

Aus der Klinik für Phoniatrie und Pädaudiologie  
Direktorin: Prof. Dr. med. Roswitha Berger  
des Fachbereichs Medizin der Philipps-Universität Marburg  
in Zusammenarbeit mit dem Universitätsklinikum Gießen und Marburg GmbH, Standort  
Marburg

**Untersuchungen zur Vergleichbarkeit von  
Ergebnissen der Stimmschallanalyse  
kommerziell verfügbarer Geräte**

**INAUGURAL-DISSERTATION**

zur

Erlangung des Doktorgrades der gesamten Humanmedizin dem  
Fachbereich Medizin der Philipps-Universität Marburg  
vorgelegt von

**Sebastian Gärtner**

aus Marburg

Marburg 2012

Angenommen vom Fachbereich Medizin der Philipps-Universität Marburg  
am: 07.11.2012.

Gedruckt mit der Genehmigung des Fachbereichs.

Dekan: Prof. Dr. med. Matthias Rothmund

Referent: Frau Prof. Dr. med. Roswitha Berger

Korreferent: Herr Prof. Dr. rer. physiol. Volker Groß

# Inhaltsverzeichnis

<b>1</b>	<b>Einleitung</b> .....	<b>1</b>
1.1	Kommunikation .....	1
1.2	Stimmfunktion .....	3
1.2.1	Physikalisch-akustische Grundlagen .....	3
1.2.2	Anatomie .....	4
1.2.3	Physiologie .....	4
1.2.4	Einflussfaktoren .....	5
1.3	Stimmstörungen.....	7
1.3.1	Einteilung der Dysphonien .....	7
1.3.2	Funktionelle Dysphonien .....	8
1.3.3	Organische Dysphonien .....	9
1.4	Diagnostische Methoden .....	10
1.4.1	Anamnese .....	11
1.4.2	Untersuchung von Kehlkopf, Atmung und Ansatzrohr .....	11
1.4.3	Auditiv-perzeptive Stimmklangbeurteilung .....	11
1.4.4	Apparative Diagnostik.....	12
1.4.4.1	Stroboskopie .....	12
1.4.4.2	Optische Kohärenztomographie .....	13
1.4.4.3	Aerodynamische Analyse .....	13
1.4.4.4	Akustische Stimmschallanalyse .....	13
1.4.5	Selbsteinschätzung durch Patienten.....	14
1.5	Computergestützte Stimmschallanalyse .....	15
1.5.1	Heiserkeitsanalyse, Phonetographie, DSI und PLI.....	15
1.5.2	Kommerziell verfügbare Stimmanalysegeräte.....	17
1.5.2.1	lingWAVES.....	17
1.5.2.2	Göttinger Heiserkeitsdiagramm (rpSzene®) .....	17
1.5.3	Vor- und Nachteile .....	18
1.5.4	Akustische Maße zur Stimmgütebeschreibung .....	19
1.5.4.1	Jitter .....	19
1.5.4.2	Shimmer .....	21
1.5.4.3	Glottal-to-Noise Excitation Ratio .....	22
1.6	Fragestellung.....	24

<b>2</b>	<b>Material und Methoden .....</b>	<b>25</b>
2.1	Studienbeschreibung .....	25
2.2	Material.....	26
2.2.1	Kommerziell verfügbare Stimmanalysegeräte.....	26
2.2.1.1	Göttinger Heiserkeitsdiagramm .....	26
2.2.1.2	lingWAVES.....	26
2.3	Methoden .....	28
2.3.1	Aufnahme .....	28
2.3.2	Statistik.....	33
2.3.2.1	Bland-Altman Plot.....	33
<b>3</b>	<b>Ergebnisse .....</b>	<b>35</b>
3.1	Studienteilnehmer.....	35
3.2	Auswertung der einzelnen Stimmparameter .....	38
3.2.1	Jitter .....	38
3.2.2	Shimmer .....	39
3.2.3	GNE .....	41
3.3	Vergleich beider Stimmschallanalysegeräte mittels Bland-Altman Plots .....	43
3.3.1	Jitter .....	43
3.3.2	Shimmer .....	46
3.3.3	GNE .....	48
<b>4</b>	<b>Diskussion .....</b>	<b>51</b>
4.1	Methodendiskussion .....	51
4.2	Ergebnisdiskussion.....	52
4.3	Diskussion der Fehlermöglichkeiten.....	58
4.4	Schlussfolgerung .....	63
<b>5</b>	<b>Zusammenfassung.....</b>	<b>65</b>
<b>6</b>	<b>Anhang .....</b>	<b>69</b>
6.1	Informationsblatt und Einverständniserklärung .....	69
6.2	Studienteilnehmer.....	73
6.3	Algorithmen .....	75
6.4	Ergebnistabellen.....	77

6.4.1	Jitter .....	77
6.4.2	Shimmer .....	78
6.4.3	GNE .....	79
6.5	MedCalc®: Bland-Altman Plots.....	80
<b>7</b>	<b>Abkürzungsverzeichnis .....</b>	<b>95</b>
<b>8</b>	<b>Verzeichnis der Akademischen Lehrer.....</b>	<b>97</b>
<b>9</b>	<b>Danksagungen.....</b>	<b>98</b>
<b>10</b>	<b>Literaturverzeichnis .....</b>	<b>99</b>

# 1 Einleitung

## 1.1 Kommunikation

Die Entwicklung einer Gesellschaft hängt stark von dem Umfang und der Geschwindigkeit des Informationsaustauschs ab. Nicht nur beim Speichern von Informationen, sondern auch bei ihrer Verbreitung werden immer wieder neue Wege und Mittel gesucht, um diese schneller, effektiver und in größerem Maße aufnehmen, verarbeiten und weitergeben zu können. Das dazu verwendete Trägermedium ist Sprache. Sie kann sowohl verbal und als auch nonverbal eingesetzt werden. Die Stimme als elektroakustischer Wandler nimmt innerhalb der verbalen Kommunikation eine wichtige Rolle ein [80]. Als Stimmgebung (Phonation) ist sie ein Teil der Sprachbildung und übersetzt Informationen in akustische Klänge. Durch die Modulation der Stimmbildung, wie zum Beispiel die Änderung der Stimmhöhe, Tonhöhe oder Lautstärke, werden zusätzlich nonverbale Informationen transportiert.

Im Zuge eines veränderten Kommunikationsbedarfes in allen Lebensbereichen moderner Gesellschaften steigen die stimmlichen Anforderungen zunehmend [20]. Der wachsende Anteil an Dienstleistungen in der Arbeitswelt und die ständig neuen Möglichkeiten der Kommunikation in der Freizeitgestaltung seien hier exemplarisch genannt. Des Weiteren spielen frühere Kommunikationsbarrieren, darunter vor allen Dingen die räumlichen Distanzen, aufgrund der technischen Entwicklungen heute kaum noch eine Rolle. Beides führt zu einer deutlich gestiegenen Stimmbelastung sowohl in Ausbildung und Beruf als auch im privaten Bereich [82].

Der Leistungs- und Erfolgsdruck im Beruf verlangt einen erheblichen Einsatz der Stimme, der letztendlich das Risiko für funktionelle Stimmstörungen deutlich erhöht. Dies betrifft vorrangig die stimmintensiven Berufe wie Sänger oder Schauspieler, bei denen aufgrund des Konkurrenzdrucks bereits kleine Veränderungen der Stimmqualität das Ende der beruflichen Laufbahn bedeuten können [23]. Auch die nach Koufman, Isaacson und Stemple zweite Stimmberufsgruppe, bestehend aus Lehrern, Geistlichen, Dozenten und Politikern, ist großen Stimmbelastungen ausgesetzt [48; 84]. Besonders Lehrer und Erzieher stehen im Mittelpunkt, weil ihre alltäglichen Anforderungen sehr hoch sind, sie

allerdings stimmtechnisch meist nicht genügend darauf vorbereitet werden [80; 82].

Zudem beeinflusst der Klang einer Stimme die Außendarstellung einer Person und damit ihren beruflichen Erfolg. Eine Studie des Netzwerks „Stimme.at“ konnte zeigen, dass angenehm klingende Stimmen als vertrauenserweckend bzw. sympathisch empfunden und trainierte Stimmen mit Durchsetzungsvermögen, Kompetenz und Autorität in Verbindung gebracht werden [43]. Dies verdeutlicht, dass sowohl die Belastbarkeit als auch die Qualität einer Stimme den Schlüssel zu einer erfolgreichen Kommunikation darstellen [34].

## 1.2 Stimmfunktion

### 1.2.1 Physikalisch-akustische Grundlagen

Für die gelungene interindividuelle Kommunikation müssen Informationen codiert werden. Dies geschieht, indem der menschliche Stimmapparat Schallwellen generiert und elektrische in akustische Informationen übersetzt. Intraindividuell findet Kommunikation häufig direkt bzw. über kurze Strecken statt (Neurone, elektromechanische Kopplung etc.). Schallwellen ist es möglich, sich über größere Distanzen und in wechselnden Medien auszubreiten und so von anderen Individuen auditiv aufgenommen, erneut decodiert und zentral verarbeitet zu werden.

Schall, physikalisch gesehen eine Druckschwankung, geht von einer Schallquelle aus, die sich als Schwingung fortpflanzt. Die einfachste Form des Schalls ist der Ton, der eine sich periodische ausbreitende Form einer Schallwelle darstellt. Die Eigenschaften eines Tons werden durch die Tonhöhe und Lautstärke beschrieben. Die Tonhöhe wird durch die Anzahl der Schwingungen pro Zeiteinheit (Schwingungen/Sekunde; Einheit: Hertz (Hz),  $1 \text{ Hz} = 1/\text{s}$ ) wiedergegeben und Frequenz ( $f$ ) genannt. Die Lautstärke ist gleichbedeutend mit dem verwendeten Schalldruck bzw. der Schallenergie und stellt die Auslenkung der Schwingungen einer Schallwelle dar. Sie wird als Amplitude bezeichnet und in der Einheit Dezibel (dB) angegeben [33; 82].

Treten mehrere periodische Schwingungen gemeinsam auf, deren Frequenzen zugleich ganzzahlige Vielfache der Grundschwingung darstellen, so spricht man von einem Klang. Die Grundschwingung wird als Grundton bezeichnet und ist für die Tonlage des Klangs verantwortlich. Die mitschwingenden, harmonischen Teiltöne werden Obertöne genannt. Ein Geräusch setzt sich aus nicht-periodischen Schwingungen zusammen und lässt sich nicht in diese Einzelkomponenten zerlegen [33; 82].

Ein Resonator ist ein schwingungsfähiger Körper, der durch eine Schallquelle in Schwingung versetzt wird. Abhängig von seinen Eigenschaften weist jeder Resonator bestimmte Eigenfrequenzen auf. Dadurch ist er in der Lage, diejenigen Teiltöne, die mit seinen Eigenfrequenzen übereinstimmen, zu verstärken und die anderen Teiltöne zu dämpfen. Dieser Vorgang nennt sich Resonanz, die verstärkten Teiltöne heißen Formanten. Die Anzahl der verstärkten Teiltöne,



ihre Lage und die Form der Verstärkung bzw. Dämpfung machen die Klangfarbe einer Stimme aus [33].

### **1.2.2 Anatomie**

Der Larynx bildet das Verbindungsstück zwischen dem Laryngopharynx und der Luftröhre [90]. Er besteht aus einem Knorpelgerüst, welches durch elastische Membranen, Bänder und Muskeln zusammengehalten wird [19]. Der Schildknorpel (Cartilago thyroidea), der Ringknorpel (Cartilago cricoidea) und die Epiglottis sind unpaarig angelegt; sie bilden das Hauptgerüst. Die Aryknorpel, die Wrisberg-Knorpel und die Santorini-Knorpel sind paarig vorhanden, wobei die Aryknorpel – auch Stellknorpel genannt – aufgrund ihrer Funktion bei der Stimmbildung eine bedeutende Rolle einnehmen. Die äußeren Kehlkopfmuskeln sind für die Hebung, Senkung und Fixation des Kehlkopfes zuständig, die inneren Muskeln dienen der Einstellung der Stimmritze [5; 90].

Innerhalb der Glottis befinden sich die Stimmlippen (Plicae vocales), die aus Schleimhautfalten inklusive elastischen Bändern und Muskeln bestehen. Um die Bedeutung des Aufbaus der Stimmlippen für die Stimmfunktion aufzuzeigen, entwickelte Hirano (1981) das „Cover-Body-Modell“ [38]. Durch die Kontraktion der Stimmuskeln werden die Stimmbänder gespannt. Die Anspannung anderer innerer Kehlkopfmuskeln führt zur Rotation der Aryknorpel und damit zum Schluss bzw. zur Öffnung der Stimmritze [90; 94].

### **1.2.3 Physiologie**

Der Larynx befindet sich in dem Bereich des Halses, in dem sich der Atem- und der Schluckweg kreuzen. Primär ist er für die Respiration und Deglutition und sekundär für die Phonation zuständig. Durch die geöffnete Epiglottis und die weitgestellte Stimmritze ist ein Ein- bzw. Ausströmen der Atemluft möglich. Ein aspirationsfreies Schlucken gelingt dadurch, dass der Larynx beim Schluckakt eine superiore-anteriore Bewegung vollzieht und durch die Epiglottis verschlossen wird [19; 33; 82].

Die Stimmbildung wird auf myoelastische und aerodynamische Wirkprinzipien zurückgeführt [92]. Demnach kommt es mit Hilfe der Expiration durch Anbla-

sen der geschlossenen Glottis zu Schwingungen der Stimmlippen [19; 33]. Der subglottische Druck überschreitet den Glottiswiderstand und bewirkt dadurch das Öffnen der Glottis. Die hindurchströmende Luft verursacht eine Sogwirkung auf die Stimmlippen und führt als aerodynamische Komponente (Bernoulli-Effekt), zusammen mit myoelastischen Rückstellkräften der Larynxmuskeln, zum Verschluss der Glottis. Dieser Zyklus wiederholt sich vielfach pro Sekunde, so dass es zu periodischen Verdichtungen und Verdünnungen der Luftsäule kommt, was sich letztendlich als Schallwellen manifestiert [19; 33]. Welche Form von Stimmlippenschwingungen auftreten, hängt von dem erzeugten subglottischen Druck und von der Einstellung der Kehlkopfmuskeln, im Speziellen von Masse, Länge und Spannung der Stimmlippen, ab. Die Tonhöhe wird durch den Spannungszustand der Stimmlippen bestimmt [19].

Der Grundton eines Klangs wird durch die Grobschwingung der Stimmlippen gebildet, während die Obertöne durch Randbewegungen, Randkantenverschiebungen und Luftverwirbelungen entstehen. Nach Erzeugung des Primärklangs durch die Stimmlippenschwingungen dienen die Luftwege oberhalb der Stimmritze, auch als Ansatzrohr bezeichnet, zur weiteren Formgebung. Das Ansatzrohr weist dabei zwei Funktionen auf: Als Resonator besitzt es eine phonatorische und als Instrument zur Lautbildung eine artikulatorische Komponente [33].

#### **1.2.4 Einflussfaktoren**

Es gibt viele Faktoren, die auf unsere Stimme einwirken. Manche von ihnen sind grundsätzlich anatomisch vorgegeben und verleihen einer Stimme ihren unveränderlichen, individuellen Charakter. Das gilt z.B. für die Indifferenzlage und die Stimmgattung sowie zum Teil für das Klangvolumen und die Leistungsfähigkeit einer Stimme [33].

Daneben spielen weitere Faktoren eine Rolle, welche im Hinblick auf ihre vokalen Effekte variieren können. Zum Beispiel die Körperhaltung sowie speziell der Spannungszustand der Schulter-Nacken-Muskulatur wirken sich direkt auf die Stimmbildung aus. Die Ursache für diese Beeinflussung liegt in der Verbindung der Larynxmuskeln mit den benachbarten Muskelketten, weshalb sie auf Fehlhaltungen und Verspannungen mit veränderten Kontraktionen reagieren. Ein weiteres Merkmal ist die Persönlichkeit, die sich besonders im Stimmklang wi-

derspiegelt. Je nach charakterlicher und soziokultureller Prägung benutzt ein Mensch unterschiedliche Mittel der Stimmmodulation. Auch eine Gesprächssituation oder der Gesprächspartner haben Einfluss auf die Stimmbildung, die neben der verbalen noch ein wesentliches Instrument der nonverbalen Kommunikation darstellt [33].

## 1.3 Stimmstörungen

Die Voraussetzung für eine gesunde Stimme ist die perfekte Koordination von Nerven und Kehlkopfmuskeln. Ohne sie wären ein gleichförmiges Schwingungsverhalten der Stimmlippen bzw. ein perfekter Glottisschluss gar nicht möglich, und folglich auch kein klarer Primärklang. Bereits kleinste Störungen, darunter geringe Masseverschiebungen, leichte Gewebeveränderungen oder asymmetrische Muskelkontraktionen, können in diesem komplexen System zum Qualitätsverlust der Stimme führen [39; 80].

Kommt es zu Stimmklangveränderungen, zu Einschränkungen der stimmlichen Leistungsfähigkeit bzw. Belastbarkeit und zu subjektiven Missempfindungen, spricht man von einer Stimmstörung [94]. Dysphonie wird als Oberbegriff für alle Arten von Stimmstörungen bezeichnet [25]. Heiserkeit ist das Kardinalsymptom und lässt sich als zusätzlicher Geräuschanteil in der Stimme nachweisen [33; 80]. Sie entsteht entweder durch einen mangelnden Glottisschluss; dabei handelt es sich um additives Rauschen bzw. auditiv gesehen um Behauchtheit. Oder sie wird durch bestehende irreguläre Stimmlippenschwingungen verursacht. Dies wird als multiplikatives Rauschen bzw. auditiv als Rauheit bezeichnet [94]. Eine verminderte Leistungsfähigkeit und Belastbarkeit der Stimme äußert sich durch Einschränkungen der Stimmlautstärke und der Dauer des Stimmgebrauchs. Zu den laryngealen Missempfindungen zählen Parästhesien, die sich vom Kratzen im Hals bis hin zu Schmerzen beim Sprechen bemerkbar machen können [94].

### 1.3.1 Einteilung der Dysphonien

Es ist bis heute nicht gelungen, eine weltweit einheitliche Klassifikation zu verankern. Einer der Gründe findet sich in der Vielfalt der Stimmstörungen: Ihre zahlreichen Übergänge und Mischformen lassen sich in der Praxis nur bedingt in ein einheitliches Schema einordnen. Die Nordamerikaner und die Europäer verwenden grundsätzlich unterschiedliche Klassifikationen [11; 94]. In Nordamerika haben sich die Einteilungen nach Morrison und Rammage (1993), Stemple et al. (2000) und Rammage et al. (2001) durchgesetzt [61; 77; 84], in Europa hingegen werden die Dysphonien klassischerweise in organische und

funktionelle Stimmstörungen unterteilt. Jedoch ist diese Einteilung nicht vollkommen korrekt, da beide Formen auch gleichzeitig auftreten bzw. ineinander übergehen können.

### **1.3.2 Funktionelle Dysphonien**

Treten bei Erkrankungen der Stimme eine Störung des Stimmklangs, der Leistungsfähigkeit und/oder subjektive Missempfindungen ohne krankhafte, primär strukturell-organische Veränderungen auf, werden diese zu den funktionellen Dysphonien gezählt. Sie resultieren aus einer Dyskoordination der Funktionsabläufe im Kehlkopf [11]. Der inadäquate Stimmgebrauch führt zu einem Ungleichgewicht zwischen glottischem Widerstand und Atemdruck und damit zu pathologischen Stimmlippenschwingungen [33; 80]. Ätiologisch geht man von einem multifaktoriellen Geschehen aus, wobei fünf Hauptkomponenten im Vordergrund stehen. Man unterscheidet somatogene, konstitutionelle, habituelle, phonogene, psychogene und symptomatische Faktoren [25; 94]. Die funktionellen Dysphonien können als hyperfunktionelle und hypofunktionelle Varianten auftreten.

Die hyperfunktionelle Dysphonie ist häufiger und durch einen vermehrten Kraftaufwand bei der Phonation gekennzeichnet. Dabei treten subjektive Missempfindungen wie Räusperzwang, Mundtrockenheit sowie eine frühzeitige Stimmermüdung auf. Die Stimme klingt rau, heiser und knarrend. Eine extreme Form stellt die Taschenfaltenstimme dar, bei der sich die Taschenfalten aufgrund enormer Anspannung der Kehlkopfmuskulatur berühren und selbst in Schwingung geraten [25; 33].

Die seltenere hypofunktionelle Dysphonie ist Folge einer Stimmermüdung, bei der sich der glottische Widerstand verringert und ein unvollständiger Glottisschluss auftritt. Als Symptome zeigen sich eine leise und behauchte Stimme [11; 25; 33].

In der Praxis lassen sich die Stimmstörungen nicht immer eindeutig zuordnen, so dass neben den gemischten Dysphonien auch sekundär hyper- und hypofunktionelle Dysphonien möglich sind. Ist eine hyperfunktionelle Dysphonie längerfristig vorhanden, kann dies sekundäre organische Veränderungen hervorru-

fen. Hyperämie, Stimmlippenknötchen, -Polypen und Kontaktgranulome seien hier genannt [11; 25].

Berufsbedingte Dysphonien werden durch endogene Faktoren wie geringe stimmliche Leistungsfähigkeit oder mangelhafte Stimmtechnik und durch exogene Faktoren wie vermehrte Stimmbelastung bei hohem Lärmpegel oder psychische Belastung verursacht [25; 80]. Sie sind häufig und sollten bereits in der Ausbildung durch prophylaktische Maßnahmen wie phoniatische Tauglichkeitsuntersuchungen und Stimmhygiene vermieden werden [25; 33].

### **1.3.3 Organische Dysphonien**

Organische Dysphonien lassen sich auf morphologische Veränderungen der Glottis und auf die Beeinträchtigung der Nervenversorgung zurückführen [94]. Sie sind entweder erworben oder angeboren. Letztere beruhen auf Anomalien des Kehlkopfs und des Atemtraktes und werden Dysplastische Dysphonien genannt.

Unter den erworbenen Stimmstörungen sind die entzündlichen Ursachen, wie etwa die akute und chronische Laryngitis, die häufigsten Erkrankungen [94]. Stimmlippenlähmungen werden durch gestörte Nervenfunktionen verursacht und sind laut Böhme (2003) am ehesten durch das Bronchialkarzinom und die Strumektomie bedingt [11]. Bei einer länger als drei Wochen bestehenden Heiserkeit muss immer an ein Larynxkarzinom gedacht werden. Es entsteht mehrheitlich an den Stimmlippen und gehört ebenso wie die gutartigen Larynxtumoren zu den weiteren Ursachen der organischen Dysphonien [25]. Neben der Strumektomie können andere operative Eingriffe wie zum Beispiel die Chordektomie bei dem oben erwähnten Glottiskarzinom zu massiven Stimmstörungen führen. Traumatische Stimmstörungen entstehen vorwiegend durch mechanische Gewalteinwirkung oder chemische Noxen [94]. Im klinischen Alltag stehen die Intubationsfolgen wie das Intubationsgranulom oder die (seltene) Aryknorpelluxation im Vordergrund [11].

## 1.4 Diagnostische Methoden

Die Stimme als zentrales Instrument der Kommunikation erfüllt mehrere Aufgaben und wird durch unterschiedlichste Faktoren beeinflusst. Eine ausführliche, morphologisch und funktionell orientierte Stimmdiagnostik zur Überprüfung der Stimmfunktion ist daher unverzichtbar [19].

Neben der Anamnese und der Untersuchung des Kehlkopfes bildet das Basisprotokoll der Europäischen Laryngologischen Gesellschaft (ELS) die Grundlage für eine systematische Stimmdiagnostik [26; 75].

	<i>normal</i>	<i>pathologisch</i>		
<b>Perzeption</b>				
Heiserkeit	0: nicht vorhanden	1: geringgradig	2: mittelgradig	3: hochgradig
Behauchtheit	0: nicht vorhanden	1: geringgradig	2: mittelgradig	3: hochgradig
Rauigkeit	0: nicht vorhanden	1: geringgradig	2: mittelgradig	3: hochgradig
<b>Videolaryngostroboskopie</b>				
Amplitude*	0: normal weit (ca. 1/3 der sichtbaren Stimmlippenbreite)	+ 1: geringgradig erweitert - 1: geringgradig verkürzt	+ 2: mittelgradig erweitert - 2: mittelgradig verkürzt	+ 3: durchschlagend - 3: aufgehoben (phonatorischer Stillstand)
Randkantenverschiebung	0: normal (mind. 1/2 der sichtbaren Stimmlippenbreite)	1: gering vermindert	2: mittelgradig vermindert	3: aufgehoben (phonatorischer Stillstand)
Symmetrie	0: normal (symm. Schwingungen nach Ort und Phase)	1: gering asymmetrisches Schwingungsmuster	2: mittelgradig asymmetrisches Schwingungsmuster	3: hochgradig asymmetrisches Schwingungsmuster
Regularität	0: regulärer Schwingungsablauf	1: gering irregulärer Schwingungsablauf	2: mittelgradig irregulärer Schwingungsablauf	3: hochgradig irregulärer Schwingungsablauf
Glottisschluss	0: vollständiger Glottisschluss	1: geringgradig unvollständiger Glottisschluss	2: mittelgradig unvollständiger Glottisschluss	3: hochgradig unvollständiger Glottisschluss
	Form des unvollst. Glottisschlusses	durchgehender Spalt ovalärer Spalt	dorsales Dreieck** sanduhrförmig	anteriorer Spalt irregulär
supraglottische Kontraktion b. d. Phonation*	0: keine supraglottischen Kontraktionen	1: geringe supraglottische Kontraktionen	2: ausgeprägte supraglottische Kontraktionen	3: supraglottische(r) Verschluss bzw. Phonation
<b>aerodynamische Messungen</b>				
Tonhaltdauer	norm.: > 15 sek.	path.: < 10 sek.		
Phonationsquotient	norm.: < 0,2 l/sek.			
<b>akustische Messungen</b>				
Jitter***	norm.: m: 0,59% w: 0,63%	path.: > 1,0%	über 5% nicht sinnvoll	
Shimmer***	norm.: m: 2,53% w: 2,0%	path.: > 4,0%	über 25% nicht sinnvoll	
leiseste Intensität	norm.: < 55 dB (A)			
lauteste Intensität*	norm.: > 90 dB (A)			
Stimmdynamik	norm.: > 40 dB (A)			
tiefste F <sub>0</sub>	norm.: m: D (73 Hz) w: e (165 Hz)			
höchste F <sub>0</sub>	norm.: m: d1 (294 Hz) w: e2 (659 Hz)			
Stimmumfang	norm.: 24–36 HT	path.: < 12 HT		
mittlere Sprechstimmlage*	norm.: m: 100 Hz – 150 Hz (G-c) w: 200 Hz – 250 Hz (g-c1)			
<b>subjektive Selbstevaluation</b>				
Stimmqualität	0: Stimme wird subjektiv als ungestört empfunden	1: Stimme wird subjektiv als geringgradig gestört empfunden	2: Stimme wird subjektiv als mittelgradig gestört empfunden	3: Stimme wird subjektiv als hochgradig gestört empfunden
kommunikative Beeinträchtigung	0: keine kommunikative Beeinträchtigung	1: geringe Beeinträchtigung bei verstärkter Stimmbelastung; keine Beeinträchtigung in der alltäglichen, sozialen Kommunikation	2: starke Beeinträchtigung bei verstärkter Stimmbelastung; geringe Beeinträchtigung auch in der alltäglichen, sozialen Kommunikation	3: starke Einschränkung auch in der alltäglichen Kommunikation; Sozialkontakte beeinträchtigt

\* nicht in Originalprotokoll der ELS enthalten; \*\* geringgradig, bei Frauen normal; \*\*\* CSL bzw. MDVP, Kay Elemetrics (gerundet).

Abbildung 1.1: Basisprotokoll der ELS, modifiziert nach Friedrich und Dejonkere (2005)

Es wurde von Dejonckere et al. 2001 als ein multidimensionales und auf minimalste Anforderungen reduziertes Protokoll entwickelt und umfasst die perzeptive, objektive und subjektive Beurteilung der Stimmfunktion (vgl. Abbildung 1.1) [16].

### **1.4.1 Anamnese**

Die Anamnese steht am Beginn jeder Stimmdiagnostik und verfolgt neben der Erhebung der typischen Punkte wie Vorerkrankungen, Stimmerkrankungen in der Familie, Stimmbelastung im Beruf und vorhandene Umweltfaktoren das Ziel, die subjektiven und objektiven Ausmaße der Stimmbeschwerden zu quantifizieren. Während des anamnestischen Gesprächs bekommt der Untersucher erste Hinweise über Stimmklang und Artikulation, sowie über Körperhaltung, Mimik und Gestik des Patienten [11; 25].

### **1.4.2 Untersuchung von Kehlkopf, Atmung und Ansatzrohr**

Eine gründliche Untersuchung beinhaltet nicht nur die Beurteilung des Kehlkopfs, sondern sämtliche Bereiche des Atemtrakts. Dazu gehört die Auskultation und Perkussion der Lungen, die Kontrolle der Atmung und die Inspektion von Mund- und Nasenhöhle. Zur Organuntersuchung des Kehlkopfs zählen die äußere Inspektion und Palpation sowie die innere Begutachtung durch die indirekte Laryngoskopie. Letztere galt lange Zeit als die bedeutsamste klinische Untersuchungsmethode, mittlerweile wird sie in Deutschland zunehmend durch endoskopische Methoden abgelöst [5; 25; 94].

### **1.4.3 Auditiv-perzeptive Stimmklangbeurteilung**

Stimmklangveränderungen äußern sich in erster Linie durch Heiserkeit. Diese lässt sich (trotz vermehrt rechnergestützter Stimmschallanalyse) nach wie vor am sensibelsten durch das Ohr eines erfahrenen Untersuchers beurteilen [26]. Um die perzeptive Bewertung des Stimmklangs zu normieren, also untersucher- und erfahrungsunabhängiger zu gestalten, wurden verschiedene Schemata entwickelt [94].



Die GRBAS-Skala von Hirano (1989) umfasst folgende fünf Dimensionen: G: „overall grade or degree“ (Grad der Heiserkeit), R: „rough“ (Rauigkeit), B: „breathy“ (Behauchtheit), A: „asthenic“ (Asthenie/Kraftlosigkeit), S: „strained quality“ (Spannung) [12; 38]. Sie findet vor allen Dingen in Japan und in den USA Verwendung.

In Deutschland hat sich das RBH-Schema etabliert, welches auf den drei Merkmalen Rauigkeit, Behauchtheit und Heiserkeit beruht [68; 76]. Während der Grad der Heiserkeit als Basiskomponente dient, beschreibt die Rauigkeit zusätzliche Geräuschbeimengungen aufgrund irregulärer Stimmlippenschwingungen. Die Behauchtheit benennt additives Rauschen aufgrund eines insuffizienten Glottisschlusses in der Stimme [25; 76]. Es hat sich gezeigt, dass die Parameter S („strain quality“) und A („asthenic“) im Gegensatz zu den anderen dreien weniger zuverlässig sind, was letztlich die Bedeutung der RBH-Skala steigert [10; 101]. Beide Skalen verwenden für die Parameter je vier Schweregrade (von 0 = nicht vorhanden bis 3 = hochgradig).

Die auditive Stimmklangbeurteilung durch den Untersucher stellt die perzeptive Komponente im Basisprotokoll der ELS dar. Im Folgenden wird näher auf die objektiven Untersuchungsmethoden eingegangen. Diese setzen sich aus der visuellen, aerodynamischen und akustischen Stimmanalyse zusammen.

## **1.4.4 Apparative Diagnostik**

### **1.4.4.1 Stroboskopie**

Die Stroboskopie ist ein visuelles Untersuchungsverfahren, das zur Routinediagnostik in der Phoniatrie gehört [25; 79]. Sie ist die wichtigste Methode zur optischen Darstellung des Kehlkopfbefundes bei Stimmstörungen [5; 16; 35; 64].

Stimmlippenschwingungen sind aufgrund ihrer hohen Frequenz mit dem bloßen Auge nicht zu erkennen. Das Stroboskop ermöglicht jedoch die Beurteilung der Stimmlippen durch eine kurze, frequenzsynchrone Belichtung der Glottis, die unter Zuhilfenahme eines Kehlkopf- oder Luftschallmikrofons erfolgt. Bei exakter Synchronisation entsteht beim Untersucher der Eindruck eines stehenden Bildes. Durch eine beabsichtigte, geringfügige Phasenverschiebung verschmel-

zen die einzelnen Bilder zu einer scheinbaren Zeitlupenbewegung [5; 11; 16; 24; 25].

Die Stroboskopie dient der Differenzierung von Dysphonien, der makroskopischen Beurteilung von benignen oder malignen Prozessen, als Hilfestellung für Therapieindikationen sowie zur Verlaufskontrolle [25]. Sie erlaubt die Beurteilung von Glottisschluss, Symmetrie, Regularität und Randkantenverschiebung [16]. Neben diesen im Basisprotokoll der ELS genannten wurden von Böhme & Gross (2001) die weiteren Kriterien Amplitude, Periodizität, Phonatorischer Stillstand und Supraglottische Engstellung eingeführt [12]. Hanschmann und Berger (2009) schlugen das so genannte ARPS-Bewertungsschema vor, welches eine Graduierung der Stroboskopiebefunde ermöglicht, indem es die pathologischen Abweichungen der Merkmale „Amplitude“ (A), „Randkantenverschiebung“ (R), „Phase“ (P) und „Schlussphase“ (S) quantifiziert und schließlich zu einem Gesamtgrad der Auffälligkeit zusammenfasst [35].

#### **1.4.4.2 Optische Kohärenztomographie**

Zur Untersuchung der Stimmlippen gibt es ein weiteres, relativ neues glottographisches Verfahren – die Optische Kohärenztomographie (OCT). Diese nicht invasive Methode funktioniert mittels Laserabtastung und liefert Informationen über den feingeweblichen Aufbau der humanen Stimmlippen. Sie kann in Zukunft als additive Option zur Stroboskopie eingesetzt werden [52].

#### **1.4.4.3 Aerodynamische Analyse**

Der Phonation-Quotient (PQ) ist ein einfaches Maß, um die aerodynamische Leistungsfähigkeit eines Stimmapparates zu beschreiben [6, 10, 18, 52]. Er setzt sich aus der Vitalkapazität (VC) und der Tonhaldedauer (MPT) zusammen und wird in Milliliter pro Sekunde angegeben. Alternativ dazu kann die mittlere phonatorische Flussrate auch direkt über einen Pneumotachographen ermittelt werden [16].

#### **1.4.4.4 Akustische Stimmschallanalyse**

Akustische Stimmschallanalyse bedeutet eine objektive und nicht-invasive computergestützte Untersuchung der Stimme. Sie beinhaltet zunächst die Auf-

nahme und digitale Speicherung eines Stimmsignals und anschließend dessen Auswertung anhand speziell dafür entwickelter Stimmparameter.

Die Pertubationsmaße Jitter und Shimmer sowie die klinisch relevantesten Maße aus der Stimmumfangsprofilmessung (höchste und tiefste Stimmtonhöhe  $F_0$ , leiseste Intensität) bilden zusammen die akustische Komponente des Basisprotokolls der ELS [16].

Die genauen Funktionen bzw. Entwicklungen der computergestützten Analysemethoden sowie ihre Vor- und Nachteile werden in einem gesonderten Kapitel behandelt (vgl. dazu Kapitel 1.5).

### **1.4.5 Selbsteinschätzung durch Patienten**

Die subjektive Beurteilung der Erkrankung durch den Patienten selbst stellt die dritte Säule der Stimmdiagnostik im Basisprotokoll der ELS dar. Im Gegensatz zu den bisher aufgeführten, weitestgehend objektiven Untersuchungsverfahren soll hier die eigene Wahrnehmung der Stimmstörung und die damit verbundene Einschränkung der Lebensqualität quantitativ untersucht werden [16; 26; 66; 94].

Aus diesem Grund wurde der Voice Handicap Index (VHI) entwickelt [42]. Er ist ein Fragenbogen, der aus den drei Kategorien „funktionell“, „physisch“ und „emotional“ besteht und jeweils zehn Fragen enthält. Mit Hilfe einer Skala von 0 bis 4 kann der Patient angeben, inwieweit jede Frage bzw. Aussage auf seine Situation zutrifft. Insgesamt sind maximal 120 Punkte möglich [33].

Der Voice Handicap Index liegt auch in deutscher Fassung vor und stimmt mit dem Original weitestgehend überein [66]. Der Stimmstörungsindex (SSI) ist ein auf 12 Items reduzierter VHI und geht auf Nawka et al. (2003) zurück [69]. Durch den Umrechnungsfaktor 2,5 ermöglicht er eine statistisch signifikante Korrelation mit dem VHI, was wiederum eine internationale Vergleichbarkeit gestattet [32]. Rosen et al. (2004) haben einen VHI-10 entwickelt, der in ihrer Studie keine signifikanten Abweichungen der Resultate im Vergleich zu dem VHI 30 aufwies [81]. In einem weiteren Schritt gelang es Nawka et al. (2009), einen internationalen VHI 9i zu entwickeln, der im Gegensatz zu den bereits genannten reduzierten VHI 12 und 10 auf Daten aus mehreren Ländern basiert und somit internationale Gültigkeit beansprucht [65].

## 1.5 Computergestützte Stimmschallanalyse

Lange Zeit bestand die Diagnostik von Stimmstörungen lediglich aus der qualitativen Beurteilung durch den Untersucher. Mit der Einführung der computergestützten Stimmschallanalyse wurde jedoch eine neue Form der Diagnostik geschaffen, die es erlaubt, eine quantitative und objektive Analyse von Stimmsignalen vorzunehmen [94].

Die Anfänge der Stimmschallanalyse durch apparative Untersuchungsmethoden gehen auf Lieberman (1961) zurück, der den Zusammenhang zwischen Funktionsstörungen des Stimmapparates und Veränderungen des Stimmschalls durch Schwankungen der Perioden aufzeigen konnte [51]. Koike (1969) gelang dies für die Schwankungen der Amplituden [47]. Die hieraus abgeleiteten Stimmparameter Jitter und Shimmer (vgl. dazu Kapitel 1.5.4) bildeten den Startpunkt einer Reihe neuentwickelter akustischer Maße, die jeweils bestimmten Merkmalen des Stimmklangs zugeordnet werden [3; 64].

### 1.5.1 Heiserkeitsanalyse, Phonetographie, DSI und PLI

Ein Großteil der Stimmschallanalyse wird klinisch durch die Heiserkeitsanalyse und die Stimmumfangsprofilmessung abgedeckt.

Heiserkeit, bestehend aus Rauigkeit und Behauchtheit, wird durch Pertubations- und Rauschmaße beschrieben [33]. Parameter, die die Veränderung des Schwingungsverhaltens einer Stimme wiedergeben, werden Pertubationsmaße genannt. Zu ihnen wird neben den genannten Jitter und Shimmer zum Beispiel der mittlere Periodenkorrelationskoeffizient gerechnet. Dieser, auch als Mean Waveform Matching Coefficient (MWMC) bezeichnet, gibt an, inwieweit die Signalform einer Periode mit der der nachfolgenden übereinstimmt [50].

Die Behauchtheit ist ein Maß für den Rauschanteil in der Stimme und wird durch die so genannten Rauschmaße ausgedrückt. Hierzu zählen der Harmonics-to-Noise-Ratio (HNR), der Signal-to-Noise-Ratio (SNR), der Normalized-Noise Energy (NNE) und der Glottal-to-Noise Excitation Ratio (GNE). Sie alle verfolgen unterschiedliche Ansätze, um das Verhältnis von Rausch- zu Energie- bzw. Signalkomponente zu beschreiben [50; 102]. Die ers-

ten drei Parameter werden international verwendet, der Glottal-to-Noise Excitation Ratio besonders in Deutschland (vgl. dazu Kapitel 1.5.4.3) [94].

Die Leistungsfähigkeit der Stimme wird durch die Phonetographie ermittelt, bei der ein Stimmumfangsprofil erstellt wird, welches aus Tonhöhenumfang und Stimmdynamik besteht [94]. Abhängig von der Stimmtonhöhe (Abszisse) wird die Lautstärke der Stimme (Ordinate) aufgezeichnet und in Form des „Stimmfeldes“ dargestellt [63; 64].

Neben der Überprüfung der Leistungsfähigkeit eignet sich die Stimmumfangsprofilmessung bei Tauglichkeitsuntersuchungen für Stimmberberufe, zur Beurteilung einer Sängerstimme sowie zur Dokumentation von Behandlungsverläufen [94]. Des Weiteren erlaubt sie die Unterscheidung von hyper- und hypofunktionellen Dysphonien einschließlich sekundär organischer Veränderungen [1; 63].

Der Dysphonia Severity Index (Deutsch: Dysphonie-Schweregrad Index; DSI) wurde entwickelt, um ein objektives und quantitatives Korrelat zu der akustisch wahrgenommenen Stimmqualität herzustellen und dieses zugleich in einer Maßzahl auszudrücken. Er basiert auf der Kombination folgender Parameter: höchstmögliche Frequenz ( $F_0$ -High in Hz), niedrigste Intensität (I-Low in dB), maximale Tonhaltedauer (MPT in s) und Jitter (%). Die Berechnung erfolgt an Hand der folgenden Formel:

$$\text{DSI} = 0,13 \times \text{MPT} + 0,0053 \times F_0\text{-High} - 0,26 \times \text{I-Low} - 1,18 \times \text{Jitter} (\%) + 12,4$$

In der Regel liegen die Werte für den DSI zwischen -5 (schwere Stimmstörung) und +5 (normale Stimme) [94; 100].

Während Wuyts (2000) zunächst von einer Korrelation des VHI mit Ergebnissen des DSI ausging, konnte dies von Woisard et al. (2007), Hsiung et al. (2002) und Henry et al. (2010) nicht bestätigt werden [37; 40; 99; 100].

