

Aus der Klinik für Audiologie und Phoniatrie  
der Medizinischen Fakultät Charité – Universitätsmedizin Berlin

## DISSERTATION

**Autofluoreszenz-Laryngoskopie, Stimmumfangsmaß, Voice Handicap Index  
Kurzformen: Zur Wertigkeit neuer diagnostischer Prozeduren und Parameter  
in der Stimmheilkunde**

zur Erlangung des akademischen Grades  
Doctor medicinae (Dr. med.)

vorgelegt der Medizinischen Fakultät  
Charité – Universitätsmedizin Berlin

von  
Matthias Seipelt  
aus Berlin

Datum der Promotion: 4. Juni 2021

# Inhaltsverzeichnis

<b>Abkürzungsverzeichnis</b> .....	<u>II</u>
<b>Abstract</b> .....	<u>III</u>
<b>Zusammenfassung</b> .....	<u>IV</u>
<b>1. Einleitung</b> .....	<u>1</u>
<b>2. Material und Methodik</b> .....	<u>3</u>
2.1 AFL in der Differentialdiagnostik laryngealer Läsionen.....	<u>3</u>
2.2 SUM und VHI-9i in der Stimmfunktionsdiagnostik bei Stimmlippenpolypen.....	<u>4</u>
2.3 VHI-12, RBH, SUM und DSI zur Therapieverlaufskontrolle bei Stimmstörungen.....	<u>5</u>
<b>3. Ergebnisse</b> .....	<u>6</u>
3.1 AFL in der Differentialdiagnostik laryngealer Läsionen.....	<u>6</u>
3.2 SUM und VHI-9i in der Stimmfunktionsdiagnostik bei Stimmlippenpolypen.....	<u>7</u>
3.3 VHI-12, RBH, SUM und DSI zur Therapieverlaufskontrolle bei Stimmstörungen.....	<u>8</u>
<b>4. Diskussion</b> .....	<u>9</u>
<b>5. Literaturverzeichnis</b> .....	<u>14</u>
<b>Eidesstattliche Versicherung</b> .....	<u>19</u>
<b>Anteilerklärung an den erfolgten Publikationen</b> .....	<u>20</u>
<b>Publikation 1</b> .....	<u>21</u>
<b>Publikation 2</b> .....	<u>28</u>
<b>Publikation 3</b> .....	<u>38</u>
<b>Lebenslauf</b> .....	<u>49</u>
<b>Publikationsliste</b> .....	<u>50</u>
<b>Danksagung</b> .....	<u>51</u>

## Abkürzungsverzeichnis

<b>AF</b>	Autofluoreszenz
<b>AFL</b>	Autofluoreszenz-Laryngoskopie
<b>AVA</b>	Automatisierte Vokalanalyse (Stimmdokumentations- und -diagnostikprogramm)
<b>B</b>	Behauchtheit
<b>DSI</b>	Dysphonie-Schweregrad-Index (englisch: <i>Dysphonia Severity Index</i> )
<b>Fa.</b>	Firma
<b>F0_max</b>	höchster Ton
<b>H</b>	Heiserkeit
<b>I_min</b>	niedrigste Intensität
<b>LAF</b>	Autofluoreszenzverlust (englisch: <i>loss of autofluorescence</i> )
<b>MPT</b>	maximale Phonationsdauer (englisch: <i>maximum phonation time</i> )
<b>MSL_dB(A)</b>	mittlere Sprechstimmlage (in Dezibel)
<b>R</b>	Rauigkeit
<b>RBH</b>	stimmlicher Rauigkeits-, Behauchtheits- und Gesamt-Heiserkeitsgrad
<b>SUM</b>	Stimmumfangsmaß (englisch: <i>vocal extent measure</i> , VEM)
<b>SUP</b>	Stimmumfangsprofil
<b>TIVA</b>	Totale intravenöse Anästhesie
<b>VFP</b>	Stimmlippenpolypen (englisch: <i>vocal fold polyps</i> )
<b>VHI</b>	Fragebogen zur Stimmbeeinträchtigung (englisch: <i>Voice Handicap Index</i> )
<b>VHI-9i</b>	internationale 9-item-Kurzform des Voice Handicap Index
<b>VHI-12</b>	12-item-Kurzform des Voice Handicap Index
<b>VLS</b>	Videolaryngostroboskopie

## Abstract

This cumulative dissertation comprises three prospective clinical trials assessing the value of autofluorescence (AF) laryngoscopy (AFL), Vocal Extent Measure (VEM) and two short form versions of the Voice Handicap Index (VHI) in larynx and voice function diagnostics of dysphonic patients. The first study evaluated AFL in differential diagnostics of benign, premalignant and malignant glottal findings. After receiving rigid high-resolution white-light videolaryngostroboscopy (VLS), a fiberoptic blue-light AFL (SAFE-3000,  $\lambda=408$  nm) was performed in 32 patients. Clinical VLS- and AFL-assessments of dignity were compared to the final histopathological results. Benign and premalignant findings revealed different AF-patterns. Distinctive AF-reductions resulted in malignant lesions, where the clinical diagnosis was already determined by VLS, but also in chronic inflammation, teleangiectatic polyps, granulomas and papillomas. In premalignant findings we were unable to detect useful differential diagnostic information with therapy-modifying relevance. In a second trial, the application of VEM and VHI-9i were examined in voice function diagnostics of 61 patients with vocal fold polyps (VFP). Calculations of VEM from voice range profiles and overall VHI-scores from shortened VHI-9i-questionnaires were done pretherapeutically and after microlaryngoscopic VFP excision. Mean VHI-9i scores decreased from  $15\pm 8$  to  $6\pm 7$ , the VEM rose from  $83\pm 28$  to  $107\pm 21$ . VEM and the established Dysphonia Severity Index (DSI) significantly correlated with each other ( $r=0.65$ ;  $p<0.01$ ). VEM and VHI-9i revealed advantages in clinical application, compared to the DSI (e.g., independence from Jitter measurement, microphone distance) and the original 30-item-VHI-questionnaire (esp. better acceptance and practicability). A third study comprising 152 patients with various organic and functional voice disorders analyzed the suitability of VEM and VHI-12 for monitoring the course of therapy. Data were gathered before and after phonosurgical or logopedic intervention, in comparison with the gold standard of auditory-perceptual voice assessment via RBH-system (R: roughness, B: breathiness, H: overall hoarseness) and with DSI. During follow-up, mean DSI rose by 1.5 and VEM by 19, while H decreased by 0.5 and VHI-12 by 5 points. The correlations of these changes were significant, with changes in H being best detected by VEM. Therefore, the VEM confirmed the auditory findings directly and thus allowed a safer diagnostics and documentation. In summary, we conclude that the AFL-associated benefit does not seem to be significant for experienced phoniaticians and laryngologists. The VEM and both VHI short forms reliably quantify the objective vocal performance and the subjective suffering; their widespread introduction into clinical voice function diagnostics appears justified and desirable.

## Zusammenfassung

Diese kumulative Dissertation umfasst drei prospektive klinische Studien zur Beurteilung der Wertigkeit von Autofluoreszenz (AF)-Laryngoskopie (AFL), dem Stimmumfangsmaß (SUM), sowie zwei Voice Handicap Index Kurzformen (VHI-12, VHI-9i) in der Larynx- und Stimmfunktionsdiagnostik dysphoner Patienten.

Die erste Studie diente der AFL-Evaluation in der Differentialdiagnostik benigner, prämaligener und maligner Glottisbefunde. Nach hochauflösender starrer Weißlicht-Videolaryngostroboskopie (VLS) erfolgte bei 32 Patienten mit gutartigen, präkanzerösen und malignen Stimmlippenläsionen eine fiberoptische Blaulicht-AFL (SAFE-3000;  $\lambda=408$  nm). Die klinischen VLS- und AFL-Dignitätsbewertungen wurden mit dem endgültigen histopathologischen Ergebnis verglichen. Es zeigte sich, dass benigne und (prä-)maligne Befunde unterschiedliche AF-Muster aufwiesen. Markante AF-Minderungen resultierten bei malignen Befunden, wenn sich die klinische Diagnose bereits in der VLS klar ergab, jedoch auch bei chronischer Entzündung, teleangiektatischen Polypen, Granulomen und Papillomen. Eine differenzialdiagnostische und therapiemodifizierende Relevanz bei prämaligen Befunden konnten wir nicht feststellen.

