Ausnahme des NNE in den unterschiedlichen Zeitintervallen am Beipsiel von Proband 17. Die Verteilung weist eine deutlich geringer Variabilität auf als bei Proband 13, der Verteilung ist in allen 5 Bildern nahezu identisch.80

Abbildung 61: Verwendeter VHI-12 Test aus dem Studienprotokoll.94

Abbildung 62: Untersuchungsprotokoll zur Beurteilung der Stroboskopie Aufnahmen aus dem Studienprotokoll.....95

Abbildung 63: Vervollständigung der Daten des Sprechstimmfeldes a) Mittlere Lautstärke der mittleren Sprechstimmlage von 17 Probanden. Da der Frequenzbereich klein gewählt ist, erscheint die SD und die Abweichung der Mittelwerte zu Beginn und am Ende der Studie groß. b) Die Mittelwerte der Dynamik zu den beiden Zeitpunkten sind sehr ähnlich. Bei Ausschluss aus der Studie ist die SD etwas höher. c) Die beiden Mittelwerte der Rufstimmfrequenz sind zu beiden Zeitpunkten ähnlich und die SD ist ähnlich verteilt. d) Ebenso verhält sich die Lautstärke der Rufstimme. Die Mittelwerte sind minimal unterschiedlich und die SD am Ende etwas größer.96

Abbildung 64: Vervollständigung der Daten des Singstimmfeldes a) Minimale Stimmfrequenz der Singstimme im Vergleich zu Beginn und am Ende der Studie. Die Daten aller Probanden liegen in einem sehr ähnlichen Bereich und die SD des Wertes ist sehr gering. b) Die Verteilung des lautesten Schallpegels der Singstimme ist in dieser Studie relativ weit. c) Die Verteilung des Dynamikbereichs der Singstimme ist nahezu identisch zu Beginn und am Ende der Studie.....96

Abbildung 65: Verteilung des Shimmers von Proband 9 im Vergleich zum Mittelwert und Standardabweichung des Gesamtkollektivs ($4,6 \pm 2,2 \%$).97

Abbildung 66: Hochnormale HNR-Werte einzelner Probanden im Vergleich zum Mittelwert des Gesamtkollektivs ($15 \pm 3,2 \text{ dB}$).97

Abbildung 67: a) Entwicklung des HNR im Studienverlauf von Proband 3 und Proband 15. b) Entwicklung des NHR im Studienverlauf von Proband 3 und Proband 15.....98

Abbildung 68: Abweichende Verteilung des H1H2 von Proband 8 im Vergleich zum Mittelwert und der Standardabweichung des Gesamtkollektivs ($-0,68 \pm 4,2 \text{ dB}$).99

Abbildung 69: Verteilung der Parameter aus dem telemedizinischen Monitoring im kleinen und großen Cluster.....101

9 Danksagung

An dieser Stelle möchte ich allen beteiligten Personen, die mich bei der Anfertigung meiner Dissertation tatkräftig unterstützt haben, meinen größten Dank aussprechen.

Mein besonderer Dank gilt Herrn Professor Dr.-Ing. Jörg Lohscheller für die hervorragende Betreuung und die enorme Unterstützung beim Anfertigen dieser Doktorarbeit. Zudem danke ich Herrn Professor Dr. med. Alessandro Bozzato für die Hilfe bei der Veröffentlichung der Arbeit. Ebenfalls möchte ich Frau Dipl. Biophys. Mona Fehling meinen Dank aussprechen, welche eine unersetzliche Hilfe bei der Lösungsfindung von Problemen war.

Herzlich bedanken möchte ich mich zudem bei Herrn Dr. med. Carl-Albert Bader aus der HNO-Klinik der Universitätsklinik des Saarlandes, der mich vor allem im klinischen Teil der Arbeit unterstützt und mir Bilder und Räumlichkeiten zur Verfügung gestellt hat. Weiterhin bedanke ich mich bei Herrn Dr. rer. nat. Dietmar Hecker, der mir bei technischen Fragen zur Seite stand.

Außerdem möchte ich mich bei allen Teilnehmern dieser Studie bedanken, ohne deren Hilfe die Arbeit nicht möglich gewesen wäre.

Zuletzt möchte ich noch meinem Freund, meiner Familie und meinen Freunden danken, die mich mit viel Geduld und ermutigenden Worten unterstützt haben.

10 Lebenslauf

„Aus datenschutzrechtlichen Gründen wird der Lebenslauf in der elektronischen Fassung der Dissertation nicht veröffentlicht.“