Mit dem Pathological Likelihood Index (Deutsch: Pathologischer Wahrscheinlichkeitsindex) haben Godino-Llorente et al. (2009) ein weiteres Maß für die Diskriminierung gesunder und pathologischer Stimmen entwickelt. Sie konnten aufzeigen, dass der PLI, welcher mit Hilfe von Methoden aus der automatischen

Spracherkennung, wie z. B. das Gaussian Mixture Model (GMM)<sup>1</sup> bzw. die Mel-Frequenz-Cepstrum-Koeffizienten<sup>2</sup> konstruiert wurde, eine sinnvolle Ergänzung zu den klassischen Stimmschallanalyseparametern darstellt [29].

## **1.5.2 Kommerziell verfügbare Stimmanalysegeräte**

Es stehen verschiedene Software-Programme zur Verfügung, die die vorgestellten Methoden bzw. die Stimmanalyse durch Stimmparameter teilweise oder komplett anbieten. In Deutschland werden u. a. lingWAVES von WEVOS - WeVOSYS, Forchheim, rpSzene<sup>®</sup> von Rheder/Partner, Hamburg, und DiVAS von Xion Medical, Berlin, eingesetzt. Zwei von ihnen sind für die vorliegende Studie relevant und werden im Folgenden näher vorgestellt.

### **1.5.2.1 lingWAVES**

lingWAVES von WEVOS - WEVOSYS stellt ein modernes Programm zur Stimmanalyse dar. Es ist eine Kombination aus herkömmlichen und neuen Methoden, welche in einem modularen Softwaresystem untergebracht sind. Die einzelnen Module werden über einen Patientenmanager verwaltet, der eine Datenbank für die patientenorientierte Verwendung beinhaltet. Dieses Konzept ermöglicht eine unkomplizierte Erweiterung der Komponenten auf Basis derselben Benutzer-Plattform. lingWAVES umfasst u. a. die Stimmumfangsprofilmessung, die Heiserkeitsanalyse, die Berechnung des Stimmstörungsindex (VHI-12), eine spektrale akustische Stimmanalyse und die kombinierte Echtzeitanalyse von Spektrographie, Tonhöhe und Energie. Es ist eine der am häufigsten verwendeten Stimmanalyse-Software im Bereich von Phoniatrie, HNO und Stimmtherapie [96; 97] .

### **1.5.2.2 Göttinger Heiserkeitsdiagramm (rpSzene<sup>®</sup>)**

Das Göttinger Heiserkeitsdiagramm (GHD), das 1998 von Michaelis, Fröhlich und Strube entwickelt wurde, beruht auf der Berechnung von Periodenkorrelation, Jitter, Shimmer und GNE [57]. Es erlaubt eine graphische Darstellung der

---

<sup>1</sup> = Gaussche Mischverteilung: Methode, zur Gruppierung (Clustering) von Regelmäßigkeiten bzw. Parametrisches Modell der Wahrscheinlichkeitsverteilung bei kontinuierlichen Messungen in biologischen Systemen, in diesem Fall bei der Spracherkennung [18; 78].

<sup>2</sup> Ein Satz von Merkmalen zur Spracherkennung [70].

Stimmgüte in zwei Ebenen und somit eine einfache Interpretation der Ergebnisse.

Die vier akustischen Größen wurden aus einer Menge von 22 akustischen Parametern ermittelt und gelten als optimale Kombination, um die Irregularität und das Rauschen der Stimme zu ermitteln. Die Irregularität wird durch die drei Größen Jitter, Shimmer und die mittlere Periodenkorrelation erfasst und bildet als so genannte Irregularitätskomponente die Abszisse im Heiserkeitsdiagramm. Der GNE als vierte Größe beschreibt das Verhältnis von Stimmlippen-schwingungen zu hörbaren Luftturbulenzen während der Stimmbildung, also von Klang zu Behauchtheit. Er dient als Grundlage für die Rauschkomponente, die als Ordinate abgebildet wird [27; 33; 58].

Die Abszisse zeigt Werte von 0 bis 10 und die Ordinate von 0 bis 5 an, wobei Werte von größer 5 bzw. 2,5 nicht mehr als normal angesehen werden. Die Irregularitätskomponente und Rauschkomponente werden als Wertepaar mit ihren Standardabweichungen im Heiserkeitsdiagramm aufgetragen. Dies ergibt eine Ellipse, deren Halbachsen von den Standardabweichungen der Irregularität und des Rauschens gebildet werden. Je größer die Ellipsen bzw. ihre Halbachsen sind, desto unbeständiger ist die aufgenommene Stimme [64]. Das Aufnahmeprotokoll beinhaltet verschiedene Phonationsmodi (normale, hohe und tiefe Stimmlage) von sechs gehaltenen Vokalen, wodurch eine hohe Aussagekraft über den Zustand einer Stimme gegeben ist [64].

### **1.5.3 Vor- und Nachteile**

Die rechnergestützte Untersuchung bietet neben der Echtzeitanalyse die Option, Sprachsignale zu speichern. Daraus ergeben sich mehrere Vorteile. Zum einen können die erhobenen Befunde dokumentiert werden und erleichtern die Diagnosestellung, Therapieplanung und Beurteilung des Behandlungserfolgs. Zum anderen eröffnen sie die Möglichkeit der intra- bzw. interindividuellen Vergleichbarkeit der Stimmgüte und schaffen somit eine gemeinsame Kommunikationsplattform für Ärzte und Therapeuten [49; 59].

Neben den genannten Vorteilen muss darauf hingewiesen werden, dass auch dieses diagnostische Instrument Grenzen aufweist. Jede Analyse repräsentiert

lediglich einen Ausschnitt der stimmlichen Leistungsfähigkeit und ist nur als eine Momentaufnahme zu bewerten [1; 25].

Nach wie vor werden akustische Stimmparameter zur Stimmschallanalyse häufig nicht durch fortlaufende Sprache, sondern durch gehaltene Vokale bestimmt. Aus diesem Grund sind die Ergebnisse einer solchen Analyse nicht direkt auf die Stimmqualität im Alltag übertragbar [28].

Die Ergebnisse der computergestützten Stimmschallanalyse sind nur bei Einhaltung von konstanten Untersuchungsbedingungen verwertbar. Dies gilt besonders für die Rahmenbedingungen wie Haltung und Position des Patienten, Ausschalten von Störgeräuschen und standardisierte Untersuchungsabläufe [59; 64].

Ein wesentlicher Faktor, der die stimmlichen Leistungen beeinflusst, ist die Motivation des Patienten. Diese lässt sich trotz aller Maßnahmen zur Objektivierung der Diagnostik nie unbeeinflusst festhalten und sollte daher bei der Interpretation von Ergebnissen immer berücksichtigt werden [94].

Grundsätzlich ist zu beachten, dass auch die computergestützte Stimmschallanalyse lediglich ein semiquantitatives Untersuchungsverfahren darstellt, da das aufgenommene Stimmsignal zur Bearbeitung in der Software durch den Untersucher markiert werden muss [82]. Sie gilt als ein wichtiges diagnostisches Mittel in der Phoniatrie, ist jedoch immer noch von der Erfahrung des Untersuchers abhängig [49; 59].

### **1.5.4 Akustische Maße zur Stimmgütebeschreibung**

Die Auswahl der im Folgenden vorgestellten akustischen Stimmparameter ist auf die für diese Studie relevanten Jitter, Shimmer und GNE beschränkt.

#### **1.5.4.1 Jitter**

Nicht-periodische Schallanteile einer Stimme entstehen u. a. durch die Irregularität der Stimmlippenschwingungen und werden als Perturbation bezeichnet. Der Jitter gehört zu den so genannten Perturbationsmaßen und lässt sich als ein Index für die Variationen in der Grundfrequenz eines Tons beschreiben [33].

Je größer die einzelnen Periodenlängen voneinander abweichen, desto größer wird der Wert des Jitters. Der Jitter eines Sinustons wäre dementsprechend



gleich null. Menschliche Stimmen weisen immer eine natürliche Variabilität auf, die auf den Herzrhythmus, die Nervenimpulse und die Struktur der Stimmlippen zurückgeführt werden [71; 86]. Übersteigen diese Periodenschwankungen einen bestimmten Wert, gilt der Jitter als pathologisch.

Um den Jitter aus einem Stimmsignal ermitteln zu können, müssen die Grundperioden erfasst werden. Hieraus wurden verschiedene Methoden entwickelt. Ein Ansatz ist das Auffinden von ausgezeichneten Schwingungszuständen wie das Zerocrossing als Durchgang des Signals durch die Nulllinie und das Peakpicking als Bestimmung des positiven bzw. negativen Maximums in einem Schwingungszyklus [56]. Neben diesen ereignisbasierten Methoden hat sich u. a. das Waveform-Matching als integrales Verfahren zur Bestimmung der Periodenlänge etabliert, welches die Messgenauigkeit durch parabolische Interpolation des Signalmaximums verbessert [88]. Letzteres wird im Göttinger Heiserkeitsdiagramm angewendet [56]. Anhand definierter Gleichungen lassen sich die Jitter-Werte aus den genannten Methoden ermitteln, wobei diese entweder als absoluter Jitter in Millisekunden oder relativ zu den verwendeten Periodenlängen als Jitter-% angegeben werden. Außerdem spielen die Fensterlänge, die Anzahl der gemittelten Perioden und die Anzahl der Analyseergebnisse eine Rolle.

Leider ist bis heute noch kein einheitliches Maß für die Grundperiodenbestimmung festgelegt, so dass die Berechnung des Jitters zwischen den einzelnen Stimmschallanalyse-Geräten variiert [25; 56; 82; 88]. Daher lassen sich in der Literatur keine allgemeingültigen Standard- bzw. Grenzwerte finden. Jitter-Werte von  $>1\%$  gelten zwar als pathologisch, doch lediglich mit der Einschränkung, dass sie im Zusammenhang mit dem Computerized Speech Lab (CSL) bzw. Multi-Dimensional Voice Program (MDVP) von Kay Elemetrics ermittelt wurden [11; 26]. Eine Studie von Uloza et al. 2005 mit 88 Probanden zeigte Jitter-%-Werte von durchschnittlich  $0,2\%$  (SD  $0,07$ ) auf, eine Studie von Petrovic-Lazic et al. 2010 (MDVP und CSL) mit 21 Probanden Werte von  $0,5\%$  (SD  $0,17$ ). Beide Studien wurden mit stimmgesunden Erwachsenen und im Frequenzbereich ihrer jeweiligen Indifferenzlage durchgeführt [74; 91].

Voraussetzung für die Berechnung des Jitters ist ein ausreichend langes Stimmsignal – je nach Analyse-Software ca.  $1000\text{ ms}$  –, wobei die Ein- und Ausschwingphase eines gehaltenen Tons nicht verwendet werden darf. Zudem

ist der Jitter von der Tonhöhe abhängig ist. Dementsprechend weisen Frauen mit einer höheren Grundfrequenz einen niedrigeren Jitter-Wert auf als Männer mit tieferen Grundfrequenzen [82].

Es gilt die generelle Einschränkung, dass extrem aperiodische akustische Signale für die akustische Analyse nicht eingesetzt werden können. Perturbationsmaße werden unzuverlässig, wenn das Sprachsignal Unterbrechungen, starke Subharmonische oder Modulationen enthält [16].

#### **1.5.4.2 Shimmer**

Der Shimmer gehört ebenfalls zu den Perturbationsmaßen. Im Gegensatz zum Jitter beschreibt er jedoch die Variationen der Schallamplituden. Das heißt, es werden die aufeinander folgenden Schwingungen bzw. die Energiemaxima auf Abweichungen verglichen. Analog zum Jitter sind geringe Amplitudenschwankungen physiologisch und lassen sich auf die gleichen Ursachen wie die oben genannten zurückführen. Mit Zunahme der Irregularität der Stimmlippenschwingungen bzw. der Heiserkeit steigt der Wert des Shimmers an [33; 82].

Auch beim Shimmer müssen die einzelnen Grundperioden zur Berechnung aufgesucht werden. Sie bilden die Grundlage, um die entsprechenden Amplitudenmaxima zu bestimmen oder deren Energie zu messen [56]. Der Shimmer wird entweder absolut in Dezibel oder relativ zu den verwendeten Periodenlängen in Prozent angegeben. Auch hier kann zurzeit kein einheitlicher Grenzwert erwähnt werden [59]. Von Friedrich (2005) wird er momentan mit 4 % angegeben, allerdings auch hier lediglich unter der Einschränkung, dass er mit CSL bzw. MDVP ermittelt wurde (vgl. 1.5.4.1 Jitter) [11; 26]. Die bereits genannte Studie von Uloza et al. 2005 ergab Shimmer-%-Werte von 1,7 % (SD 0,73) und die Studie von Petrovic-Lazic et al. 2010 Werte von 1,8 % (SD 0,44) [74; 91].

Brockmann et al. 2010 konnten zeigen, dass sowohl Jitter als auch Shimmer unterschiedlich stark von verschiedenen Faktoren beeinflusst wurden. Während der Schalldruckpegel (SPL) mehr Einfluss auf die beiden Stimmparameter hatte, war die Grundfrequenz  $f_0$  von geringer Bedeutung. Das Geschlecht und der gewählte Vokal wirkten sich nur wenig, jedoch immer noch klinisch relevant, auf Jitter und Shimmer aus. Die Autoren weisen darauf hin, dass dies durch die aktuellen Leitlinien noch nicht ausreichend berücksichtigt würde [14].

### 1.5.4.3 Glottal-to-Noise Excitation Ratio

Um den Grad der Heiserkeit einer Stimme zu erfassen, wird neben der Irregularität auch der Rauschanteil gemessen. Dieser Rauschanteil lässt sich durch den Glottal-to-Noise Excitation Ratio beschreiben, welcher 1997 von Michaelis et al. eingeführt wurde [58].

Gemeinsam mit Kollegen vom Dritten Physikalischen Institut der Georg-August-Universität in Göttingen entwickelte er ein Maß, das das Verhältnis von Stimmlippenanregung (Klang) zu turbulentem Rauschen (Behauchtheit) in der Stimm- bildung widerspiegelt. Ein wichtiges Kriterium war die Unabhängigkeit von den anderen bereits bestehenden Maßen zur Beschreibung der Aperiodizität einer Stimme [56; 82].

In diesem neuen Ansatz zur Ermittlung des Rauschverhaltens eines Stimmsignals wird dargestellt, inwieweit Hilbert-Einhüllende in unterschiedlichen Frequenzbändern miteinander korrelieren. Einhüllende werden in der Mathematik als Kurven beschrieben, die eine Schar gleichartiger Kurven oder Geraden einhüllen und diese per definitionem nur einmal berühren. Die Hilbert-Transformation geht auf David Hilbert zurück und ist eine lineare Integraltransformation, die aus einem reellen Signal ein analytisches Signal bildet. Die Hilbert-Einhüllenden werden aus dem Betrag dieses analytischen Signals berechnet [56; 58].

Da sich reelle Stimmsignale von gepulsten Signalen unterscheiden, liegt der Korrelationswert einer gesunden Stimme z. B. nur bei 0,7 statt 1,0. Durch eine Inverse Filterung, das heißt Glättung des Signalspektrums, und die Berechnung des Maximums einer Korrelationsfunktion zwischen -0,3 ms und 0,3 ms, lassen sich Korrelationswerte von nahezu 1 (in diesem Beispiel 0,98) erzielen. Für unregelmäßige bzw. aphone Stimmsignale (Flüsterstimme = starkes Rauschen) sinken die Werte deutlich bzw. sehr deutlich ab, so dass sich insgesamt drei Grade der Stimmqualitäten voneinander abgrenzen lassen. Durch Erhöhung der Anzahl an zu vergleichenden Hilbert-Einhüllenden und der Verwendung des Maximums aller Korrelationskoeffizienten lässt sich die Trennschärfe zwischen Glottisanregung und Anregung durch turbulentes Rauschen deutlich verbessern [56; 58].

Der GNE weist im Gegensatz zum Cepstral Harmonic to Noise Ratio (CHNR) und Normalized Noise Energy (NNE) den Vorteil auf, dass er kaum von Jitter-

oder Shimmer-Messungen abhängt [56–58]. Zudem liefert er zuverlässige Messungen in Bezug auf die Diskriminierung zwischen normalen und pathologischen Stimmen, vergleichbar mit den genannten Rauschmaßen NNE oder HNR, so dass dieser Parameter als Screening-Methode geeignet ist [30].

Des Weiteren ist er der einzige Parameter, der noch bei schwerer Heiserkeit und selbst bei aphonem Stimmen verwertbare Ergebnisse liefert [94].

## 1.6 Fragestellung

Die Anzahl der Patienten mit Stimmfunktionsstörungen nimmt merklich zu. Da eine Stimmstörung noch immer häufig zu spät diagnostiziert wird, findet auch die stimmliche Behandlung zu spät statt. Oft ergibt sich aus dieser Tatsache die Notwendigkeit des Berufswechsels oder der Berufsunfähigkeit. Es erhöht sich auch die Zahl der Stimmtherapien, insbesondere zur Erhaltung der Berufsfähigkeit bei Patienten, die beruflich einer hohen Stimmbelastung unterliegen [67].

Die Bewertung stimmlicher Leistungsparameter wird durch zur Verfügung stehende computergestützte Systeme erleichtert bzw. objektivierbar gemacht. Es existieren mehrere kommerziell verfügbare computergestützte Systeme zur Stimmschallanalyse. Leider ist bis heute noch kein einheitliches Maß für die Grundperiodenbestimmung festgelegt, so dass vor allem die Berechnung des Jitters und Shimmers zwischen den einzelnen Stimmschallanalyse-Geräten variieren kann. Dies stellt gerade in der klinischen Anwendung ein Problem dar, da die gewonnenen Untersuchungsergebnisse weder intraindividuell noch interindividuell sicher vergleichbar sind, wenn nicht jeweils das gleiche Untersuchungsgerät eingesetzt worden ist.

In der vorliegenden Studie wurden die Programme lingWAVES von WEVOS - WEVOSYS und rpSzene<sup>®</sup> von Rheder/Partner unter klinischen Bedingungen anhand dreier gängiger Stimmparameter miteinander verglichen, um mögliche Unterschiede zwischen zwei in Deutschland verbreiteten, kommerziell verfügbaren Stimmschallanalysegeräten aufzeigen zu können.

Dabei wurden folgende Fragestellungen untersucht:

1. Weisen die Analyseparameter Jitter, Shimmer und GNE von lingWAVES signifikante Unterschiede zu denen von rpSzene<sup>®</sup> berechneten Werten auf, wenn identische Stimmsequenzen verwendet werden?
2. Ergeben sich Hinweise auf systematische Abweichungen oder lineare Abhängigkeiten von den Ergebnissen beider Geräte und lassen sich Rückschlüsse auf mögliche Ursachen wie z. B. die unterschiedliche Grundperiodenbestimmung ziehen?

## 2 Material und Methoden

### 2.1 Studienbeschreibung

Die vorliegende prospektive Studie wurde im Zeitraum von Februar 2009 bis Juni 2010 durchgeführt. 60 Studienteilnehmer – davon 30 Stimmpatienten und 30 Probanden<sup>3</sup> – unterzogen sich einer Stimmschallanalyse, bei der die folgenden sechs Vokale [a], [ä], [e], [i], [o] und [u] als gehaltene Vokale synchron in zwei Stimmschallanalysegeräten phoniert wurden. Dabei handelte es sich um zwei in Deutschland gängige computergestützte Stimmschall-Analysegeräte – lingWAVES von WEVOS - WEVOSYS, Forchheim, und rpSzene<sup>®</sup> von Rheder/Partner, Hamburg.

Die gewonnenen Daten wurden anonymisiert abgespeichert. Im Anschluss erfolgte die parallele Auswertung der Stimmsequenzen in den beiden zu untersuchenden computergestützten Stimmschall-Analysegeräten und in einem weiteren Schritt die Gegenüberstellung der gewonnenen Ergebnisse.

Für die Stimmschallanalyse wurden die drei gängigen Stimmanalyseparameter Jitter, Shimmer und Glottal-to-Noise Excitation Ratio verwendet. Die Ergebnisse wurden in dem Programm MS Excel<sup>®</sup> – Redmond, USA – tabellarisch gespeichert und in Form von Säulendiagrammen graphisch dargestellt. Die statistische Auswertung erfolgte mit Hilfe von Bland-Altman Plots in dem Programm MedCalc<sup>®</sup> – Mariakerke, Belgien.

---

<sup>3</sup> Zur besseren Lesbarkeit wird im weiteren Verlauf der Arbeit die männliche Form verwendet. Die jeweiligen Begriffe beziehen die weibliche Form jeweils mit ein.

## 2.2 Material

### 2.2.1 Kommerziell verfügbare Stimmanalysegeräte

#### 2.2.1.1 Göttinger Heiserkeitsdiagramm

Für die Aufzeichnung des Göttinger Heiserkeitsdiagramms wird in der Klinik für Phoniatrie und Pädaudiologie in Marburg ein von der Firma Rehder/Partner GmbH, Hamburg, gelieferter handelsüblicher PC (Betriebssystem: Windows XP) mit vorinstallierter Software „rpSzene®“ in der Version 7.0e eingesetzt. Zu dem System gehört neben der Software ein Headset-Mikrofon mit Minimixer (AKG B29L), welcher über den Line-in Eingang der Soundkarte (Creative Sound Blaster-PCI) mit dem PC verbunden ist.

Die Aufnahmen erfolgten über den Monokanal des Minimixers. Für die Analyse der Stimmufnahmen waren die Abtastrate auf eine Frequenz von 44100 Hz und das Analysefenster auf eine Länge von 4 Sekunden bzw. auf einen Vor-schub von 0,25 Sekunden eingestellt. Der Faktor für die Unterabtastung diente der groben Vorberechnung der Periodenlänge und betrug 4.

#### 2.2.1.2 lingWAVES

Das lingWAVES Voice Diagnostic Center (VDC) + beinhaltet eine spezielle Hard- und Software der Firma WEVOS - WEVOSYS, Forchheim, und findet Anwendung in den Bereichen Phoniatrie, HNO-Heilkunde und Logopädie. Zur Ausstattung gehört neben der Software ein Mikrofon mit integriertem Schallpegelmessers (SPL meter, IEC 651 Typ2 / ANSI S1.4 Type 2). Die Übertragung der Daten des Schallpegelmessers zum PC erfolgt über einen seriellen Anschluss des Rechners, die Übertragung des Stimmschalls über den Line-in Eingang der Soundkarte [4; 96; 97].

Während der Studie wurde die Software in der Version 2.50.0043 verwendet, die auf einem handelsüblichen Computer (Betriebssystem: Windows XP) installiert war. Für die Studie wurden zwei Tools des VDC in Anspruch genommen.

Das Aufnahme-Tool „Recorder“ ist ein Oszillogramm mit Editier-Funktionen und dient zur Aufnahme und Weiterverarbeitung des Stimmsignals. Mit Hilfe der Zoom-Funktion konnten die Stimmsequenzen mit einer Präzision von 1/44100 Sekunden geschnitten werden.

Der „Vospector“ als Analyse-Tool basiert auf dem RBH-Index (Rauheit – Behauchtheit – Heiserkeit) und ermöglicht nach Einfügen eines Stimmsignals neben der Berechnung der Grundfrequenz auch die der akustischen Messwerte Jitter, Shimmer und GNE. Außerdem zeigt er auf Grundlage dieser Werte die Irregularität, das Rauschen und den Gesamtgrad einer Stimmstörung an [96; 97].

Für die Studie wurden die Aufnahmeeigenschaften auf eine Abtastrate von 44100 Hz und eine Abtasttiefe von 16 Bit sowie auf den Mono-Kanal eingestellt.



## 2.3 Methoden

### 2.3.1 Aufnahme

Nach der Aufklärung und dem Einverständnis der Patienten und Probanden zur Studienteilnahme wurde mit den Vorbereitungen für die Stimmaufnahmen begonnen. Die Mikrofone, d.h. das Headset und der Schallpegelmesser, wurden eingeschaltet und die Software – sofern noch nicht geschehen – auf beiden PCs gestartet.

Die Mikrofon-Abstände bis zum Mund des Studienteilnehmers wurden kontrolliert, da hier am ehesten mit systematischen Fehlern zu rechnen ist. Bei dem Headset-Mikrofon von rpSzene<sup>®</sup> war der Mund-Mikrofonabstand mit ca. 3 cm automatisch vorgegeben. Um die geforderten 30 cm zum Schallpegelmesser von lingWAVES einzuhalten, wurde ein Abstandhalter installiert, der die Position der Testperson festlegte.

Alle Aufnahmen wurden im Stehen aufgezeichnet. Das Mikrofon und der Abstandhalter waren jeweils auf einem höhenverstellbaren Stativ angebracht, so dass eine individuelle Anpassung an die Körperhöhe des jeweiligen Studienteilnehmers problemlos möglich war.

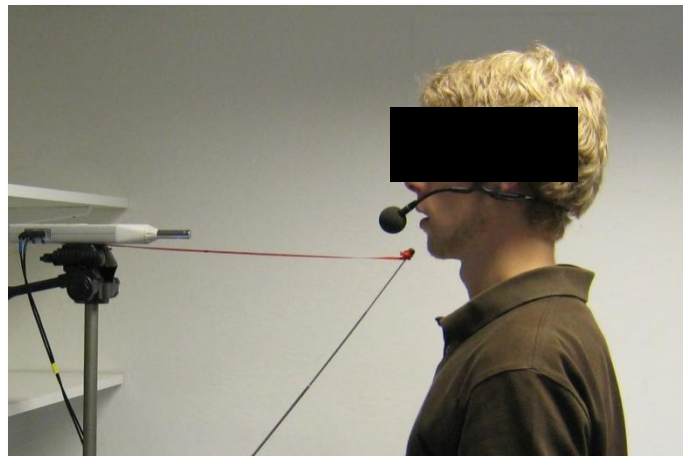


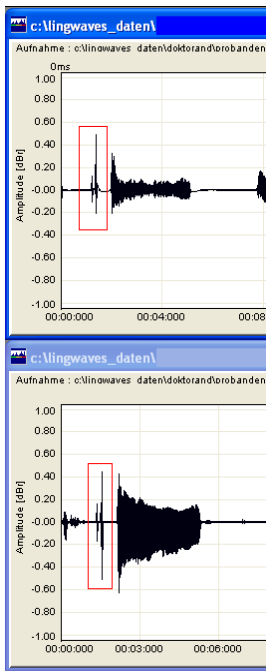
Abbildung 2.1: Versuchsaufbau mit Abstandhalter und beiden Mikrofonen

Es galt, sämtliche Störgeräusche zu minimieren, um einen Rauschpegel am Schallpegelmesser von < 45 dB zu erreichen. Die Tonhöhe und -lautstärke der

phonierten Vokale mussten vor der Aufnahme getestet werden. Die Patienten bzw. Probanden sollten die Vokale im Frequenzbereich ihrer jeweiligen Indifferenzlage phonieren. Bei den Männern liegt dieser Bereich bei 120 - 190 Hz, bei den Frauen um eine Oktave höher. Die Lautstärke sollte, wenn möglich, zwischen 65 und 75 dB liegen und wurde vor der Aufnahme kontrolliert.

Die Stimmaufnahmen wurden in beiden Programmen parallel gestartet. Vor den

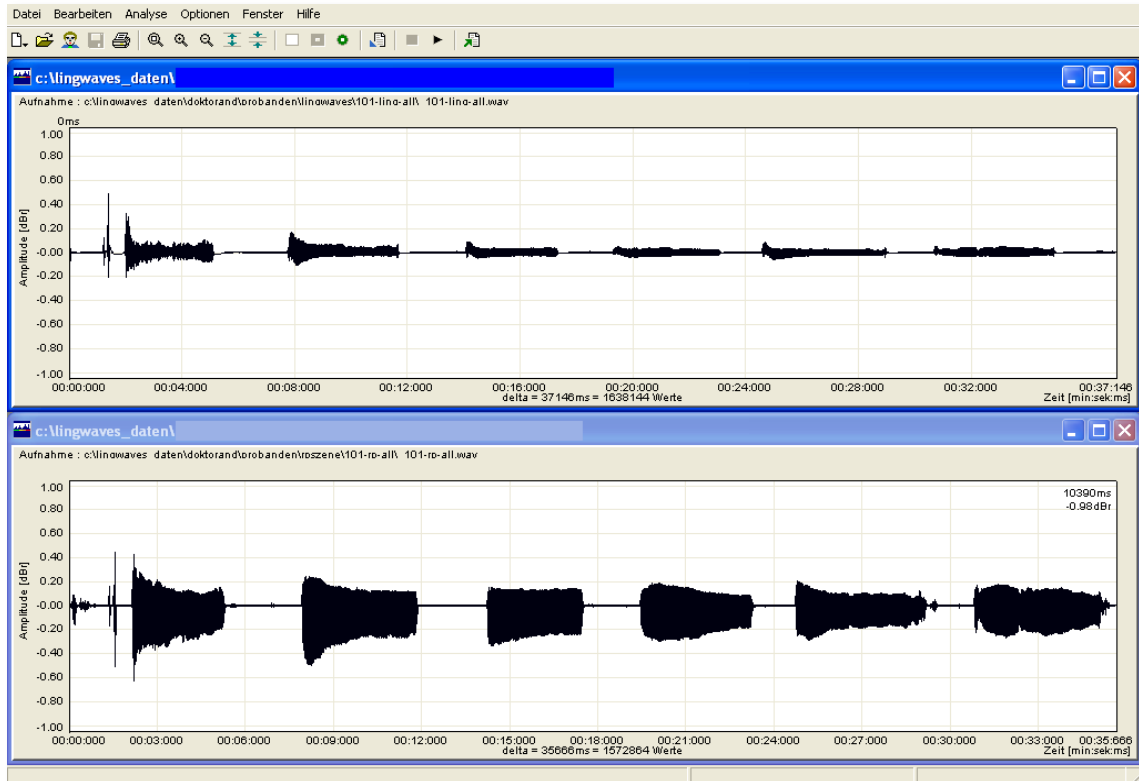
Phonationen wurde als erstes ein Signalton aufgenommen, der als Referenzpunkt dienen sollte. Dadurch konnten später die gleichen Stimmsequenzen aus beiden Geräten miteinander verglichen werden. Zur Erzeugung des Signaltons wurde ein Kugelschreiber verwendet, der durch das Einrasten der Mine ein kurzes Klickgeräusch verursachte. Dieses war als charakteristischer Ausschlag in beiden Aufnahmen sichtbar (vgl. Abbildung 2.2). Der Kugelschreiber musste in die Nähe des Mundes der Testperson gehalten werden, damit der Laufzeitunterschied der Signalaufnahme zwischen den beiden Mikrofonen denen des späteren Stimmschalls nachempfunden wurde.



**Abbildung 2.2:** Charakteristischer Signalton in beiden Aufnahmen als Referenzpunkt

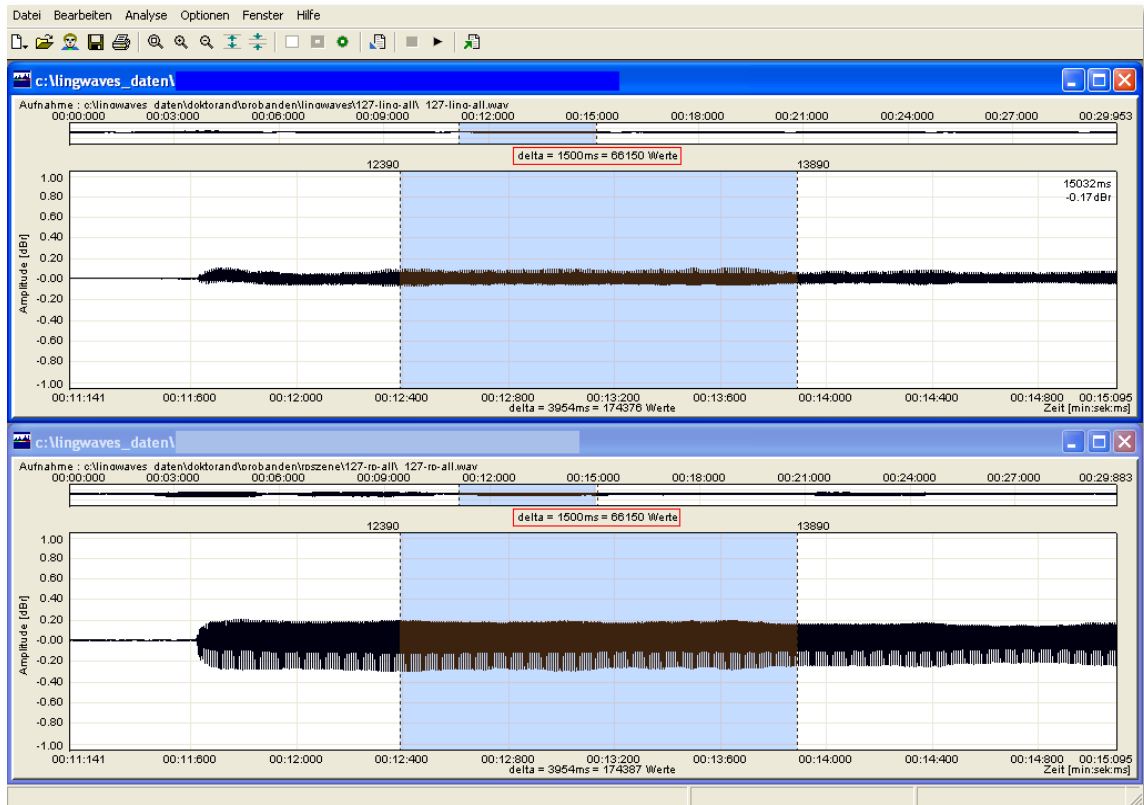
Die Testperson musste die sechs Vokale, [a], [ä], [e], [i], [o] und [u], synchron in beide Mikrofone phonieren und dabei jeden Vokal mindestens 3 Sekunden aushalten. Die Aufnahmen – jeweils ca. 30 Sekunden lang – wurden als Wave-Datei abgespeichert.

Zur Bearbeitung wurden die Aufnahmen von lingWAVES und rpSzene<sup>®</sup> in das Aufnahme-Tool von lingWAVES importiert (vgl. Abbildung 2.3).



**Abbildung 2.3:** Die synchron aufgenommenen Vokale derselben Testperson im Aufnahmemodul von lingWAVES; oben lingWAVES-Aufnahme, unten rpSzene<sup>®</sup>-Aufnahme

Sowohl der Signalton zu Beginn der Aufnahme als auch das Aufnahme-Tool mit seiner Präzision von 1/44100 Sekunden ermöglichten es, die Vokale in nahezu identische Stimmsequenzen mit einer Länge von 1.500 ms zu schneiden. Die Gefahr, dass die Aufnahmen dabei in zeitlich verschobenen Sequenzen hätten geschnitten werden können, war aufgrund des oben genannten Verfahrens mit einer Präzision von deutlich weniger als 1 Millisekunde vernachlässigbar gering. Bezogen auf eine Sequenzlänge von 1500 ms entsprach dies einem Wert von  $1,5 \cdot 10^{-6}$  (0,00015 %), so dass die Stimmsequenzen in diesem Fall praktisch als identisch bezeichnet werden konnten (vgl. Abbildung 2.4).



**Abbildung 2.4: Identische Stimmsequenz eines Vokals aus beiden Stimmchallanalysegeräten; oben lingWAVES-Aufnahme, unten rpSzene®-Aufnahme**

Bei 60 Aufnahmen mit jeweils sechs Vokalen ergaben sich für jedes Stimmchallanalysegerät 360 Stimmsequenzen, die einzeln als Wave-Dateien exportiert wurden.

Im nächsten Schritt wurden diese Stimmsequenzen im Analysemodus der jeweiligen Stimmchallanalyse-Software auf ihre Stimmgüte hin untersucht. Dazu wurden die geschnittenen Aufnahmen dem ursprünglichen Aufnahmegerät zugeordnet und dort anhand der Parameter Jitter und Shimmer, beide in Prozent angegeben, sowie des GNE analysiert (vgl. Abbildung 2.5 und 2.6).

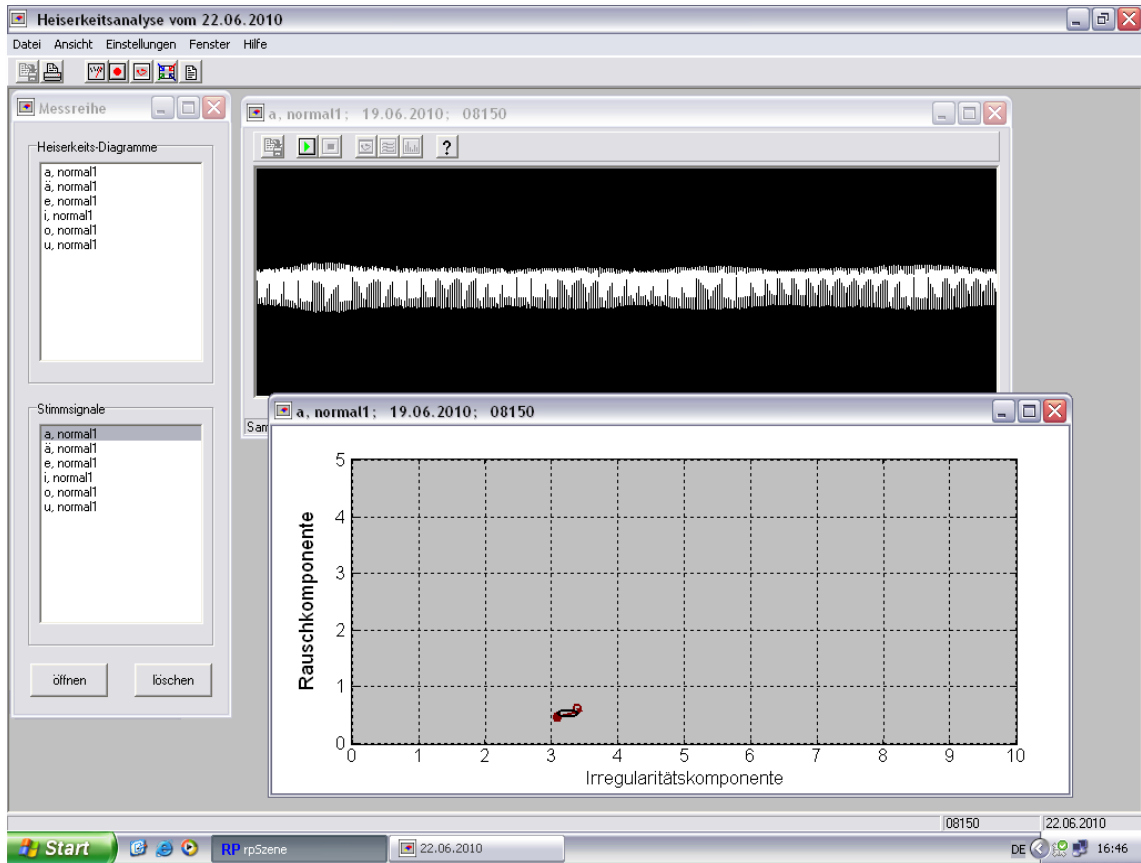


Abbildung 2.5: Stimmanalyse in rpSzene®

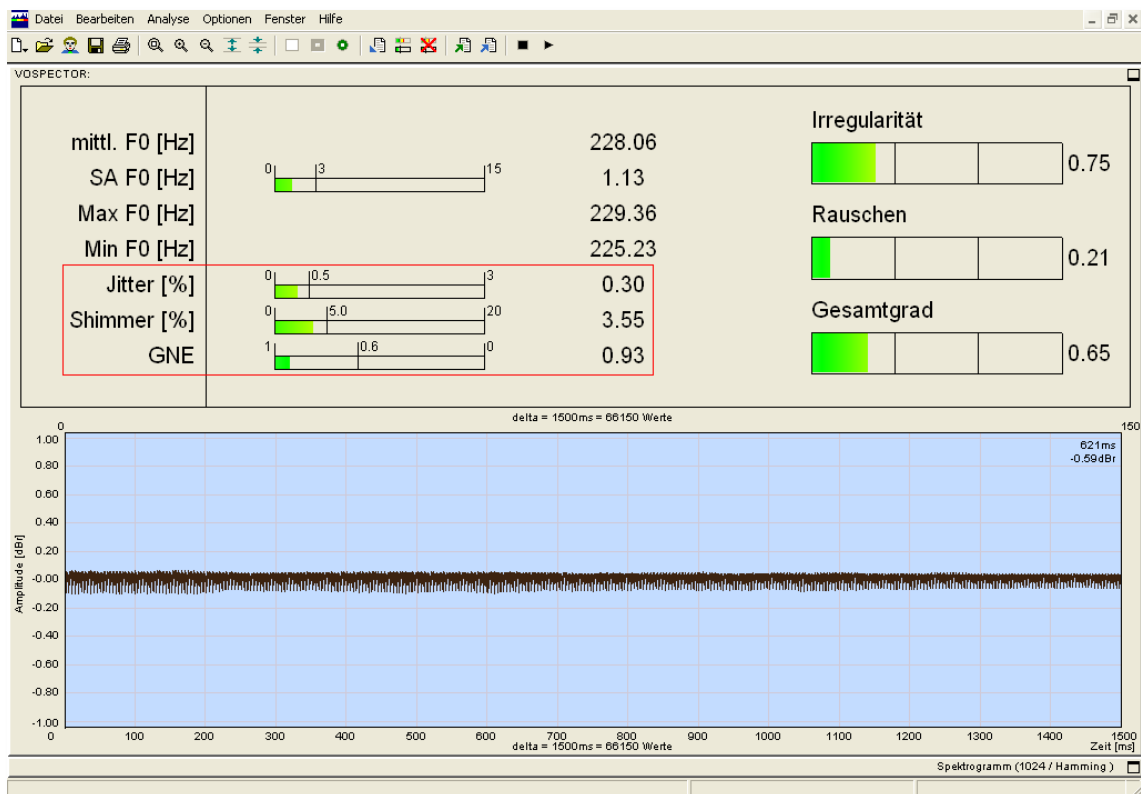


Abbildung 2.6: Stimmanalyse im Vospector von lingWAVES

Für jeden Parameter wurde eine Excel-Tabelle angelegt und die je 360 Ergebnisse eingetragen (vgl. dazu Anhang, Kapitel 6.4). Weiterhin erfolgte die graphische Darstellung in Form von sechs Säulendiagrammen (vgl. dazu Kapitel 3.2).