In einer zweiten Studie wurden der Einsatz des SUM und des VHI-9i in der Stimmfunktionsdiagnostik bei 61 Patienten mit Stimmlippenpolypen (VFP) untersucht. Die Berechnungen des SUM aus den Stimmumfangsprofilen sowie der Gesamt-Scores aus den verkürzten VHI-9i-Fragebögen wurden prätherapeutisch und nach mikrolaryngoskopischer VFP-Abtragung durchgeführt. Der durchschnittliche VHI-9i-Score verringerte sich von  $15\pm 8$  auf  $6\pm 7$ , der SUM stieg von  $83\pm 28$  auf  $107\pm 21$ . Das SUM und der etablierte Dysphonia Severity Index (DSI) korrelierten signifikant miteinander ( $r=0,65$ ;  $p<0,01$ ). Die Parameter SUM und VHI-9i wiesen eine Reihe von Vorteilen im klinischen Einsatz gegenüber dem DSI (z.B. Unabhängigkeit von Jitter-Messung, Mikrofon-abstandsbedingt abweichenden Schallpegeln) bzw. der ausführlichen 30-item Originalversion des VHI auf (v.a. bessere Akzeptanz und Praktikabilität).

In einer dritten Studie an 152 Patienten mit verschiedenen organischen und funktionellen Stimmstörungen wurde die Eignung der Parameter SUM und VHI-12 zur Therapieverlaufskontrolle analysiert. Die Datenerfassung erfolgte prätherapeutisch und nach phonochirurgischer oder logopädischer Intervention, im Vergleich mit dem Goldstandard der auditiv-perzeptiven Stimmbeurteilung mittels RBH-System (R: Rauigkeit, B: Behauchtheit, H: Gesamtheiserkeit) sowie mit dem DSI. Im Therapieverlauf stiegen der mittlere DSI um 1,5 und das SUM um 19, während H um durchschnittlich 0,5 und der VHI-12 um 5 Punkte sanken. Die Korrelationen dieser Veränderungen waren signifikant, wobei Änderungen im Heiserkeitsgrad am besten mit

dem SUM erkannt wurden. Das SUM stützte demnach den auditiven Befund direkt und erlaubte so eine sicherere Diagnostik und Dokumentation.

Zusammenfassend schlussfolgern wir, dass der AF-laryngoskopische Erkenntnisgewinn für erfahrene Phoniater und Laryngologen nicht bedeutsam zu sein scheint. Die Parameter SUM und beide VHI-Kurzformen quantifizieren verlässlich die objektive stimmliche Leistungsfähigkeit bzw. den subjektiven Leidensdruck; ihre breite Einführung in die klinische Stimmfunktionsdiagnostik erscheint gerechtfertigt und wünschenswert.

## 1. Einleitung

Die Phoniatrie bzw. „Stimmheilkunde“ ist das medizinische Fachgebiet, das sich diagnostisch, therapeutisch und wissenschaftlich mit Beeinträchtigungen der Stimme, des Sprechens, der Sprache und des Schluckens befasst. Nach der weltweiten Begründung der Phoniatrie 1905 in Berlin von Hermann Gutzmann sen. führten zahlreiche Impulse zur globalen Ausbreitung und Etablierung dieses Fachgebietes in der Medizin [1]. Heute ist die Phoniatrie in Verbindung mit der Pädaudiologie die medizinische Spezialdisziplin, die sich auf die Pathophysiologie der Kommunikation gründet und die Vorbeugung, Erkennung, konservative und operative Behandlung sowie Rehabilitation von krankheitsbedingten Störungen der peripheren und zentralen Hörfunktion, der Sprech- und Sprachfunktion, der Laut- und Schriftsprache, der Stimm-, Kau- und Schluckfunktion einschließlich psychosomatischer Begleiterkrankungen und musikermedizinischer Erkrankungen umfasst [2,3].

Als vergleichsweise junge, in Deutschland seit 1993 selbständige Disziplin handelt es sich erst seit 2018 um ein eigenständig anerkanntes Gebiet mit einer neuen, autonomen Weiterbildungsordnung [4]. Die Phoniatrie und Pädaudiologie basiert auf den anatomisch-physiologischen, diagnostischen und therapeutischen Grundlagen verschiedenster medizinischer Arbeitsbereiche mit engen interdisziplinären Beziehungen zur Hals-Nasen-Ohrenheilkunde, Pädiatrie, Neurologie, Psychiatrie, Geriatrie, Mund-Kiefer-Gesichtschirurgie, Zahn- und Rehabilitationsmedizin [5]. Zudem bestehen enge Verknüpfungen mit vielen para- und nicht-medizinischen Disziplinen, wie z.B. Logopädie, Linguistik, Phonetik, Hörgeräteakustik, (Sonder-)Pädagogik, Psychologie und Akustik. Global und volkswirtschaftlich betrachtet ist die Phoniatrie und Pädaudiologie von wachsender Bedeutung: Während in früheren Generationen die Unfähigkeit zu manueller Tätigkeit den Krankheitsbegriff wesentlich geprägt hat, so gewinnen heute kommunikative Fähigkeiten stetig an Bedeutung. Hierbei spielen sowohl impressive (Hören) als auch expressive (Sprache, Sprechen, Stimme) Kommunikationsbeeinträchtigungen eine wesentliche Rolle. Untersuchungen von Ruben zeigten, dass der jährliche volkswirtschaftliche Verlust durch eine beeinträchtigte Kommunikationsfähigkeit allein in den USA in einer Größenordnung von > 100 Mrd. US-Dollar liegt [6].

Vor diesem Hintergrund besteht für das hochspezialisierte Fachgebiet der Phoniatrie und Pädaudiologie die alternativlose Notwendigkeit, sich diagnostisch und therapeutisch fortwährend weiterzuentwickeln, um die eingeschränkte Kommunikationsfähigkeit der ihr anvertrauten Patienten zukünftig immer besser erkennen und auch versorgen zu können. Hierbei steht

insbesondere die Diagnostik der gestörten Stimmfunktion im Fokus, bei der die Eigenschaften des mehrdimensionalen und subjektiven „Phänomens Stimme“ exakt erfasst sowie die stimmliche Qualität und Leistungsfähigkeit genau gemessen werden müssen [7,8]. In der phoniatischen Routine sind diverse Instrumente im Einsatz, die eine Vielzahl akustischer und aerodynamischer Parameter bei der Stimmproduktion quantifizieren [9-11]. Durch Fortschritte in der Messtechnik und Signalverarbeitung, aber auch durch kritische Auseinandersetzung mit bereits Bestehendem kann dieser Pool an stimmbezogenen Messungen kontinuierlich erweitert werden. Das Ziel darf jedoch nicht dem Selbstzweck der bloßen Schaffung eines zusätzlichen Parameters dienen, sondern muss dem Anspruch gerecht werden, den Ist-Zustand und damit die umfassende Beurteilung der stimmlichen Leistungsfähigkeit besser zu objektivieren und zu standardisieren. Deshalb ist jeder neue Parameter kritisch zu evaluieren, auf seinen Neuwert zu prüfen und in seiner Notwendigkeit zu begründen. Gleiches gilt für den diagnostischen Einsatz neu entwickelter Technologien, wobei der zu erzielende Erkenntnisgewinn für in der Differentialdiagnostik laryngealer Läsionen erfahrene Phoniater und Laryngologen kritisch hinterfragt werden muss.

Die erste Arbeit der vorliegenden kumulativen Dissertation beschäftigte sich mit der Autofluoreszenz-Laryngoskopie (AFL). Die Nutzung von Autofluoreszenz (AF) und Schmalbandlicht als neuen Technologien zur Frühdiagnostik maligner Schleimhautveränderungen werden an den unterschiedlichsten Lokalisationen in zunehmend vielen medizinischen Fachdisziplinen eingesetzt [12-16]. Der Vorteil der AF-Diagnostik mit Blaulicht gegenüber der konventionellen Weißlichtendoskopie soll darin bestehen, dass sich bereits prä- und frühmaligne Läsionen aufgrund einer verminderten AF kontrastreich gegenüber dem umgebenden gesunden Gewebe darstellen und voneinander differenzieren lassen. Grundlage der endoskopisch nachweisbaren unterschiedlichen AF sind Veränderungen der Konzentration endogener Fluorophore, der Morphologie des Gewebes und gewebeoptischer Parameter [17-21]. In unserer prospektiven Studie sollte der diagnostische Wert der AFL im Vergleich zur etablierten Weißlicht-Videolaryngostroboskopie (VLS) und gegenüber dem Goldstandard der Mikrolaryngoskopie mit histopathologischer Untersuchung in der Differentialdiagnostik von Larynxläsionen evaluiert werden.