## 2.3.2 Statistik

### 2.3.2.1 Bland-Altman Plot

Für die Auswertung der Ergebnisse wurde zunächst eine Analyse nach Bland und Altman mit Hilfe des Computerprogramms MedCalc<sup>®</sup> durchgeführt. Diese eignet sich dazu, zwei quantitative Messverfahren auf die Übereinstimmung ihrer Messergebnisse hin zu untersuchen und graphisch als Bland-Altman Plot darzustellen. In dem Plot werden die Mittelwerte  $(x_i + y_i)/2$  auf der Abszisse und die Differenzen  $(x_i - y_i)$  auf der Ordinate aufgetragen [13; 93].

J. Martin Bland und Douglas G. Altman entwickelten 1986 diese statistische Methode, um die Nachteile des in der medizinischen Statistik sehr beliebten Korrelationskoeffizienten auszugleichen. Im Gegensatz zum Korrelationskoeffizienten zeigt der Bland-Altman Plot nicht nur einen möglichen linearen Zusammenhang (Korrelation), sondern eine tatsächliche Gleichwertigkeit der Daten auf. Dies lässt sich graphisch daran erläutern, dass der Korrelationskoeffizient zum Beispiel bereits dann einen Wert nahe 1 bzw. -1 anzeigen würde, wenn die Punkte der gegeneinander aufgetragenen Methoden annähernd auf einer Geraden liegen. Im Gegensatz dazu müssten die Punkte in der Graphik bei der Bland-Altman Analyse auf der Winkelhalbierenden ( $Y = X$ ) liegen, um eine hohe Übereinstimmung zu zeigen [2; 7; 9; 13].

Wie bereits erwähnt, wird die Analyse standardmäßig als Plot dargestellt und erlaubt damit in erster Linie eine optische Beurteilung der dargestellten Ergebnisse. Weisen zwei Verfahren tatsächlich eine hohe Übereinstimmung auf, dann bewegen sich die Punkte als Punktwolke um den Mittelwert der auf der Ordinate aufgetragenen Differenzen (vgl. dazu Abbildung 2.7).

In Studien wie der vorliegenden, in denen kein Standardverfahren als Referenz genommen werden kann, dienen die Mittelwerte auf der Abszisse als Annäherungswerte an den „wahren“ Wert. Die Differenzen bzw. deren Mittelwert auf der Ordinate lassen Rückschlüsse auf systematische Fehler (Bias) zu. Demnach lässt sich ablesen, ob ein Verfahren grundsätzlich höhere oder niedrigere

Messergebnisse liefert als ein anderes und ob die Abweichung der Methoden bzw. die Streuung der Abweichungen im Zusammenhang mit der Größe der Werte steht. Dies ist eine der Stärken des Bland-Altman Plots. Die doppelte Standardabweichung ( $\pm 1,96 \cdot SD$ ) umfasst den Bereich, in dem 95 % aller Wertepaarabweichungen liegen und gilt als Maß für die Übereinstimmung (Präzision) beider Verfahren. Durch sie wird die Wiederholungsgenauigkeit zukünftiger Vergleichsmessungen beschrieben [8; 9].

Sollte sich bei steigenden Werten auf der Abszisse eine Zunahme der Varianz der Differenzen zweier Methoden zeigen, dann ist es sinnvoll, entweder eine logarithmische Transformation der Daten vorzunehmen oder die Differenzen relativ zu dem Mittelwert als Prozentwert auf der Abszisse im Plot aufzutragen. Letzteres Vorgehen hat den Vorteil, dass die Daten anschließend nicht mehr rücktransformiert werden müssen, sondern direkt abgelesen werden können. Anstelle der logarithmischen Transformation besteht auch die Möglichkeit, Ratios anstatt der Differenzen auf der Ordinate aufzutragen, was prinzipiell zu ähnlichen Ergebnissen führt [2; 21; 36].

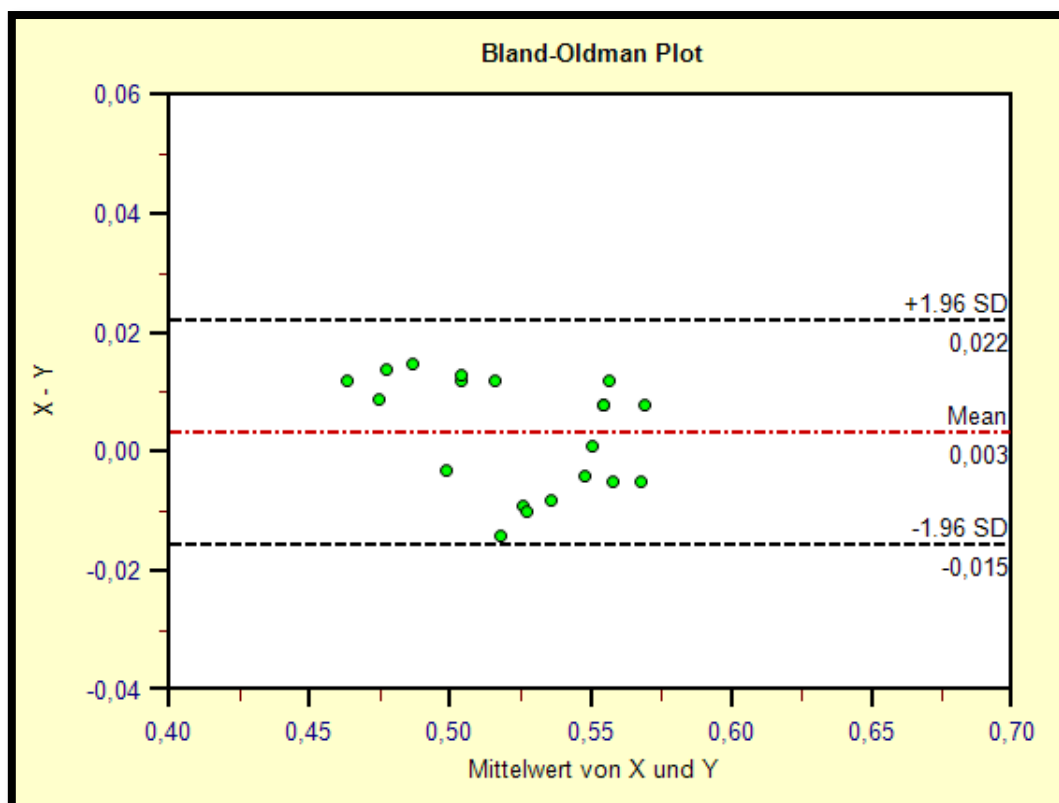


Abbildung 2.7: Beispiel eines Bland-Altman Plots mit fiktiven Daten

## 3 Ergebnisse

### 3.1 Studienteilnehmer

An der Studie nahmen 30 Patienten sowie 30 Probanden teil. Die Patienten konnten im Rahmen der phoniatischen Sprechstunde der Klinik für Phoniatrie und Pädaudiologie in Marburg, die Probanden durch direkte Einladungen für die Teilnahme an der Studie gewonnen werden.

Als Einschlusskriterien dienten Patienten mit Stimmstörungen und bestehender Heiserkeit sowie Probanden ohne aktuell behandlungsbedürftige Stimmerkrankungen. Waren Patienten bzw. Probanden unter 18 Jahren oder konnten Patienten den Vokalhaltversuch nicht über mehrere Sekunden umsetzen, galt dies als Ausschlusskriterium.

Der Anteil von weiblichen zu männlichen Studienteilnehmern lag in beiden Gruppen bei je 15 Frauen und 15 Männer 50 % (siehe Tabelle 3.1).

Teilnehmer	Männer	Frauen
Patienten	15	15
Probanden	15	15

Tabelle 3.1: Geschlechtsverteilung der Studienteilnehmer

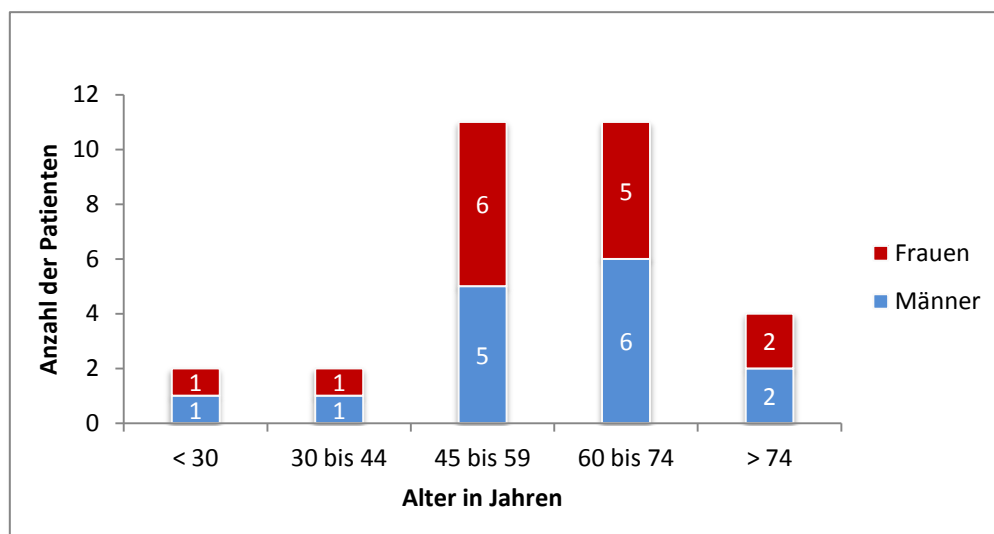


Abbildung 3.1: Altersverteilung der Patienten



Das durchschnittliche Alter in der Gruppe der Patienten lag bei 58,7 Jahren; der jüngste Teilnehmer war 28 Jahre und der älteste 81 Jahre alt (siehe Abbildung 3.1). In der Gruppe der Probanden zeigte sich eine Altersverteilung von 18 bis 56 Jahren, bei einem mittleren Alter von 29,1 Jahren (siehe Abbildung 3.2).

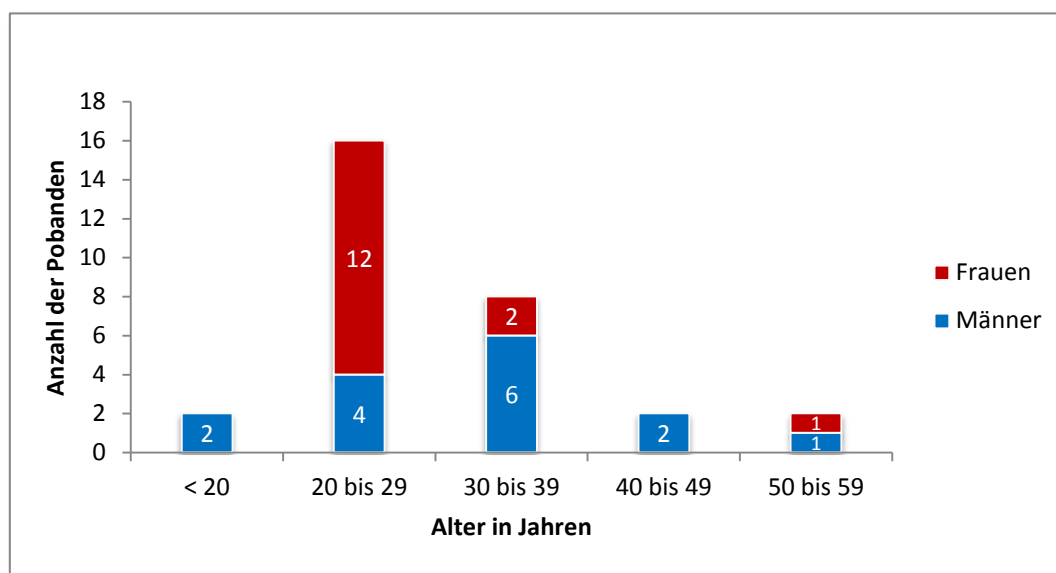


Abbildung 3.2: Altersverteilung der Probanden

30 der 60 Studienteilnehmer hatten zum Zeitpunkt der Stimmschallanalyse eine vorbekannte Erkrankung der Stimme. Diese Patienten wiesen insgesamt 14 verschiedene Erkrankungen auf, wobei die Funktionelle Dysphonie mit einer Häufigkeit von 11 (36,7 %) und die Stimmlippenparese mit 4 (13,3 %) den größten Anteil ausmachten. Die chronische Laryngitis kam in 3 (10 %) und das Stimmlippenödem in 2 (6,7 %) Fällen vor. Die folgenden zehn Erkrankungen waren jeweils bei einem (3,3 %) der Patienten diagnostiziert worden: Z. n. Laryngitis, Z. n. Stimmlippenparese, N. laryngeus superior-Schaden, Phoniationsverdickung, Z. n. Stimmlippenknötchen, Stimmlippenpolyp, Z. n. Stimmlippenpolyp, Larynxstenose, Stimmlippen-Missbildung und Z. n. Stimmlippen-OP (siehe Abbildung 3.3).

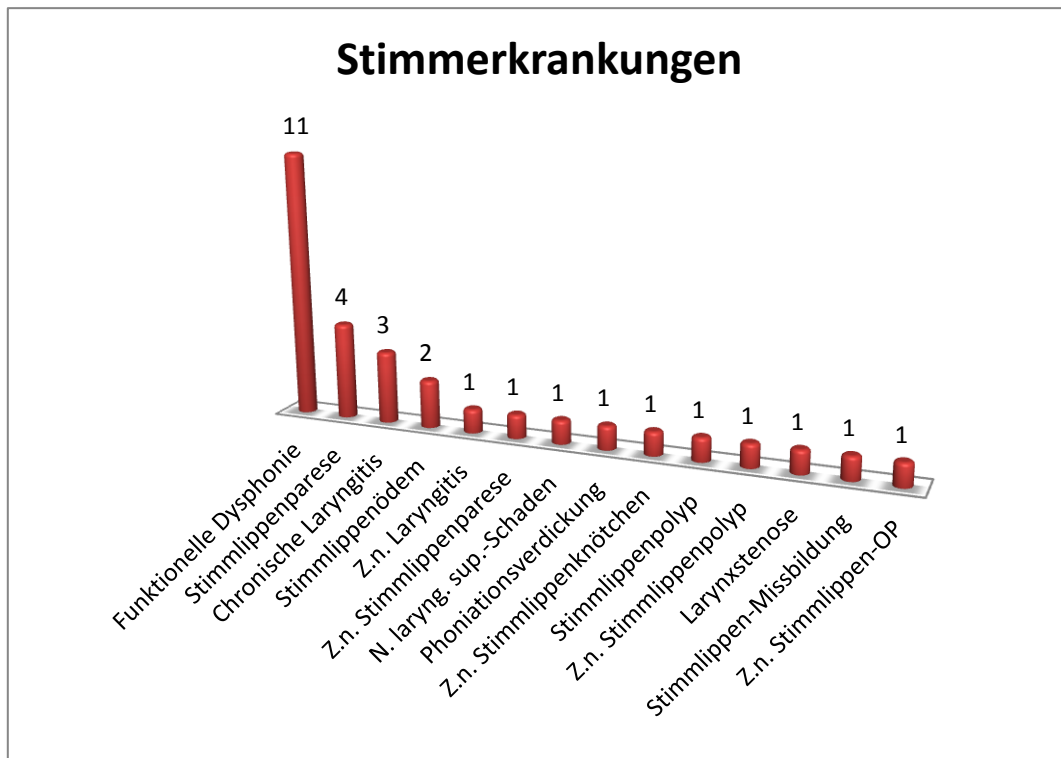


Abbildung 3.3: Stimmerkrankungen in der Patientengruppe

Bei den Probanden gab es eine Person mit einer Gemischt funktionellen Dysphonie (3,3 %) in der Vorgeschichte, die übrigen 29 Probanden hatten keine Vorerkrankungen im Zusammenhang mit der Stimme.

## 3.2 Auswertung der einzelnen Stimmparameter

Betrachtet man beide Stimmschallanalysegeräte gemeinsam, so wurden für die drei Stimmparameter (mit 6 Vokalen pro Studienteilnehmer) insgesamt 2160 Werte ermittelt. Für jeden Parameter lagen 360 Ergebniswerte pro Gerät vor, also 60 für jeden Vokal. Aufgrund der zum Teil erheblichen Differenzen zwischen den Werten der einzelnen Vokale innerhalb derselben Stimmsequenz wurden die Werte nicht gemittelt, sondern separat ausgewertet.

Bevor die Resultate des Vergleichs beider Stimmschallanalysegeräte vorgestellt werden, erfolgt eine kurze Zusammenfassung der Ergebnisse der Stimmparameter. Die Abbildungen 3.4 bis 3.9 geben einen Überblick über die Verteilung sämtlicher Werte für die Stimmparameter. Zur besseren Beurteilung sind die sechs Vokale farblich unterschiedlich gekennzeichnet. Für weitere Einzelheiten wird auf die Ergebnistabellen im Anhang verwiesen (vgl. dazu Kapitel 6.4).

### 3.2.1 Jitter

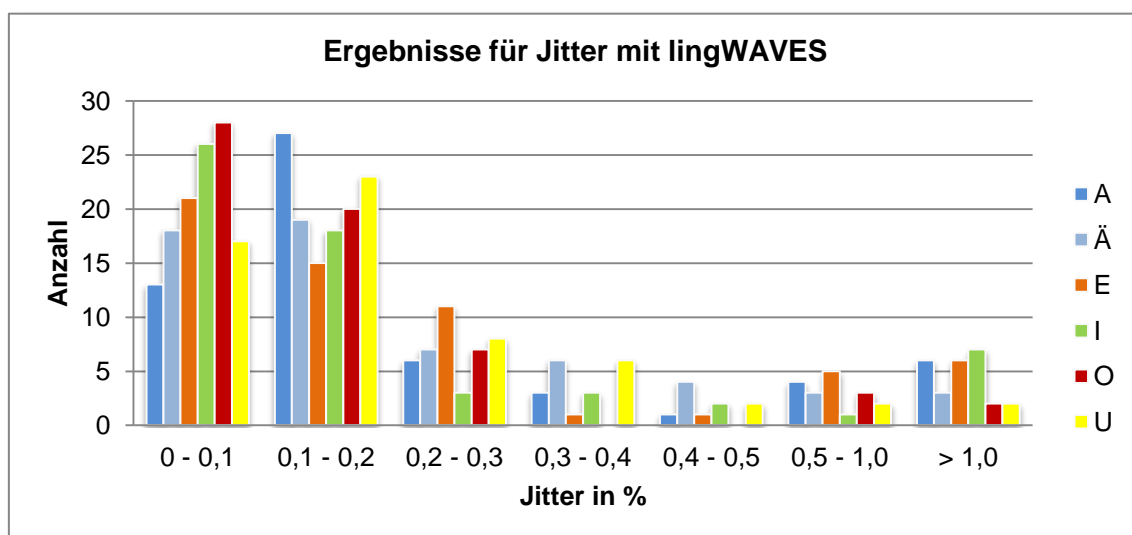


Abbildung 3.4: Jitter mit lingWAVES

Bei der Analyse mit lingWAVES variierten die Ergebnisse für den Jitter zwischen 0,03 % und 14,36 %. 245 (68,1 %), also mehr als zwei Drittel der Werte, lagen im Bereich von 0 bis 0,2 %. 26 Werte (7,2 %) waren größer als 1,0 %,

wobei für den Vokal [a] mit 6,2 %, 7,1 % und 14,36 % drei Werte ermittelt wurden, die deutlich über 5 % lagen.

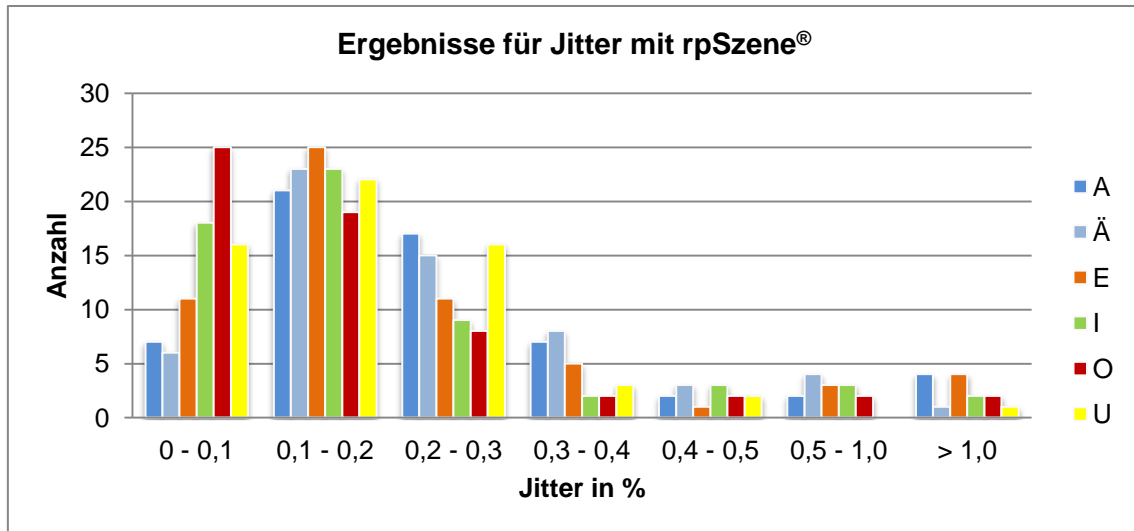


Abbildung 3.5: Jitter mit rpSzene®

Die Jitter-Berechnung mit rpSzene® ergab Werte zwischen 0,02 % und 10,19 %. Mit 216 (60,0 %) der 360 Ergebnisse waren mehr als die Hälfte der Werte kleiner oder gleich 0,2 %. Jenseits der 1 %-Marke befanden sich 14 Werte (3,9 %), von denen lediglich einer (ebenfalls Vokal [a]) mit 10,19 % außerhalb des 5 %-Bereichs lag.

### 3.2.2 Shimmer

Die Auswertung der Stimmsequenzen mit lingWAVES erzielte Shimmer-Werte im Bereich von 1,6 % bis 40,1 %. Der Großteil von ihnen, nämlich 285 Werte (79,2 %), lag zwischen 4 % und 20 %, lediglich 63 (17,5 %) Resultate wiesen geringere Werte auf. 12 (3,3 %) der 360 ermittelten Ergebnisse waren größer als 20 %, wobei fünf (einer von Vokal [i] und vier von Vokal [a]) über 30 % lagen.

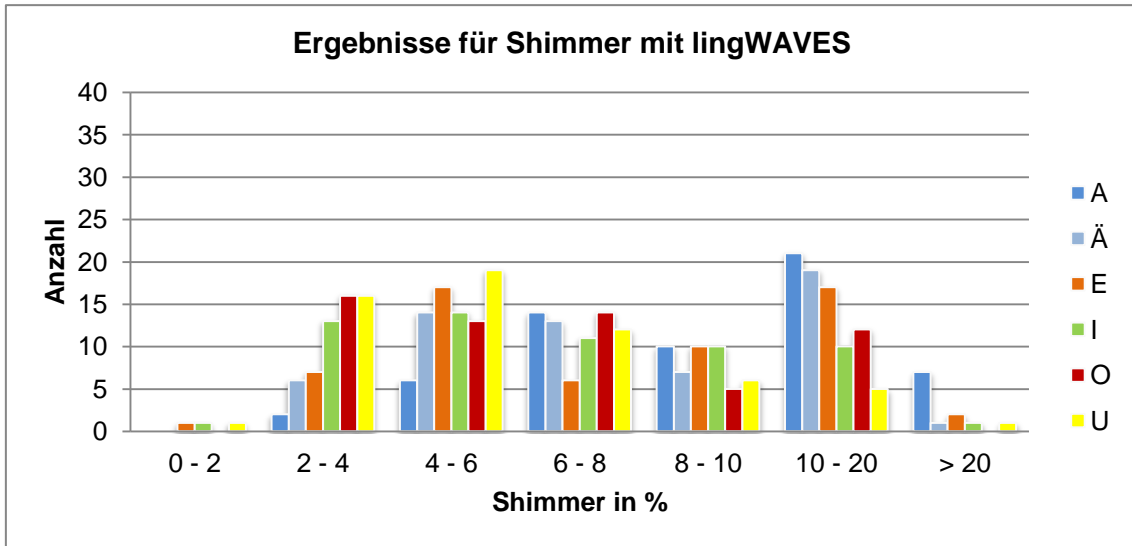


Abbildung 3.6: Shimmer mit lingWAVES

Die Stimmschallanalyse mit rpSzene<sup>®</sup> wies einen Minimalwert von 0,55 % bzw. Maximalwert von 37,3 % auf. Mit insgesamt 313 (86,9 %) Werten entfiel der überwiegende Teil auf den Bereich von 0 bis 4 %. Lediglich 43 (11,9 %) lagen zwischen 4 % und 20 %. Drei der vier Shimmer-Werte jenseits der 20 %-Marke waren mit 37,2 %, 43,5 % sowie 47,5 % ebenfalls deutlich über dem Wert von 30 % angesiedelt.

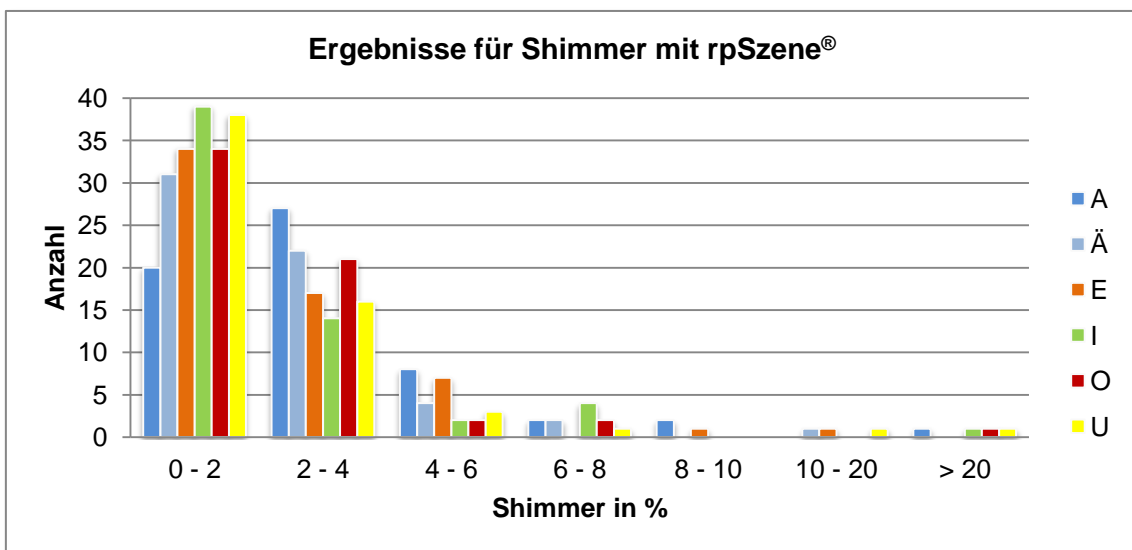


Abbildung 3.7: Shimmer mit rpSzene<sup>®</sup>

### 3.2.3 GNE

Wie sich bereits aus der Bezeichnung „Glottal-to-Noise Excitation Ratio“ ableiten lässt, stellt der GNE einen Verhältnis-Wert dar. Daher können die Werte für den GNE immer nur in dem Bereich von 0 bis 1 liegen.

Die in dieser Studie mit lingWAVES erzielten Ergebnisse des GNE deckten bei Werten von 0,04 bis 0,96 annähernd das gesamte Spektrum der theoretisch möglichen Werte ab. Mit 196 (54,4 %) lagen mehr als die Hälfte von ihnen im Bereich von 0,6 bis 1,0. Während die Anzahl der Werte für die Vokale [a] bis [i] mit steigendem GNE-Wert tendenziell zunahm (Ausnahme Vokal [i] bei 0,8 bis 1,0; siehe Abbildung 3.8.), verhielt es sich mit den Vokalen [o] und [u] umgekehrt. Mit 17 (28,3 %) bzw. 37 (61,7 %) Werten waren diese am stärksten in der Gruppe 0 bis 0,2 vertreten.

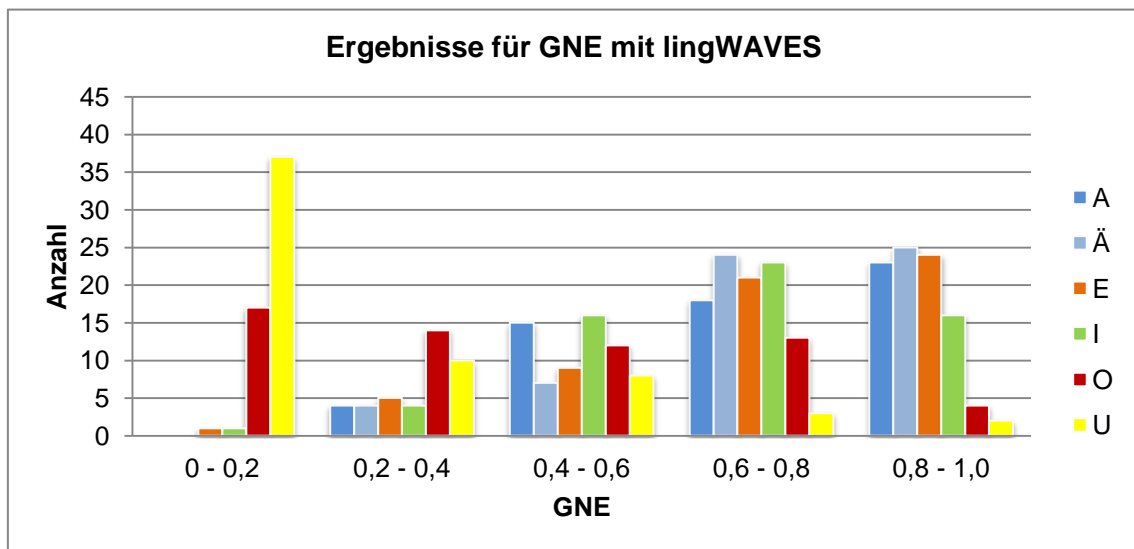


Abbildung 3.8: GNE mit lingWAVES

Auch die mit rpSzene<sup>®</sup> ermittelten Ergebnisse für den GNE zeigten bei Werten von 0,08 bis 0,98 eine breite Verteilung. Im Gegensatz zu der Analyse mit lingWAVES war die Mehrheit von ihnen jedoch im oberen Bereich angesiedelt. 261 (72,5 %) Ergebnisse wiesen GNE-Werte von über 0,6 auf, immerhin 191 (53,1 %) von über 0,8.

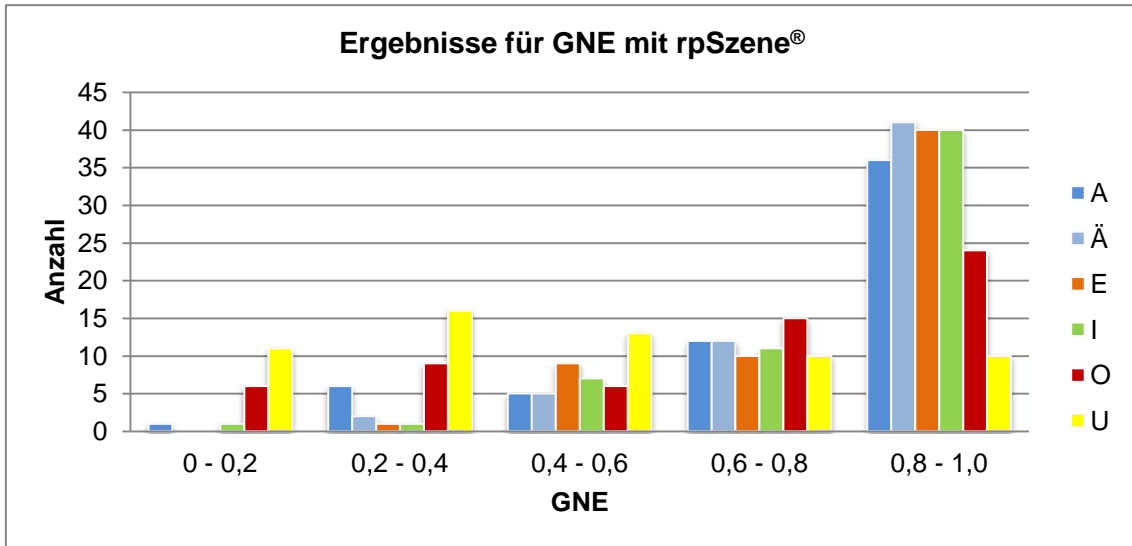


Abbildung 3.9: GNE mit rpSzene®

### 3.3 Vergleich beider Stimmschallanalysegeräte mittels Bland-Altman Plots

Für jeden der Vokale [a], [ä], [e], [i], [o] und [u] wurde pro Stimmparameter ein Bland-Altman Plot mit MedCalc<sup>®</sup> erstellt. Die Mittelwerte und Differenzen der 60 Wertepaare von lingWAVES bzw. rpSzene<sup>®</sup> wurden dabei grafisch gegeneinander aufgetragen. Aufgrund einer besseren Vergleichbarkeit der verschiedenen Plots wurde zur Berechnung der Differenz grundsätzlich der rpSzene<sup>®</sup>-Wert von dem lingWAVES-Wert abgezogen. Die Differenzwerte ergaben daher sowohl positive als auch negative Resultate. Ergänzend zu dieser Form der Bland-Altman Plots erfolgte im Fall des Jitters jeweils eine zusätzliche grafische Darstellung, bei der die Differenzen im Verhältnis zu den Mittelwerten ( $(\text{lingWAVES} - \text{rpSzene}^{\text{®}}) / \text{Mittelwert}$ ) als Prozentwert auf der Ordinate aufgetragen wurden. Im Falle des Shimmers wurden ergänzend Bland-Altman Plots mit Ratios angefertigt.

#### 3.3.1 Jitter

Aufgrund der ähnlichen Ergebnisse und somit auch ähnlichen Aussagekraft der sechs Bland-Altman Plots für die Jitterwerte wird in diesem Abschnitt lediglich exemplarisch der des Vokals [ä] näher beschrieben. Im Anhang befinden sich sämtliche Bland-Altman Plots (vgl. dazu Kapitel 6.5).

Wie bereits in Kapitel 3.2.1 beschrieben, wiesen einzelne Ergebnisse der Stimmschallanalyse zum Teil sehr hohe Jitterwerte von über 5 % auf. Dies spiegelte sich bei der Verteilung der Wertepaare in den Bland-Altman Plots insofern wider, dass ein Großteil der Punkte sehr eng beieinander lag, wohingegen einzelne Punkte eine hohe Streuung zeigten. Dieses Verteilungsmuster trat in allen sechs Bland-Altman Plots auf.

Im Fall des Vokals [ä] waren die Differenzen beider Stimmschallanalysegeräte im Durchschnitt nahezu null (siehe Abbildung 3.10, rote Linie). Dies bestätigten auch die anderen Ergebnisse, wobei der Plot für den Vokal [a] mit 0,3 % die größte mittlere Differenz aufwies. Die auf der Abszisse aufgetragenen Mittelwerte lagen beim Beispiel des Vokals [ä] bei 0,06 % bis 1,57 %, ansonsten grundsätzlich zwischen 0,03 % und 4,08 %. Eine Ausnahme bildete der 10,28 %-Wert, der im Rahmen der Bland-Altman-Analyse für den Vokal [a] er-



mittelt wurde.

Die von MedCalc<sup>®</sup> standardmäßig mit angegebenen Übereinstimmungsintervalle (schwarze Linien) waren in diesen Bland-Altman Plots nicht aussagekräftig, da sie bei Anordnung der Punktwerte in Form eines nach rechts offenen Kegels nicht die tatsächlichen Grenzen widerspiegeln (vgl. dazu Kap. 2.3.2.1 bzw. Kap. 4.2).

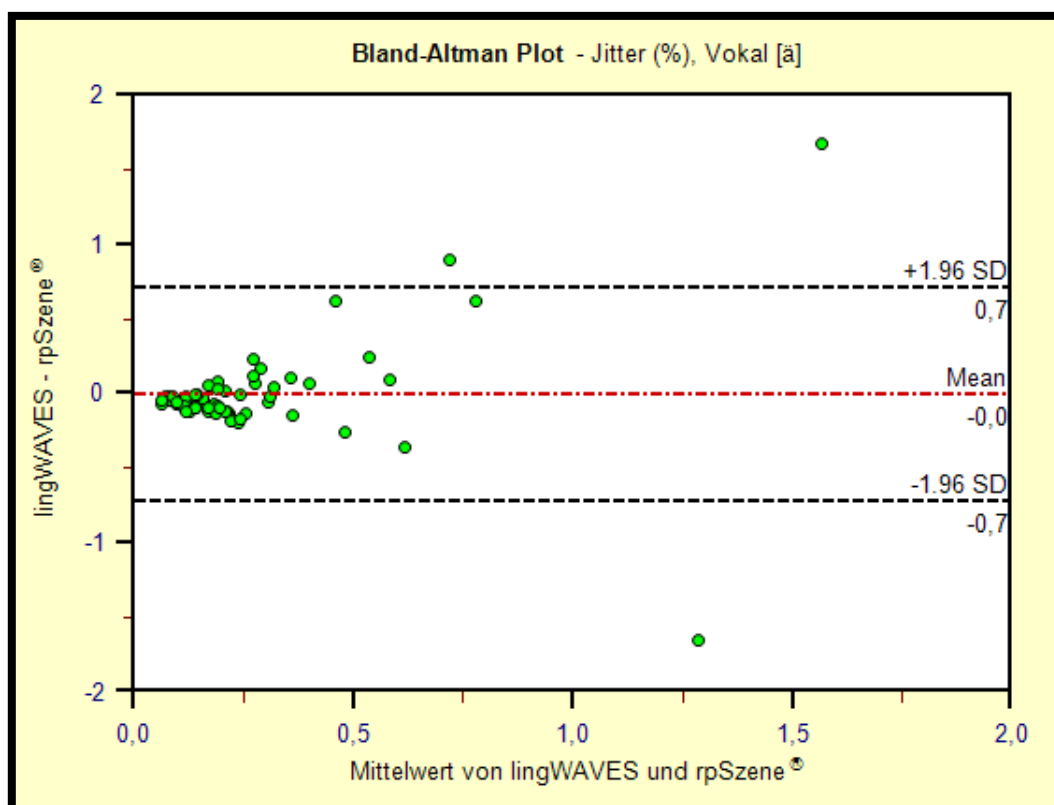


Abbildung 3.10: Bland-Altman Plot für Jitterwerte des Vokals [ä]

Analog verhielt es sich mit der Varianz. In diesem Fall erwies sich die Streuung der Werte um den Mittelwert der Differenzen als problematisch, da diese mit der Zunahme des Mittelwerts von lingWAVES und rpSzene<sup>®</sup> erheblich anstieg. Diese Entwicklung ließ sich für alle sechs Vokale nachweisen, wurde jedoch am Beispiel des Vokals [ä] am deutlichsten erkennbar (siehe Abbildung 3.10).

Die daraufhin zusätzlich angefertigten Bland-Altman Plots, die die relativ auf den Mittelwert bezogenen Differenzen beinhalteten, zeigten eine gleichmäßigere Anordnung der Punkte um den Mittelwert (siehe Abbildung 3.11) und stellten eine bessere Grundlage für die Berechnung des 95 %igen Übereinstimmungs-

intervalls bzw. der 1,96fachen Standardabweichungen dar. Die mittleren Differenzen der so zusätzlich erstellten Bland-Altman Plots sind mit den entsprechenden Standardabweichungen in Tabelle 3.2 aufgeführt.

Vokale	[a]	[ä]	[e]	[i]	[o]	[u]
Mittlere Differenz/Mittelwerte (%)	-8,2	-23,4	-4,5	-8,1	-11,7	4,3
Übereinstimmungsintervall (%)	113,2 -129,5	89,3 -136,1	115,9 -125,0	114,8 -131,0	93,0 -116,4	100,9 -92,2

Tabelle 3.2: Mittlere Differenzen / Mittelwerte mit Übereinstimmungsintervall (+/- 1,96SD) für Bland-Altman Plots des Jitter

Hier ist zu erkennen, dass die 100 %-Marke in neun von zwölf Fällen innerhalb der Intervallgrenzen lag und in den übrigen drei Fällen immerhin noch die 90 %-Marke einschloss.