In einer zweiten klinischen Studie wurden der Einsatz des neuen Stimmumfangsmaßes (SUM) sowie der 9-item Kurzform des Voice Handicap Index (VHI-9i) in der Stimmfunktionsdiagnostik bei Patienten mit Stimmlippenpolypen (VFP) untersucht [22]. VFP sind gutartige oberflächliche Schwellungen im Bereich der Lamina propria, die die Stimmlippenschwingungen mechanisch behindern und dadurch die Phonation erschweren [23-25]. Sie entstehen vornehmlich an

Lokalisationen, die während der Stimmgebung den stärksten muskulären und aerodynamischen Kräften ausgesetzt sind [26,27]. Die Standardtherapie besteht in einer kompletten chirurgischen Entfernung, über direkte Laryngoskopie in Intubationsnarkose oder am wachen, sitzenden Patienten in Lokalanästhesie [28-30]. Ziel unserer Untersuchung war die Beurteilung des Ausmaßes der operativen Stimmverbesserung, unter Beurteilung der Anwendbarkeit und Praktikabilität der in der phoniatischen Routinediagnostik noch nicht allgemein etablierten Stimmfunktionsparameter SUM und VHI-9i.

Die dritte Arbeit dieser kumulativen Dissertation widmete sich Patienten mit verschiedenen organischen und funktionellen Stimmstörungen, bei denen die Eignung der Parameter SUM und der 12-item Kurzform des Voice Handicap Index (VHI-12) zur Therapieverlaufskontrolle analysiert wurde [22]. Die häufigsten Ursachen organischer Dysphonien sind benigne und maligne Neubildungen, Entzündungen sowie Lähmungen [31,32]. Funktionelle Dysphonien sind gekennzeichnet durch Veränderungen des Stimmklanges, Einschränkungen der stimmlichen Leistungsfähigkeit sowie laryngeale Missempfindungen, ohne dass erkennbare organischen Veränderungen vorliegen [33,34]. Je nach klinischer Erscheinungsform werden hyper- und hypofunktionelle Dysphonien unterschieden, wobei in der VLS ein gestörtes Schwingungs- und Schließungsverhalten der Stimmlippen sichtbar werden kann [35]. Für eine umfassende Dokumentation phonochirurgischer und konservativer stimmtherapeutischer Verfahren ist es erforderlich, dass Änderungen der Stimmqualität ausreichend gut detektiert und quantifiziert werden. Ziel war es daher, die messbaren Veränderungen der objektiven Parameter inklusive des SUM im Therapieverlauf eines jeden Patienten zu untersuchen und mit den Änderungen der subjektiven Stimmparameter unter Einbeziehung des VHI-12 zu vergleichen.

## **2. Material und Methodik**

### ***2.1 AFL in der Differentialdiagnostik laryngealer Läsionen***

In dieser prospektiven klinischen Studie wurde der AF-laryngoskopische Erkenntnisgewinn in der Differentialdiagnostik benigner, prä-maligner und maligner Glottisbefunde untersucht. Über die Ambulanzen der Kliniken für HNO-Heilkunde sowie Audiologie und Phoniatrie wurden 32 Patienten mit Verdacht auf maligne und benigne Stimmlippenläsionen rekrutiert (27 Männer und 5 Frauen, Durchschnittsalter  $56 \pm 9$  Jahre [Mittelwert  $\pm$  Standardabweichung]).

Bei Erhebung des Kopf-Hals-Status erfolgte zunächst eine konventionelle Weißlichtendoskopie unter Nutzung einer hochauflösenden, starren 90°-Optik (EndoSTROB System, Fa. XION).

Neben der Lupenlaryngoskopie kam die digitale Videostroboskopie zum Einsatz, wobei die Parameter phonatorische Stimmlippenbeweglichkeit und Randkantenverschieblichkeit besondere Berücksichtigung bei der Stellung der klinischen Verdachtsdiagnose fanden.

Die AF-Laryngoskopie wurde von einem unabhängigen, erfahrenen Bronchoskopiker mittels digitalem Video-AF-Bronchoskop verblindet durchgeführt (SAFE-3000 System, Fa. PENTAX). Nach Ausleuchtung des Endolarynx im Weißlichtmodus mittels 300 W Xenon-Lampe diente eine Laserdiode der Wellenlänge  $\lambda = 408$  nm als AF-Lichtquelle zur Erzeugung des speziellen Blaulichtes für die Fluoreszenzanregung. Per Knopfdruck wurde nahtlos zwischen Video- und AF-Modus umgeschaltet. Zwecks genauer Visualisierung und sofortigem Vergleich relevanter Gewebestrukturen erfolgte per Twin-Modus die simultane Darstellung des jeweiligen endoskopischen Weißlicht- und korrespondierenden AF-Bildes auf dem Monitor. Die Befunde wurden digital gespeichert und ausgedruckt. Anschließend erfolgte eine Dignitätsbewertung der Aufnahmen durch 4 erfahrene Laryngologen und Phoniater. Der AF-Verlust LAF ("loss of autofluorescence") wurde dabei klassifiziert in die Grade LAF 0, LAF 1, und LAF 2 (0 = kein, 1 = milder, 2 = deutlicher AF-Verlust).

Im Rahmen einer diagnostischen oder therapeutischen Mikrolaryngoskopie wurden die Läsionen je nach Befund biopsiert bzw. exzidiert und zwecks Diagnosesicherung feingeweblich untersucht. Um die diagnostische Wertigkeit der AF-Endoskopie im Vergleich zur Weißlichtendoskopie sowie zur histopathologischen Diagnostik als Goldstandard zu evaluieren, wurden Kennwerte für die Qualität diagnostischer Tests berechnet (Sensitivität, Spezifität, falsch-positiver und falsch-negativer Wert).

## ***2.2 SUM und VHI-9i in der Stimmfunktionsdiagnostik bei Stimmlippenpolypen***

In dieser zweiten prospektiv-klinischen Studie wurden bei 61 Patienten (28 Männer und 33 Frauen, Durchschnittsalter 45 Jahre, Altersspanne 21 bis 84 Jahre) die funktionellen Ergebnisse der mikrolaryngoskopischen VFP-Abtragung untersucht. Das intraoperative therapeutische Vorgehen bestand in einer kompletten Exzision, bei der nach Fassen des Polypen mit einem geeigneten Miniaturzängchen zunächst die kraniale, anschließend die kaudale Anheftungsstelle der Polypenbasis mit einem Mikroscherchen tangential durchtrennt und abgeschnitten wurde. Dabei blieb unter Schonung der Randkante die oberflächliche Schicht der Lamina propria erhalten.

Die Datenerfassung erfolgte prä-, intra- und 3 Monate posttherapeutisch. Die Befunde und Behandlungsergebnisse wurden anhand prä- und postoperativer VLS, intraoperativer Aufnahmen, sowie mittels prä- und postoperativer Stimmfunktionsdiagnostik dokumentiert und

ausgewertet. Untersuchungsinstrumente waren allgemein etablierte objektive und subjektive Verfahren: die auditiv-perzeptive Stimmbewertung mittels RBH-Systematik (Bewertung von stimmlicher Rauigkeit R, Behauchtheit B, und dem Gesamt-Heiserkeitsgrad H), die Stimmumfangsprofil (SUP) - Messung, sowie akustisch-aerodynamische Analysen. Zusätzlich erfolgte die Selbsteinschätzung der Stimme durch die Patienten mit dem VHI-9i.

Außerdem wurde der auf der Basis von Stimmfeldfläche und Stimmfeldform errechnete Parameter SUM eingesetzt. Das SUM quantifiziert die durch die SUP-Messung dokumentierte stimmliche Leistungsfähigkeit. Ein ideales Stimmfeld im Sinne einer hohen stimmlichen Leistungsfähigkeit zeichnet sich dadurch aus, dass über einen weiten Tonumfang die Kurven der leisesten und lautesten Töne gleich weit voneinander entfernt sind. Beim SUM wird ein Stimmfeld als Bezugsgröße im Stimmfelddiagramm mit skalierten Abszissen durch einen Kreis idealisiert. Für die Berechnung des SUM wird der Flächeninhalt des Stimmfeldes mit dem Quotienten aus dem theoretischen Umfang eines Kreises mit gleicher Fläche wie dem Stimmfeld und dem tatsächlichen Umfang des Stimmfeldes multipliziert. Mittels des Stimm-dokumentations- und -diagnostikprogramms „Automatisierte Vokalanalyse“ (AVA) konnten alle benötigten Daten aus den Dateien des Registrierungsprogramms DiVAS (Fa. XION) ausgelesen und das SUM automatisiert und noch während der Untersuchung - ohne Mehraufwand für den Untersucher - angezeigt werden. Als Ergebnis der SUM-Berechnung resultierte eine dimensionslose Zahl im Bereich zwischen -150 und 150, wobei ein größerer Wert eine größere stimmliche Leistungsfähigkeit widerspiegelt. Das SUM sollte in der Auswertung mit dem etablierten Dysphonie Schweregrad Index (DSI) verglichen und bezüglich seiner diagnostischen Eignung bewertet werden.