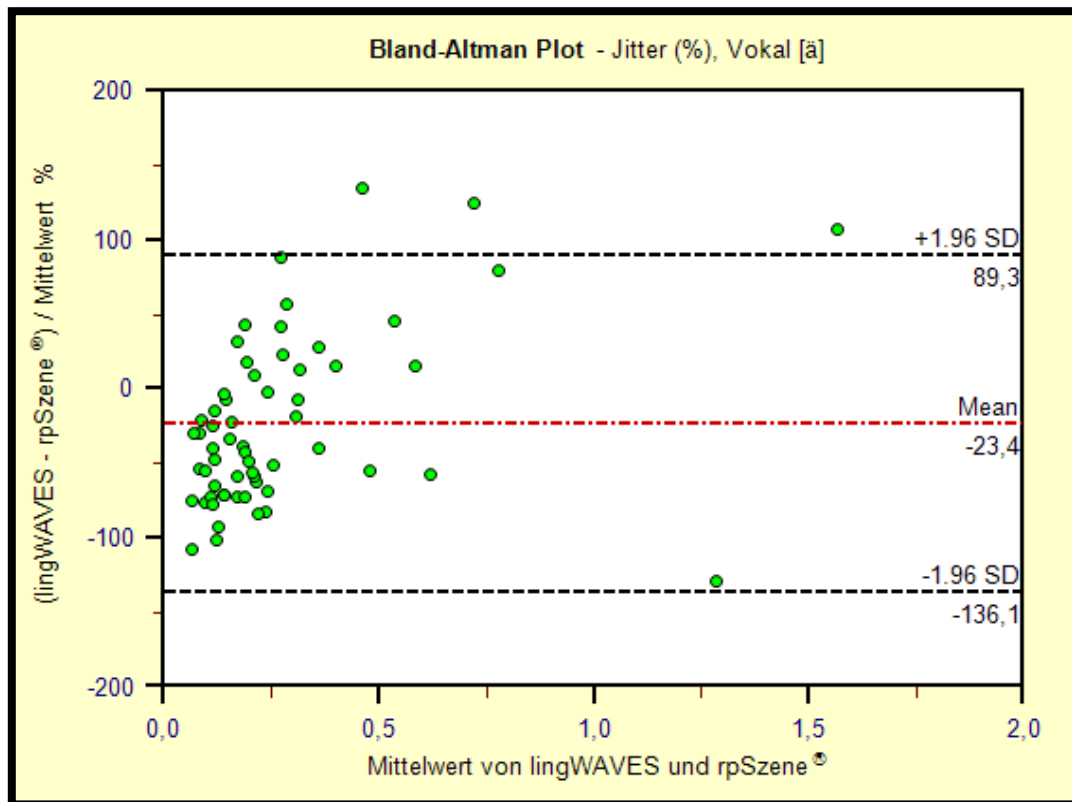


Abbildung 3.11: Bland-Altman Plot für Jitterwerte des Vokals [ä] mit relativen Differenzen

### 3.3.2 Shimmer

Analog zu der Betrachtung der Bland-Altman Plots für die Jitter-Ergebnisse wird der Plot des Vokals [a] exemplarisch für die Resultate des Shimmer abgebildet (siehe Abbildung 3.12). Die Plots für die weiteren Vokale befinden sich im Anhang (vgl. dazu Kapitel 6.5).

Die Bland-Altman Plots für die Shimmer-Werte wiesen zunächst den klassischen Aufbau mit den auf der Abszisse aufgetragenen Mittelwerten bzw. den auf der Ordinate aufgetragenen absoluten Differenzen auf. Die Mittelwerte auf der Abszisse lagen im Fall des Vokals [a] bei 2,40 % bis 34,63 %, für alle sechs Plots insgesamt zwischen 1,19 % und 43,71 %.

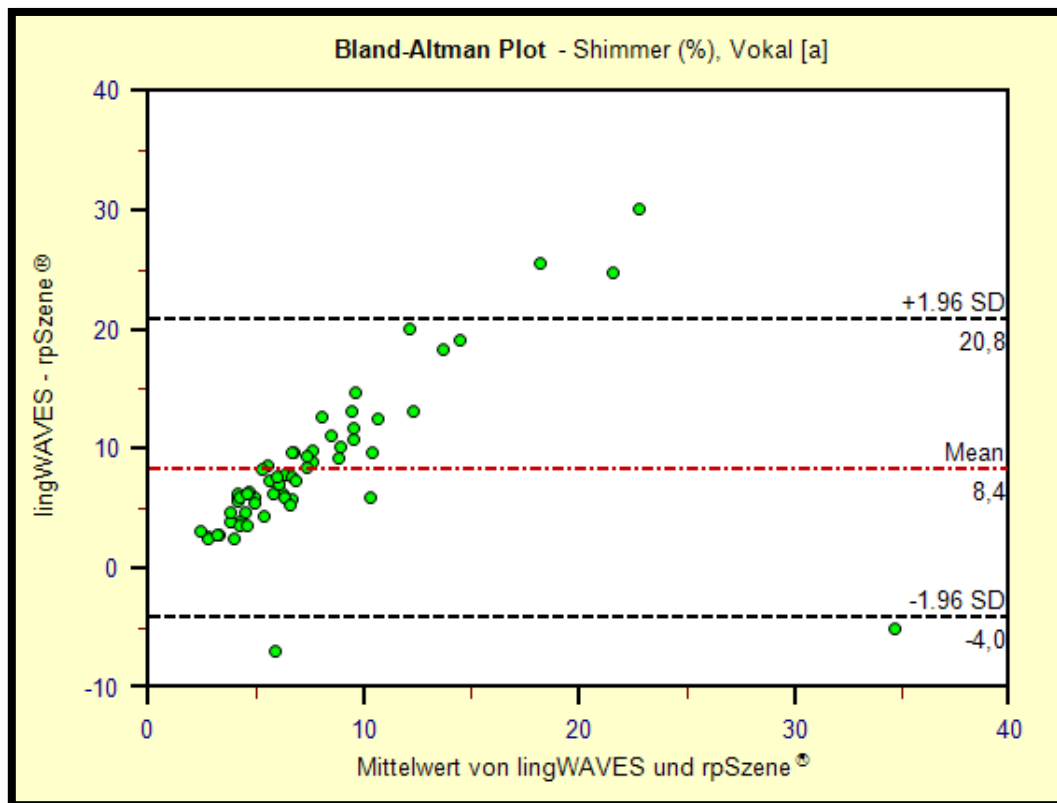


Abbildung 3.12: Bland-Altman Plot für Shimmerwerte des Vokals [a]

Die mittlere Differenz für den Vokal [a] lag bei 8,4 %. Diese reduzierte sich bei Betrachtung der übrigen Vokale in alphabetischer Reihenfolge bis auf letztlich 3,3 % für den Vokal [u] (siehe Tabelle 3.3). Die Grenzen des 95 %igen Übereinstimmungsintervalls befanden sich zwischen -4,0 % und 20,8 % für den

abgebildeten Plot des Vokals [a]. Für die anderen zeigten sich Werte von minimal -4,7 % (Vokal [u]) bis maximal 20,8 % (Vokal [a]). Insgesamt fiel auf, dass die unteren Grenzwerte jeweils negative und die oberen jeweils positive Werte annahmen.

Vokale	[a]	[ä]	[e]	[i]	[o]	[u]
Mittlere Differenz (%)	8,4	6,0	5,8	4,7	4,5	3,3
Übereinstimmungsintervall (%)	20,8 -4,0	12,5 -0,5	12,7 -1,2	11,4 -2,1	12,9 -4,0	11,2 -4,7

**Tabelle 3.3: Mittlere Differenzen mit Übereinstimmungsintervall (+/- 1,96SD) für Bland-Altman Plots des Shimmer**

Anhand der Abbildung 3.12 lässt sich exemplarisch erklären, was grundsätzlich bei allen Bland-Altman Plots für die Shimmerwerte galt: Mit steigenden Mittelwerten von lingWAVES und rpSzene<sup>®</sup> zeigte sich jeweils eine Zunahme ihrer Differenzen. Mit anderen Worten, die Mittelwerte und Differenzen waren proportional zueinander. Auch wenn sich das Ausmaß der Differenzzunahme zwischen den einzelnen Plots unterschied, war die annähernd lineare Anordnung der Punkte in den Darstellungen deutlich sichtbar.

Da Bland-Altman Plots mit dieser Konstellation grundsätzlich nicht zu beurteilen sind, wurden in einem weiteren Schritt die Ratios (Verhältniswerte) gebildet und anstatt der Differenzen auf der Ordinate aufgetragen. Diese neuangefertigten Plots offenbarten nun wie gewünscht eine gleichmäßigere Verteilung der Punkte um die mittleren Ratios (vgl. Abbildung 3.13 sowie Anhang, Kapitel 6.5).

In Tabelle 3.4 sind die mittleren Ratios sowie die Intervallgrenzen für den Übereinstimmungsbereich aufgelistet. Hierbei zeigt sich, dass die Ratios für die Shimmerwerte nicht nur alle deutlich über null lagen, sondern mit 3,0 bis 4,3 auch jeweils Werte in ähnlicher Größenordnung aufwiesen. Die Grenzwerte für das 95 %ige Übereinstimmungsintervall waren mit Ausnahme von Vokal [o] (-1,1) alle im positiven Bereich.

Vokale	[a]	[ä]	[e]	[i]	[o]	[u]
Mittlerer Ratio	4,3	3,8	4,0	3,9	3,5	3,0
Übereinstimmungsintervall (%)	8,1 0,4	7,0 0,7	7,9 0,1	7,3 0,5	8,0 -1,1	5,5 0,4

Tabelle 3.4: Mittlere Ratios mit Übereinstimmungsintervall (+/- 1,96SD) für Bland-Altman Plots des Shimmer

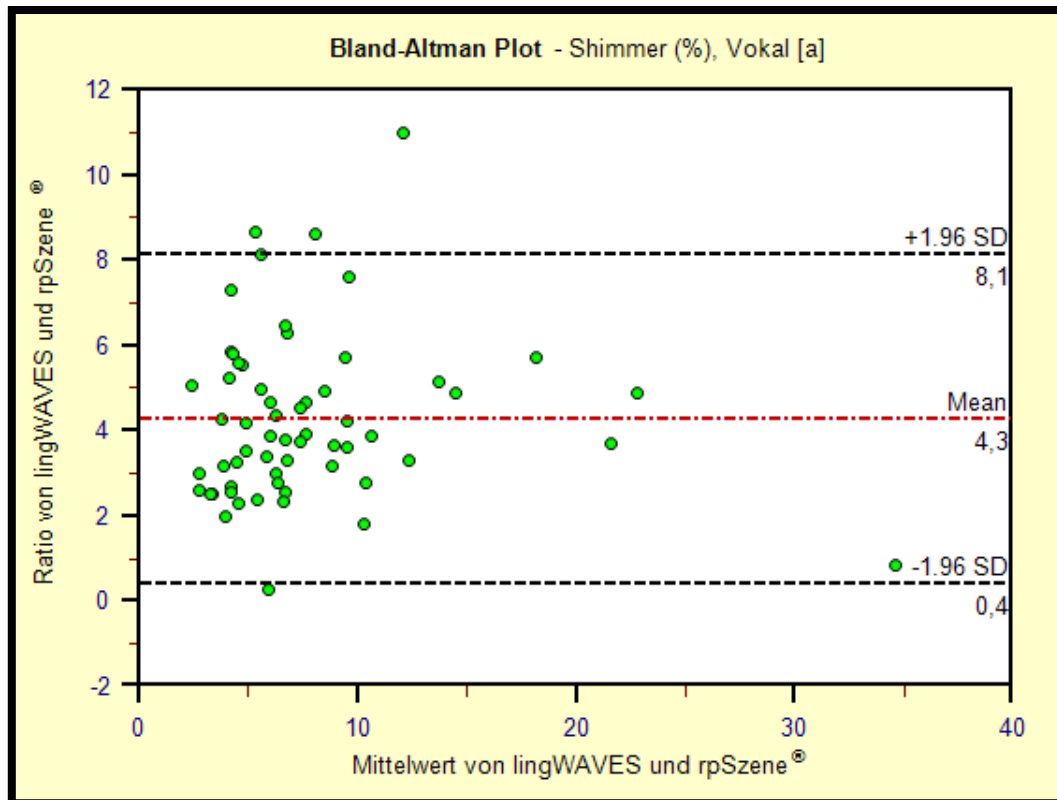


Abbildung 3.13: Bland-Altman Plot für Shimmerwerte des Vokals [a]

### 3.3.3 GNE

Im Gegensatz zu den Bland-Altman Plots für die Jitter- und Shimmerwerte verteilten sich die Ergebnisse für den GNE als Punktwolke um die mittleren Differenzen. Das Beispiel des Vokals [a] (siehe Abbildung 3.14) zeigte ein regelmäßiges Verteilungsmuster mit einer Akkumulation im Bereich von 0,8 bis 1,0. Für die Vokale [ä], [e] und [i] fand sich eine ähnliche Anordnung der Punktwerte (vgl. dazu Anhang, Kapitel 6.5). In den Bland-Altman Plots der Vokale [o] und [u] fehlte diese Konzentrierung der Punktwolke im rechten Teil der Grafik, da beide Vokale überproportional viele Ergebnisse mit geringeren GNE-Werten

bzw. Mittelwerten aufwiesen (vgl. dazu Kapitel 3.2.3). Im Gegenteil, im Fall des Vokals [u] war sogar eine Verlagerung hin zu den kleineren Werten zu erkennen (siehe Abbildung 3.15).

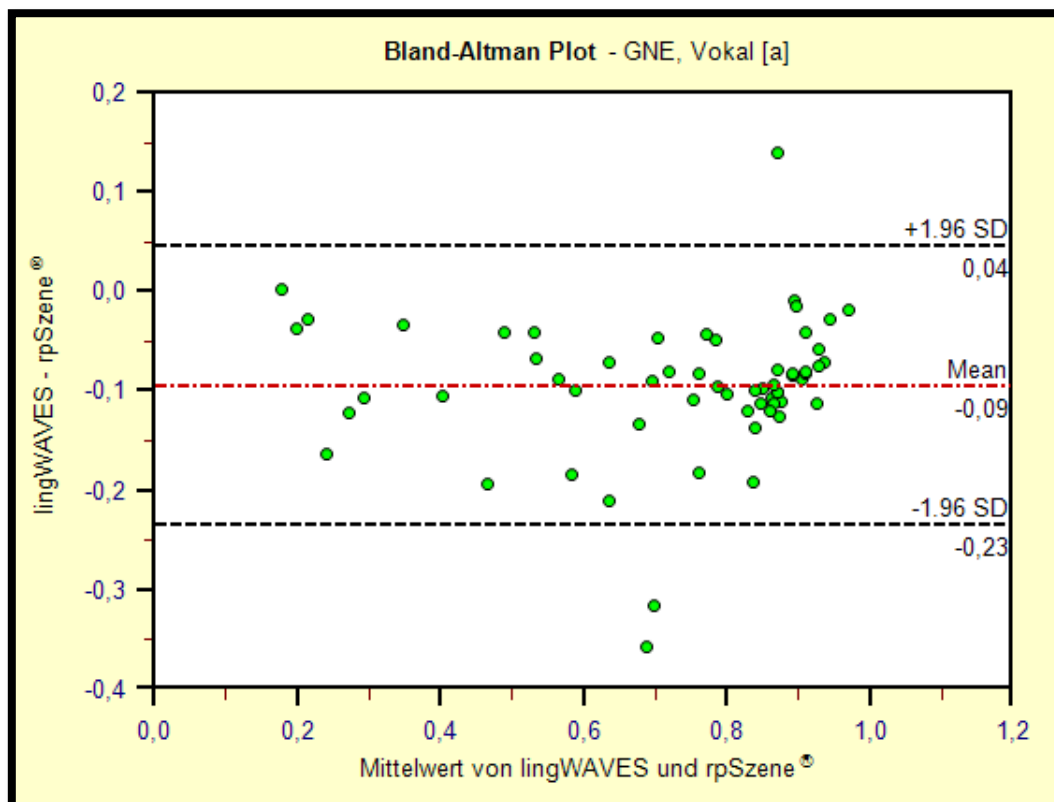


Abbildung 3.14: Bland-Altman Plot für GNE-Werte des Vokals [a]

Die Mittelwerte für lingWAVES und rpSzene® bewegten sich zwischen 0,06 und 0,97. Charakteristischerweise lagen die Minimal- bzw. Maximalwerte für den Vokal [a] mit 0,18 und 0,97 deutlich über denen des Vokals [u], der Werte von 0,06 bis 0,83 aufwies.

Vokale	[a]	[ä]	[e]	[i]	[o]	[u]
Mittlere Differenz (%)	-0,09	-0,10	-0,12	-0,14	-0,22	-0,23
Übereinstimmungsintervall (%)	0,04 -0,23	0,01 -0,19	0,03 -0,26	0,02 -0,30	0,16 -0,60	0,11 -0,57

Tabelle 3.5: Mittlere Differenzen mit Übereinstimmungsintervall (+/- 1,96SD) für Bland-Altman Plots des GNE

Die Bland-Altman-Analyse für den Stimmparameter GNE ergab mittlere Differenzen von -0,09 bis -0,23, wobei die Ergebnisse in alphabetischer Reihenfolge jeweils immer negativere Werte annahmen (siehe Tabelle 3.5). Ähnlich wie zuvor bei den Shimmerwerten gesehen, zeigten auch hier die unteren Grenzwerte der Übereinstimmungsintervalle jeweils negative und die oberen jeweils positive Werte.

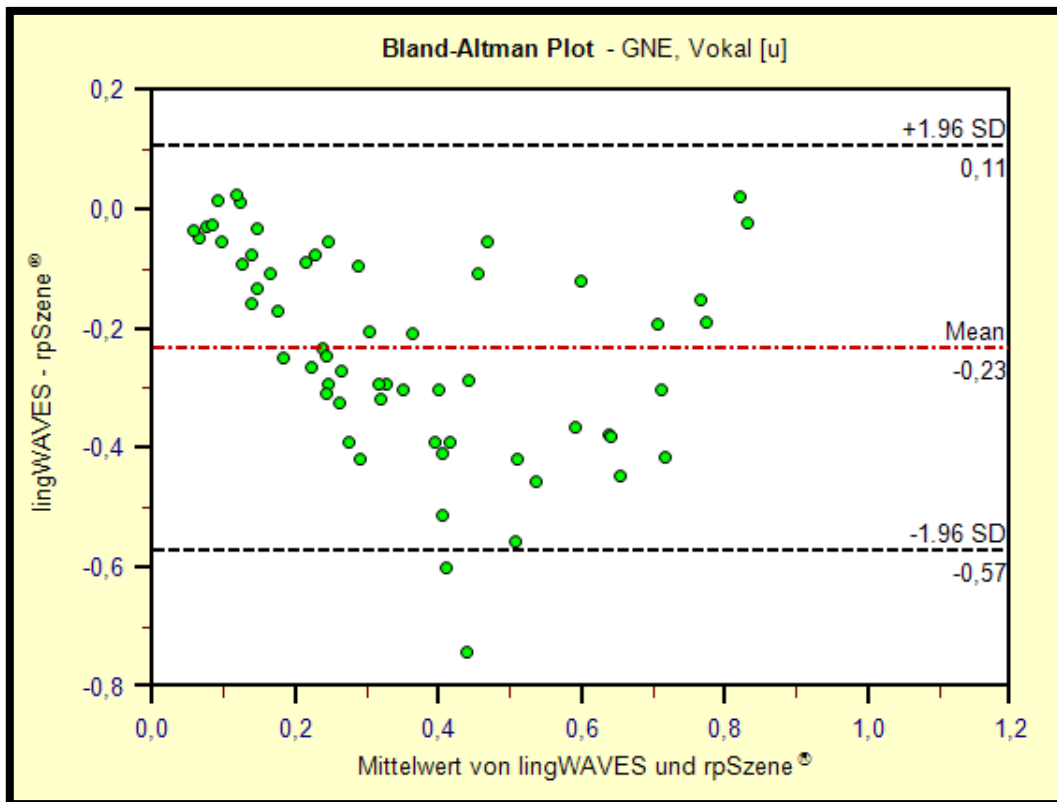


Abbildung 3.15: Bland-Altman Plot für GNE-Werte des Vokals [u]

## 4 Diskussion

### 4.1 Methodendiskussion

In die vorliegende Studie wurden nicht nur 30 Patienten mit Stimmstörungen, sondern auch 30 gesunde Probanden ohne bzw. ohne aktuell behandlungsbedürftige Stimmerkrankungen aufgenommen. Das Ziel war ein möglichst breites Spektrum verschiedener Stimmqualitäten, um sowohl pathologische als auch physiologische Werte für die jeweils berechneten Stimmanalyseparameter zu erhalten. Die Auswertung der einzelnen Stimmparameter (vgl. dazu Kapitel 3.2) ergab, dass dieses Ziel insofern erreicht wurde, da die in Form von Säulendiagrammen dargestellten Ergebnisse (vgl. Tabellen 3.4 bis 3.9) in allen Fällen neben einer zum Teil starken Häufung Streuungen von sehr niedrigen bis deutlich erhöhten Werten aufwiesen.

In der nachfolgenden Vergleichsanalyse wurden die Ergebnisse der Patienten- bzw. Probandengruppe zusammen ausgewertet. Auf eine Subgruppenanalyse wurde verzichtet, da in der vorliegenden Studie zwei Stimmanalysegeräte auf ihre Unterschiede hinsichtlich der Auswertungsergebnisse untersucht und keine Differenzierungen bezüglich der verschiedenen Stimmqualitäten vorgenommen werden sollten. Die geschlechtsspezifische Auswertung fand nicht statt. Dadurch wurde möglicherweise ein wichtiger Confounder außer Acht gelassen.

Als Voraussetzung für eine möglichst genaue Beurteilung der Stimmqualität musste für jeden Studienteilnehmer ein ausreichend langes Stimmsignal aufgenommen werden. Die für diese Studie ausgewählte Signallänge von 1500 ms konnte allerdings nur realisiert werden, wenn die jeweiligen Studienteilnehmer einen Tonhalteversuch von mindestens 2500 ms bis 3000 ms absolvieren konnten, da die Ein- bzw. Ausschwingungsanteile des jeweiligen Stimmsignals nicht für die Analyse geeignet sind. Von daher mussten Patienten mit hochgradigen Stimmstörungen und damit stark eingeschränktem Phonationsvermögen von der Teilnahme an der Studie ausgeschlossen werden. Hieraus ergab sich jedoch keine verfälschende Selektion des Patientenguts, da solche Patienten in der klinischen Praxis sowieso keiner aussagekräftigen Stimmschallanalyse unterzogen werden können.



## 4.2 Ergebnisdiskussion

Wie in Kapitel 1.5.4 beschrieben, unterliegt die Bestimmung der drei akustischen Stimmparameter grundsätzlich verschiedenen Einflussfaktoren. Aus diesem Grund waren bereits im Vorfeld jeweils unterschiedliche Ergebnisse für die Vergleichsanalyse erwartet worden. Nach Abschluss der Auswertung bestätigte sich diese Annahme. Dies war ein Grund dafür, warum die Beurteilung der Bland-Altman Plots für jeden Stimmparameter getrennt erfolgte. Analog dazu wiesen auch die Ergebnisse der einzelnen Vokale zum Teil deutliche Unterschiede auf, so dass außerdem eine vokalspezifische Auswertung erfolgte.

Wie im Folgenden ausgeführt, ergaben sich jedoch aus dieser selektiven Betrachtung rückblickend keine Unterschiede in der Interpretation der Ergebnisse.

Die Auswertung eines Bland-Altman Plots geschieht, wie bereits erwähnt, mittels optischer Interpretation der Diagramme. Sie erlaubt eine klare Aussage bezüglich der Gleichwertigkeit zweier Methoden. Beim Fehlen eines signifikanten Unterschieds gelten beide Methoden theoretisch als gleichwertig und damit als austauschbar. Zeigt sich jedoch eine systematische Abweichung (Bias), d. h. ein positiver oder negativer Wert für die mittleren Differenzen, dann sollte eine mögliche Abhängigkeit des Messfehlers von dem geschätzten „wahren“ Wert (Mittelwert auf der Abszisse) sowie die klinische Relevanz des Unterschieds ermittelt werden [9; 36].

Die Berechnungen der Jitterwerte ergaben auf den ersten Blick annähernd gleiche Ergebnisse für beide Stimmanalysegeräte. Dafür sprach auch der optische Eindruck nach Betrachtung der beiden entsprechenden Säulendiagramme (vgl. Abbildung 3.4 und 3.5). Bei mittleren Differenzen von nur 0,0 bis 0,1 % (Ausnahme Vokal [a]: 0,3 %) und damit nur geringen Abweichungen beider Messverfahren schien sich dieser Eindruck auch in den Bland-Altman Plots zu bestätigen. Lediglich für den Vokal [a] zeigte sich eine größere Abweichung, was jedoch problemlos auf die Auswirkungen der drei zum Teil erheblichen Ausreißer zurückzuführen war.

Doch wie bereits im Ergebnisteil erläutert, war diese erste Annahme insofern nicht korrekt, da sich die Punkteverteilungen in den Plots als ein nach rechts

offener Kegel darstellten. Dies entsprach einer Zunahme der Schwankungsbreite der Differenzwerte bei steigenden „wahren“ Werten. In diesem Fall ist eine Beurteilung des Plots nicht aussagekräftig [2; 21; 36]. Die anschließend dargestellten Bland-Altman Plots mit relativen Differenzwerten (auf die Mittelwerte bzw. den „wahren“ Wert bezogen) ergaben Werte von ca. +/- 90 % bis weit über +/- 100 % für die 95%igen Übereinstimmungsintervalle (siehe dazu Tabelle 3.2). Anders gesprochen, unter diesen Bedingungen würden die Ergebnisse von zukünftigen Jitter-Messungen in 95% der Fälle Werteschwankungen in Größenordnungen des „wahren“ Wertes – also ca. 100 % – um die mittlere Differenz aufweisen.

Daraus ließe sich schlussfolgern, dass die Größe des möglichen Bias im Falle des Jitters jeweils vom „wahren“ Wert abhinge und die Ergebnisse der zwei Messverfahren keine systematischen Abweichungen zeigten. Letztendlich sprach dieses Ergebnis eher für ein Problem in der Berechnung des Jitters selbst als für eine gerätespezifische Ursache. Eine klare Aussage bezüglich der Gleichwertigkeit beider untersuchter Stimmschallanalysegeräte im Hinblick auf die Jitter-Berechnungen ließ sich anhand dieser Ergebnisse nicht treffen.

Aus klinischer Sicht sind daher weniger die unterschiedlichen Ergebnisse der zwei Stimmanalysegeräte als vielmehr die große Schwankungsbreite des Jitters selbst als problematisch einzuordnen.

Betrachtet man die jeweiligen Ergebnisse des Shimmers aus rpSzene<sup>®</sup> und lingWAVES, war hier im Gegensatz zu den Jitter-Werten ein deutlicher Unterschied zu erkennen. Die als Säulendiagramm dargestellten Shimmerwerte von lingWAVES wiesen deutlich größere Werte auf als die von rpSzene<sup>®</sup> (vgl. Abbildung 3.6 und 3.7). Während bei lingWAVES zum Beispiel nur 63 Werte im Bereich von 0 bis 4 % lagen, waren es bei rpSzene<sup>®</sup> 313 Shimmerwerte.

Die Auswertung mittels Bland-Altman Plots ergab eine deutlich sichtbare, annähernd lineare Anordnung der Punkte im Sinne einer ansteigenden Gerade. Das bedeutet, dass sich die mittlere Differenz und damit der Bias jeweils proportional zu dem „wahren“ Wert verhielten. Aufgrund dieser Tatsache war zunächst keine Aussage bezüglich der Gleichwertigkeit beider Messverfahren möglich. Erst nach Bildung von Ratios (Verhältniswerten) und Auftragen derselben auf der Ordinate, anstatt wie üblich der Differenzen von lingWAVES und rpSzene<sup>®</sup>,

entstanden aussagekräftige Plots. Diese wiesen eine typische Punkteverteilung in Form einer Wolke um den mittleren Ratio-Wert auf.

Da in allen sechs Fällen positive Werte für den mittleren Ratio (3,0 bis 4,3) sowie mit Ausnahme des Vokals [o] auch für deren 95 %ige Übereinstimmungsintervalle ermittelt wurden, kann man die Unterschiede zwischen beiden Messgeräten als systematische Abweichungen bezeichnen. Allerdings muss beachtet werden, dass es sich nicht um einen klassischen Messfehler im Sinne eines Differenzwertes handelte, sondern um einen Multiplikationsfaktor. Das bedeutet, dass der jeweilige Faktor zur Umrechnung der Messergebnisse des einen Stimmschallanalysegerätes in die des anderen berücksichtigt werden müsste. Diese theoretische Umrechnung ist jedoch nur dann sinnvoll, wenn die Streuung der Ratios und damit die Standardabweichung SD ein gewisses Maß nicht überschreiten. Dass dies jedoch in diesem Fall zutraf, war anhand der mit Ratios erstellten Bland-Altman Plots deutlich zu erkennen (vgl. dazu Tabelle 3.4). Bei Werten von jeweils knapp über 0 (Ausnahme: Vokal [o]) bis Werten im Bereich von 5,5 bzw. 8,1 lagen die berechneten Grenzwerte der Übereinstimmungsintervalle insgesamt so weit auseinander, dass die ermittelten Multiplikationswerte nur eine geringe Genauigkeit aufwiesen. Mit anderen Worten, die Ergebnisse zeigten zwar deutliche und auch klinisch relevante Unterschiede, waren allerdings aufgrund der fehlenden Genauigkeit nicht durch einen Umrechnungsfaktor quantifizierbar.

Die Ergebnisse für die GNE-Bestimmung ergaben eine breit gestreute Wertverteilung für beide Stimmschallanalysegeräte (vgl. Abbildung 3.8 und 3.9). Insgesamt ergaben sich für rpSzene<sup>®</sup> größere Werte als für lingWAVES. Die sechs Bland-Altman Plots zeigten allesamt aussagekräftige Plots; die Ergebnisse waren jeweils als klassische Punktwolke um die mittleren Differenzen abgebildet (vgl. Abbildung 3.14 und 3.15 sowie Anhang, Kapitel 6.5). Bei mittleren Differenzen von -0,09 bis -0,23 zeigte sich ein Bias von ca. 0,1 bis 0,2 zugunsten der GNE-Werte von „rpSzene<sup>®</sup>“. Da die in Tabelle 3.4 aufgeführten Übereinstimmungsintervalle allerdings bei allen Vokalen die Null mit einschlossen, kann für diesen Stimmparameter keine quantitative Aussage bezüglich der systematischen Unterschiede für zukünftige Vergleichsmessungen gemacht werden. Selbst wenn die Null außerhalb des Übereinstimmungsintervalls gelegen hätte,

wären die Ergebnisse für die Vergleichsanalyse einzelner Vokale aufgrund der immer noch zu großen Unterschiede als klinisch relevant und damit nicht als vergleichbar einzustufen gewesen.

Auch wenn die soeben beschriebenen Ergebnisse nicht als mathematisch signifikant unterschiedlich bezeichnet werden können, ist abschließend festzustellen, dass keiner der untersuchten Stimmparameter Jitter, Shimmer oder GNE Ergebnisse lieferte, welche eine Vergleichbarkeit von lingWAVES und rpSzene<sup>®</sup> gestattet hätte. Beim Jitter zeigte sich eine große Schwankungsbreite in Abhängigkeit von dem „wahren“ Wert, sodass eine klare Aussage in Hinblick auf die Gleichwertigkeit beider Stimmschallanalysegeräte ebenso wenig möglich war wie eine mögliche Beurteilung bezüglich der klinischen Relevanz der Geräteunterschiede. Der Shimmer konnte erst nach Modulation der Bland-Altman Plots mittels Berechnung der Ratios ausgewertet werden und zeigte dabei nicht nur einen systematischen, sondern auch einen klinisch relevanten Unterschied bei fünf Vokalen (Ausnahme Vokal [o]) auf. Die dabei ermittelten Umrechnungsfaktoren in Form der mittleren Ratios müssen jedoch aufgrund der recht breiten Übereinstimmungsintervalle für die praktische Anwendung als zu ungenau bewertet werden. Die Auswertung der GNE-Werte ergab zwar klassische Bland-Altman Plots mit typischer Punkteverteilung und im Vergleich zu den vorherigen Parametern auch deutlich bessere Übereinstimmung, ließ jedoch schließlich aufgrund der Lage der Intervallgrenzen für den Übereinstimmungsbereich keine eindeutige Aussage über die Vergleichbarkeit von lingWAVES und rpSzene<sup>®</sup> zu.

Internationale Studien, die ebenfalls die Vergleichbarkeit von Stimmanalysesystemen untersuchten, wiesen größtenteils ähnliche Ergebnisse auf.

Bielamowicz et al. (1996) kamen in ihrer Studie, in der die Grundfrequenz  $f_0$ , die Pertubationsmaße Jitter, Shimmer und HNR bzw. SNR in vier Analysesystemen – CSpeech, Computerized Speech Laboratory (CLS), SoundScope und ein manuell zu bedienendes Stimmanalysesystem – miteinander verglichen wurden, zu folgenden Ergebnissen: Während für die Grundfrequenz  $f_0$  zwar keine optimalen, jedoch relativ gute Vergleichsergebnisse erzielt wurden, zeigte die Auswertung der Pertubationsmaße nur schwache Korrelationen zwischen den ver-

wendeten Analysegeräten. Letztere wurden u. a. auf die Anwendung von verschiedenen Analyse-Algorithmen in den Stimmanalysesystemen zurückgeführt [6].

Karnell et al. führten 1991 eine Studie zum Vergleich von Jitter- und Shimmer-Daten anhand von drei verschiedenen Analysesystemen durch. Nach Auswertung mit einem Visi-Pitch PC System (Pine Brook) und zwei Systemen auf Basis von Minicomputern (Chicago und Denver) kamen sie zu dem Ergebnis, dass die gewonnenen Pertubationsmaße-Daten nicht vergleichbar waren [45].

In einer weiteren Studie von Karnell et al. (1995) wurden die drei Stimmschallanalysesysteme Voice Analysis Program (Kay Elemetrics) auf einem Kay 5500 DSP Sona-Graph, CSpeech program auf einem AT-kompatiblen Computer (Zenith model Z-200) und AUDED/SEG Programm auf einem Digital LSI 11-23 Computer mithilfe derselben Pertubationsmaße Jitter und Shimmer miteinander verglichen. Die Ergebnisse für Jitter und Shimmer wiesen lediglich geringe Korrelationen bzw. mehrheitlich sogar signifikante Unterschiede auf. Lediglich für die zusätzlich bestimmten Grundfrequenzen  $f_0$  konnte eine Übereinstimmung zwischen den Stimmschallanalysesystemen aufgezeigt werden. Die Autoren dieser Studie sahen die mangelnde Standardisierung der Pertubationsmaße als entscheidende Ursache für die unterschiedlichen Ergebnisse an. Daher wiesen sie auf die Notwendigkeit einer engen Kooperation zwischen Entwicklern von Analysesystemen und Klinikern bei der Entwicklung neuer Pertubationsmaßen sowie deren Überprüfung auf klinische Anwendbarkeit hin [44].

In einer Studie zum Vergleich von Doctor Speech (DRS, Tiger Electronics, Neu-Anspach, Deutschland) und Computerized Speech Lab (CSL, Kay Elemetrics Corporation, Lincoln Park, USA) konnten Smits et al. (2005) eine signifikante Korrelation für  $f_0$ , HNR und relative Shimmerwerte nachweisen. Letztere waren allerdings nur mithilfe eines Umrechnungsfaktors vergleichbar. Für die relativen und absoluten Jitterwerte sowie für den  $F_0SD^4$  ließen sich lediglich schwache Korrelationen aufzeigen. In Übereinstimmung mit den zuvor beschriebenen Studien von Bielamowicz et al. und Karnell et al. wiesen auch die Autoren dieser Studie auf die unterschiedlichen Analyse-Algorithmen als Ursache für die unterschiedlichen Ergebnisse der Pertubationsmaße hin. Davon grundsätzlich

---

<sup>4</sup> Standardabweichung von  $F_0$ .

ausgenommen waren die Ergebnisse der Grundfrequenz  $f_0$ , die in allen genannten Studien vergleichbare Werte ergaben [6; 44; 55; 83].

Auch die Ergebnisse der von Morris und Brown (1996) verglichenen Grundfrequenzen  $F_0$  in sechs verschiedenen Aufnahmesystemen zeigten eine gute Reliabilität für alle Geräte, wenn auch bei einer Variation der Übereinstimmungen auf unterschiedlichen Niveaus [60].

Untersuchungen von Maryn et al. (2009), in denen CSL mit Multi-Dimensional Voice Program (MDVP) und ein PC mit Praat verglichen wurden, ergaben im Inter-Programm-Vergleich schwache bis moderate Korrelationen für frequenzabhängige Pertubationsmaße bzw. schwache bis starke Korrelationen für amplitudenabhängige Pertubationsmaße. Im Inter-System-Vergleich zeigten sich in beiden Fällen nur schwache bis moderate Korrelationen, d. h. mit Ausnahme der amplitudenabhängigen Pertubationsmaße im Inter-Programm-Vergleich war keine tatsächliche Vergleichbarkeit zwischen beiden verwendeten System bzw. Programmen gegeben [55].

Die Ursache für die unterschiedlichen Ergebnisse von amplituden- und frequenzabhängigen Pertubationsmaßen lag nach Bielamowicz et al. (1996) daran, dass die Abhängigkeit des Jitters von der exakten Grundperioden-Bestimmung größer war als bei dem Shimmer (vgl. Kapitel 1.5.4) [6].

Das Ziel einer weiteren Studie (2006) war die Entwicklung von Normalwerten für die Parameter Grundfrequenz  $f_0$ , Jitter, Shimmer und HNR bei Erwachsenen im Alter von 20 bis 45 Jahren mit dem CSL-4300 (Kay Elemetrics Corporation, Lincoln Park, USA). Naufel de Filippe et al. kamen zu dem Ergebnis, dass eine Entwicklung von Normalwerten unter den aktuellen Bedingungen nur für jedes Analysesystem getrennt stattfinden könne. Neben den unterschiedlichen Analyse-Algorithmen würden auch die Verwendungen von verschiedenen Aufnahmekriterien, Computersystemen und Mikrofonen eine einheitliche Normierung so gut wie unmöglich machen [62].

### 4.3 Diskussion der Fehlermöglichkeiten

Nachdem in Kapitel 4.2 Teil 1 der im Vorfeld dargelegten Fragestellungen (vgl. dazu Kap. 1.6) schwerpunktmäßig unter Einbindung anderer Studien erörtert wurde, wird im Folgenden vermehrt auf den Teil 2 eingegangen.

Im Rahmen der Ursachenforschung für die zum Teil erheblichen Ergebnisunterschiede bei dem Vergleich der zwei Stimmschallanalysegeräte ist eine Differenzierung zwischen den möglichen Fehlern in der Durchführung der Studie und denen in der Berechnung der Stimmschallparameter durch die Analysegeräte selbst notwendig.

Zum Ausschluss von systematischen Fehlern im Ablauf der Studie war der Aufnahme- und Bearbeitungsalgorithmus (vgl. Anhang, Kapitel 6.3) bereits im Vorfeld zum einen auf die Einhaltung der Vorgaben von den Geräteherstellern und zum anderen auf die Herstellung von identischen Stimmsequenzen abgestimmt worden (vgl. Kapitel 2.3.1). Die Versuchsdurchführung erforderte des Weiteren ein ausreichendes Kooperationsverhalten durch die Teilnehmer. Die unterschiedliche Qualität der Phonationen als Folge ihrer möglicherweise unterschiedlichen Motivation stellte für die Ergebnisse der Studie hingegen weniger ein Problem dar. Für den Vergleich von den zwei synchron aufgenommenen Stimmsequenzen war die Qualität des Stimmsignals sekundär.

Besonderes Augenmerk galt dem Einhalten der vorgegebenen Abstände der Mikrofone zu den Studienteilnehmern, da hier bereits im Vorfeld eine nicht unerhebliche Fehlerquelle vermutet wurde. Der vom Hersteller WEVOS - WEVOSYS geforderte Abstand von 30 cm zu dem Versuchsteilnehmer ist unter klinischen Bedingungen bei den technischen Gegebenheiten nur schwer einzuhalten. Um diesem jedoch bestmöglich nachzukommen, wurde für diese Studie, wie bereits erwähnt, ein Abstandhalter installiert [97].

Aus anderen Arbeiten ist bekannt, dass sich der Mikrofonabstand auf die Ergebnisse der Jitter- und Shimmerwerte auswirkt. Titze und Winholtz [89] konnten belegen, dass sich neben der Empfindlichkeit der verwendeten Mikrofone auch der Abstand bzw. Winkel von der Signalquelle zum Mikrofon wesentlich auf die Berechnung von Pertubationsmaßen auswirkt. Dass es bei Vergrößerung des Mikrofonabstands zu einem Anstieg der Pertubationsmaße kommt,

kann nach Ansicht der Autoren mit der Verringerung des signal-to-noise-Ratios<sup>5</sup> begründet werden. In einer weiteren Studie von Winholtz und Titze [98] ließen sich für Headset-Mikrofone und Stativ-Mikrofone bei konstantem Abstand und unbeweglicher Kopfhaltung vergleichbare Ergebnisse erzielen. Bewegungsartefakte, verursacht durch Veränderungen des Quelle-Mikrofon-Abstands, führten jedoch zu Abweichungen des Amplitudenkoeffizienten um bis zu dem Faktor zwei bis drei.

Parsa, Jamieson und Pretty (2001) untersuchten, wie sich die Wahl des Mikrofons auf wichtige Stimmanalyseparameter auswirkt und wie man die Auswirkungen durch Filterung oder durch zusätzliche Parameter kompensieren könnte. Hierbei schlossen sie im Gegensatz zu der Arbeit von Titze und Winholtz (1993) die Messmethoden des glottalen Rauschverhaltens in ihre Untersuchungen mit ein. Diese hatten in einer vorangegangenen Arbeit von Parsa und Jamieson (2000) bessere Ergebnisse für die Einteilung von normalen und pathologischen Stimmen gegenüber den amplituden- bzw. grundfrequenzabhängigen Perturbationsmaßen ergeben. In der Tat konnte die Mikrofon-Charakteristik als Ursache für veränderte Werte der klinisch relevanten Stimmparameter durch Modifikation des Stimmsignals ausgemacht werden. Die Folge war eine verminderte Trennschärfe dieser Parameter, also eine schlechtere Unterscheidung zwischen normalen und pathologischen Stimmen. Unter Verwendung eines Equalizers (Gleichungsfilters) konnte die Klassifikationsgenauigkeit verbessert werden. Die Kombination von akustischen Messmethoden ließ eine bessere Trennschärfe erreichen [72; 73].

Um den Einfluss der Aufnahmequalität (Datenerfassung und Mikrofone) auf die Genauigkeit von akustischer Stimmschallanalyse zu erfassen, haben Deliyski et al. (2005) eine Studie mit sechs unterschiedlichen Mikrofonen und drei gängigen Aufnahmesystemen (MDVP, PRAAT, und TF32 - früher CSpeech) durchgeführt. Die Ergebnisse dieser durch Kombination geschaffenen 18 Datenerfassungssysteme wurden mithilfe des discretization error<sup>6</sup>, des effective dynamic range, und der ANOVA<sup>7</sup> analysiert. Die Auswertung ihrer Arbeit bestätigte den

---

<sup>5</sup> Das Verhältnis von Nutzsignal zu Umgebungsrauschen bzw. der Geräuschspannungsabstand [31].

<sup>6</sup> = Diskretisierungsfehler: Die Ungenauigkeit bei der Erzeugung einer Sampling-Rate durch einen Oszillator [17].

<sup>7</sup> „analysis of variance“ = Varianzanalyse.



vermuteten signifikanten Effekt der Aufnahmequalität auf die Berechnung von Stimmanalyseparametern. Es konnte gezeigt werden, dass der Diskretisierungsfehler der wichtigste Faktor für die Grundfrequenz  $F_0$  im Bereich der Datenerfassungshardware darstellt (im Gegensatz zu Umgebungslärm, Mikrofon-Type oder Software). Die Ergebnisse deuten zudem darauf hin, dass der effective dynamic range im Bereich der Datenerfassung die größte Bedeutung für die Messgenauigkeit des Shimmers aufweist und auch den Jitter als wichtiger Faktor beeinflusst [17].

Svec und Granqvist veröffentlichten 2010 einen Artikel über Richtlinien bezüglich der Auswahl von Mikrofonen für die Stimmschallanalyse. Nach Überprüfung und Analyse der Literatur kamen sie zu folgenden Empfehlungen:

Die Frequenzantwort der Mikrofone sollte gering ausfallen, d. h. eine Variation von weniger als 2 dB aufweisen. Der Frequenzbereich sollte zwischen der am niedrigsten zu erwartenden Grundfrequenz des Stimmsignals und der am meisten von Interesse seienden Spektral-Komponente liegen. Es wird ein Geräuschpegel der Mikrofone von mindestens 15 dB unterhalb des Schallpegels der leisesten Phonation empfohlen. Die obere Grenze des dynamischen Bereichs des Mikrofons sollte mindestens auf dem Level der lautesten Phonationen liegen. Richtmikrofone, die zu SPL- und spektralen Messmethoden verwendet werden, sollten nur in der Entfernung, innerhalb derer ihre maximale Frequenz gering ausfällt, platziert werden, um den Proximity-Effekt<sup>8</sup> zu vermeiden. Andernfalls seien sie für spektrale und Füllstandmessungen ungeeignet. Diese numerischen Werte für die Empfehlungen wurden für Mikrofone mit den Entfernungen 30 cm und 5 cm abgeleitet [85].

Ingrisano et al. (1998) untersuchten den Einfluss des Umgebungslärms auf die Berechnung von Stimmanalyseparametern. Hierzu kombinierten sie die Aufnahme eines lärmverursachenden Computerlüfters mit der eines natürlichen Stimmsignals bzw. einer künstlichen Schallwelle. Die Auswertung ergab in beiden Fällen erhöhte Werte für die Pertubationsmaße und könnte als lärm-induzierte Fehlermessungen beschrieben werden [41].

---

<sup>8</sup> = Nahbesprechungseffekt: Beschreibung der Anhebung der tiefen Frequenzen bei Richtmikrofonen im Nahfeld einer Schallquelle [22].

Bei der Ursachenforschung für die fehlende Vergleichbarkeit von Stimmanalysegeräten und deren Parametern spielt demnach neben den Anwendungsfehlern in erster Linie die technische Komponente eine bedeutende Rolle. Auf der Ebene der Hardware sind das zum einen die bereits genannten Faktoren Qualität und Positionierung der Mikrofone und zum anderen die eingesetzten Soundkarten bzw. Datenerfassungssysteme. Im Bereich der Software ist die Aufbereitung eines Stimmsignals im Sinne der Anwendung von z. B. Filtertechniken sowie deren Auswertung mittels Analyseprogrammen von Bedeutung. Dass die Berechnung der Pertubationsmaße im Besonderen von einer möglichst exakten Grundperiodenbestimmung abhängt und dass hier eine bedeutende Fehlerquelle vermutet wird, wurde bereits in den Kapiteln 1.5.4 und 4.2 ausgeführt.

Ma et al. (2005) haben im Rahmen einer Studie, in welcher 112 krankhafte und 41 gesunde Stimmen in drei Gruppen (periodisch/nahezu periodisch, teils subharmonisch und aperiodisch) unterteilt wurden, die Frage untersucht, ob die häufig verwendeten akustischen Pertubationsmaße tatsächlich für die Analyse von pathologischen Stimmsequenzen geeignet seien. Obwohl nur die erste Gruppe (periodische bzw. nahezu periodische Stimmsignale) untersucht wurde, wiesen lediglich drei der vier untersuchten Stimmparameter (Grundfrequenz, relatives amplitudenabhängiges Perturbationmaß, Shimmer (%) und Noise-to-harmonic Ratio), und zwar für die Frauen, signifikante Unterschiede auf. Einschränkung muss allerdings bedacht werden, dass die untersuchten männlichen Stimmen der ersten Gruppe eine geringe Fallzahl (vier bzw. zwei) aufwiesen. Trotzdem gelangten die Autoren zu der Aussage, dass die klinische Anwendung der akustischen Pertubationsmaße unter diesen Bedingungen überdacht werden müsse [53].

Eine retrospektive Studie von Werth et al. (2010) umfasste die Analyse von ausgewählten klinischen Pertubationsmaßen, des Dysphonie Schweregrad Index (DSI) und der subjektiv wahrgenommenen Heiserkeit bei 580 gesunden bzw. 1700 krankhaften Stimmen. Es konnten zwar Unterschiede zwischen den klinischen Gruppen aufgezeigt werden, eine individuelle Stimmbeurteilung war jedoch aufgrund der hohen Schwankung der Parameterwerte nicht möglich [95]. Brockmann-Bauser et al. (2011) kamen in einer Arbeit, welche auf Grundlage einer Medline-Recherche der letzten zwei Jahre entstand, ebenfalls zu dem

Ergebnis, dass die aktuell am meisten verwendeten Pertubationsmaße – Jitter und Shimmer – nicht die erhofften unabhängigen und objektiven Parameter darstellen, wie häufig angenommen. Laut ihrer Recherche sind sie weder für Stimmlippenveränderungen noch für die Stimmqualität bzw. den Heiserkeitsgrad einer Stimme als Indikatoren verwendbar. Zum Nachweis des Letzteren führten sie u. a. Daten einer zuvor unveröffentlichten Vergleichsstudie mit perceptiven Heiserkeitsmessungen (GRBAS-Skala) und Jitter- bzw. Shimmer-Messungen (PRAAT) auf. Diese ergab keinen Zusammenhang zwischen den genannten Pertubationsmaßen und dem Schweregrad einer Stimmstörung. Sie wiesen außerdem auf die Bedeutung möglicher Confounder wie Geschlecht, Art des Vokals und Phonationslautstärke bzw. -höhe hin. Erst nach Berücksichtigung bzw. Vereinheitlichung der Analyseverfahren bestünde die Möglichkeit, das Potenzial von akustischen Analysemethoden besser zu erkennen bzw. effektiver auszuschöpfen [14; 15].

Die Art des Vokals als möglicher Confounder wurde u. a. von MacCallum et al. (2011) und Kiliç et al. (2004) beschrieben. In beiden Fällen ließen sich abweichende Ergebnisse zwischen unterschiedlichen Vokalen durch keine entsprechende Werte gängiger Pertubationsmaße wiedergegeben [46; 54].

## 4.4 Schlussfolgerung

Vergleicht man die Ergebnisse der vorliegenden Studie mit den oben genannten Arbeiten, lässt sich übereinstimmend feststellen, dass die computergestützte Stimmschallanalyse zum aktuellen Zeitpunkt noch nicht als ein so zuverlässiges und optimales Instrument zur Beurteilung von Stimmstörungen bezeichnet werden kann, wie erhofft.

Die bereits auf diesem Gebiet geleistete wissenschaftliche Arbeit soll jedoch an dieser Stelle – gerade vor dem Hintergrund der fehlenden diagnostischen Alternativen zu unserem menschlichen Ohr – nicht geschmälert werden. Allerdings ist es ratsam, bevor in weiteren Studien neue Grenzwerte für Pertubationsmaße getestet oder die Frage nach einer optimaleren Kombination der technischen Ausrüstung diskutiert werden, eine grundsätzliche Strukturierung der modernen Stimmschallanalyse vorzunehmen. Wie die oben aufgeführten Studien zeigen, sind die Probleme der fehlenden Vergleichbarkeit unterschiedlicher Datenerfassungs- und Analysesysteme auf zwei wesentliche Ursachen zurückzuführen.

Zunächst muss gewährleistet sein, dass die bei der akustischen Stimmschallanalyse zur Verfügung stehenden Komponenten, bestehend aus Hardware, Software sowie Know-how des Anwenders, eine hinreichende Qualität nachweisen. Jeder Bestandteil dieses Prozesses muss demnach auf seine Eignung bzw. Auswirkungen für die Ergebnisse untersucht werden. Beginnend mit Ort der Aufnahme und Quantifizierung des Umgebungslärms über Qualität und Position des Mikrofons bis hin zu Qualität der Datenerfassungssysteme inkl. Soundkarten sowie richtige Anwendung der Aufnahmealgorithmen bzw. Analysesoftware sind viele Stellschrauben zu beachten, die auf die Qualität der Stimmschallanalyse Einfluss haben [87]. Hinzu kommt ein weiterer bedeutender Punkt, der in dieser Arbeit mehrfach betont wurde: Neben der Optimierung der soeben aufgelisteten Einflussfaktoren muss die Frage nach den geeigneten Stimmschallanalyseparametern gestellt werden. Gerade in neueren Arbeiten wurden Zweifel im Hinblick auf die klinische Anwendbarkeit der am häufigsten gebrauchten Pertubationsmaße laut. Neben der fraglichen Eignung wurde auch auf die Auswirkung weiterer Confounder (Geschlecht, Art des Vokals und Phonationslautstärke bzw. –höhe) hingewiesen [15; 53].

Im Rahmen dieser beschriebenen „Qualitätssicherung“ ließe sich ein weiterer Sachverhalt wesentlich leichter umsetzen: die Vergleichbarkeit der verschiedenen computergestützten Systeme zur Stimmschallanalyse. Lägen einheitliche Kriterien für die Verwendung von Hard- und Software sowie Aufnahme- und Analysealgorithmen vor, könnten mögliche Komponenten leichter kombiniert bzw. ganze Datenerfassungs- und Analysesysteme leichter miteinander verglichen werden. So wie Svec und Granqvist (2010) eine Richtlinie für die Verwendung von Mikrofonen in der Stimmschallanalyse entwickelten, wäre ein international einheitliches Protokoll für die Durchführung der gesamten computerbasierten Stimmschallanalyse vorteilhaft [85; 87].

## 5 Zusammenfassung

Die Kommunikation nimmt in unserer Gesellschaft eine immer größere Bedeutung ein, so dass die Stimme als wichtiges Instrument zur verbalen Verständigung dadurch vermehrt gefordert wird. Zwangsläufig führt die höhere Belastung zu einer größeren Anzahl an Stimmfunktionsstörungen und damit zu einer zunehmenden Inanspruchnahme der Stimmdiagnostik. Hierfür stehen verschiedene diagnostische Methoden zur Verfügung, die im Wesentlichen im Basisprotokoll der Europäischen Laryngologischen Gesellschaft (ELS) aufgeführt werden. Neben der perzeptiven und subjektiven Beurteilung der Stimmfunktion hat sich in den letzten Jahrzehnten die objektive Beurteilung im Sinne der computergestützten Stimmschallanalyse durchgesetzt. Hierfür werden z. B. aufgenommene Stimmsignale auf verschiedene Parameter hin untersucht.

Ziel dieser Studie war es, zwei kommerziell verfügbare Analysegeräte (lingWAVES von WEVOS - WEVOSYS, Forchheim, und rpSzene<sup>®</sup> von Rheder/Partner, Hamburg) auf die Vergleichbarkeit ihrer Ergebnisse (Stimmschallparameter) hin zu überprüfen. Hierfür unterzogen sich im Zeitraum von Februar 2009 bis Juni 2010 60 Studienteilnehmer – davon 30 Stimmpatienten und 30 Probanden im Alter von 18 bis 81 Jahren – einer Stimmschallanalyse. Es wurden die sechs Vokale [a], [ä], [e], [i], [o] und [u] als gehaltene Vokale synchron in die zwei Stimmschallanalysegeräte phoniert und jeweils getrennt anhand der drei Stimmschallparameter Jitter, Shimmer und Glottal-to-Noise Excitation Ratio ausgewertet. Anschließend erfolgte die statistische Auswertung der Ergebnisse mithilfe des Programms MS Excel<sup>®</sup> – Redmond, USA, und mit Bland-Altman Plots in dem Programm MedCalc<sup>®</sup> – Mariakerke, Belgien.

Die Auswertung ergab keine Vergleichbarkeit von lingWAVES und rpSzene<sup>®</sup> für die untersuchten Stimmparameter Jitter, Shimmer oder GNE. Der Jitter zeigte eine große Schwankungsbreite in Abhängigkeit von dem „wahren“ Wert, sodass eine klare Aussage im Hinblick auf die Gleichwertigkeit beider Stimmschallanalysegeräte ebenso wenig möglich war wie eine mögliche Beurteilung bezüglich der klinischen Relevanz der Geräteunterschiede. Der Shimmer konnte erst nach Modulation der Bland-Altman Plots mittels Berechnung der Ratios ausgewertet werden und zeigte nicht nur einen systematischen, sondern auch einen

klinisch relevanten Unterschied bei fünf Vokalen (Ausnahme Vokal [o]). Die dabei ermittelten Umrechnungsfaktoren in Form der mittleren Ratios müssen jedoch aufgrund der recht breiten Übereinstimmungsintervalle für die praktische Anwendung als zu ungenau bewertet werden. Die Auswertung der GNE-Werte ergab zwar klassische Bland-Altman Plots mit typischer Punkteverteilung und eine im Vergleich zu den vorherigen Parametern auch deutlich bessere Übereinstimmung, ließ jedoch letztlich aufgrund der Lage der Intervallgrenzen für den Übereinstimmungsbereich keine eindeutige Aussage über die Vergleichbarkeit von lingWAVES und rpSzene<sup>®</sup> zu.

Da sich diese Ergebnisse auch mit Resultaten aus anderen Studien decken und im Rahmen der Ursachenforschung essentielle, zum Teil ungelöste Einflussfaktoren zu finden sind, kann die computergestützte Stimmschallanalyse zum aktuellen Zeitpunkt noch nicht als ein so zuverlässiges und optimales Instrument zur Beurteilung von Stimmstörungen bezeichnet werden, wie erhofft. Es wird eine Überprüfung der Qualität aller einzelnen Bestandteile, die Standardisierung der Algorithmen sowie die Überarbeitung der gängigsten Stimmanalyseparameter in der computergestützten Stimmschallanalyse empfohlen.

## Summary

Communication in our society has become progressively important, so that voice as an important instrument for verbal communication is demanded increasingly. Inevitably, the increased strain leads to a greater number of vocal dysfunctions and thus to a growing demand for voice diagnostics. There are various diagnostic methods available which are listed essentially in the basic protocol of the European Laryngological Society (ELS). In addition to the perceptual and subjective assessment of voice function, the objective assessment in accordance with computerized voice sound analysis has been very successful in recent decades. To achieve this, for example, recorded voice signals are investigated on various parameters.

The aim of this study was to examine two commercially available analyzers (lingWAVES by Wevosys, Forchheim and rpScene<sup>®</sup> by Rehder / Partner, Hamburg) on the comparability of its results (vocal acoustic parameters). For this project during the period of February 2009 until June 2010 60 study participants – including 30 voice patients and 30 subjects aged 18 to 81 years – were investigated with vocal sound analysis. The following six vowels [a] [ae], [e], [i], [o] and [u] were phonated as synchronously sustained vowels in the two vocal sound analyzer and were analyzed separately on the basis of the three vocal acoustic parameters jitter, shimmer and Glottal-to-Noise Excitation Ratio. Subsequently, the results were evaluated statistically by using Microsoft Excel<sup>©</sup> – Redmond, U.S., and Bland-Altman plots in the program MedCalc<sup>®</sup> – Mariakerke, Belgium.

The evaluation showed no comparability of lingWAVES and rpScene<sup>®</sup> for the investigated vocal parameters jitter, shimmer, or GNE. Jitter showed a large variation in dependence to the "true" value, so that a clear statement with a view to the equivalence of both vocal sound analyzers was just as impossible as a possible assessment regarding the clinical relevance of the differences between both analyzers. Only after modulation of the Bland-Altman plots by calculating the ratios could Shimmer be evaluated. Not only did it show a systematic, but also a clinically relevant difference of five vowels (except for vowel [o]). The determined conversion factors in the form of average ratios however must be evaluated as too vague for the practical application because of the fair-



ly wide limits of agreement. Although the analysis of the GNE-values showed classical Bland-Altman plots with typical distribution of points and a clearly better agreement than the previous parameters, it was however impossible to give an ultimately clear statement about the comparability of lingWAVES and rpSzene<sup>®</sup> because of the location of the interval limits of agreement.

Because these results correspond with results of other studies and because essential factors, which are partly unresolved, can be found as a part of causality research, computer-based vocal sound analysis cannot be defined as such a reliable and useful tool for the evaluation of voice disorders at the present time, as expected. It is recommended to review the quality of all the individual components, the standardization of the algorithms and the revision of the popular voice analysis parameters in the computer-assisted analysis of vocal sound.

## 6 Anhang

### 6.1 Informationsblatt und Einverständniserklärung

	<b>Philipps Universität Marburg</b>
<p>Philipps-Universität, Fachbereich Medizin, 35033 Marburg</p>	<p>Fachbereich Medizin Klinik für Phoniatrie und Pädaudiologie <b>Prof. Dr. med. Roswitha Berger</b> Tel.: (0 64 21) 58 6-64 39 Fax: (0 64 21) 58 6-28 24 E-Mail: <a href="mailto:bergerr@med.uni-marburg.de">bergerr@med.uni-marburg.de</a> Anschrift: Deutschhausstr. 3 35037 Marburg Web: <a href="http://www.uni-marburg.de/phoniatrie">www.uni-marburg.de/phoniatrie</a> Az.: Prof. Berger Marburg, 21.09.2010</p>
<h3>INFORMATIONSBLETT</h3> <p>ZUR STUDIE:</p> <p>„Untersuchungen zur Vergleichbarkeit von Ergebnissen der Stimmchallanalyse kommerziell verfügbarer Geräte“</p>	
<p><b>Sehr geehrte Patientin, sehr geehrter Patient,</b></p>	
<p>Erkrankungen der menschlichen Stimme haben in den letzten Jahren zunehmend soziale Bedeutung erlangt, da in den meisten Berufen die stimmlichen Anforderungen gestiegen sind.</p>	
<p>Stimmstörungen führen zu einer Beeinträchtigung des Stimmklanges in Form von Heiserkeit und zur Einschränkung der stimmlichen Leistungsfähigkeit. Gerade diese Beeinträchtigungen haben Auswirkungen auf die Berufsausübung, besonders wenn eine Tätigkeit mit hoher Sprech- und Stimmbelastung vorliegt. Da eine Stimmstörung noch immer zu spät diagnostiziert wird, findet auch die stimmliche Behandlung zu spät statt. Häufig ergibt sich aus dieser Tatsache die Notwendigkeit des Berufswechsels oder der Berufsunfähigkeit.</p>	
<p>Über die noch verbleibende stimmliche Leistung einer bereits erkrankten Stimme gibt es bisher kaum Untersuchungen. Die objektive Diagnostik von Stimmstörungen wird durch mehrere Faktoren erschwert:</p>	
<ul style="list-style-type: none"><li>▶ durch die große individuelle Vielfalt der menschlichen Stimme</li><li>▶ dadurch, dass die Stimmfunktion bewusst krankhaft verändert werden kann, und</li><li>▶ durch die Tatsache, dass die Stimmfunktion an drei unterschiedliche Funktionsbereiche (Atmungssystem, Stimmlippenschwingung und Resonanzräume) gekoppelt ist.</li></ul>	

- 2 -

Ein „pathologischer Stimmklang“ wird als Heiserkeit wahrgenommen. Als Ursache dafür kommt eine Glottisschlussinsuffizienz und/oder ein irreguläres Schwingungsverhalten der Stimmlippen in Betracht. Heiserkeit kann in drei Schweregraden vorliegen und wird auditiv nach „Behauchtheit und Rauigkeit“ quantifiziert.

Für die objektive Diagnostik einer Stimmstörung hat die akustische Schallanalyse in den letzten Jahren zunehmend an Bedeutung gewonnen, da mit den zur Verfügung stehenden Programmen der Heiserkeitsgrad ermittelt werden kann und dadurch gesunde und krankhaftveränderte Stimmen beurteilt werden können.

Es existieren mehrere kommerziell verfügbarere computergestützte Systeme zur Analyse des Stimmschalls. Allerdings ist bisher nicht bekannt, ob die verschiedenen Systeme untereinander vergleichbar sind.

Wir möchten jetzt der Fragestellung nachgehen, ob identische Stimmaufnahmen mit 2 verschiedenen Analyseprogrammen identische Ergebnisse des Schallsignals liefern.

Für die vorgelegte Studie sollen akustische Stimmparameter bestimmt und miteinander verglichen werden. Wir möchten dafür die „kurzfristige Schwankung der Grundfrequenz“ den Jitter in %, die „kurzfristige Amplitudenmodulation den Shimmer in % und das Maß für das Signal/Rausch Verhältnis des Stimmsignals den Glottal-to Noise Excitation Ratio, GNE genannt, nutzen.

Dazu werden die üblichen Schallanalyse-Untersuchungscomputer mit dem „Göttinger Heiserkeitsdiagramm“ in rp scene und dem „lingWaves“ von „WEVOS- WEVOSYS“ genutzt.

Geplant ist eine Stimmaufnahme, bei der Sie als Patientin bzw. Patient 6 Vokale für jeweils 3 Sekunden simultan in 2 verschiedene Mikrophone phonieren. Die Aufnahme mit dem Schallpegelmesser wird von der Analysesoftware „lingWaves“, die Aufnahme mit dem Headset von dem „Göttinger Heiserkeitsdiagramm“ aufgezeichnet.

Zu Beginn der Aufnahme werden die vom Hersteller angegebenen Abstände zu den beiden Mikrofonen und die Einhaltung eines maximalen Rauschpegels in Ruhe von 45 dB überprüft. Die Probeaufnahme dient zur Ermittlung der geeigneten Tonhöhe in der jeweiligen Indifferenzlage (Männer: 120 – 190Hz, Frauen: eine Oktave höher), und der vorgegebenen Tonlautstärke von 70 +/- 5 dB mit Hilfe des Schallpegelmessers.

Ein vorgeschalteter Signalton ermöglicht das Schneiden der Aufnahmen in zeitlich identische Sequenzen, welche von beiden Analyseprogrammen an Hand der Parameter Jitter, Shimmer und GNE ausgewertet werden.

- 3 -

Aus den Ergebnissen dieser Studie erhoffen wir uns neue Erkenntnisse über die Vergleichbarkeit der beiden genannten computergestützten Systeme zur Stimmschallanalyse und für die daraus resultierende klinische Konsequenz zu erhalten.

Wir würden uns freuen, wenn Sie an unserer Studie teilnehmen.

Ihre Daten werden vertraulich behandelt und pseudonymisiert bearbeitet, d. h. in Form einer Schlüsselliste erfasst, welche bis zum Ende der Studie bei den Bearbeitern verbleibt. Danach werden Ihre Daten gelöscht.

Die Teilnahme an der Studie ist freiwillig. Sie können jederzeit und ohne Angabe von Gründen Ihre Einwilligung widerrufen bzw. zurückziehen, ohne dass das Vertrauensverhältnis zu Ihrem behandelnden Arzt in irgendeiner Weise leidet oder dieses nachteilige Folgen für Ihre ärztliche Behandlung hat.

Wir bedanken uns für Ihre Mithilfe

Prof. Dr. R. Berger

**Abbildung 6.1:** Informationsblatt zur Studie für die Patienten

Philipps



Universität  
Marburg

Fachbereich Medizin  
Klinik für Phoniatrie und Pädaudiologie  
Anschrift: Deutschhausstr. 3, 35037 Marburg

**Prof. Dr. med. Roswitha Berger**  
Tel.: (0 64 21) 28 6-64 39  
Fax: (0 64 21) 28 6-28 24  
E-Mail: bergerr@med.uni-marburg.de  
Web: www.uni-marburg.de/phoniatrie

## EINVERSTÄNDNISERKLÄRUNG

ZUR STUDIE:

„Untersuchungen zur Vergleichbarkeit von Ergebnissen der Stimmschallanalyse  
kommerziell verfügbarer Geräte“

Ich, \_\_\_\_\_ (Name des Patienten), wurde von meinem  
Arzt vollständig über Wesen, Bedeutung und Tragweite der klinischen Prüfung mit dem o.g. Titel  
aufgeklärt. Ich habe den Aufklärungstext gelesen und verstanden. Ich hatte die Möglichkeit Fragen  
zu stellen, und habe die Antworten verstanden und akzeptiere sie. Mein Arzt hat mich über die mit  
der Teilnahme an der Studie verbundenen Risiken und den möglichen Nutzen informiert.

Ich hatte ausreichend Zeit, mich zur Teilnahme an dieser Studie zu entscheiden und weiß, dass  
die Teilnahme an dieser klinischen Studie freiwillig ist. Ich weiß, dass ich jederzeit und ohne  
Angaben von Gründen diese Zustimmung widerrufen kann, ohne dass sich dieser Entschluss  
nachteilig auf die spätere Behandlung durch meinen Arzt auswirken wird.

Mir ist bekannt, dass meine persönlichen Daten in verschlüsselter Form gespeichert werden.

Ich habe eine Kopie der Patienteninformation und dieser Einwilligungserklärung erhalten. Ich  
erkläre hiermit meine freiwillige Teilnahme an dieser klinischen Studie.

\_\_\_\_\_  
Ort und Datum

\_\_\_\_\_  
Unterschrift des Patienten

\_\_\_\_\_  
Ort und Datum

\_\_\_\_\_  
Unterschrift des aufklärenden Arztes

**Abbildung 6.2:** Einverständniserklärung für die Patienten

## 6.2 Studienteilnehmer

Probandenliste					
Nr:	Initialen:	Geschlecht:	Alter:	Vorerkrankungen der Stimme:	Aufgeklärt:
101	XX	m	30	-	✓
102	XX	m	22	-	✓
103	XX	w	23	-	✓
104	XX	m	23	-	✓
105	XX	m	29	-	✓
106	XX	w	23	-	✓
107	XX	w	24	-	✓
108	XX	w	55	-	✓
109	XX	m	19	-	✓
110	XX	w	23	-	✓
111	XX	w	35	gem. funkt. Dysphonie	✓
112	XX	w	24	-	✓
113	XX	w	29	-	✓
114	XX	m	30	-	✓
115	XX	m	31	-	✓
116	XX	m	56	-	✓
117	XX	m	23	-	✓
118	XX	w	24	-	✓
119	XX	w	21	-	✓
120	XX	w	30	-	✓
121	XX	w	20	-	✓
122	XX	w	21	-	✓
123	XX	m	18	-	✓
124	XX	m	36	-	✓
125	XX	m	33	-	✓
126	XX	m	35	-	✓
127	XX	m	44	-	✓
128	XX	m	49	-	✓
129	XX	w	21	-	✓
130	XX	w	23	-	✓

Tabelle 6.1: 30 Probandenteilnehmer mit Geschlecht, Alter und Vorerkrankungen der Stimme; Blau = Männer; Rot = Frauen

Patientenliste					
Nr:	PID:	Geschlecht:	Alter:	Diagnose:	Aufgeklärt:
201	...xxx	w	68	Funktionelle Dysphonie	✓
202	...xxx	w	66	Funktionelle Dysphonie	✓
203	...xxx	w	51	Funktionelle Dysphonie	✓
204	...xxx	w	78	Z. n. Stimmlippen - OP	✓
205	...xxx	m	75	Funktionelle Dysphonie	✓
206	...xxx	w	68	Funktionelle Dysphonie	✓
207	...xxx	w	46	Stimmlippenödem	✓
208	...xxx	m	32	Z. n. Laryngitis	✓
209	...xxx	w	70	Z. n. Stimmlippenparese	✓
210	...xxx	m	81	Funktionelle Dysphonie	✓
211	...xxx	w	67	Funktionelle Dysphonie	✓
212	...xxx	m	61	Stimmlippen - Missbildung	✓
213	...xxx	m	69	Stimmlippenparese	✓
214	...xxx	w	59	Stimmlippenparese	✓
215	...xxx	w	57	Stimmlippenödem	✓
216	...xxx	m	29	Larynxstenose	✓
217	...xxx	w	58	Stimmlippenpolyp	✓
218	...xxx	w	80	Funktionelle Dysphonie	✓
219	...xxx	w	28	Phoniationsverdickung	✓
220	...xxx	m	62	N. laryng. sup. - Schaden	✓
221	...xxx	w	54	Stimmlippenparese	✓
222	...xxx	m	61	Funktionelle Dysphonie	✓
223	...xxx	w	42	Chronische Laryngitis	✓
224	...xxx	m	55	Chronische Laryngitis	✓
225	...xxx	m	48	Chronische Laryngitis	✓
226	...xxx	m	55	Stimmlippenparese rechts	✓
227	...xxx	m	70	Funktionelle Dysphonie	✓
228	...xxx	m	52	Funktionelle Dysphonie	✓
229	...xxx	m	47	Z. n. Stimmlippenknötchen	✓
230	...xxx	m	71	Z. n. Stimmlippenpolyp	✓

Tabelle 6.2: 30 Patiententeilnehmer mit Geschlecht, Alter und Vorerkrankungen der Stimme; Blau = Männer; Rot = Frauen

## 6.3 Algorithmen

### AUFNAHME-MODUS

- ✓ **Aufklärung !!!**
- ✓ Mikrophone **einschalten**
- ✓ Mikrofon**abstände / -höhe** kontrollieren (lingWaves: 30cm; rpSzene®: Headset)
- ✓ **Rauschpegel** ablesen (< 45dB)
- ✓ **Tonhöhe** vorgeben: Indifferenzlage (Männer: 120 – 190Hz, Frauen: Oktave drüber)
- ✓ **Lautstärke** testen: Probe-Vokal (70 +/-5 dB)
- ✓ Aufnahmen **starten** und bis zum Ende durchlaufen lassen
- ✓ **Signalton** einspielen (Kuli-„Klick“; nur zu Beginn; direkt am Mund des Probanden)
- ✓ Vokale (s. u.) jeweils **mindestens 3 Sekunden** aushalten
- ✓ Aufnahme **beenden**
- ✓ Stimmsignal **exportieren**:
  - In lingWaves nach: C:\lingWAVES\_Daten\Doktorand\...
  - *(letzter Button in der Mitte oben „als WAV-Datei speichern“)*
  - In rpSzene® nach: C:\Dokumente und Einstellungen\Administrator\...
  - *(Rechtsklick im Signalfeld → „Exportieren...“)*
- ✓ **Initialen/PID, Geschlecht, Alter, phonetische Vorerkrankungen/Diagnose** in Liste eintragen

“a” “ä” “e” “i” “o” “u”

Abbildung 6.3: Aufnahme-Modus mit abgedruckten Vokalen zum Ablesen für die Teilnehmer



## AUFNAHME- UND BEARBEITUNGSGRUNDRISS

1. Beide Mikrophone und die beiden PCs einschalten
2. Die Programme **lingWAVES** und **rpSzene**<sup>®</sup> aufrufen
3. Aufnahme-Modus: *(siehe dazu gleichnamiges Extrablatt)*
  - a. Rauschpegel ablesen
  - b. Beide Aufnahmen starten
  - c. Signalton (Kuli-Klick) erzeugen
  - d. Vokale durch Probanden aufsagen lassen:
    - ✓ Möglichst in gleicher Tonlage & Lautstärke
    - ✓ Mikrophonabstände  
*(lingWAVES: 30cm; rpSzene<sup>®</sup>: Headset)*
    - ✓ **Mindestens 3 Sekunden**
  - e. Aufnahme beenden
4. **rpSzene**<sup>®</sup>:
  - a. Die 6 Aufnahmen (Vokal a – u) in den entsprechenden Ordner bzw. auf USB-Stick exportieren (z. B. als a.rp.wav, ä.rp.wav ...)
  - b. Die Projektdateien ebenfalls speichern *(nicht zwingend erforderlich)*
5. **lingWAVES**:
  - a. USB-Stick anschließen und die erste Wave-Datei von **rpSzene**<sup>®</sup> (z. B. Vokal a) ins Aufnahme-Fenster importieren
  - b. Diese mit der entsprechenden Vokalaufnahme von **lingWAVES** vergleichen und identische Abstandsmaße für den Analyseabschnitt festlegen (z. B. Beginn des Analyseabschnittes 1000ms nach Start des Signaltons)
  - c. Den Analyseabschnitt auf eine Länge von **1500ms** markieren, in ein neues Aufnahmedokument einfügen und als Wave-Datei abspeichern bzw. auf den USB-Stick exportieren (z. B. als a-schnitt.rp.wav, a-schnitt.ling.wav ...)
  - d. Wie in b. und c. beschrieben jeweils mit allen weiteren Vokalaufnahmen von **rpSzene**<sup>®</sup> und **lingWAVES** verfahren
  - e. Wave-Dateien von **lingWAVES** im VOSPECTOR öffnen und Jitter, Shimmer sowie GNE berechnen, danach Ergebnisse in Tabelle übertragen
6. **rpSzene**<sup>®</sup>:
  - a. Den USB-Stick wieder anschließen und die bearbeiteten Wave-Dateien in den **rpSzene**<sup>®</sup>-Bearbeitungsmodus importieren
  - b. Die Aufnahmen komplett markieren und einzeln mit dem Heiserkeits-Diagramm analysieren
  - c. Die Heiserkeits-Diagramme zusammenführen, nach Excel exportieren und die Werte für Jitter, Shimmer und GNE in o. g. Tabelle übernehmen
7. Analyse: Auswertung der Ergebnisse der einzelnen Parameter mit MS Excel<sup>®</sup> bzw. Medcalc<sup>®</sup>

Abbildung 6.4: Gesamter Bearbeitungsalgorithmus mit integriertem Aufnahmemodus

## 6.4 Ergebnistabellen

### 6.4.1 Jitter

lingWAVES		Jitter (%)					
	A	Ä	E	I	O	U	
1	0,15	0,08	0,10	0,11	0,09	0,09	
2	0,59	0,13	0,84	0,39	0,10	0,12	
3	0,18	0,15	0,15	0,20	0,14	0,23	
4	0,12	0,09	0,09	0,07	0,08	0,12	
5	0,17	0,19	0,75	0,07	0,20	0,16	
6	0,11	1,17	0,22	2,03	0,26	0,22	
7	0,20	0,14	0,22	0,44	3,40	0,53	
8	0,05	0,07	0,05	0,05	0,03	0,05	
9	0,19	0,09	0,13	0,12	0,09	0,12	
10	0,21	0,23	0,19	0,08	0,04	0,10	
11	0,19	0,14	0,06	0,08	0,11	0,14	
12	0,06	0,14	0,09	0,07	0,14	0,27	
13	0,10	0,22	0,19	0,10	0,22	0,24	
14	0,09	0,11	0,07	0,07	0,05	0,10	
15	0,12	0,43	0,19	0,05	0,08	0,07	
16	0,18	0,15	0,11	0,08	0,10	0,09	
17	0,22	0,77	0,19	0,15	0,13	0,08	
18	0,07	0,06	0,12	0,08	0,04	0,05	
19	0,14	0,06	0,09	0,10	0,14	0,28	
20	0,11	0,15	0,23	0,06	0,05	0,09	
21	0,11	0,12	0,06	0,13	0,12	0,11	
22	0,06	0,03	0,10	0,11	0,04	0,12	
23	0,18	0,15	0,17	0,17	0,08	0,12	
24	0,08	0,13	0,10	0,11	0,07	0,10	
25	0,13	0,07	0,11	0,13	0,07	0,18	
26	0,07	0,06	0,06	0,05	0,06	0,04	
27	0,07	0,04	0,03	0,07	0,05	0,07	
28	0,19	0,28	0,23	0,16	0,19	0,16	
29	0,33	0,46	0,14	0,15	0,09	0,10	
30	0,07	0,10	0,05	0,05	0,13	0,17	
31	0,20	0,30	0,23	0,15	0,15	0,33	
32	0,30	0,63	1,77	1,27	0,27	0,57	
33	0,28	0,31	0,23	0,11	0,08	0,22	
34	0,14	0,09	0,15	0,21	0,07	0,35	
35	0,17	0,15	0,10	0,10	0,14	0,10	
36	2,51	0,20	0,08	1,83	0,05	2,88	
37	6,24	1,09	3,14	2,28	0,86	0,48	
38	0,14	0,15	0,45	0,55	0,66	0,38	
39	0,72	0,16	0,14	0,06	0,20	0,14	
40	0,33	0,37	0,22	0,12	0,29	0,14	
41	0,24	0,21	0,11	0,36	0,08	0,07	
42	3,43	2,40	0,77	3,77	0,13	0,38	
43	7,12	0,33	3,56	0,31	0,92	4,00	
44	0,37	0,66	0,60	2,14	0,29	0,42	
45	0,13	0,08	0,07	0,05	0,07	0,12	
46	1,16	0,41	0,58	0,45	0,12	0,36	
47	0,13	0,11	0,05	0,03	0,06	0,15	
48	0,16	0,07	0,14	0,10	0,17	0,28	
49	0,55	0,34	0,27	0,25	0,19	0,38	
50	0,05	0,07	0,07	0,05	0,05	0,06	
51	0,24	0,14	1,29	0,17	0,29	0,16	
52	0,16	0,29	0,24	0,18	0,12	0,16	
53	0,88	0,44	2,19	0,13	0,13	0,13	
54	0,45	0,24	0,30	0,29	0,12	0,17	
55	0,11	0,07	0,06	0,05	0,06	0,09	
56	14,36	0,35	3,09	1,36	1,03	0,23	
57	0,15	0,39	0,33	0,08	0,13	0,12	
58	0,11	0,12	0,29	0,14	0,25	0,14	
59	0,07	0,09	0,07	0,10	0,05	0,11	
60	0,07	0,06	0,06	0,08	0,09	0,12	

Tabelle 6.3: Jitterwerte für lingWAVES

rp-Szene®		Jitter (%)					
	A	Ä	E	I	O	U	
1	0,20	0,16	0,18	0,13	0,15	0,11	
2	0,30	0,18	0,20	0,50	0,24	0,27	
3	0,30	0,28	0,15	0,38	0,26	0,23	
4	0,14	0,15	0,15	0,08	0,10	0,08	
5	0,31	0,32	0,51	0,12	0,17	0,13	
6	0,27	0,27	0,24	0,28	0,68	0,11	
7	0,25	0,34	0,24	0,24	0,47	0,29	
8	0,08	0,09	0,07	0,08	0,02	0,12	
9	0,25	0,19	0,17	0,15	0,12	0,10	
10	0,14	0,15	0,11	0,10	0,05	0,12	
11	0,11	0,18	0,11	0,04	0,10	0,07	
12	0,15	0,15	0,12	0,10	0,09	0,18	
13	0,20	0,20	0,29	0,23	0,19	0,30	
14	0,24	0,23	0,14	0,10	0,08	0,08	
15	0,30	0,37	0,31	0,09	0,16	0,15	
16	0,20	0,27	0,20	0,15	0,12	0,09	
17	0,25	0,15	0,19	0,14	0,26	0,11	
18	0,11	0,13	0,15	0,17	0,06	0,10	
19	0,31	0,10	0,07	0,11	0,07	0,20	
20	0,20	0,22	0,23	0,12	0,09	0,15	
21	0,19	0,26	1,41	0,31	0,23	0,21	
22	0,13	0,10	0,07	0,09	0,05	0,12	
23	0,34	0,23	0,23	0,22	0,13	0,16	
24	0,22	0,31	0,16	0,14	0,13	0,12	
25	0,11	0,15	0,15	0,15	0,09	0,26	
26	0,08	0,08	0,10	0,10	0,08	0,09	
27	0,09	0,09	0,09	0,12	0,09	0,05	
28	0,21	0,34	0,30	0,19	0,23	0,21	
29	0,43	2,11	0,19	0,11	0,19	0,22	
30	0,07	0,13	0,07	0,09	0,08	0,08	
31	0,11	0,32	0,20	0,11	0,12	0,20	
32	0,24	0,54	1,77	1,98	0,19	0,42	
33	0,17	0,25	0,19	0,11	0,08	0,12	
34	0,28	0,14	0,17	0,12	0,08	0,33	
35	0,28	0,27	0,23	0,15	0,22	0,11	
36	2,81	0,14	0,10	0,11	0,06	0,23	
37	1,03	0,47	0,67	0,87	0,26	0,19	
38	0,37	0,25	0,20	0,19	0,16	0,30	
39	0,22	0,33	0,21	0,10	0,13	0,19	
40	0,20	0,21	0,16	0,05	0,23	0,11	
41	0,23	0,18	0,10	0,07	0,07	0,08	
42	0,65	0,73	0,51	0,26	0,15	0,24	
43	1,05	0,22	0,34	0,52	0,62	3,49	
44	0,21	0,41	0,37	4,01	0,34	0,39	
45	0,14	0,10	0,10	0,06	0,08	0,06	
46	0,30	0,31	0,43	0,50	3,00	0,44	
47	0,14	0,13	0,08	0,04	0,04	0,09	
48	0,17	0,16	0,13	0,12	0,13	0,24	
49	0,33	0,30	0,30	0,46	0,19	0,39	
50	0,06	0,19	0,18	0,08	0,08	0,06	
51	0,40	0,15	0,27	0,17	0,35	0,26	
52	0,38	0,43	0,38	0,27	0,18	0,23	
53	0,65	0,80	1,13	0,21	0,13	0,30	
54	0,31	0,25	0,21	0,25	0,13	0,15	
55	0,14	0,12	0,11	0,08	0,08	0,06	
56	10,19	0,61	1,02	1,00	1,85	0,27	
57	0,08	0,15	0,11	0,13	0,08	0,07	
58	0,18	0,22	0,38	0,14	0,48	0,18	
59	0,13	0,19	0,08	0,07	0,07	0,07	
60	0,10	0,18	0,11	0,25	0,10	0,11	