### ***2.3 VHI-12, RBH, SUM und DSI zur Therapieverlaufskontrolle bei Stimmstörungen***

In diese prospektiv-klinische Studie wurden insgesamt 152 Patienten (50 Männer [16–75 Jahre, Median 42] und 102 Frauen [17–70 Jahre, Median 48]) mit verschiedenen Stimmproblemen eingeschlossen. Alle Daten wurden beim prätherapeutischen Besuch und drei Monate post-interventionell erhoben, sodass 304 Datensätze gesammelt wurden. Entsprechend der diagnostizierten klinischen Pathologie und dem vorherigen Krankheitsverlauf erhielten die Patienten entweder eine chirurgische Behandlung oder konservative Stimmübungen. Die logopädische Stimmübungstherapie wurde von qualifizierten Logopäden durchgeführt und umfasste 20 Sitzungen (2-mal pro Woche à 45 Minuten). Die phonomikrochirurgische Befundabtragung erfolgte durch 3 erfahrene Phonochirurgen mittels direkter Mikrolaryngoskopie in Vollnarkose (TIVA mit Propofol/ Remifentanyl).

Zur Beurteilung der Behandlungsergebnisse kamen etablierte Untersuchungsinstrumente zum Einsatz. Die digitale VLS wurde mit einem hochauflösenden starren Videolaryngoskop (10 mm; 70°) mit integriertem Mikrofon (Fa. XION) durchgeführt. Die Laryngoskopie diente der Unterscheidung zwischen organischer und funktioneller Genese der Dysphonie. Die Stroboskopie machte die Stimmlippenvibrationen während der Phonation sichtbar und zeigte Beeinträchtigungen der Randkantenverschieblichkeit bzw. der phonatorischen Stimmlippenbeweglichkeit an. Das LingWAVES-Programm (Fa. WEVOSYS) wurde zur standardisierten SUP-Registrierung und zur akustisch-aerodynamischen Analyse eingesetzt. Basierend auf den Parametern niedrigste Intensität (I\_min), höchster Ton (F0\_max), maximale Phonationsdauer (MPT) und Jitter wurde der DSI berechnet. Die Berechnung des noch nicht allgemein etablierten SUM erfolgte mit dem proprietären Softwareprogramm AVA. Der VHI-12 diente zur subjektiven Selbsteinschätzung der eigenen Stimmfunktion durch die Patienten. Die externe auditiv-perzeptive Stimmbeurteilung erfolgte anhand des vorzulesenden Mustertextes „Der Nordwind und die Sonne“ (Äsop-Fabel) mittels RBH-System, wobei der Gesamtheiserkeitsgrad H als Goldstandard zugrunde gelegt wurde, um den Therapieerfolg anzuzeigen.

Die Ergebnisanalyse basierte auf der prä- und posttherapeutischen Stimmfunktionsdiagnostik sowie auf der VLS. Die Parameter H, VHI-12, DSI und SUM wurden vor und 3 Monate nach der jeweiligen Intervention miteinander verglichen. Es wurde geprüft, ob die Therapie zu einem signifikanten Unterschied der erhobenen Parameter führte. Zusätzlich wurden die Messwerte vor und nach der Behandlung sowie auch die jeweiligen Änderungen miteinander korreliert. Statistisch kamen die Berechnungen der Rangkorrelationskoeffizienten nach Spearman sowie der t-Test für verbundene Stichproben zur Anwendung. Das Signifikanzniveau betrug  $\alpha = 0,05$ .

### **3. Ergebnisse**

#### ***3.1 AFL in der Differentialdiagnostik laryngealer Läsionen***

Bei den 32 eingeschlossenen Patienten waren bezüglich der Befundlokalisation bei Erstvorstellung die Glottispathologien in 26 Fällen einseitig (rechte Stimmlippe n=12, linke Stimmlippe n=14) und in 6 Fällen beidseitig lokalisiert. Hinsichtlich der Dignität dieser Befunde wurde mittels lupenlaryngoskopischer und videostroboskopischer Weißlichtdiagnostik der klinische Verdacht auf 15/32 benigne und 17/32 maligne Läsionen gestellt. Bei Beurteilung der AF-Diagnostik erschienen 13/32 Befunde benigne und 19/32 malignitätssuspekt. Übereinstimmung zwischen beiden Untersuchungsverfahren hinsichtlich der prädiktiven

Dignitätsannahme bestand bei 28/32 Patienten (87,5%). Die endgültige Histopathologie ergab jeweils 16/32 benigne bzw. maligne Befunde.

In der AF-Laryngoskopie traten die verschiedenen LAF-Grade in unterschiedlicher Häufigkeit auf: LAF 2 bei n=20, LAF 1 bei n=4, und LAF 0 bei n=8 Befunden. Benigne und (prä-)maligne Läsionen wiesen dabei unterschiedliche AF-Muster auf. Ein ausbleibender AF-Verlust (LAF 0) resultierte bei normaler Schleimhaut, Leukoplakien (Hyperkeratosen) und benignen Veränderungen des Epithels und der Lamina propria (z.B. Reinke-Ödem, Polypen). Typische Beispiele für milde AF-Minderungen (LAF 1) waren Hyperplasien und geringgradige Epitheldysplasien. Ein markanter AF-Verlust (LAF 2) resultierte einerseits bei malignen Befunden (invasive Plattenepithelkarzinome, Carcinomata in situ), wenn sich die klinische Diagnose bereits in der Weißlicht-VLS klar ergab, jedoch auch bei chronischer Entzündung, mittel- und hochgradigen Dysplasien, teleangiektatischen Polypen, Granulomen und Papillomen. Im Vergleich mit der pathologisch gesicherten Diagnose der endgültigen Histologie zeigte sich, dass die vorangegangene klinische Diagnosestellung mittels Weißlicht bei einem Patienten (1/32 falsch positiv), mittels Blaulicht bei 5 Patienten falsch war (4/32 falsch positiv, 1/32 falsch negativ). Die Sensitivität der AF-Diagnostik betrug 94%, die Spezifität 69%. Der falsch-positive Wert ergab 0,31; d.h. mit einer Wahrscheinlichkeit von 31% lag bei einem benignen Befund ein malignitätssuspekter deutlicher AF-Verlust (LAF 2) vor. Der falsch-negative Wert betrug 0,06; d.h. mit einer Wahrscheinlichkeit von 6% ergab sich bei einem malignen Befund eine auf Benignität hinweisende, fehlende (LAF 0) oder milde AF-Minderung (LAF 1).

### ***3.2 SUM und VHI-9i in der Stimmfunktionsdiagnostik bei Stimmlippenpolypen***

Hinsichtlich der präoperativen Charakterisierung waren 32 VFP hämorrhagischen und 29 VFP hyalinen Aussehens. Die überwiegende Mehrzahl saß den Stimmlippen breitbasig auf (n=55), nur 6 Polypen waren gestielt. 54 Patienten hatten einen einseitigen Stimmlippenbefund (links: n=28, rechts: n=26), bei 7 Patienten zeigten sich beidseitige Stimmlippenpolypen. Bezüglich des Ausmaßes der individuellen präoperativen Stimmbelastung hatten 37 Patienten einen vorwiegend nicht-professionellen Stimmgebrauch. 24 Patienten hatten einen intensiveren, überwiegend professionellen Stimmgebrauch, hiervon waren 9 Betroffene als Sänger und Schauspieler sogenannte „elite vocal performer“.

Alle VFP konnten intraoperativ wie geplant komplett entfernt werden. Die laryngoskopischen Kontrolluntersuchungen zeigten, dass sich die Heilung der Stimmlippen regelrecht durch narbenfreies Überwachsen des Defektes mit stabilem Epithel auf der Basis der erhaltenen Lamina propria vollzog. Stroboskopisch konnte postoperativ eine wiederhergestellte regelrechte

Randkantenverschieblichkeit nachgewiesen werden. Innerhalb des Beobachtungszeitraumes (275±288 Tage) wurden keine relevanten Nebenwirkungen oder Rezidive beobachtet. Lediglich bei einer Patientin kam es im Verlauf zu einer erneuten Stimmlippenverdickung im Bereich des ehemaligen abgetragenen Polypen.