Tabelle 6.4: Jitterwerte für rpScene®

## 6.4.2 Shimmer

	lingWAVES		Shimmer (%)			
	A	Ä	E	I	O	U
1	22,89	5,49	2,70	2,90	4,88	3,25
2	11,64	8,12	6,91	7,53	2,73	3,91
3	9,39	12,53	8,29	5,48	4,05	3,09
4	22,18	18,74	6,41	8,56	5,69	5,02
5	16,05	8,09	5,78	6,87	7,65	4,63
6	6,88	7,50	5,40	15,01	12,70	8,24
7	7,91	4,97	8,85	11,89	15,75	10,83
8	7,13	3,33	8,12	5,42	6,34	5,16
9	12,13	10,54	9,90	8,67	7,34	4,86
10	9,85	12,07	15,26	4,97	2,21	4,47
11	6,93	5,30	2,59	2,38	3,09	2,91
12	4,13	2,88	3,65	2,17	2,52	4,19
13	6,18	3,04	5,72	4,00	3,11	2,70
14	5,81	4,24	3,87	11,69	6,28	7,29
15	10,22	10,82	10,07	7,38	4,24	4,16
16	14,94	13,09	11,99	9,46	10,58	8,29
17	24,04	16,91	16,63	8,68	8,56	8,59
18	4,00	3,25	3,66	4,62	3,40	1,83
19	7,59	4,60	7,58	7,07	6,62	6,72
20	6,15	4,11	4,01	3,96	3,11	5,12
21	4,76	4,69	10,83	3,47	4,01	3,45
22	7,32	7,22	5,65	3,13	2,16	2,42
23	10,58	9,43	6,02	11,95	6,08	6,42
24	9,33	7,14	11,62	6,89	9,87	3,83
25	9,56	11,26	10,77	7,18	17,05	8,58
26	7,95	6,11	4,97	3,88	3,94	4,09
27	4,01	4,09	5,22	4,32	7,50	4,69
28	16,89	15,06	10,56	11,45	13,79	5,34
29	9,02	8,85	4,38	7,03	3,50	4,40
30	5,25	5,09	5,02	5,69	4,85	9,11
31	2,49	7,15	5,92	5,93	4,20	7,66
32	12,61	10,22	21,65	11,87	12,57	6,45
33	14,40	4,43	5,54	5,37	5,38	4,54
34	9,27	6,24	10,08	4,28	7,50	4,43
35	18,95	15,18	9,87	7,32	17,47	17,04
36	15,28	9,00	8,58	5,02	2,54	6,55
37	33,96	15,76	14,73	18,27	10,07	12,40
38	9,27	10,02	14,25	16,21	14,00	16,58
39	4,62	6,53	8,11	6,59	3,66	5,06
40	15,38	15,20	6,08	3,18	9,30	5,25
41	9,58	3,83	3,69	5,63	2,72	3,42
42	30,97	13,19	12,57	10,72	6,49	7,86
43	37,88	15,38	22,05	8,70	15,38	22,92
44	9,87	7,72	10,46	40,10	6,83	7,29
45	10,47	7,81	8,36	4,02	9,18	4,09
46	14,02	10,15	7,93	8,27	7,84	13,27
47	7,29	2,64	1,58	1,88	4,30	2,75
48	11,56	5,74	2,95	3,13	2,76	2,44
49	13,47	9,11	10,18	3,81	6,35	6,59
50	7,74	5,42	5,78	4,45	5,60	2,46
51	11,61	4,81	5,19	3,75	3,97	3,82
52	7,67	7,82	5,63	7,49	4,89	6,39
53	13,30	22,93	15,31	8,85	4,52	3,37
54	14,07	6,67	11,44	8,85	8,92	8,76
55	6,07	6,21	4,85	4,92	6,35	4,90
56	32,11	10,74	12,18	12,10	15,30	6,31
57	9,50	9,42	5,18	3,88	5,39	3,85
58	12,10	11,80	9,91	9,60	10,35	4,84
59	16,96	7,78	8,87	7,75	7,49	6,60
60	6,34	4,24	4,38	8,16	3,52	3,05

Tabelle 6.5: Shimmerwerte für lingWAVES

	rp-Szene®		Shimmer (%)			
	A	Ä	E	I	O	U
1	4,45	1,06	1,07	0,65	1,98	2,66
2	1,86	1,75	2,34	1,92	1,95	2,08
3	3,13	2,27	1,53	1,76	1,24	1,34
4	2,02	2,79	2,31	1,42	1,69	1,77
5	2,81	2,35	2,11	1,33	2,17	1,40
6	2,11	1,77	1,88	1,62	2,32	1,33
7	1,90	1,78	1,68	2,00	3,71	2,99
8	1,21	1,06	0,90	1,57	1,27	1,19
9	3,09	2,40	2,37	2,11	1,90	2,22
10	1,21	1,25	1,56	1,25	1,34	1,60
11	1,32	0,94	0,65	0,55	0,87	0,92
12	1,37	1,35	1,03	0,92	1,49	1,44
13	2,29	1,58	1,33	1,84	0,96	1,38
14	1,84	1,52	1,64	1,22	1,26	1,13
15	2,35	3,07	2,70	2,87	2,03	1,50
16	4,11	2,90	1,65	2,20	1,83	2,43
17	4,92	6,59	3,68	2,95	2,41	2,72
18	1,52	1,18	0,85	1,03	0,98	1,01
19	3,16	3,18	2,64	1,65	1,95	1,85
20	1,43	1,40	1,45	1,41	2,01	2,23
21	1,90	2,30	3,17	0,80	1,14	0,95
22	1,00	0,91	1,40	0,74	0,79	1,09
23	2,79	2,80	3,28	2,58	2,81	1,85
24	1,88	1,95	1,32	1,21	2,14	2,41
25	3,76	2,04	1,78	4,80	0,98	3,42
26	1,43	1,49	1,31	1,29	1,08	1,21
27	0,79	2,23	1,38	2,59	2,75	1,55
28	4,33	2,41	2,21	1,67	3,20	3,09
29	2,66	5,27	2,21	1,92	1,82	2,00
30	2,66	1,43	1,27	1,42	1,26	1,78
31	9,34	1,90	1,31	0,93	1,77	1,28
32	2,69	3,40	5,83	5,78	2,72	1,60
33	1,67	1,52	1,77	1,29	1,64	1,21
34	3,34	4,34	4,58	1,27	3,51	1,79
35	5,70	3,86	3,14	2,07	2,25	4,42
36	5,51	1,48	0,98	1,01	1,00	1,26
37	9,17	3,00	4,82	7,23	2,94	3,05
38	3,94	2,12	2,88	3,66	3,27	3,60
39	1,83	1,86	1,17	1,40	1,11	1,26
40	3,63	3,82	2,34	1,24	2,97	1,84
41	2,48	1,23	1,46	1,06	1,35	1,25
42	5,40	4,81	5,37	2,06	2,18	2,60
43	7,74	6,94	5,18	7,38	6,59	43,49
44	2,12	1,71	1,80	47,31	2,30	5,51
45	3,15	2,54	4,56	1,07	1,95	1,71
46	3,81	2,74	5,26	6,36	22,82	10,44
47	1,26	0,98	0,80	0,56	1,14	0,94
48	1,78	1,17	0,91	0,84	1,18	1,30
49	4,26	4,48	8,84	7,26	5,29	6,08
50	1,38	1,48	0,94	0,79	1,35	0,69
51	3,08	1,56	1,47	1,94	2,32	4,07
52	2,18	1,57	1,85	2,45	1,94	1,59
53	7,26	13,64	10,88	2,89	2,68	1,87
54	2,86	2,09	3,00	1,75	3,40	3,99
55	2,36	1,46	1,23	0,99	1,87	1,04
56	37,15	3,71	3,53	3,95	6,31	3,30
57	1,09	1,53	1,90	1,35	1,50	1,21
58	2,66	3,06	3,97	2,65	4,19	3,39
59	2,22	1,35	1,03	1,59	0,96	1,66
60	2,75	1,70	1,46	3,03	1,86	1,81

Tabelle 6.6: Shimmerwerte für rpScene®

## 6.4.3 GNE

	lingWAVES						GNE					
	A	Ä	E	I	O	U	A	Ä	E	I	O	U
1	0,51	0,49	0,81	0,75	0,13	0,12						
2	0,89	0,88	0,91	0,85	0,86	0,61						
3	0,74	0,81	0,76	0,86	0,76	0,23						
4	0,72	0,83	0,84	0,74	0,38	0,04						
5	0,81	0,82	0,77	0,60	0,18	0,10						
6	0,51	0,69	0,51	0,42	0,12	0,07						
7	0,53	0,75	0,56	0,42	0,09	0,09						
8	0,67	0,69	0,87	0,79	0,59	0,68						
9	0,80	0,87	0,88	0,86	0,61	0,45						
10	0,81	0,54	0,67	0,60	0,36	0,20						
11	0,77	0,70	0,80	0,82	0,66	0,09						
12	0,83	0,80	0,85	0,86	0,54	0,18						
13	0,82	0,89	0,89	0,82	0,74	0,40						
14	0,87	0,90	0,93	0,84	0,79	0,20						
15	0,86	0,91	0,87	0,83	0,74	0,56						
16	0,50	0,77	0,66	0,85	0,30	0,26						
17	0,74	0,72	0,74	0,73	0,60	0,16						
18	0,70	0,76	0,78	0,79	0,47	0,08						
19	0,87	0,87	0,80	0,79	0,13	0,10						
20	0,90	0,90	0,84	0,74	0,27	0,11						
21	0,79	0,79	0,73	0,67	0,47	0,31						
22	0,89	0,92	0,88	0,76	0,82	0,43						
23	0,82	0,87	0,82	0,65	0,58	0,06						
24	0,89	0,89	0,88	0,74	0,66	0,15						
25	0,75	0,86	0,75	0,71	0,52	0,20						
26	0,89	0,84	0,88	0,83	0,78	0,69						
27	0,85	0,87	0,88	0,68	0,49	0,25						
28	0,85	0,82	0,86	0,67	0,48	0,22						
29	0,79	0,71	0,67	0,54	0,37	0,07						
30	0,82	0,87	0,86	0,81	0,33	0,08						
31	0,94	0,82	0,40	0,55	0,10	0,06						
32	0,76	0,85	0,78	0,41	0,07	0,13						
33	0,65	0,67	0,43	0,74	0,18	0,17						
34	0,54	0,72	0,44	0,37	0,27	0,13						
35	0,87	0,83	0,89	0,91	0,80	0,41						
36	0,60	0,80	0,74	0,55	0,50	0,24						
37	0,24	0,32	0,20	0,12	0,07	0,06						
38	0,54	0,78	0,61	0,51	0,20	0,08						
39	0,82	0,80	0,82	0,77	0,48	0,45						
40	0,37	0,60	0,57	0,36	0,13	0,19						
41	0,68	0,79	0,80	0,44	0,22	0,12						
42	0,18	0,31	0,35	0,24	0,04	0,13						
43	0,35	0,30	0,41	0,31	0,24	0,22						
44	0,16	0,45	0,37	0,42	0,05	0,07						
45	0,47	0,70	0,49	0,68	0,18	0,20						
46	0,52	0,58	0,69	0,66	0,75	0,54						
47	0,81	0,92	0,88	0,91	0,62	0,30						
48	0,49	0,85	0,63	0,71	0,27	0,04						
49	0,21	0,22	0,32	0,50	0,11	0,09						
50	0,90	0,74	0,90	0,86	0,75	0,51						
51	0,77	0,78	0,82	0,58	0,40	0,08						
52	0,68	0,70	0,47	0,43	0,07	0,10						
53	0,93	0,84	0,89	0,77	0,88	0,82						
54	0,18	0,63	0,30	0,67	0,31	0,13						
55	0,80	0,67	0,62	0,84	0,70	0,44						
56	0,20	0,71	0,74	0,42	0,29	0,17						
57	0,33	0,47	0,45	0,65	0,18	0,10						
58	0,96	0,92	0,96	0,90	0,88	0,83						
59	0,61	0,71	0,75	0,73	0,38	0,11						
60	0,75	0,50	0,66	0,60	0,51	0,30						

Tabelle 6.7: GNE-Werte für lingWAVES

	rp-Szene®						GNE					
	A	Ä	E	I	O	U	A	Ä	E	I	O	U
1	0,55	0,61	0,90	0,86	0,39	0,35						
2	0,90	0,97	0,97	0,96	0,90	0,80						
3	0,93	0,95	0,95	0,91	0,92	0,79						
4	0,80	0,91	0,93	0,91	0,76	0,09						
5	0,94	0,92	0,88	0,88	0,77	0,39						
6	0,87	0,80	0,86	0,87	0,66	0,13						
7	0,74	0,82	0,79	0,76	0,34	0,26						
8	0,85	0,87	0,94	0,92	0,85	0,87						
9	0,90	0,97	0,96	0,96	0,86	0,83						
10	0,92	0,67	0,86	0,87	0,80	0,59						
11	0,89	0,89	0,96	0,96	0,89	0,40						
12	0,91	0,94	0,92	0,95	0,82	0,47						
13	0,93	0,97	0,96	0,94	0,91	0,51						
14	0,98	0,98	0,98	0,95	0,94	0,61						
15	0,95	0,97	0,95	0,96	0,92	0,86						
16	0,57	0,88	0,79	0,96	0,65	0,47						
17	0,84	0,84	0,86	0,86	0,83	0,48						
18	0,81	0,77	0,90	0,95	0,68	0,50						
19	0,95	0,97	0,89	0,90	0,47	0,42						
20	0,97	0,97	0,96	0,94	0,93	0,71						
21	0,90	0,92	0,87	0,85	0,90	0,76						
22	0,93	0,97	0,94	0,95	0,90	0,88						
23	0,92	0,95	0,94	0,88	0,90	0,31						
24	0,97	0,97	0,98	0,89	0,91	0,66						
25	0,85	0,96	0,91	0,88	0,70	0,40						
26	0,90	0,92	0,95	0,92	0,81	0,84						
27	0,93	0,95	0,87	0,82	0,73	0,55						
28	0,93	0,95	0,97	0,81	0,87	0,61						
29	0,89	0,80	0,84	0,54	0,16	0,81						
30	0,92	0,96	0,94	0,93	0,86	0,47						
31	0,80	0,86	0,56	0,77	0,30	0,22						
32	0,81	0,91	0,84	0,63	0,16	0,12						
33	0,74	0,82	0,84	0,92	0,72	0,46						
34	0,64	0,71	0,57	0,50	0,40	0,40						
35	0,95	0,88	0,95	0,96	0,88	0,77						
36	0,67	0,88	0,84	0,57	0,68	0,33						
37	0,35	0,45	0,40	0,20	0,11	0,09						
38	0,86	0,86	0,80	0,68	0,41	0,21						
39	0,91	0,94	0,97	0,96	0,93	0,83						
40	0,56	0,77	0,69	0,63	0,31	0,27						
41	0,73	0,90	0,89	0,66	0,71	0,36						
42	0,22	0,39	0,46	0,27	0,14	0,16						
43	0,46	0,54	0,44	0,48	0,20	0,27						
44	0,32	0,54	0,55	0,58	0,15	0,10						
45	0,51	0,75	0,68	0,87	0,57	0,50						
46	0,61	0,69	0,80	0,80	0,84	0,66						
47	0,92	0,97	0,93	0,95	0,85	0,59						
48	0,67	0,91	0,76	0,85	0,28	0,08						
49	0,33	0,35	0,44	0,69	0,60	0,36						
50	0,96	0,88	0,94	0,83	0,95	0,92						
51	0,91	0,94	0,91	0,64	0,39	0,17						
52	0,76	0,76	0,54	0,43	0,61	0,18						
53	0,96	0,90	0,91	0,78	0,59	0,84						
54	0,18	0,72	0,44	0,82	0,31	0,10						
55	0,92	0,87	0,76	0,96	0,89	0,49						
56	0,23	0,76	0,74	0,55	0,47	0,26						
57	0,36	0,53	0,58	0,79	0,38	0,09						
58	0,98	0,94	0,95	0,93	0,74	0,81						
59	0,74	0,81	0,93	0,89	0,67	0,22						
60	0,79	0,57	0,74	0,81	0,61	0,72						

Tabelle 6.8: GNE-Werte für rpSzene®

## 6.5 MedCalc®: Bland-Altman Plots

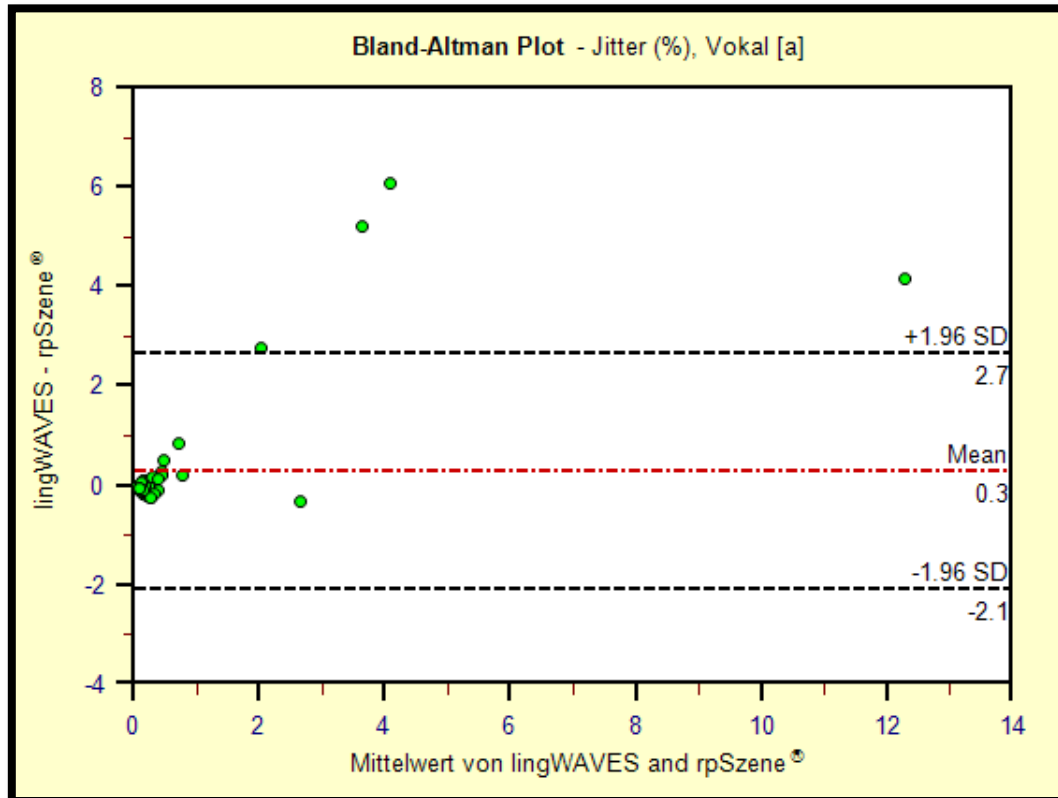


Abbildung 6.5: Bland-Altman Plot für Jitterwerte des Vokals [a]

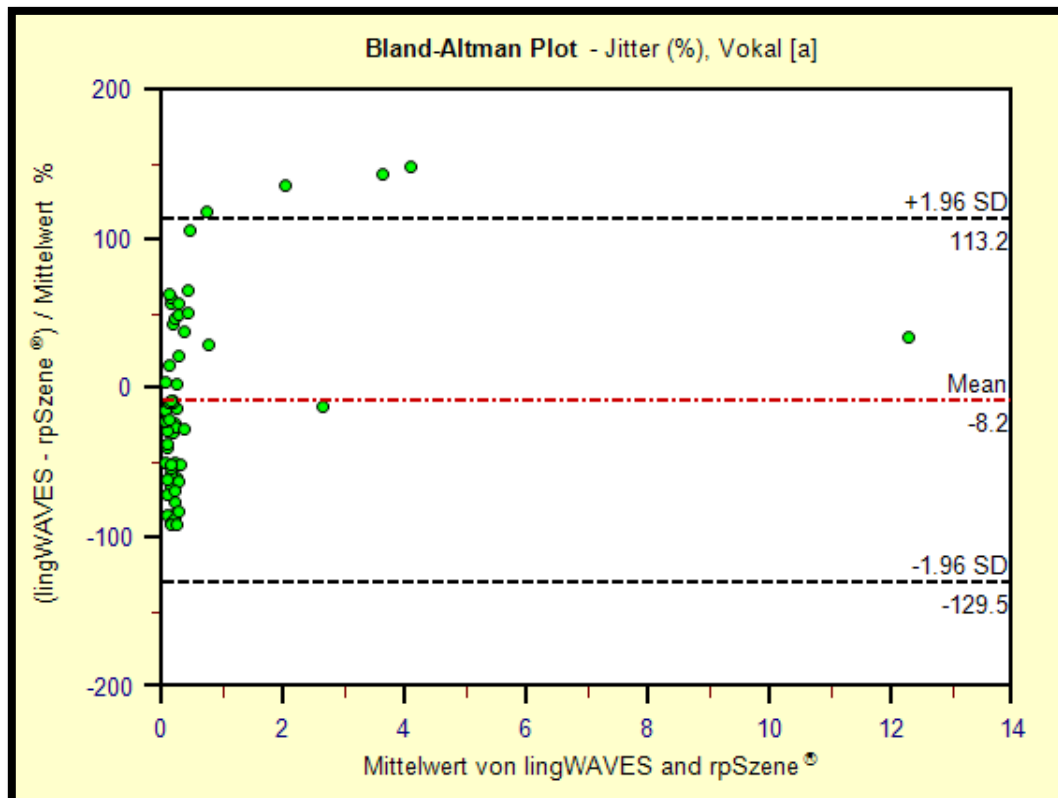


Abbildung 6.6: Bland-Altman Plot für Jitterwerte des Vokals [a] mit Differenzen in %

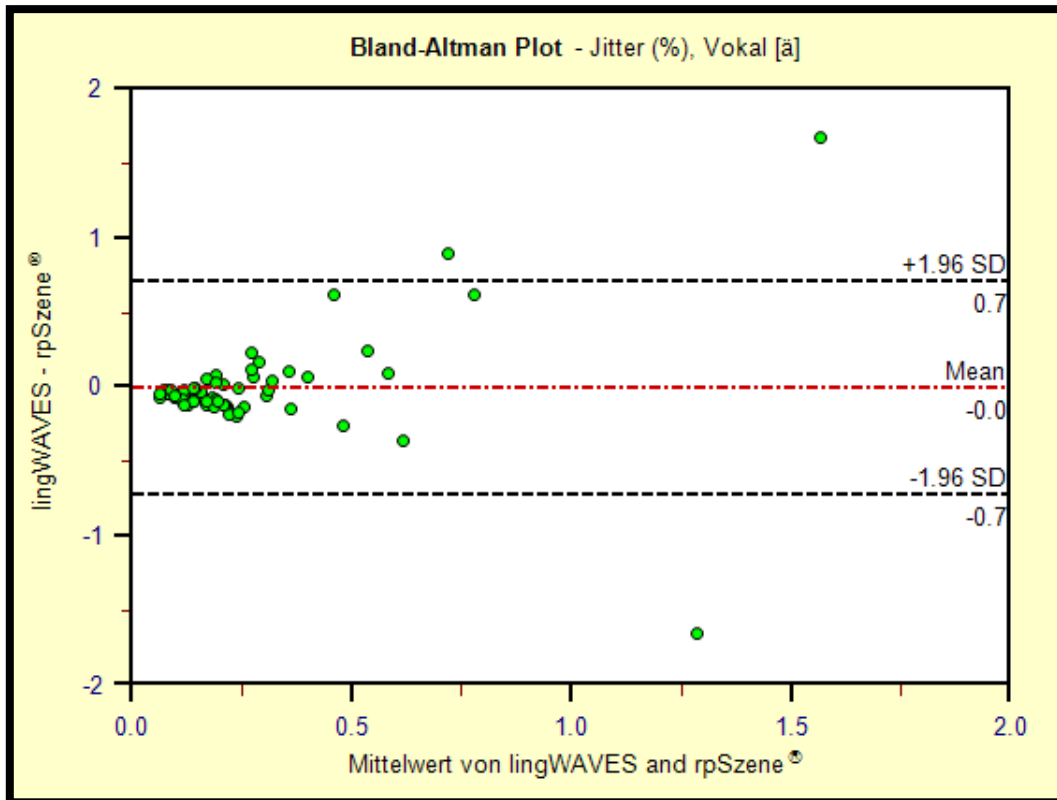


Abbildung 6.7: Bland-Altman Plot für Jitterwerte des Vokals [ä]

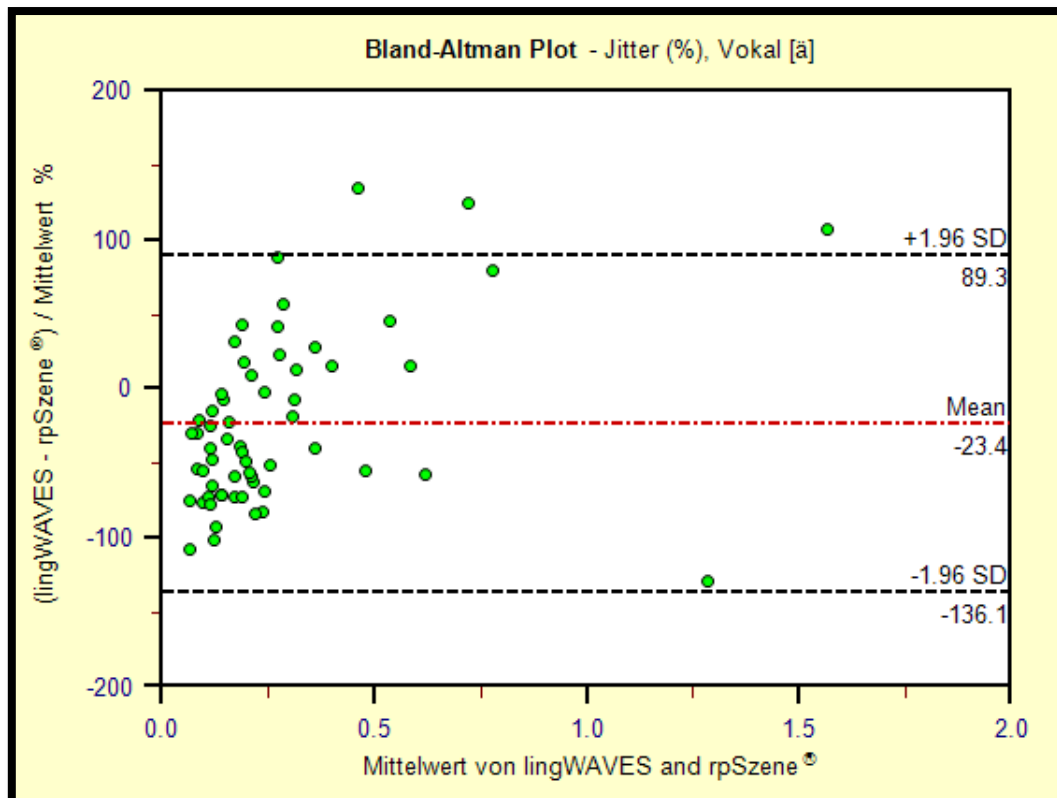


Abbildung 6.8: Bland-Altman Plot für Jitterwerte des Vokals [ä] mit Differenzen in %

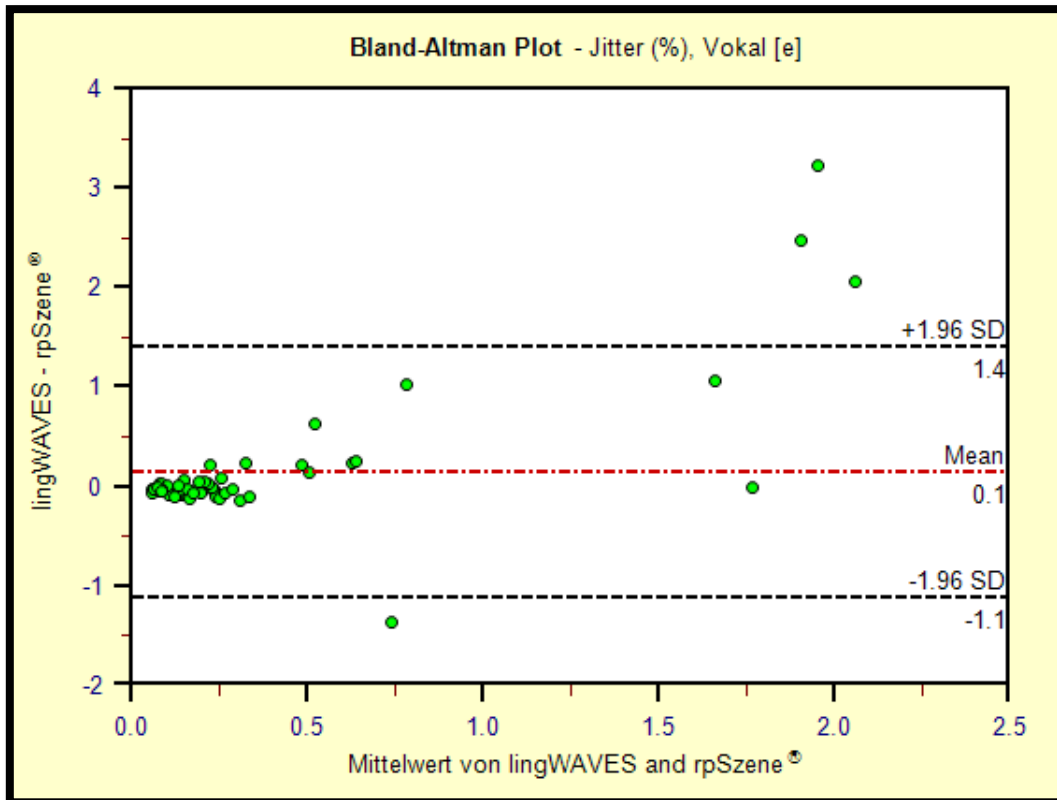


Abbildung 6.9: Bland-Altman Plot für Jitterwerte des Vokals [e]

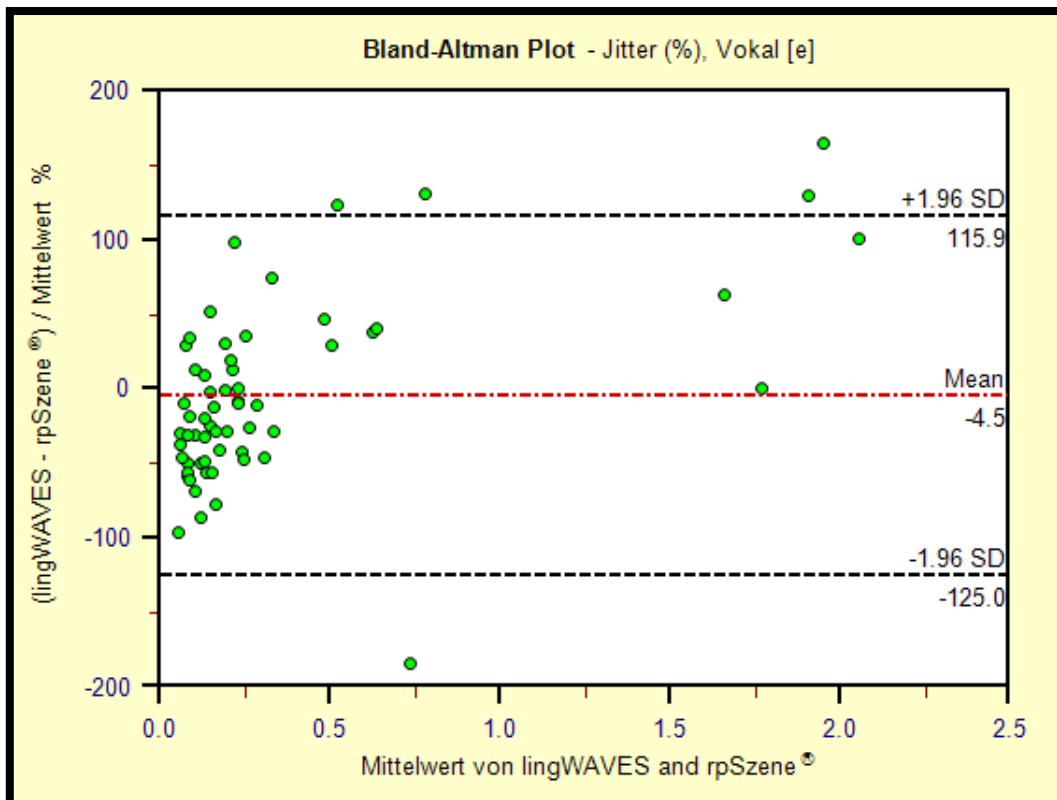


Abbildung 6.10: Bland-Altman Plot für Jitterwerte des Vokals [e] mit Differenzen in %

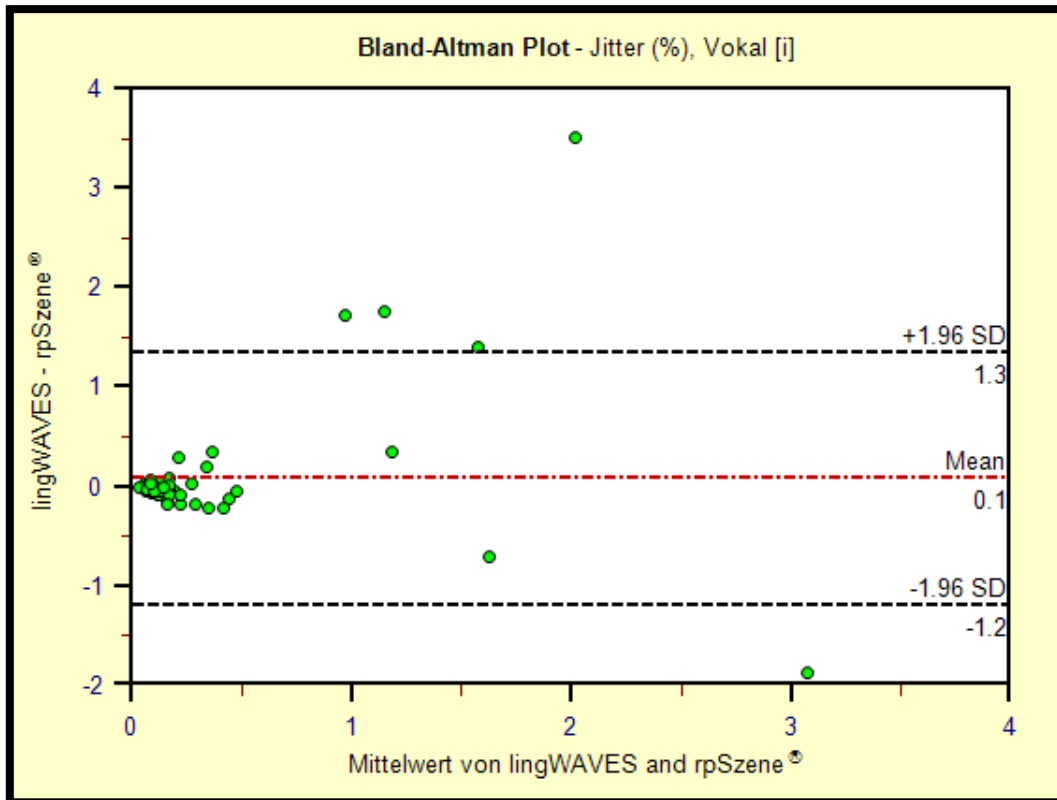


Abbildung 6.11: Bland-Altman Plot für Jitterwerte des Vokals [i]

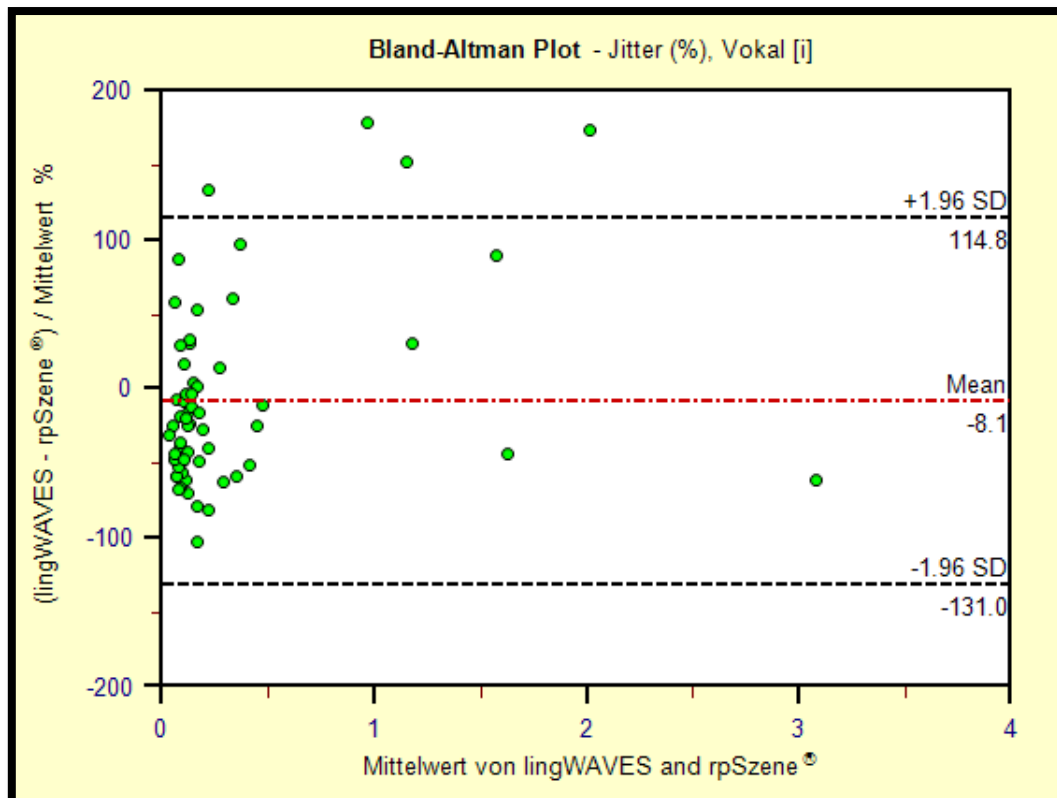


Abbildung 6.12: Bland-Altman Plot für Jitterwerte des Vokals [i] mit Differenzen in %



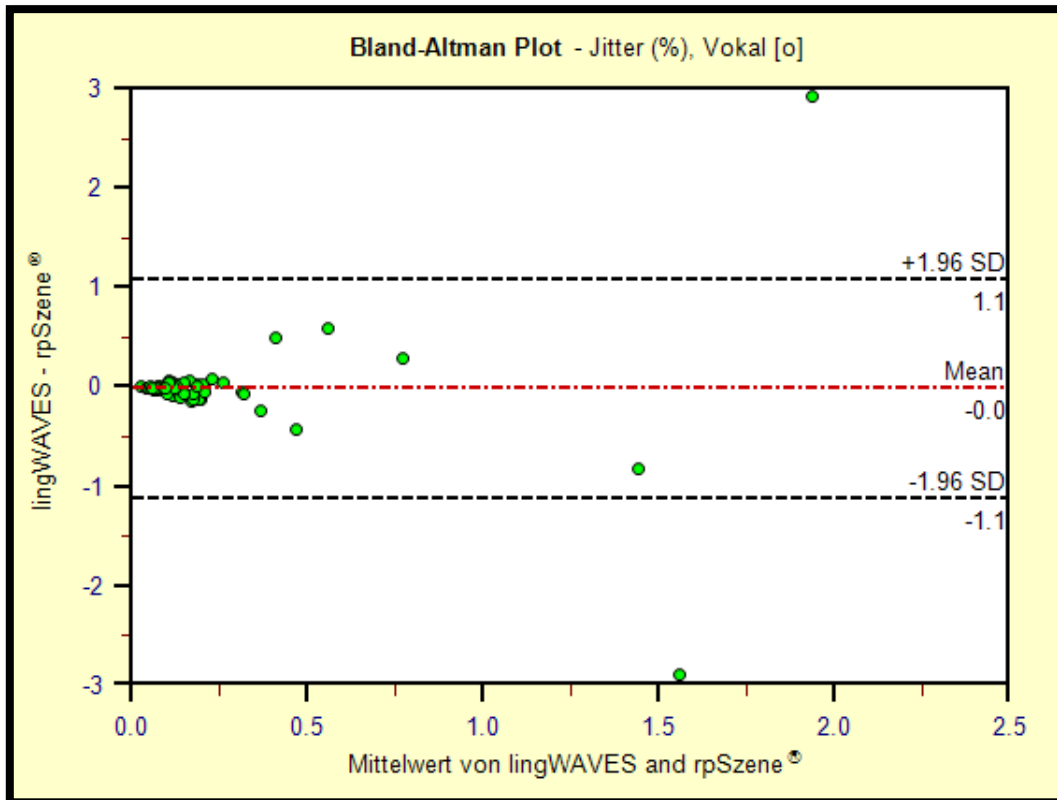


Abbildung 6.13: Bland-Altman Plot für Jitterwerte des Vokals [o]

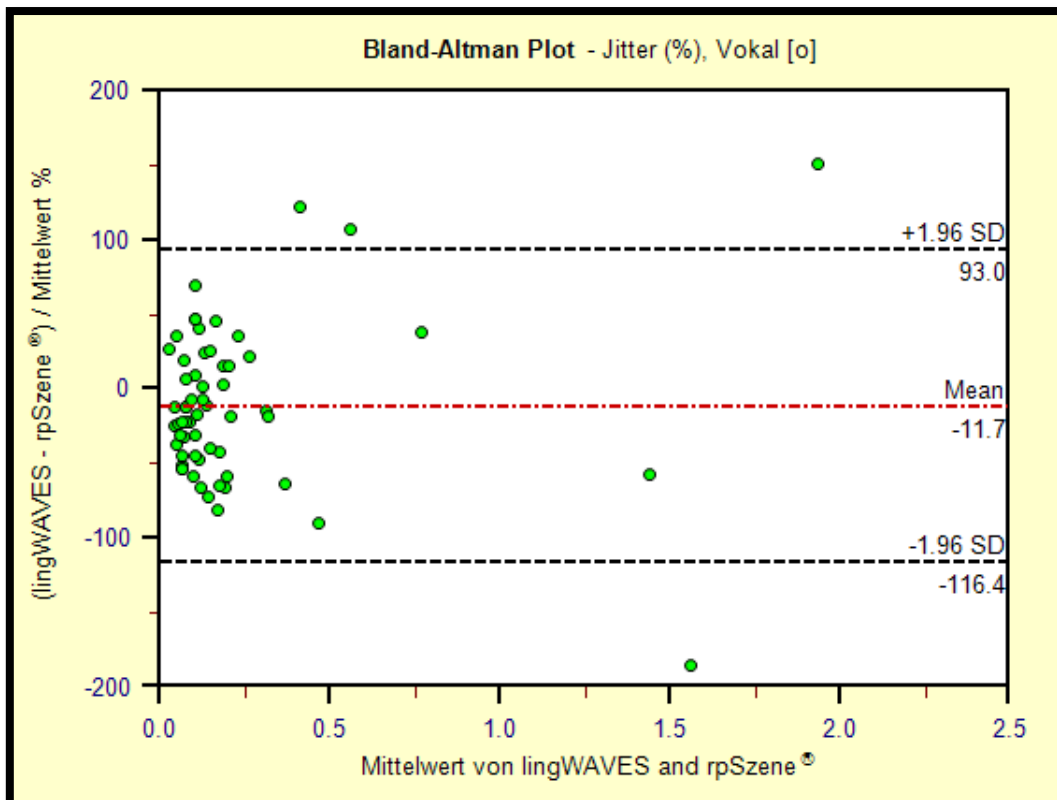


Abbildung 6.14: Bland-Altman Plot für Jitterwerte des Vokals [o] mit Differenzen in %

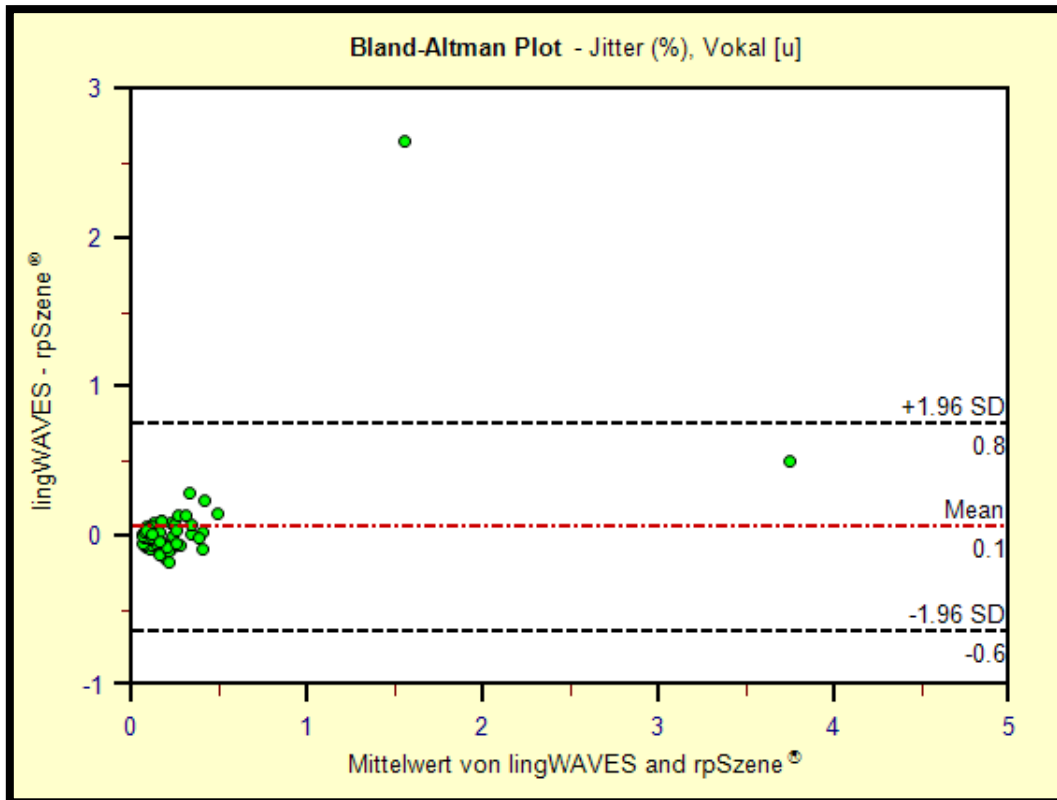


Abbildung 6.15: Bland-Altman Plot für Jitterwerte des Vokals [u]

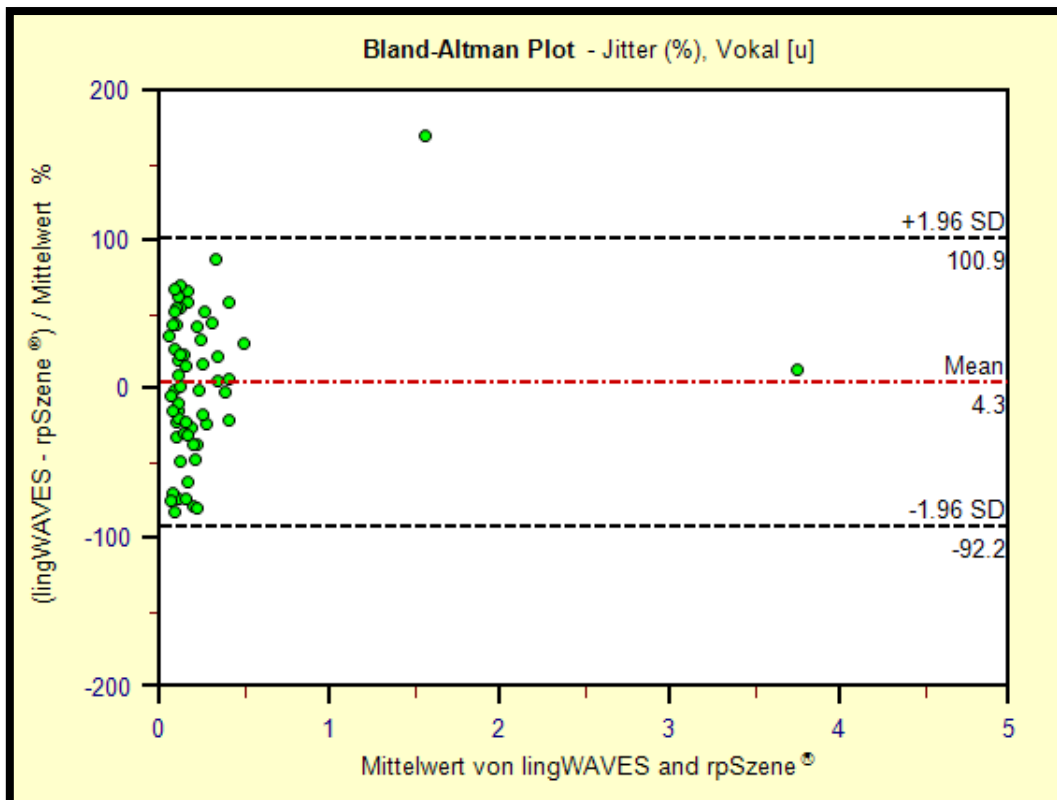


Abbildung 6.16: Bland-Altman Plot für Jitterwerte des Vokals [u] mit Differenzen in %

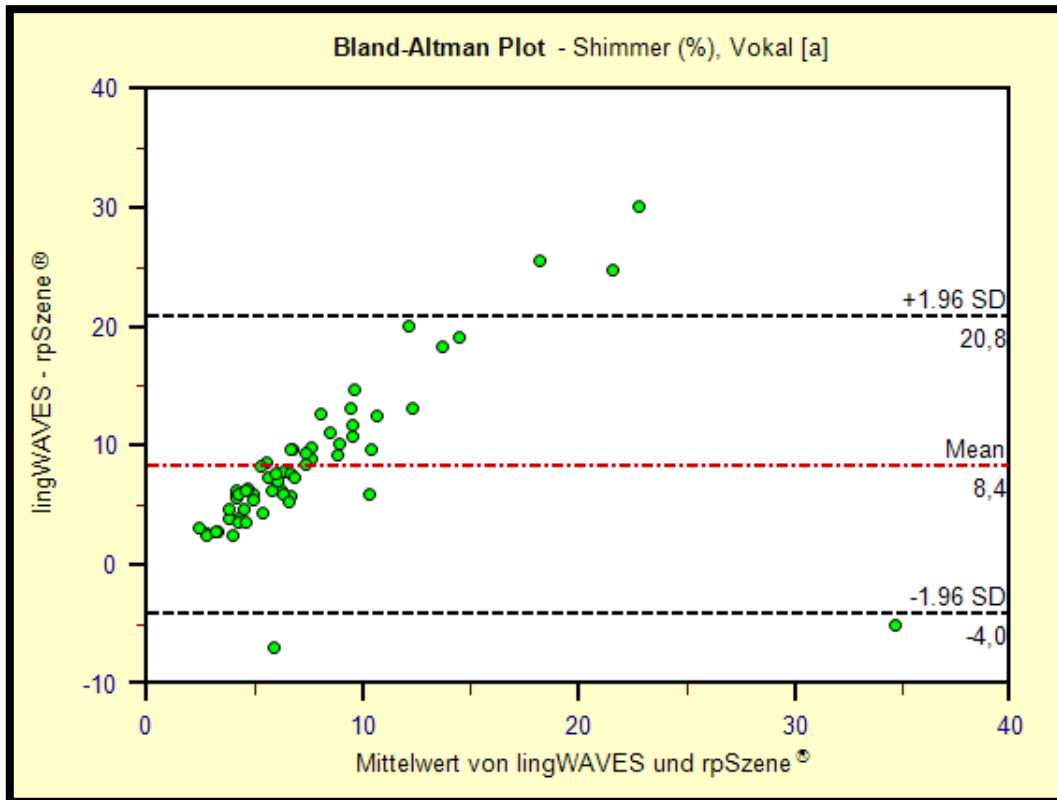


Abbildung 6.17: Bland-Altman Plot für Shimmerwerte des Vokals [a]

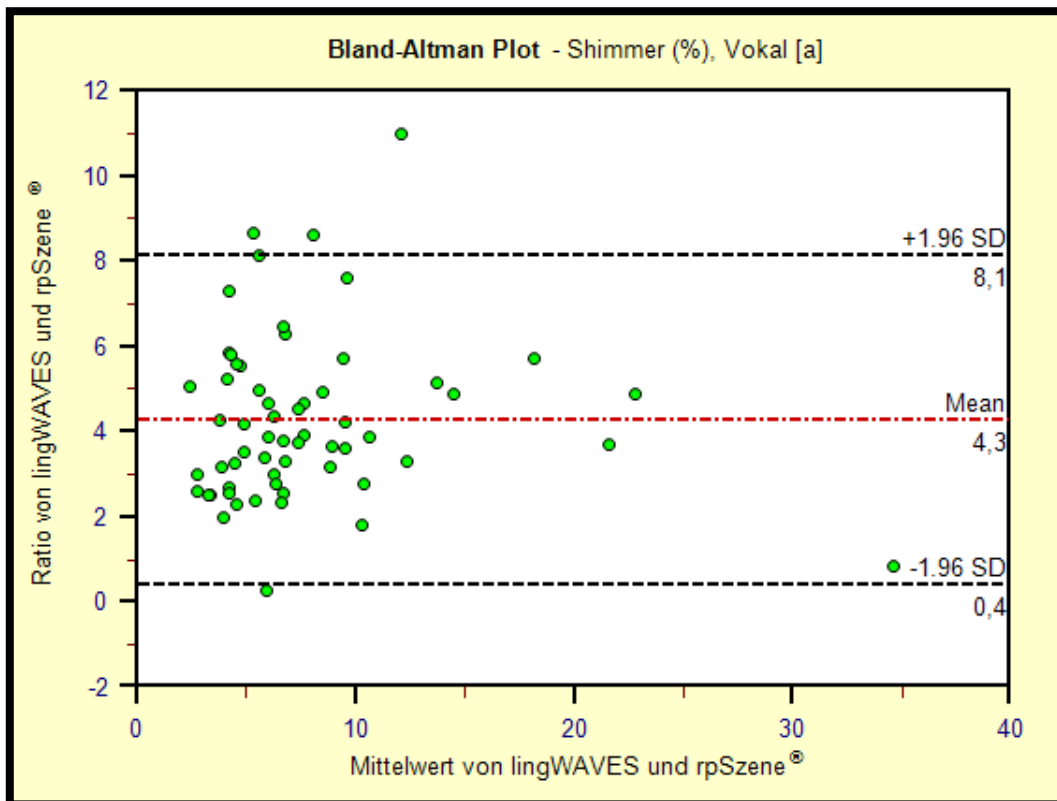


Abbildung 6.18: Bland-Altman Plot für Shimmerwerte des Vokals [a] mit Ratios

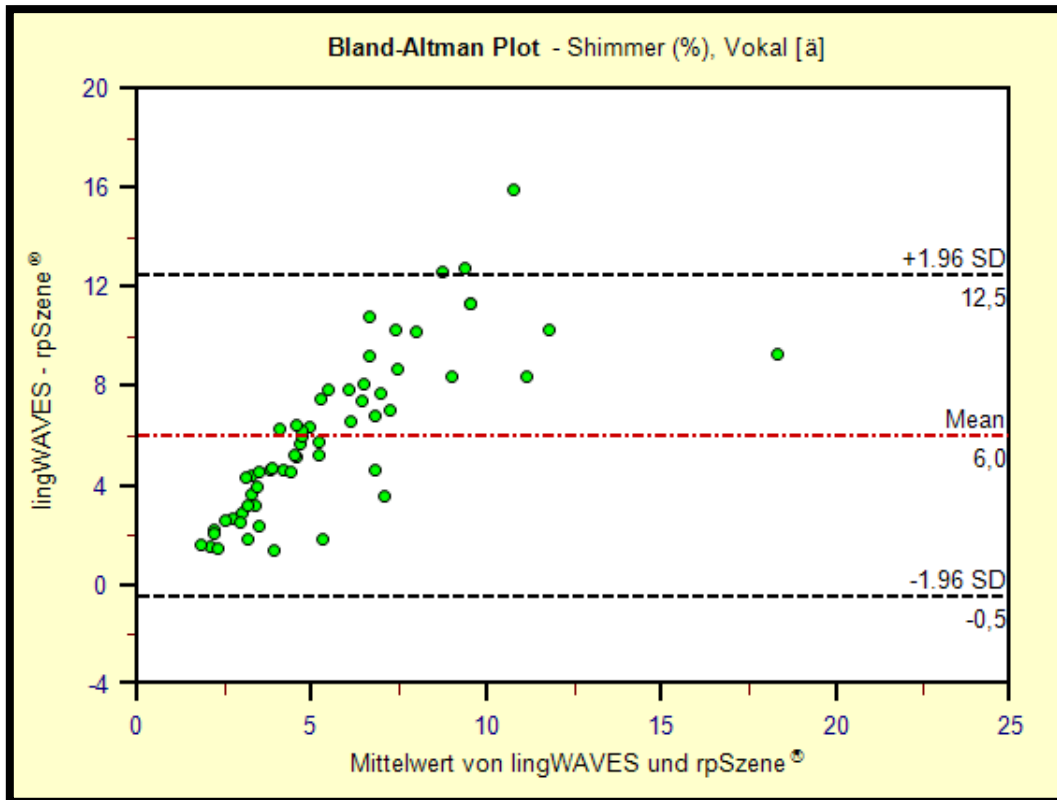


Abbildung 6.19: Bland-Altman Plot für Shimmerwerte des Vokals [ä]

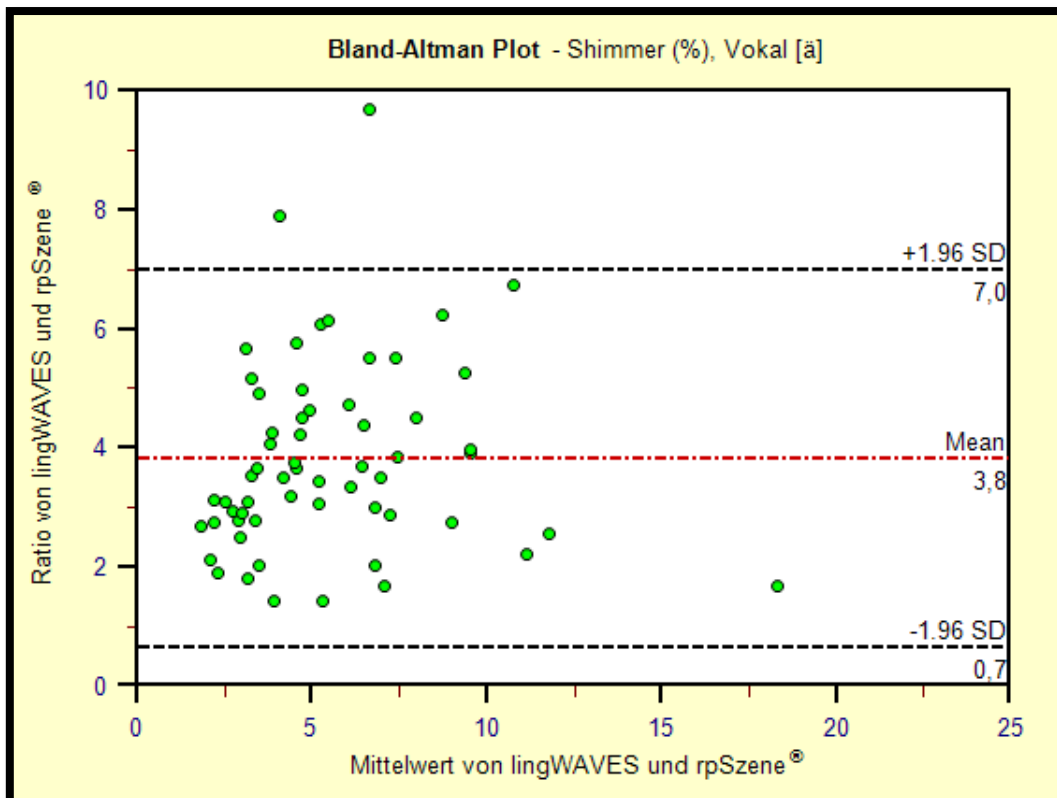


Abbildung 6.20: Bland-Altman Plot für Shimmerwerte des Vokals [ä] mit Ratios

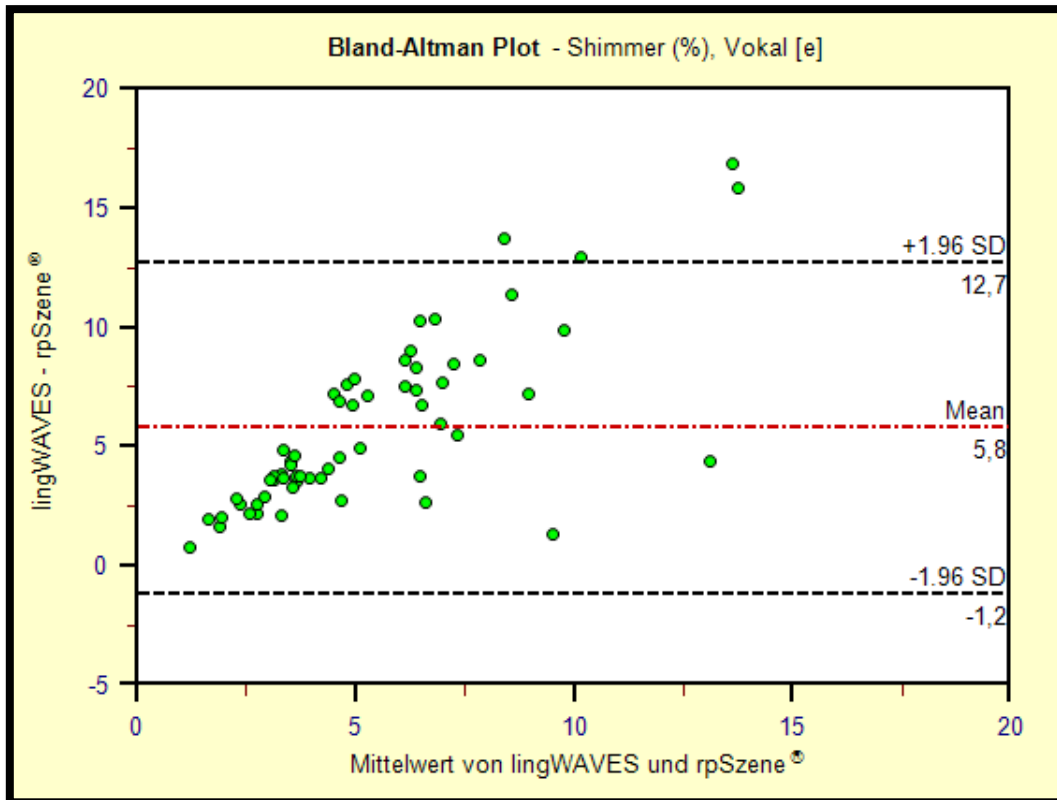


Abbildung 6.21: Bland-Altman Plot für Shimmerwerte des Vokals [e]

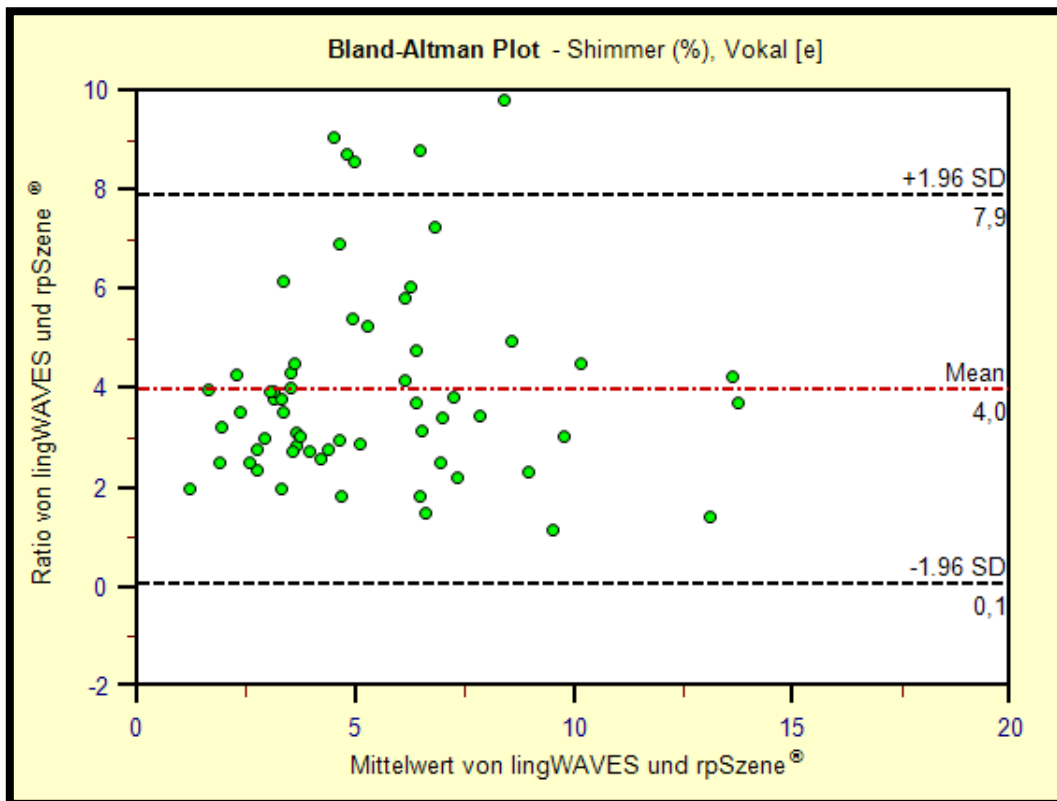


Abbildung 6.22: Bland-Altman Plot für Shimmerwerte des Vokals [e] mit Ratios

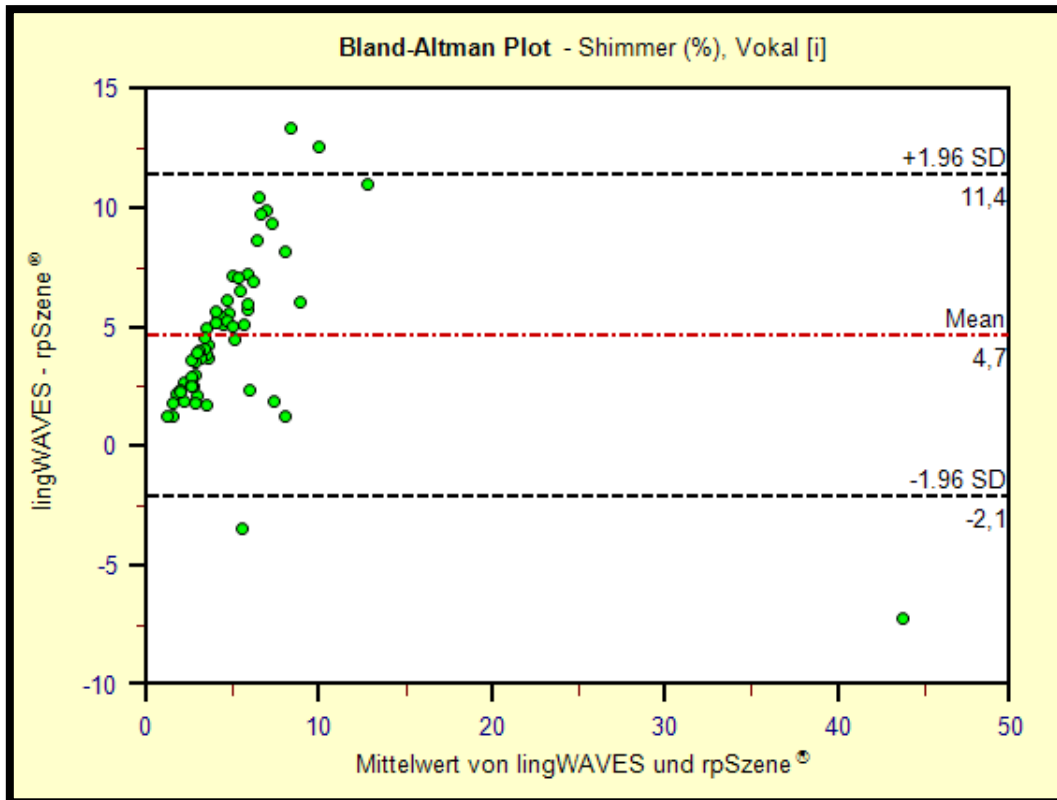


Abbildung 6.23: Bland-Altman Plot für Shimmerwerte des Vokals [i]

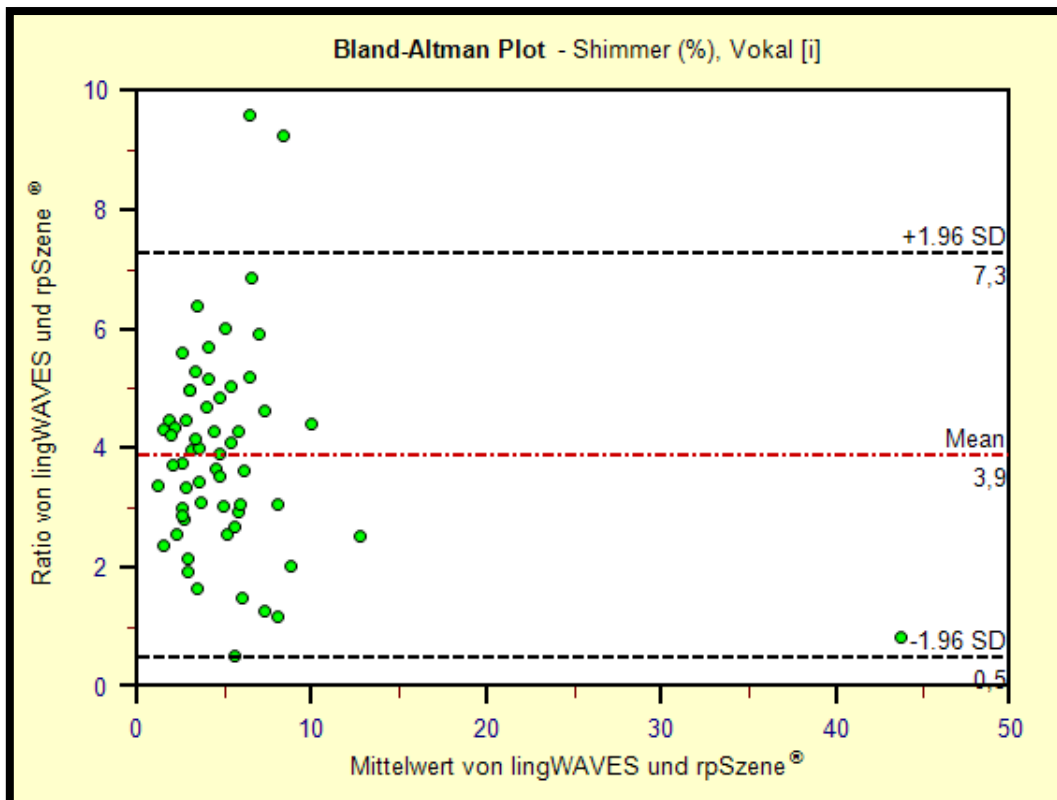


Abbildung 6.24: Bland-Altman Plot für Shimmerwerte des Vokals [i] mit Ratios

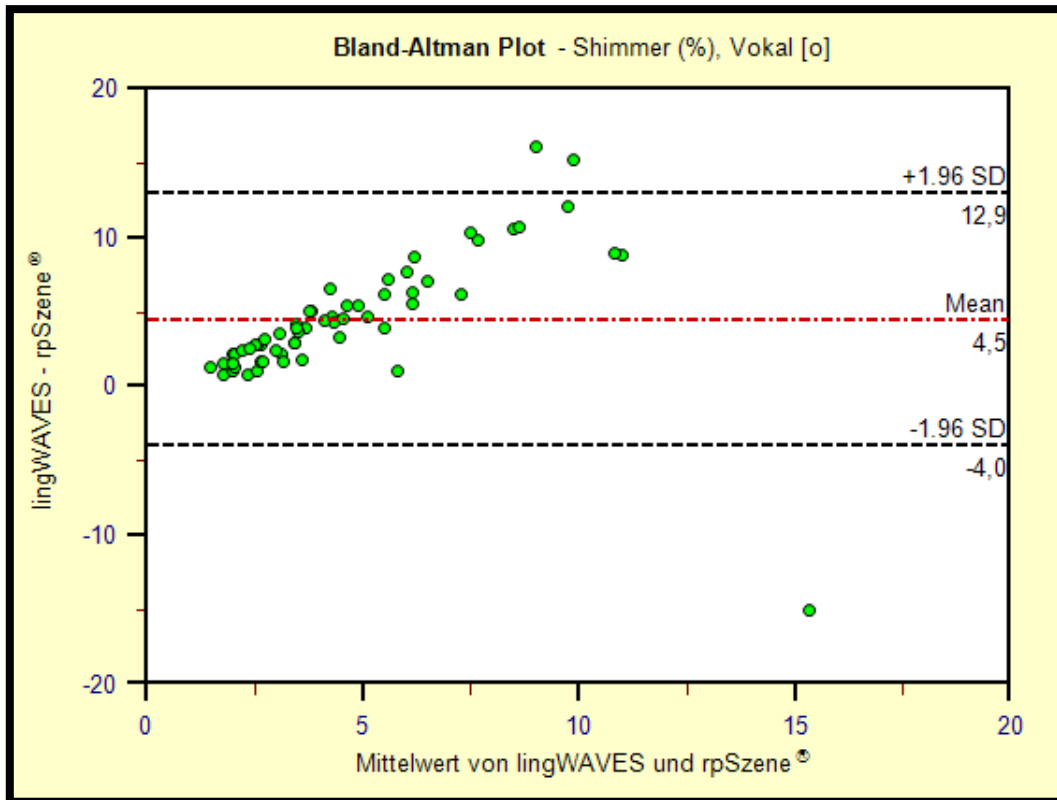


Abbildung 6.25: Bland-Altman Plot für Shimmerwerte des Vokals [o]

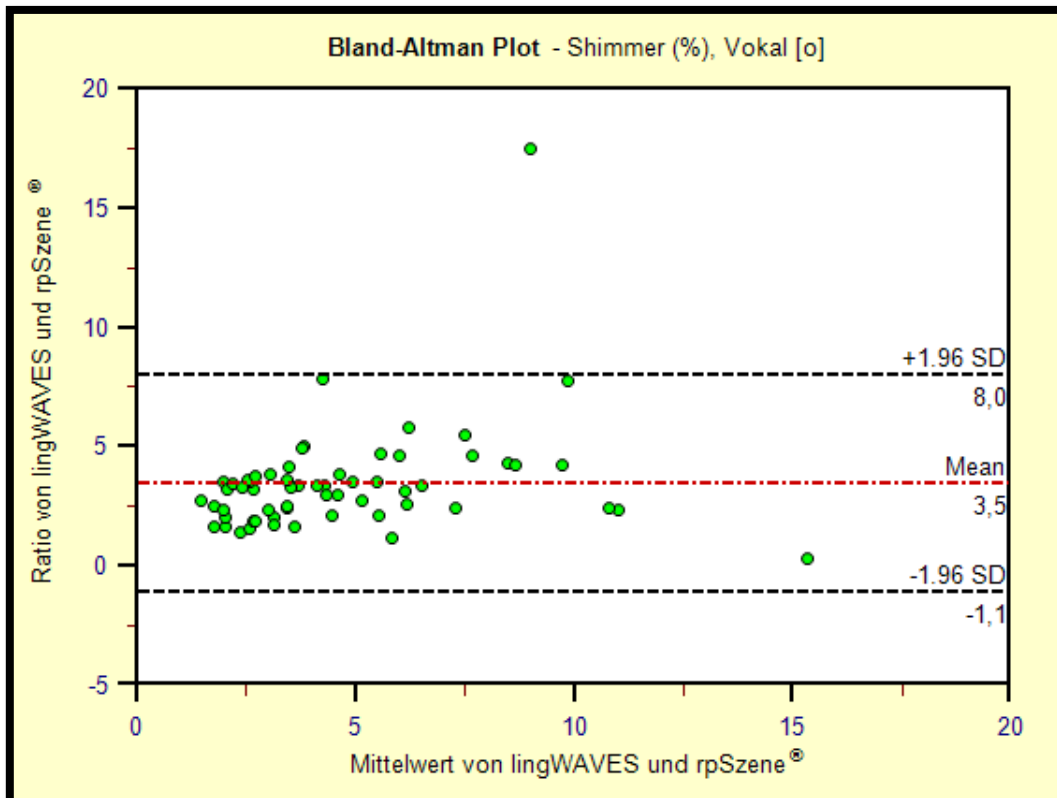


Abbildung 6.26: Bland-Altman Plot für Shimmerwerte des Vokals [o] mit Ratios

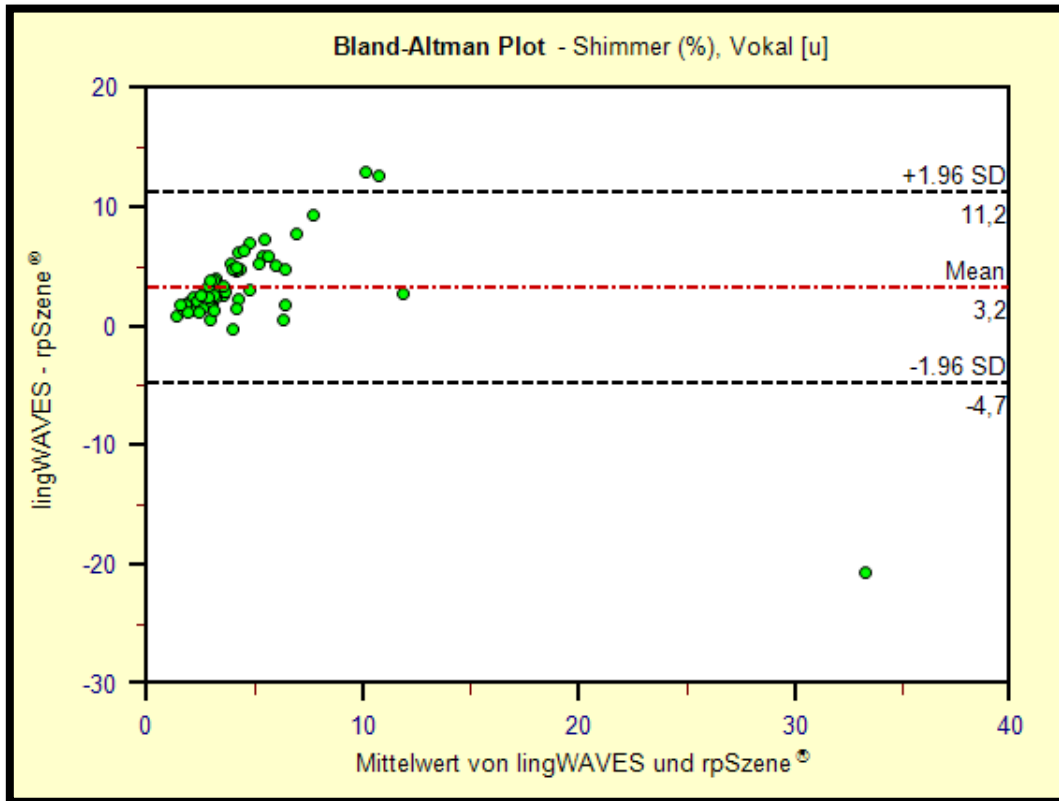


Abbildung 6.27: Bland-Altman Plot für Shimmerwerte des Vokals [u]

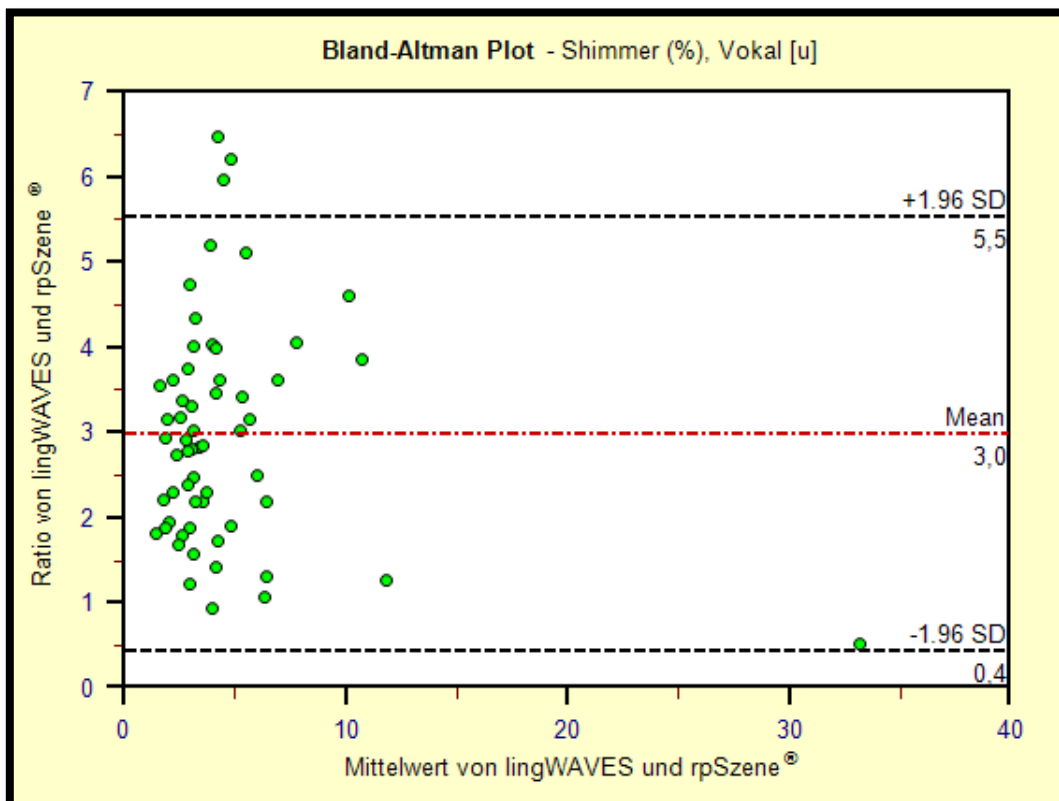


Abbildung 6.28: Bland-Altman Plot für Shimmerwerte des Vokals [u] mit Ratios



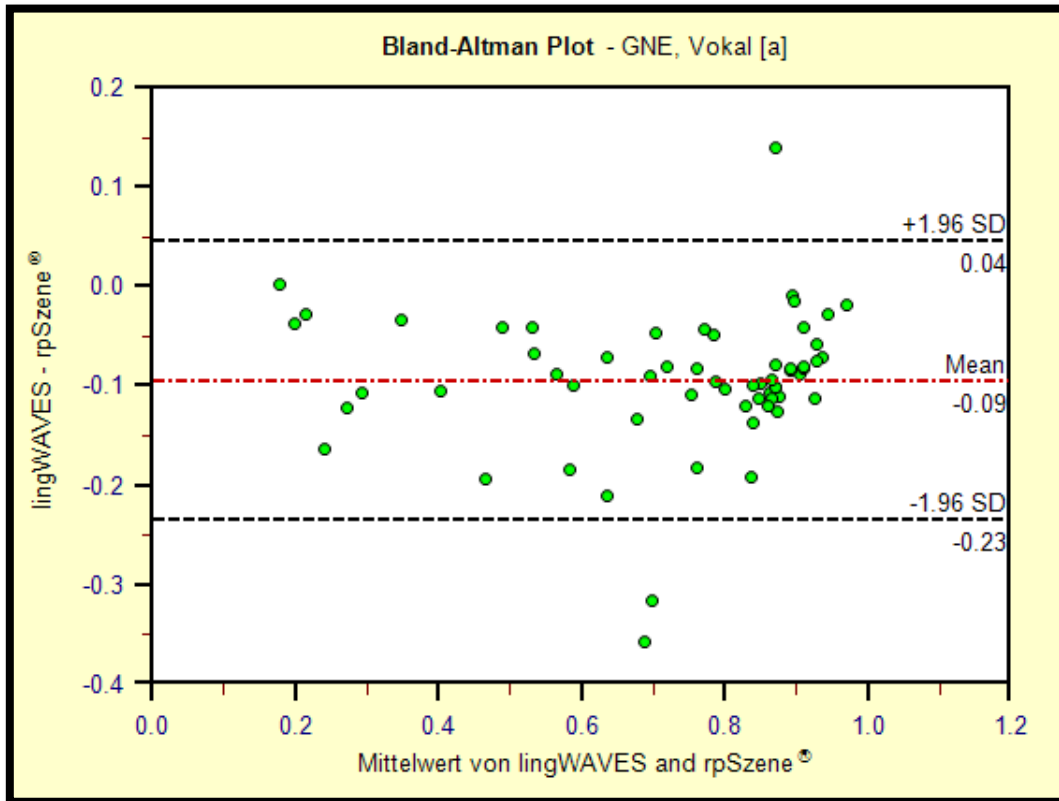


Abbildung 6.29: Bland-Altman Plot für GNE-Werte des Vokals [a]