Die Ergebnisse zeigten postoperativ eine größtenteils deutliche Verbesserung der Stimmfunktion im Vergleich zum präoperativen Befund. In der auditiven Beurteilung mittels RBH-System waren die Stimmen im verblindeten prä- und posttherapeutischen Vergleich signifikant weniger rau, behaucht und heiser ( $1,1 \pm 0,8$  vs.  $0,4 \pm 0,5$ ). Die subjektive Selbsteinschätzung der Stimme ergab im VHI-9i ein deutliches Absinken des Beschwerde-Scores von durchschnittlich  $15 \pm 8$  auf  $6 \pm 7$  Punkte ( $p=0,000$ ). Bis auf den Jitter zeigten sich auch sämtliche objektiven akustischen und aerodynamischen Parameter signifikant gebessert. Der durchschnittliche Gesamtstimmumfang erweiterte sich um  $4 \pm 5$  Halbtöne, die mittlere ungespannte Sprechstimmlage sank im Mittel um  $1 \pm 2$  Halbtöne. Die maximale Phonationsdauer (MPT) erhöhte sich von durchschnittlich  $12 \pm 4$  Sekunden präoperativ auf  $16 \pm 6$  Sekunden postoperativ ( $p=0,000$ ). Der Parameter DSI stieg von  $2,6 \pm 2,1$  auf  $4,0 \pm 2,2$  ( $p=0,000$ ), der Parameter SUM von  $83 \pm 28$  auf  $107 \pm 21$  ( $p=0,000$ ). SUM und DSI sowie ihre prä- und posttherapeutischen Differenzen korrelierten jeweils hoch signifikant miteinander.

### ***3.3 VHI-12, RBH, SUM und DSI zur Therapieverlaufskontrolle bei Stimmstörungen***

Die Studienkohorte umfasste 66 Individuen (43%) mit nicht-professionellem Stimmgebrauch und 86 Patienten (57%) mit hoher beruflicher Stimmbelastung. Die Studienteilnehmer zeigten prätherapeutisch verschiedene klinische Störungsbilder. Mittels VLS wurden bei 101 Probanden (66%) organische Erkrankungen auf Stimmlippenebene festgestellt. Die Einteilung entsprechend der zugrundeliegenden Pathologien ergab bei 41 Patienten (27%) Erkrankungen der Lamina propria, bei 24 Patienten (16%) laryngeale Bewegungsstörungen, bei 19 Patienten (12%) epitheliale Erkrankungen und bei 5 Patienten (3%) Pathologien im Arytenoid-Bereich. Einundfünfzig Individuen (34%) hatten eine normale Kehlkopf-anatomie, litten jedoch an einer stimmbelastungsinduzierten funktionellen Dysphonie. Initial hatten 46 Probanden (30%) keine Heiserkeit (H0), darunter 29 Patienten mit funktioneller Dysphonie, 13 Patienten mit sehr kleinen organischen Glottisbefunden, sowie 4 Patienten mit Stimmlippen-ferneren Gewebeveränderungen. Bezüglich der Therapieverlaufskontrolle zeigten sich alle untersuchten Stimmparameter posttherapeutisch verbessert. Der mittlere RBH-Status offenbarte eine geringere Rauigkeit, Behauchtheit und Gesamt-Heiserkeit ( $p<0,001$ ). H nahm von 1,2 auf 0,7 ab und reduzierte sich bei den meisten Patienten von H1 auf H0 ( $n=29$ ; d.h. 19%) bzw. von H2 auf H1 ( $n=22$ ; d.h.

14%). Der VHI-12 verringerte sich im Durchschnitt von 15 auf 10, was einer Änderung von ursprünglich moderater hin zu milder Beeinträchtigung entspricht ( $p < 0,001$ ). Der DSI war signifikant verbessert mit einem mittleren Anstieg von 2,2 auf 3,7 ( $p < 0,001$ ). Das SUM stieg ebenfalls von 60 auf 79 und spiegelt eine signifikant größere Stimmkapazität wider ( $p < 0,001$ ). Der Vergleich beider Behandlungsmodalitäten zeigte, dass phonochirurgische Interventionen den größten Einfluss auf die Stimmfunktion hatten, mit einer höheren numerischen Verbesserung fast aller Parameter. Konservative Therapien lieferten quantitativ kleinere Verbesserungen, führten aber auch zur qualitativen Wiederherstellung der Stimmfunktion. Alter und Geschlecht hatten keinen signifikanten Einfluss auf das Behandlungsergebnis.

Die Korrelation der Parameter H, VHI-12, DSI und SUM untereinander zeigte eine schwache bis moderate lineare Beziehung. Die Stärke der Beziehung änderte sich aufgrund der Therapie meist nur geringfügig. Das SUM korrelierte mit VHI-12 bei  $r = -0,4$  und mit H bei  $r = -0,7$ , was unabhängig vom Therapiestatus eine moderate bzw. starke negative Beziehung aufzeigte. Die Untersuchung der therapiebedingten Veränderungen ( $\Delta$ ) in den einzelnen Messdaten ergab, dass Korrelationen dieser Veränderungen für alle Parameter zu eher kleinen Koeffizienten führten.  $\Delta$  VHI-12 zeigte keine signifikante Beziehung zu  $\Delta$  DSI, dafür jedoch zu  $\Delta$  SUM ( $r = -0,2$ ) und zu  $\Delta$  H ( $r = 0,3$ ). Die Beziehung zwischen  $\Delta$  SUM und  $\Delta$  DSI war moderat ( $r = 0,5$ ). Alle diese Korrelationen erwiesen sich als hoch signifikant ( $p < 0,01$ ).

#### 4. Diskussion

Zusammenfassend zeigten die Ergebnisse der in diese kumulative Dissertation eingeflossenen Untersuchungen, dass nicht die Autofluoreszenz-Laryngoskopie mit dem verwendeten SAFE-3000 System, jedoch das Stimmumfangsmaß SUM sowie die eingesetzten Voice Handicap Index Kurzformen VHI-9i und VHI-12 ein großes Potenzial besitzen, die spezifische Funktionsdiagnostik in der Stimmheilkunde zu verbessern.

Unsere erste Studie demonstrierte, dass die AF-Laryngoskopie als nicht-invasives bildgebendes Verfahren unsere weißlichtendoskopischen Untersuchungsergebnisse in 87,5% der Fälle bestätigte. Ein zusätzlicher Benefit im Sinne einer gesteigerten Sicherheit bzw. Eindeutigkeit bei der Dignitätsbeurteilung von Glottispathologien zeigte sich nicht. Eine differenzialdiagnostische und therapiemodifizierende Relevanz insbesondere bei prämaligen Befunden konnten wir nicht feststellen. Im Vergleich zum Weißlicht resultierten mit dem monochromatischen Licht des

verwendeten AF-Systems eine reduzierte Spezifität (69%) und ein hoher falsch-positiver Wert (31%). Falsch-positive Ergebnisse führen dazu, dass Patienten mit unklaren Befunden eher einer histologischen Diagnosesicherung in ITN zugeführt werden als notwendig, falsch-negative Ergebnisse durch Abwarten zu einer Größenzunahme des Befundes und höherem Exzisions-trauma. Andere AF-Systeme (z.B. D-Light, Fa. STORZ) zeigen eine vergleichsweise höhere Spezifität, was durch das breitere Anregungsspektrum bedingt sein kann [36,37].

Ein weiterer Nachteil der AFE mittels SAFE-3000 System war die Bildqualität, die generell leicht reduziert erschien im Vergleich zum Weißlichtmodus. Unsere endolaryngealen AF-Bilder waren deutlich schärfer und klarer im Vergleich zum SAFE-1000 System [38,39], aber nicht so gut wie die Visualisierung mit dem D-Light System [40,41]. Negativen Einfluss auf unsere Bildqualität hatten auch Patientenbewegungen (z.B. beim Schlucken oder Husten) und die reduzierte Lichtleistung. So imitierten Omega-förmige oder retrograd positionierte Kehldeckel AF-Aussparungen in der vorderen Kommissur bzw. im vorderen Stimmlippendrittel aufgrund von Schattenbildung im Bereich dieser Strukturen. Artifiziell reduzierte AF-Signale resultierten auch bei hypertrophen Taschenfalten, wenn sich Schattenstrukturen über den lateralen Teil der Stimmlippen ausbreiteten. Allerdings gelang es, durch distalere Neupositionierung des flexiblen Endoskops näher an der jeweiligen Kehlkopfpathologie die störende Schattenbildung proximaler Strukturen weitgehend zu verringern bzw. zu vermeiden. Unabhängig davon gelang es im AF-Modus oft nicht suffizient, Tumorränder sowie prämaligne Gewebeveränderungen ausreichend scharf vom gesunden Gewebe abzugrenzen. Im Gegensatz zu unseren Ergebnissen zeigen Veröffentlichungen mit dem D-Light AF-System bessere Kontraste zur Abgrenzung von Vorläuferläsionen und dysplastischen Karzinomrändern [40,41]. Generell erfordert die Interpretation von AF-Bildern ausreichend Kenntnisse über die Grundprinzipien und Fallstricke dieser Technik, wobei eine Lernkurve zu erwarten ist, die sich positiv auf die klinische Anwendungssicherheit auswirken sollte.