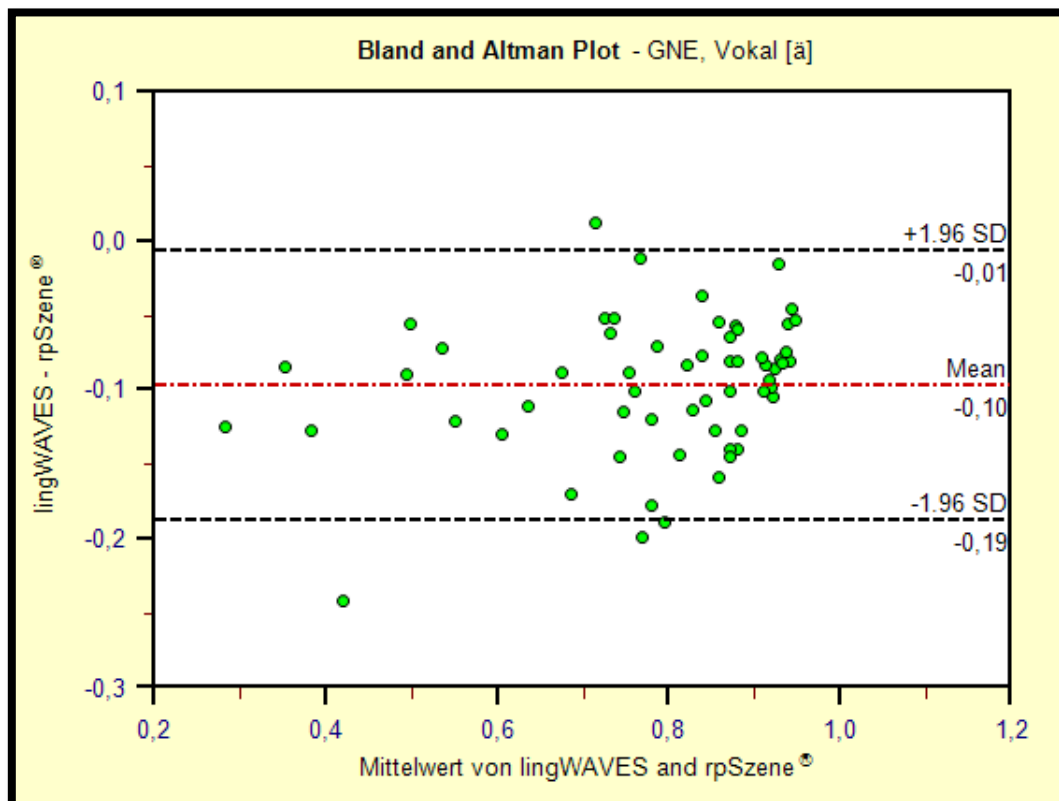


Abbildung 6.30: Bland-Altman Plot für GNE-Werte des Vokals [ä]

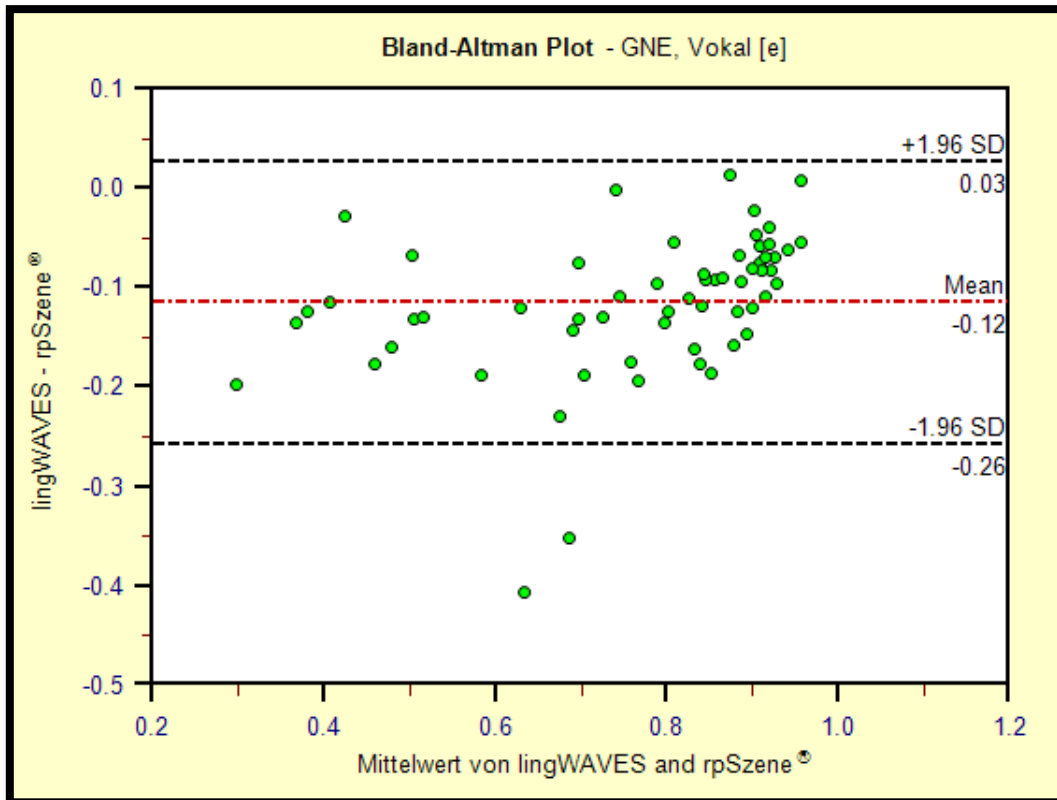


Abbildung 6.31: Bland-Altman Plot für GNE-Werte des Vokals [e]

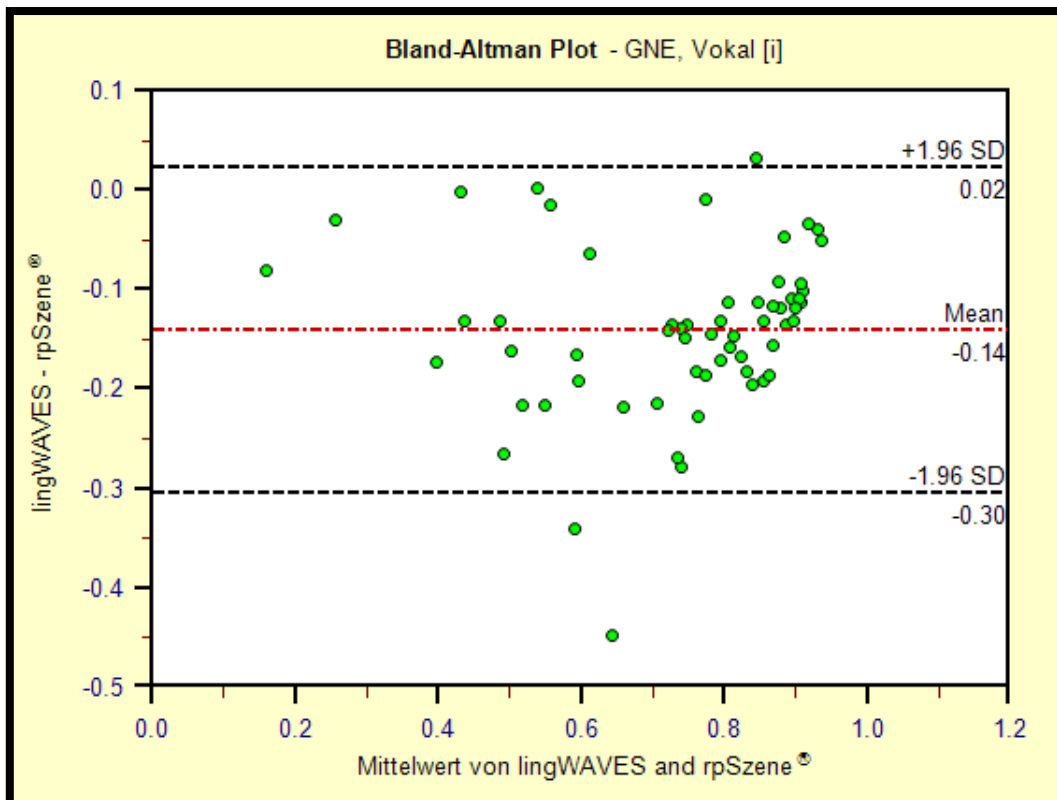


Abbildung 6.32: Bland-Altman Plot für GNE-Werte des Vokals [i]

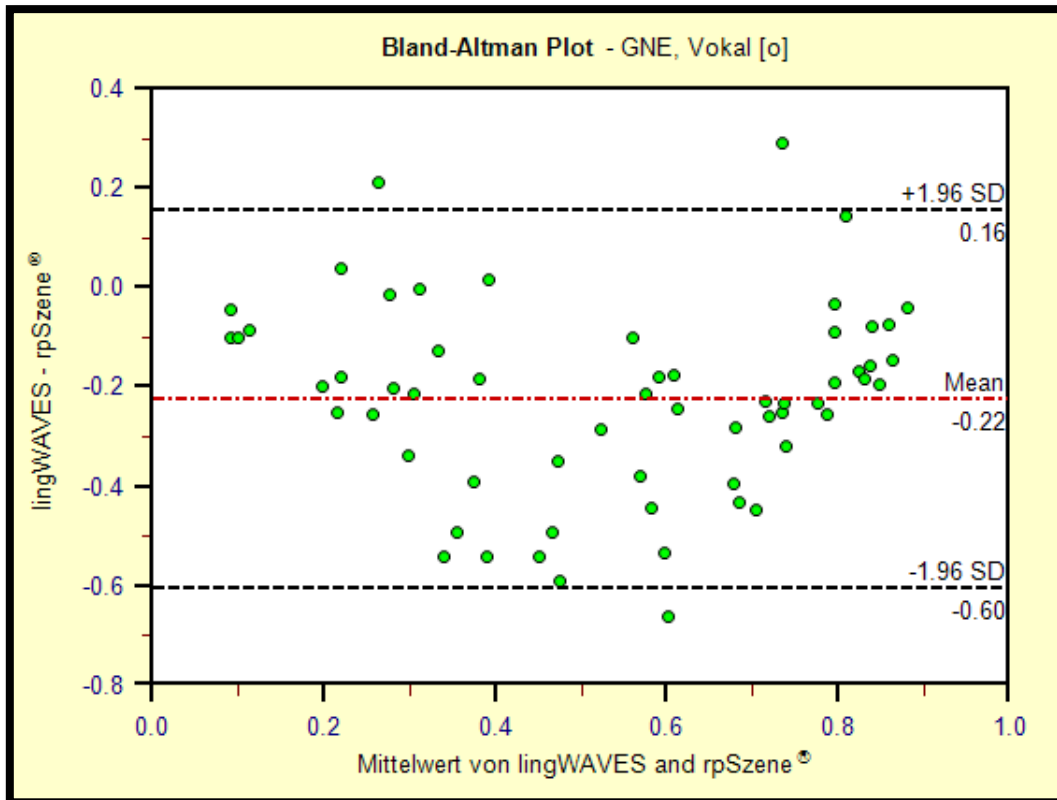


Abbildung 6.33: Bland-Altman Plot für GNE-Werte des Vokals [o]

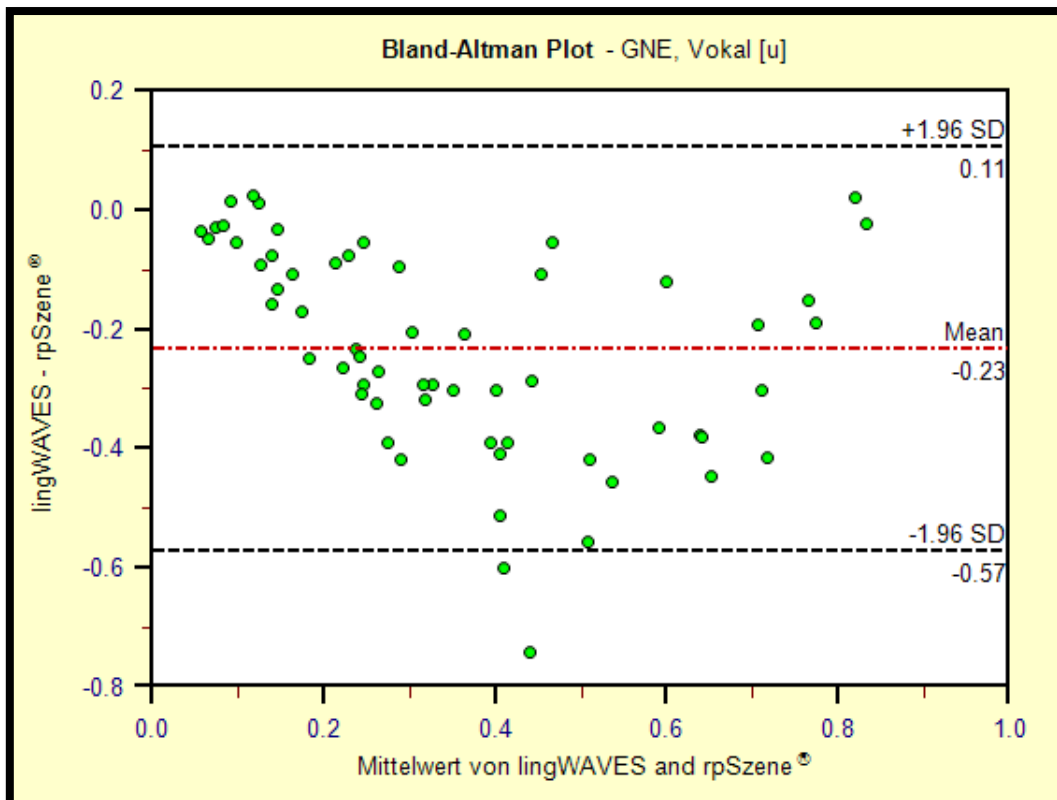


Abbildung 6.34: Bland-Altman Plot für GNE-Werte des Vokals [u]

## 7 Abkürzungsverzeichnis

%	= Prozent
[a]	= Vokal „a“ (gilt auch für „ä“, „e“, „i“, „o“, „u“)
ANOVA	= analysis of variance
ARPS-Schema	= Amplitude, Randkantenverschiebung Phase und Schlussphase
CHNR	= Cepstral Harmonic to Noise Ratio
CSL	= Computerized Speech Lab
dB	= Dezibel
DSI	= Dysphonia Severity Index
ELS	= Europäische Laryngologische Gesellschaft
F <sub>0</sub>	= Grundfrequenz
F <sub>0</sub> SD	= Standardabweichung von der Grundfrequenz f <sub>0</sub>
GDH	= Göttinger Heiserkeitsdiagramm
GMM	= Gaussian Mixture Model
GNE	= Glottal-to-Noise Excitation Ratio
GRBAS-Skala	= overall grade or degree, rough, breathy, asthenic und strained quality
HNR	= Harmonics-to-Noise-Ratio
HZ	= Hertz
I-Low	= Niedrigste Intensität
MDVP	= Multi-Dimensional Voice Program
MPT	= Tonhaldedauer
MS Excel©	= Microsoft Excel©
MWMC	= Mean Waveform Matching Coefficient
NNE	= Normalized-Noise Energy
OTC	= Optische Kohärenztomographie
PLI	= Pathological Likelihood Index
PQ	= Phonation-Quotient
RBH-Schema	= Rauigkeit, Behauchtheit und Heiserkeit
SD	= Standardabweichung
SNR	= Signal-to-Noise-Ratio

SPL	= Schalldruckpegel
SSI	= Stimmstörungsindex
VC	= Vitalkapazität
VDC	= Voice Diagnostic Center
vgl.	= vergleiche
VHI	= Voice Handicap Index
Z. n.	= Zustand nach

## **8 Verzeichnis der Akademischen Lehrer**

Meine akademischen Lehrer waren die Damen und Herren Universitätsprofessorinnen, Universitätsprofessoren und Dozentinnen und Dozenten in Marburg:

Adamkiewicz, Aumüller, Basler, Barth, Basler, Baum, Berger, Braun, Cetin, Czubayko, Daut, Engenhard-Cabillic, Feuser, Gerdes, Görg, Görg, Grundmann, Gundermann, Hasilik, Hermann-Lingen, Hertl, Hofmann, Höffgen,, Jungclas, Kalder, Kann, Kill, Koolmann, Klose, Krieg, Kroll, Lang, Lenz, Leonhard, Lill, Löffler, Lohoff, Maier, Maisch, Mandrek, Moll, Mueller, Mutters, Neubauer, Oertel, Plant, Prokop, Renz, Reuser, Richter, Röper, Röhm, Rothmund, Schäfer, Schmidt, Schrader, Seitz, Steiniger, Tibesku, Vogelmeier, Voigt, Wagner, Weihe, Werner, Westermann, Wilhelm, Wulf, Zemlin.

## 9 Danksagungen

Ich möchte mich besonders bei Frau Prof. Dr. med. Roswitha Berger für die Möglichkeit der Anfertigung einer Dissertation bedanken. Darüber hinaus danke ich ihr für die Anregungen und Vorschläge sowie die Unterstützung in der Gewinnung von Patienten für die Studie.

Besonderer Dank gilt ebenso Herrn Dr. med. Holger Hanschmann, der mich während der gesamten Arbeit, von Planung der Studie über die praktische Durchführung bis hin zur Auswertung stets in sympathischer und konstruktiver Art und Weise und mit wertvollen Ratschlägen begleitet hat. Weiterhin danke ich ihm dafür, dass mit seiner Hilfe viele Patienten und Probanden für die Studien gewonnen werden konnten.

Frau Dr. rer. nat. Tina Timmesfeld danke ich für ihre umfassende statistische Beratung in der Planung und Auswertung der Studie.

Bedanken möchte ich mich außerdem bei allen Mitarbeitern der Klinik für Phonetik und Pädaudiologie des Universitätsklinikums Gießen und Marburg GmbH, Standort Marburg, für die fachliche und stets freundliche Unterstützung.

Ein großer Dank gilt natürlich allen Versuchspersonen, die sich zur Teilnahme an dieser wissenschaftlichen Studie bereit erklärten und damit deren Durchführung erst ermöglichten.

Bedanken möchte ich mich ebenfalls bei Anne Höhne, Fiona Weiß, Helen Höchsmann und Martijn Schneider für die orthographischen und stilistischen Korrekturen bzw. Verbesserungsvorschläge.

Ich danke meiner Familie für Ihre Geduld und für die Unterstützung bzw. Motivation während des Studiums und der Fertigstellung dieser Arbeit.

## 10 Literaturverzeichnis

1. Airainer R, Klingholz F: Computergestützte Stimmfeldauswertung als Diagnostikhilfe bei funktionellen Dysphonien. Computer-assisted phonetography as a diagnostic aid in functional dysphonia. *Laryngo-Rhino-Otologie* 1991. Band 70 (7): 362–366.
2. Altman DG, Bland JM: Measurement in medicine: The analysis of method comparison studies. *Journal of the Royal Statistical Society. Series D (The Statistician)* 1983. Band 32 (3): 307–317.
3. Baken RJ, Orlikoff RF: *Clinical measurement of speech and voice*. 2. Auflage, Singular Thomson Learning, San Diego 2000.
4. Barfs A: *Die Veränderungen des Sprachsignals beim Schreien - eine akustische Analyse ausgewählter phonetischer Parameter*. Fachbereich II Sprach- und Literaturwissenschaften. Universität Trier 2005.
5. Behrbohm H, Kaschke O, Nawka T: *Kurzlehrbuch Hals-Nasen-Ohren-Heilkunde*. 33 Tabellen. Thieme, Stuttgart 2009.
6. Bielamowicz S, Kreiman J, Gerratt BR, et al.: Comparison of voice analysis systems for perturbation measurement. *Journal of Speech and Hearing Research* 1996. Band 39 (1): 126–134.
7. Bland JM, Altman DG: Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet* 1986. Band i: 307–310.
8. Bland JM, Altman DG: Comparing methods of measurement: why plotting difference against standard method is misleading. *Lancet* 1995. Band 346: 1085–1087.
9. Bland JM, Altman DG: Measuring agreement in method comparison studies. *Statistical methods in medical research* 1999. Band 8 (2): 135–160.



10. Bodt MS de, Wuyts FL, van de Heyning PH, Croux C: Test-retest study of the GRBAS scale: influence of experience and professional background on perceptual rating of voice quality. *Journal of Voice* 1997. Band 11 (1): 74–80.
11. Böhme G [Hrg]: Sprach-, Sprech-, Stimm- und Schluckstörungen. Band 1: Klinik. 4., aktualisierte und erw. Auflage, Urban & Fischer, München 2003.
12. Böhme G, Gross M: Stroboskopie und andere Verfahren zur Analyse von Stimmlippenschwingungen. Lehrbuch. Median-Verlag, Heidelberg 2001.
13. Brazdionyte J, Macas A: Bland-Altman analysis as an alternative approach for statistical evaluation of agreement between two methods for measuring hemodynamics during acute myocardial infarction. *Medicina (Kaunas, Lithuania)* 2007. Band 43 (3): 208–214.
14. Brockmann M, Drinnan MJ, Storck C, Carding PN: 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* 2011. Band 25 (1): 44–53.
15. Brockmann-Bauser M, Drinnan MJ: Routine acoustic voice analysis: time to think again? *Current Opinion in Otolaryngology & Head and Neck Surgery* 2011. Band 19 (3): 165–170.
16. Dejonckere PH, Bradley P, Clemente P, et al.: 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 Phoniatics of the European Laryngological Society (ELS). *European Archives of Oto-Rhino-Laryngology* 2001. Band 258 (2): 77–82.
17. Deliyski DD, Evans MK, Shaw HS: Influence of data acquisition environment on accuracy of acoustic voice quality measurements. *Journal of Voice* 2005. Band 19 (2): 176–186.

18. Demleitner M: Einführung in die Statistische Sprachverarbeitung. Exkurs: Gaussian Mixtures. <http://www.cl.uni-heidelberg.de/kurs/skripte/stat/html/page032.html>. [Stand: 09.06.2012].
19. Denk DM, Swoboda H, Steiner E: Physiology of the larynx. Der Radiologe 1998. Band 38 (2): 63–70.
20. Deuster D, Am Zehnhoff-Dinnesen A, Schmidt C, et al.: Development and evaluation of the new module 'communication disorders' in medical education. Medical Teacher 2008. Band 30 (8): e225-31.
21. Dewitte K, Fierens C, Stockl D, Thienpont LM: Application of the Bland-Altman Plot for interpretation of method-comparison studies: A critical investigation of its practice. Clinical Chemistry 2002. Band 48 (5): 799–801.
22. Dickreiter M: Handbuch der Tonstudioteknik. 7. Auflage, Saur, München 2008.
23. Echternach M, Arndt S, Zander MF, Richter B: Stimmdiagnostik bei professionellen Sopranistinnen. Anwendung des Protokolls der Europäischen Laryngologischen Gesellschaft (ELS). HNO 2009. Band 57 (3): 266–272.
24. Fleischer S, Hess M: Die Bedeutung der Stroboskopie in der laryngologischen Praxis. HNO 2006. Band 54 (8): 628–634.
25. Friedrich G, Bigenzahn W, Zorowka P, Brunner E: Phoniatrie und Pädaudiologie. Einführung in die medizinischen, psychologischen und linguistischen Grundlagen von Stimme, Sprache und Gehör. 4., korr. Auflage, Huber, Bern 2008.
26. Friedrich G, Dejonckere PH: Das Stimmdiagnostik-Protokoll der European Laryngological Society (ELS). Erste Erfahrungen im Rahmen einer Multizenterstudie. Laryngo-Rhino-Otologie 2005. Band 84 (10): 744–752.
27. Fröhlich M, Michaelis D, Kruse E: Objektive Beschreibung der Stimmgüte unter Verwendung des Heiserkeitsdiagramms. HNO 1998. Band 46 (7): 684–689.

28. Fröhlich M, Strube HW, Kruse E: Akustische Parameter zur Stimmgütebeschreibung aus fortlaufender Sprache. In: Gross M, Eysholdt U (Hrsg): Aktuelle phoniatisch-pädaudiologische Aspekte 1996. Band 4, Verlag Abteilung Phoniatrie, Göttingen 1997, 22–24.
29. Godino-Llorente JI, Gómez-Vilda P, Cruz-Roldán F, et al.: Pathological Likelihood Index as a measurement of the degree of voice normality and perceived hoarseness. *Journal of Voice* 2010. Band 24 (6): 667–677.
30. Godino-Llorente JI, Osma-Ruiz V, Sáenz-Lechón N, et al.: The effectiveness of the Glottal to Noise Excitation Ratio for the screening of voice disorders. *Journal of Voice* 2010. Band 24 (1): 47–56.
31. Görne T: Mikrofone in Theorie und Praxis. 6. Auflage, Elektor-Verlag, Aachen 2002.
32. Gugatschka M, Rechenmacher J, Chibidziura J, Friedrich G: Vergleichbarkeit und Umrechnung von Stimmstörungsindex (SSI) und Voice Handicap Index (VHI). *Laryngo-Rhino-Otologie* 2007. Band 86 (11): 785–788.
33. Hammer SS: Stimmtherapie mit Erwachsenen. Was Stimmtherapeuten wissen sollten. *Praxiswissen Logopädie*. 4. Auflage, Springer, Heidelberg 2009.
34. Hanschmann H, Müller J, Berger R: Schallanalytische Beurteilung in der Stimmdiagnostik. In: Bose I (Hrsg): *Sprechwissenschaft*. Hallesche Schriften zur Sprechwissenschaft und Phonetik. Band 22, Lang, Frankfurt am Main 2007.
35. Hanschmann H, Berger R: Quantifizierung des videostroboskopischen Stimmlippenbefundes. *Laryngo-Rhino-Otologie* 2009. Band 88 (1): 6–8.
36. Held L, Rufibach K, Seifert B: Einführung in die Biostatistik. Vorlesungsskript. Institut für Sozial- und Präventivmedizin, Abteilung Biostatistik. Universität Zürich 2010.

37. Henry LR, Helou LB, Solomon NP, et al.: Functional voice outcomes after thyroidectomy: An assessment of the Dysphonia Severity Index (DSI) after thyroidectomy. *Surgery* 2010. Band 147 (6): 861–870.
38. Hirano M: Clinical examination of voice. *Disorders of human communication*. Band 5, Springer, Wien 1981.
39. Hoppe U.: Mechanismen der Heiserkeit – Visualisierung und Interpretation durch nicht lineare Dynamik. *Laryngo-Rhino-Otologie* 2002. Band 81 (12): 906–907.
40. Hsiung MW, Pai L, Wang HW: Correlation between voice handicap index and voice laboratory measurements in dysphonic patients. *European Archives of Oto-Rhino-Laryngology* 2002. Band 259 (2): 97–99.
41. Ingrisano DR, Perry CK, Jepson KR: Environmental noise: A threat to automatic voice analysis. *American Journal of Speech-Language Pathology* 1998. Band 7 (1): 91–96.
42. Jacobson BH, Johnson A, Grywalski C: The Voice Handicap Index (VHI): Development and validation. *American Journal of Speech-Language Pathology* 1997. Band 1 (6): 66–70.
43. Karmasin H: Wer sich im Beruf durchsetzt, braucht Stimme. <http://www.dozentenscout.de/News/details/Wer-sich-im-Beruf-durchsetzt-braucht-Stimme>. [Stand: 09.06.2012].
44. Karnell MP, Hall KD, Landahl KL: Comparison of fundamental frequency and perturbation measurements among three analysis systems. *Journal of Voice* 1995. Band 9 (4): 383–393.
45. Karnell MP, Scherer RS, Fischer LB: Comparison of acoustic voice perturbation measures among three independent voice laboratories. *Journal of Speech and Hearing Research* 1991. Band 34 (4): 781–790.
46. Kiliç MA, Oğüt F, Dursun G, et al.: The effects of vowels on voice perturbation measures. *Journal of Voice* 2004. Band 18 (3): 318–324.

47. Koike Y: Vowel amplitude modulations in patients with laryngeal diseases. *Journal of the Acoustical Society of America* 1969. Band 45 (4): 839–844.
48. Koufman JA, Isaacson G: The spectrum of vocal dysfunction. *Otolaryngologic Clinics of North America* 1991. Band 24 (5): 985–988.
49. Kramer J: Computergestützte Stimmanalyse im Rahmen der Stimmarbeit nach dem Konzept Schlaffhorst-Andersen. *Forum Logopädie* 2009. Band 23 (6): 26–32.
50. Lessing J: Entwicklung einer Klassifikationsmethode zur akustischen Analyse fortlaufender Sprache unterschiedlicher Stimmgüte mittels Neuronaler Netze und deren Anwendung. Dissertation. Universität Göttingen 2007.
51. Lieberman P: Perturbations in vocal pitch. *Journal of the Acoustical Society of America* 1961. Band 33 (5): 597–603.
52. Lüerssen K, Lubatschowski H, Ursinus K, et al.: Charakterisierung von Stimmlippen mittels optischer Kohärenztomographie. *HNO* 2006. Band 54 (8): 611–615.
53. Ma EP, Yiu EM: Suitability of acoustic perturbation measures in analysing periodic and nearly periodic voice signals. *Folia Phoniatica et Logopaedica* 2005. Band 57 (1): 38–47.
54. MacCallum JK, Zhang Y, Jiang JJ: Vowel selection and its effects on perturbation and nonlinear dynamic measures. *Folia Phoniatica et Logopaedica* 2011. Band 63 (2): 88–97.
55. Maryn Y, Corthals P, Bodt M de, et al.: Perturbation measures of voice: a comparative study between Multi-Dimensional Voice Program and Praat. *Folia Phoniatica et Logopaedica* 2009. Band 61 (4): 217–226.
56. Michaelis D: Das Göttinger Heiserkeits-Diagramm - Entwicklung und Prüfung eines akustischen Verfahrens zur objektiven Stimmgütebeurteilung pathologischer Stimmen. Dissertation. Universität Göttingen 2000.

57. Michaelis D, Fröhlich M, Strube HW: Selection and combination of acoustic features for the description of pathologic voices. *Journal of the Acoustical Society of America* 1998. Band 103 (3): 1628–1639.
58. Michaelis D, Gramss T, Strube HW: Glottal-to-Noise Excitation Ratio - a new measure for describing pathological voices. *Acustica / acta acustica* 1997. Band 83 (4): 700–706.
59. Minnema W, Stoll HC: Objektive computergestützte Stimmanalyse mit „Praat“. [http://wevosys.com/knowledge/\\_data\\_knowledge/13.pdf](http://wevosys.com/knowledge/_data_knowledge/13.pdf). [Stand: 09.10.2010].
60. Morris RJ, Brown WS: Comparison of various automatic means for measuring mean fundamental frequency. *Journal of Voice* 1996. Band 10 (2): 159–165.
61. Morrison MD, Rammage LA: Muscle misuse voice disorders: description and classification. *Acta Oto-Laryngologica* 1993. Band 113 (3): 428–434.
62. Naufel Felipe AC de, Grillo MH, Grechi TH: Standardization of acoustic measures for normal voice patterns. *Brazilian Journal of Otorhinolaryngology* 2006. Band 72 (5): 659–664.
63. Nawka T: Auf der Suche nach dem verlorenen Ton - Stimmfeldmessung in der Praxis. [http://www.hno-nachrichten.de/archiv/2006/02/hno0602\\_28.pdf](http://www.hno-nachrichten.de/archiv/2006/02/hno0602_28.pdf). [Stand: 09.10.2010].
64. Nawka T, Franke I., Galkin E.: Objektive Messverfahren in der Stimm-diagnostik. *Forum Logopädie* 2006. Band 20 (4): 14–21.
65. Nawka T, Verdonck-de Leeuw IM, Bodt M de, et al.: Item reduction of the voice handicap index based on the original version and on European translations. *Folia Phoniatria et Logopaedica* 2009. Band 61 (1): 37–48.
66. Nawka T, Wiesmann U, Gonnermann U: Validierung des Voice Handicap Index (VHI) in der deutschen Fassung. *HNO* 2003. Band 51 (11): 921–930.

67. Nawka T, Wirth G, Anders L: Stimmstörungen. Für Ärzte, Logopäden, Sprachheilpädagogen und Sprechwissenschaftler; mit 30 Tabellen. 5., völlig überarb. Auflage, Dt. Ärzte-Verlag, Köln 2008.
68. Nawka T, Anders LC: Die auditive Bewertung heiserer Stimmen nach dem RBH-System. Doppel-Audio-CD mit Stimmbeispielen. Thieme, Stuttgart 1996.
69. Nawka T, Gonnermann U: Stimmstörungsindex (SSI). <http://www.egms.de/static/en/meetings/dgpp2003/03dgpp034.shtml>. [Stand: 09.06.2012].
70. Niemann H: Klassifikation von Mustern. 2. Auflage, Springer-Verlag, Berlin 2003. <http://www5.informatik.uni-erlangen.de/fileadmin/Persons/NiemannHeinrich/klassifikation-von-mustern/m00-www.pdf>. [Stand: 09.06.2012].
71. Orlikoff RF, Baken RJ: The effect of the heartbeat on vocal fundamental frequency perturbation. *Journal of Speech and Hearing Research* 1989. Band 32 (3): 576–582.
72. Parsa V, Jamieson DG, Pretty BR: Effects of microphone type on acoustic measures of voice. *Journal of Voice* 2001. Band 15 (3): 331–343.
73. Parsa V, Jamieson DG: Identification of pathological voices using glottal noise measures. *Journal of Speech, Language and Hearing Research* 2000. Band 43 (2): 469–485.
74. Petrovic-Lazic M, Babac S, Vukovic M, et al.: Acoustic voice analysis of patients with vocal fold polyp. *Journal of Voice* 2011. Band 25 (1): 94–97.
75. Pribuisienė R, Uloza V, Saferis V: Multidimensional voice analysis of reflux laryngitis patients. *European Archives of Oto-Rhino-Laryngology* 2005. Band 262 (1): 35–40.

76. Ptok M, Iven C, Jessen M, Schwemmler C: Objektiv gemessene Stimmlippenschwingungsirregularität vs. subjektiver Eindruck der Rauigkeit. HNO 2006. Band 54 (2): 132–138.
77. Rammage L, Morrison MD, Nichol H: Management of the voice and its disorders. 2 edition. Singular/Thomson Learning, San Diego CA 2001.
78. Reynolds D: Gaussian Mixture Models. [http://extwebprod.ll.mit.edu/mision/communications/publications/publication-files/full\\_papers/0802\\_Reynolds\\_Biometrics-GMM.pdf](http://extwebprod.ll.mit.edu/mision/communications/publications/publication-files/full_papers/0802_Reynolds_Biometrics-GMM.pdf). [Stand: 09.06.2012].
79. Richter B, Echternach M: State of the Art in der Kehlkopffunktionsdiagnostik - Kombination endoskopischer, akustischer und bildgebender Verfahren. Endoskopie heute 2008. Band 21: 20–33.
80. Richter B, Echternach M: Stimm diagnostik und -therapie bei Angehörigen stimmintensiver Berufe. HNO 2010. Band 58 (4): 389–398.
81. Rosen CA, Lee AS, Osborne J, et al.: Development and validation of the Voice Handicap Index-10. The Laryngoscope 2004. Band 114 (9): 1549–1556.
82. Schneider B, Bigenzahn W: Stimm diagnostik. Ein Leit faden für die Praxis. Springer, Wien 2007.
83. Smits I, Ceuppens P, Bodt MS de: A comparative study of acoustic voice measurements by means of Dr. Speech and Computerized Speech Lab. Journal of Voice 2005. Band 19 (2): 187–196.
84. Stemple JC, Glaze LE, Gerdeman BK: Clinical voice pathology. Theory and management. 3rd ed. Singular Pub. Group, San Diego 2000.
85. Svec JG, Granqvist S: Guidelines for selecting microphones for human voice production research. American Journal of Speech-Language Pathology 2010. Band 19 (4): 356–368.
86. Titze IR: A model for neurologic sources of aperiodicity in vocal fold vibration. Journal of Speech and Hearing Research 1991. Band 34 (3): 460–472.



87. Titze IR: The G. Paul Moore Lecture. Toward standards in acoustic analysis of voice. *Journal of Voice* 1994. Band 8 (1): 1–7.
88. Titze IR, Liang: Comparison of Fo extraction methods for high-precision voice perturbation measurements. *Journal of Speech and Hearing Research* 1993. Band 36 (6): 1120–1133.
89. Titze IR, Winholtz WS: Effect of microphone type and placement on voice perturbation measurements. *Journal of Speech and Hearing Research* 1993. Band 36 (6): 1177–1190.
90. Tortora GJ, Derrickson BH, Pries AR: *Anatomie und Physiologie*. Wiley-VCH, Weinheim 2006.
91. Uloza V, Saferis V, Uloziene I: Perceptual and acoustic assessment of voice pathology and the efficacy of endolaryngeal phonosurgery. *Journal of Voice* 2005. Band 19 (1): 138–145.
92. van den Berg J: Myoelastic-aerodynamic theory of voice production. *Journal of Speech and Hearing Research* 1958. Band 1 (3): 227–244.
93. Weiss C: *Basiswissen medizinische Statistik*. 5. überarb. Auflage, Springer, Heidelberg 2010. [Stand: 09.06.2012].
94. Wendler J, Seidner W, Eysholdt U, Appel H: *Lehrbuch der Phoniatrie und Pädaudiologie*. 4., völlig überarb. Auflage, Thieme, Stuttgart 2005.
95. Werth K, Voigt D, Döllinger M, et al.: Clinical value of acoustic voice measures: A retrospective study. *European Archives of Oto-Rhino-Laryngology* 2010. Band 267 (8): 1261–1271.
96. WEVOS - WEVOSYS – Germany: lingWAVES. Standardisierte klinische Stimmdiagnostik, visuelles Biofeedback und Sprachanalyse. <http://www.wevosys.de/produkte/lingwaves/lingwaves.html>. [Stand: 25.10.2011].
97. WEVOS - WEVOSYS – Germany: lingWAVES Voice Diagnostic Center. Handbuch, Version 2.5. [http://www.wevosys.de/download/handbook\\_GE\\_lingWAVES\\_Voice\\_Diagnostic\\_Center.pdf](http://www.wevosys.de/download/handbook_GE_lingWAVES_Voice_Diagnostic_Center.pdf). [Stand: 09.06.2012].

98. Winholtz WS, Titze IR: Miniature head-mounted microphone for voice perturbation analysis. *Journal of Speech, Language and Hearing Research* 1997. Band 40 (4): 894–899.
99. Woisard V, Bodin S, Yardeni E, Puech M: The Voice Handicap Index: Correlation between subjective patient response and quantitative assessment of voice. *Journal of Voice* 2007. Band 21 (5): 623–631.
100. Wuyts FL, Bodt MS de, Molenberghs G, et al.: The dysphonia severity index: An objective measure of vocal quality based on a multiparameter approach. *Journal of Speech, Language and Hearing Research* 2000. Band 43 (3): 796–809.
101. Yamaguchi H, Shrivastav R, Andrews ML, Niimi S: A comparison of voice quality ratings made by Japanese and American listeners using the GRBAS scale. *Folia Phoniatria et Logopaedica* 2003. Band 55 (3): 147–157.
102. Zwirner P, Michaelis D, Fröhlich M, Kruse E: Korrelationen zwischen perzeptueller Beurteilung von Stimmen nach dem RBH-System und akustischen Parametern. In: Gross M (Hrsg): *Aktuelle phoniatriisch-pädaudiologische Aspekte* 1997. Band 5, Median-Verlag, Heidelberg 1998.