Unsere Ergebnisse lassen darauf schließen, dass der mit dem verwendeten AF-System zu erzielende Erkenntnisgewinn für in der Differentialdiagnostik laryngealer Läsionen erfahrene Phoniater und Laryngologen nicht bedeutsam ist. Vor diesem Hintergrund erscheint zum gegenwärtigen Zeitpunkt die klinische Erfahrung des Untersuchers in der Beurteilung der weißlichtendoskopischen Lupenlaryngoskopie und vor allem der Videolaryngostroboskopie verlässlicher als die AF-Endoskopie mit dem SAFE-3000 System als relativ unspezifische Methode zur Diagnostik von Schleimhautveränderungen.

In unserer zweiten Studie zeigte sich, dass sowohl das SUM als auch der VHI-9i erfolgreich in

der Stimmfunktionsdiagnostik bei VFP Patienten eingesetzt werden konnten. Bei stark dysphonen Stimmen mit sehr kleinem Stimmumfang und großen Abweichungen vom Ideal des kreisförmigen Modells resultierten erwartungsgemäß deutlich negative SUM-Werte. Der VHI-9i maß verlässlich die subjektive Stimmfunktionseinschränkung, vergleichbar zur früher verwendeten, ausführlichen 30-item VHI-Originalversion, bei jedoch offensichtlich besserer klinischer Akzeptanz und Praktikabilität. Die gesteigerte Effizienz der Phonation nach Befundabtragung führte größtenteils zu einer deutlichen Stimmverbesserung. Dies bestätigt Studien aus der Literatur, die ebenfalls eine Besserung perzeptiver und akustischer Stimmparameter nach phonomikrochirurgischer VFP Entfernung dokumentierten. Zeitels et al. [42] untersuchten 71 VFP-betroffene Sänger und Schauspieler, die in der postoperativen Kontrolle eine objektiv und subjektiv nachweisbare bessere Stimmfunktion aufwiesen. Studien von Karasu et al. [29], Rosen et al. [43], Uloza et al. [30] und Petrovic-Lazic et al. [25] bekräftigten diese Ergebnisse. Bei unseren Patienten verbesserten sich alle subjektiven und objektiven Parameter außer Jitter und MSL<sub>dB(A)</sub> signifikant. SUM und DSI korrelierten wie erwartet negativ mit dem Alter und dem RBH-Status, aber positiv mit MPT. Unsere Ergebnisse zeigten auch eine Korrelation zwischen R bzw. H mit der Polypengröße, woraus erkennbar ist, dass die Rauigkeit den Gesamteindruck der Heiserkeit bestimmt. MPT korrelierte am deutlichsten mit B und bestätigte damit den klinischen Eindruck, dass stärkere Behauchtheit für einen höheren Luftverbrauch und eine kürzere Phonationszeit steht. Bezüglich der Polypengröße befanden Cho et al. [44], dass sie der entscheidendste Einflussfaktor für die Stimmqualität ist. Unseren Ergebnissen zufolge beeinflusste das Ausmaß des präoperativen Befundes das operative Outcome jedoch nicht. Wichtigste Voraussetzung zur Vermeidung einer postoperativen Dysphonie scheint unseres Erachtens nicht die präoperative Polypengröße, sondern die Anwendung der geeigneten mikrochirurgischen Technik zu sein.

Bezüglich des Vergleichs von SUM und DSI korrelierten beide Maße sowie deren prä- und posttherapeutischen Differenzen hochsignifikant miteinander. Somit können SUM und DSI als vergleichbar geeignete Parameter angesehen werden, die eng miteinander in Beziehung stehen. Eine Dysphonie reduziert die Stimmkapazität, was mit beiden Parametern gemessen werden kann. Während der DSI darauf abzielt, die Schwere der Dysphonie als negatives Kriterium zu beschreiben, spiegelt das SUM die stimmlichen Fähigkeiten wider und ermöglicht eine Klassifizierung der stimmlichen Leistungsfähigkeit als positives Kriterium. Vorteile des SUM gegenüber dem DSI bestehen z.B. in der Unabhängigkeit von der Jitter-Messung sowie Mikrophon-abstandsbedingt abweichenden Schallpegeln. Wie bereits in anderen Studien beschrieben, fanden auch wir keine Korrelation zwischen SUM bzw. DSI mit dem VHI [45,46].

Wir schlussfolgern aus unseren Ergebnissen, dass die mikrolaryngoskopische Abtragung von Stimmlippenpolypen bei korrekter Diagnosestellung, fachgemäßem phonomikrochirurgischem Vorgehen, sowie bei regelrechtem postoperativem Verlauf eine sichere, subjektiv und objektiv höchst zufriedenstellende Therapie zur Stimmverbesserung darstellt. Die Parameter SUM und VHI-9i erscheinen überaus geeignet, die stimmliche Beeinträchtigung bzw. die stimmliche Leistungsfähigkeit zu quantifizieren und weisen eine Reihe von Vorteilen gegenüber den bisher etablierten Verfahren auf.

Unsere dritte Studie bestätigte, dass die Parameter VHI-12, RBH, SUM und DSI zur Therapieverlaufskontrolle bei Patienten mit Stimmstörungen verschiedenster Genese geeignet sind. Die untersuchten objektiven und subjektiven Parameter verbesserten sich aufgrund ihrer individuellen Konstruktion und Intention jedoch unterschiedlich. Wie angenommen zeigte sich im etablierten DSI, dass die meisten unserer Patienten DSI-Werte zwischen -5 und +5 aufwiesen, wobei -5 einer hochgradig dysphonen und +5 einer perzeptiv normalen Stimme entsprach. Aufgrund der besonderen Struktur unserer Patientenkohorte, die einerseits eine beträchtliche Anzahl hochausgebildeter Profisänger und andererseits Probanden mit extrem schlechten Stimmen einschloss, hatten mehr Studienteilnehmer als erwartet (34%) anfängliche DSI-Werte, die die o.g. Grenzen an beiden Enden überschritten. Nach der Therapie verbesserten sich die mittleren DSI-Messdaten zwar nicht um durchschnittlich einen Schweregrad [47], zeigten aber einen signifikanten Anstieg. Dies bestätigt frühere Studien, die den DSI als nützlichen Parameter zur Messung des Schweregrads der Dysphonie sowie der Stimmverbesserung nach Therapie beschreiben [45, 48-50]. Andere Studien demonstrierten, dass der DSI durch Messunterschiede der Registrierungsprogramme sowie durch Alter oder Geschlecht beeinflusst wird [51-54]. Aus diesem Grund haben wir das SUM als neuen objektiven Stimmparameter entwickelt und untersucht, der durch diese Faktoren nicht beeinträchtigt wird. Diesbezüglich lagen die SUM-Werte initial meist zwischen 0 und 120. Erwartungsgemäß wurden diese Grenzwerte überschritten: (1.) am oberen Ende bei funktionell gestörten Profisängern mit „großen“ Stimmen und sehr großem SUP, (2.) am unteren Ende bei hochgradig dysphonen Stimmen mit sehr kleinem SUP. Unsere Ergebnisse bestätigten, dass ein größeres und „glatteres“ SUP ohne relevante Intensitätssprünge zu höheren Werten führte. Die auf der Basis von SUP-Fläche und -Umfang quantifizierte Stimmkapazität zeigte im posttherapeutischen Verlauf einen sehr deutlichen Anstieg. Dies steht im Einklang mit den Ergebnissen phonochirurgischer Studien bei Patienten mit Reinke-Ödemen [55], VFP [56] und Stimmlippenknötchen [57], die auch jeweils 3 Monate nach der operativen Befundabtragung signifikant größere SUM-Werte aufwiesen.

Bezüglich der subjektiven Parameter quantifizierte der VHI-12 das selbst erlebte Ausmaß des Stimmproblems erfolgreich. Durch die Therapie kam es bei den meisten Patienten zu einer Verbesserung von mäßiger hin zu leichtgradiger Beeinträchtigung. Diese Ergebnisse stimmen mit anderen Studien überein, in denen VHI-Fragebögen zur Untersuchung des chirurgischen und konservativen Behandlungserfolgs bei organischer und funktioneller Dysphonie verwendet wurden [55-60]. Vergleichbar zu anderen Studien wurde die auditiv-perzeptive Stimmbeurteilung als Hauptindikator zur Beurteilung des Therapieerfolgs herangezogen [61-64]. Der mittlere RBH-Status der Stimmen unserer Patienten zeigte sich deutlich verbessert und bestätigte die Ergebnisse früherer Untersuchungen [56-59]. Änderungen im Heiserkeitsgrad H wurden mit dem SUM am besten erkannt. Somit wird das SUM zusätzlich zur Quantifizierung der Stimmkapazität durch die auditiven Befunde gestützt und validiert, und bietet zudem einen intervallskalierten Parameter für die Dokumentation.

Wir schlussfolgern, dass SUM, DSI, RBH sowie VHI (in der originalen Lang- und beiden Kurzversionen) offensichtlich verschiedene Aspekte der Stimmfunktion repräsentieren und als sich einander ergänzende, objektive oder subjektive Maße entweder für die Bewertung der Stimmqualität, der Stimmleistung oder der wahrgenommenen Stimmeinschränkung betrachtet werden können. Die Parameter SUM und beide untersuchten VHI-Kurzformen quantifizieren verlässlich die objektive stimmliche Leistungsfähigkeit bzw. den subjektiven Leidensdruck; ihre breite Einführung in die klinische Stimmfunktionsdiagnostik erscheint gerechtfertigt und wünschenswert.

## 5. Literaturverzeichnis

1. Scholz W. Die Bedeutung Hermann Gutzmanns in seiner Zeit und heute. *Folia Phoniatr* 1956; 8:58-62.
2. Wendler J, Seidner W, Eysholdt U. *Lehrbuch der Phoniatrie und Pädaudiologie*. Stuttgart: Thieme; 2014.
3. Bundesärztekammer. *Musterweiterbildungsordnung 2018: Gebiet Phoniatrie und Pädaudiologie*. [https://www.bundesaerztekammer.de/fileadmin/user\\_upload/downloads/pdf-Ordner/Weiterbildung/MWBO-2018.pdf](https://www.bundesaerztekammer.de/fileadmin/user_upload/downloads/pdf-Ordner/Weiterbildung/MWBO-2018.pdf) (abgerufen am 11. August 2019), S. 239-244.
4. Korzilius H. (Muster-)Weiterbildungsordnung: Novelle einstimmig beschlossen. *Dtsch Arztebl* 2018; 115:A970-A973.
5. DGPP 2019. Website der Deutschen Gesellschaft für Phoniatrie und Pädaudiologie. <http://www.dgpp.de/cms/pages/de/historie.php> (abgerufen am 11. August 2019).
6. Ruben RJ. Redefining the survival of the fittest: communication disorders in the 21st century. *Laryngoscope* 2000; 110:241-245.
7. Dejonckere PH, Bradley P, Clemente P, Cornut G, Crevier-Buchman L, Friedrich G, Van De Heyning P, Remacle M, Woisard V; Committee on Phoniatics of the European Laryngological Society (ELS). 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). *Eur Arch Otorhinolaryngol* 2001; 258:77-82.
8. Friedrich G, Dejonckere PH. Das Stimmdiagnostik-Protokoll der European Laryngological Society (ELS) – erste Erfahrungen im Rahmen einer Multizenterstudie. *LaryngoRhinoOtol* 2005; 84:744-752.
9. Dejonckere PH. Assessment of Voice and Respiratory Function. In: Remacle M, Eckel HE (Hrsg). *Surgery of Larynx and Trachea*. Berlin: Springer; 2010.
10. Hanschmann H, Lohmann A, Berger R. Comparison of subjective assessment of voice disorders and objective voice measurement. *Folia Phoniatr Logop* 2011; 63: 83–87.
11. Schneider-Stickler B, Bigenzahn W. *Stimmdiagnostik*. Wien: Springer; 2013.
12. Horiuchi H, Tamai N, Kamba S, Inomata H, Ohya TR, Sumiyama K. Real-time computer-aided diagnosis of diminutive rectosigmoid polyps using an auto-fluorescence imaging system and novel color intensity analysis software. *Scand J Gastroenterol* 2019; 54:800-805.
13. Bochenek K, Aebisher D, Międzybrodzka A, Cieślak G, Kawczyk-Krupka A. Methods for bladder cancer diagnosis - The role of autofluorescence and photodynamic diagnosis. *Photodiagnosis Photodyn Ther* 2019; 27:141-148.
14. Lee YN, Moon JH, Choi HJ. Role of Image-Enhanced Endoscopy in Pancreatobiliary Diseases. *Clin*

- Endosc 2018; 51:541-546.
15. Andreev VY, Yanev NA, Stanimirov SK, Ivanov IN, Popovska SL, Hristova PA, Ivanov YY. Autofluorescence and White Light Bronchoscopy in the Diagnosis of Endobronchial Malignant Lesions. *Folia Med (Plovdiv)*. 2018; 60:439-446.
  16. Sieroń-Stołtny K, Kwiatek S, Latos W, Kawczyk-Krupka A, Cieślar G, Stanek A, Ziaja D, Bugaj AM, Sieroń A. Autofluorescence endoscopy with "real-time" digital image processing in differential diagnostics of selected benign and malignant lesions in the oesophagus. *Photodiagnosis Photodyn Ther* 2012; 9:5-10.
  17. Cho JH. Advanced Imaging Technology Other than Narrow Band Imaging. *Clin Endosc*. 2015; 48:503-510.
  18. Kraft M, Arens C, Betz C, Fostiropoulos K. Fluorescence imaging in laryngology: Physical principles, clinical applications and study results. *HNO* 2016; 64:4-12.
  19. Arens C, Glanz H, Wonckhaus J, Hersemeyer K, Kraft M. Histologic assessment of epithelial thickness in early laryngeal cancer or precursor lesions and its impact on endoscopic imaging. *Eur Arch Otorhinolaryngol* 2007; 264:645-649.
  20. Uppal A, Gupta PK. Measurement of NADH concentration in normal and malignant human tissues from breast and oral cavity. *Biotechnol Appl Biochem* 2003; 37(pt 1):45-50.
  21. Stepp H, Sroka R, Baumgartner R. Fluorescence endoscopy of gastrointestinal diseases: basic principles, techniques, and clinical experience. *Endoscopy* 1998; 30:379-386.
  22. Nawka T, Verdonck-de Leeuw IM, De Bodt M, Guimaraes I, Holmberg EB, Rosen CA, Schindler A, Woisard V, Whurr R, Konerding U. Item reduction of the voice handicap index based on the original version and on European translations. *Folia Phoniater Logop* 2009; 61:37-48.
  23. Pickhard A, Reiter R. Benign vocal fold lesions. *Laryngorhinootologie* 2013; 92:304-312.
  24. Dikkers FG, Nikkels PG. Lamina propria of the mucosa of benign lesions of the vocal folds. *Laryngoscope* 1999; 109:1684-1689.
  25. Petrović-Lazić M, Babac S, Vuković M, Kosanović R, Ivanković Z. Acoustic voice analysis of patients with vocal fold polyp. *J Voice* 2011; 25:94-97.
  26. Martins RH, Defaveri J, Domingues MA, de Albuquerque e Silva R. Vocal polyps: clinical, morphological, and immunohistochemical aspects. *J Voice*. 2011; 25:98-106.
  27. Nawka T, Martin A, Caffier PP. Microlaryngoscopy and phonomicrosurgery. *HNO*. 2013; 61:108-116.
  28. Wang CT, Liao LJ, Huang TW, Lo WC, Cheng PW. Comparison of treatment outcomes of transnasal vocal fold polypectomy versus microlaryngoscopic surgery. *Laryngoscope* 2015; 125:1155-1160.
  29. Karasu MF, Gundogdu R, Cagli S, Aydin M, Arli T, Aydemir S, Yuce I. Comparison of effects on voice of diode laser and cold knife microlaryngology techniques for vocal fold polyps. *J Voice* 2014; 8:387-392.
  30. Uloza V, Saferis V, Uloziene I. Perceptual and acoustic assessment of voice pathology and the

- efficacy of endolaryngeal phonosurgery. *J Voice* 2005; 19:138-145.
31. Voigt-Zimmermann S, Lampe K, Arens C. Differential diagnosis of hoarseness. *Laryngorhinootologie* 2014; 93:263-284.
  32. Van Houtte E, Van Lierde K, Claeys S. Pathophysiology and treatment of muscle tension dysphonia: a review of the current knowledge. *J Voice* 2011; 25:202-207.
  33. Hamdan AL, Khalifee E, Ghanem A, Mansour H, Yammine E. Predictive value of globus pharyngeus in patients with functional dysphonia versus organic dysphonia. *Laryngoscope* 2019; 129:930-934.
  34. Reiter R, Hoffmann TK, Pickhard A, Brosch S. Hoarseness-causes and treatments. *Dtsch Arztebl Int* 2015; 112:329-337.
  35. Schneider B, Wendler J, Seidner W. The relevance of stroboscopy in functional dysphonias. *Folia Phoniatr Logop* 2002; 54:44-54.
  36. Arens C, Reussner D, Woenckhaus J, Leunig A, Betz CS, Glanz H. Indirect fluorescence laryngoscopy in the diagnosis of precancerous and cancerous laryngeal lesions. *Eur Arch Otorhinolaryngol* 2007; 264(6):621-626.
  37. Kriegmair MC, Honeck P, Theuring M, Bolenz C, Ritter M. Wide-field autofluorescence-guided TUR-B for the detection of bladder cancer: a pilot study. *World J Urol* 2018; 36(5):745-751.
  38. Baletic N, Malicevic H, Petrovic Z, Marinkovic-Eric J, Peric A. Advantages and limitations of the autofluorescent diagnostics of the laryngeal cancer and precancerosis. *Eur Arch Otorhinolaryngol* 2010; 267:925-931.
  39. Mostafa BE, Shafik AG, Fawaz S. The role of flexible autofluorescence laryngoscopy in the diagnosis of malignant lesions of the larynx. *Acta Otolaryngol* 2007; 127:175-179.
  40. Kraft M, Betz CS, Leunig A, Arens C. Value of fluorescence endoscopy for the early diagnosis of laryngeal cancer and its precursor lesions. *Head Neck* 2011; 33:941-948.
  41. Arens C, Glanz H, Woenckhaus J, Hersemeyer K, Kraft M. Histologic assessment of epithelial thickness in early laryngeal cancer or precursor lesions and its impact on endoscopic imaging. *Eur Arch Otorhinolaryngol* 2007; 264:645-649.
  42. Zeitels SM, Hillman RE, Desloge R, Mauri M, Doyle PB. Phonosurgery in singers and performing artists: treatment outcomes, management theories, and future directions. *Ann Otol Rhinol Laryngol Suppl* 2002; 190:21-40.
  43. Rosen CA, Murry T, Zinn A, Zullo T, Sonbolian M. Voice handicap index change following treatment of voice disorders. *J Voice* 2000; 14(4):619-623.
  44. Cho KJ, Nam IC, Hwang YS, Shim MR, Park JO, Cho JH, Joo YH, Kim MS, Sun DI. Analysis of factors influencing voice quality and therapeutic approaches in vocal polyp patients. *Eur Arch Otorhinolaryngol* 2011; 268(9):1321-1327.
  45. Hakkesteeft MM, Brocaar MP, Wieringa MH. The applicability of the dysphonia severity index and the voice handicap index in evaluating effects of voice therapy and phonosurgery. *J Voice* 2010; 24(2):199-205.

46. Freymann ML, Mathmann P, Rummich J, Müller C, Neumann K, Nawka T, Caffier PP. Gender-specific reference ranges of the vocal extent measure in young and healthy adults. *Logoped Phoniatr Vocol* 2019; Jun 3:1-9. doi: 10.1080/14015439.2019.1617894. [Epub ahead of print]
47. Gonnermann U. *Quantifizierbare Verfahren zur Bewertung von Dysphonien*. Peter Lang, Frankfurt/Main; 2007.
48. Nemr K, Simões-Zenari M, de Souza GS, Hachiya A, Tsuji DH. Correlation of the Dysphonia Severity Index (DSI), Consensus Auditory-Perceptual Evaluation of Voice (CAPE-V), and Gender in Brazilians With and Without Voice Disorders. *J Voice* 2016; 30(6):765.e7-765.e11.
49. Uloza V, Latoszek BBV, Ulozaitė-Staniene N, Petrauskas T, Maryn Y. A comparison of Dysphonia Severity Index and Acoustic Voice Quality Index measures in differentiating normal and dysphonic voices. *Eur Arch Otorhinolaryngol* 2018; 275(4):949-958.
50. Brockmann-Bauser M, Balandat B, Bohlender JE. Immediate Lip Trill Effects on the Standard Diagnostic Measures Voice Range Profile, Jitter, Maximum Phonation Time, and Dysphonia Severity Index. *J Voice* 2019 Jun 10. pii: S0892-1997(18)30573-3. doi: 10.1016/j.jvoice.2019.04.011. [Epub ahead of print]
51. Caffier PP, Möller A, Forbes E, Müller C, Freymann ML, Nawka T. The Vocal Extent Measure: Development of a Novel Parameter in Voice Diagnostics and Initial Clinical Experience. *Biomed Res Int*. 2018 Mar 4;2018:3836714.
52. Hakkesteeft MM, Brocaar MP, Wieringa MH, Feenstra L. Influence of age and gender on the dysphonia severity index. A study of normative values. *Folia Phoniatr Logop* 2006; 58:264-273.
53. Aichinger P, Feichter F, Aichstill B, Bigenzahn W, Schneider-Stickler B. Inter-device reliability of DSI measurement. *Logoped Phoniatr Vocol* 2012; 37(4):167-173.
54. Barsties V, Latoszek B, Ulozaitė-Staniene N, Maryn Y, Petrauskas T, Uloza V. The Influence of Gender and Age on the Acoustic Voice Quality Index and Dysphonia Severity Index: A Normative Study. *J Voice* 2019; 33(3):340-345.
55. Salmen T, Ermakova T, Schindler A, Ko SR, Göktas Ö, Gross M, Nawka T, Caffier PP. Efficacy of microsurgery in Reinke's oedema evaluated by traditional voice assessment integrated with the Vocal Extent Measure (VEM). *Acta Otorhinolaryngol Ital* 2018; 38:194-203.
56. Salmen T, Ermakova T, Möller A, Seipelt M, Weikert S, Rummich J, Gross M, Nawka T, Caffier PP. The Value of Vocal Extent Measure (VEM) Assessing Phonomicrosurgical Outcomes in Vocal Fold Polyps. *J Voice* 2017; 31:114.e7-114.e15.
57. Caffier PP, Salmen T, Ermakova T, Forbes E, Ko SR, Song W, Gross M, Nawka T. Phonomicrosurgery in Vocal Fold Nodules: Quantification of Outcomes in Professional and Non-Professional Voice Users. *Med Probl Perform Art* 2017; 32(4):187-194.
58. Ropero Rendón MDM, Ermakova T, Freymann ML, Ruschin A, Nawka T, Caffier PP. Efficacy of Phonosurgery, Logopedic Voice Treatment and Vocal Pedagogy in Common Voice Problems of Singers. *Adv Ther* 2018; 35(7):1069-1086.

59. Reetz S, Bohlender JE, Brockmann-Bauser M. Do Standard Instrumental Acoustic, Perceptual, and Subjective Voice Outcomes Indicate Therapy Success in Patients With Functional Dysphonia? *J Voice* 2019; 33(3):317-324.
60. Yılmaz T. Surgical treatment of glottic web using butterfly mucosal flap technique: Experience on 12 patients. *Laryngoscope* 2019; 129(6):1423-1427.
61. Lee Y, Kim G, Kwon S. The Usefulness of Auditory Perceptual Assessment and Acoustic Analysis for Classifying the Voice Severity. *J Voice* 2019 May 17. pii: S0892-1997(19)30087-6. doi: 10.1016/j.jvoice.2019.04.013. [Epub ahead of print]
62. Kempster GB, Gerratt BR, Verdolini Abbott K, Barkmeier-Kraemer J, Hillman RE. Consensus auditory-perceptual evaluation of voice: development of a standardized clinical protocol. *Am J Speech Lang Pathol* 2009; 18:124-132.
63. Schönweiler R, Wübbelt P, Hess M, Ptok M. Psychoacoustic scaling of acoustic voice parameters by multicenter voice ratings. *Laryngorhinotologie* 2001; 80:117-122.
64. Ptok M, Schwemmler C, Iven C, Jessen M, Nawka T. On the auditory evaluation of voice quality. *HNO*. 2006; 54:793-802.

## **Eidesstattliche Versicherung**

„Ich, Matthias Seipelt, versichere an Eides statt durch meine eigenhändige Unterschrift, dass ich die vorgelegte Dissertation mit dem Thema: „Autofluoreszenz-Laryngoskopie, Stimmumfangsmaß, Voice Handicap Index Kurzformen: Zur Wertigkeit neuer diagnostischer Prozeduren und Parameter in der Stimmheilkunde“ selbstständig und ohne nicht offengelegte Hilfe Dritter verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel genutzt habe.

Alle Stellen, die wörtlich oder dem Sinne nach auf Publikationen oder Vorträgen anderer Autoren beruhen, sind als solche in korrekter Zitierung (siehe „Uniform Requirements for Manuscripts (URM)“ des ICMJE -[www.icmje.org](http://www.icmje.org)) kenntlich gemacht. Die Abschnitte zu Methodik (insbesondere praktische Arbeiten, Laborbestimmungen, statistische Aufarbeitung) und Resultaten (insbesondere Abbildungen, Graphiken und Tabellen) entsprechen den URM (s.o.) und werden von mir verantwortet.

Meine Anteile an den ausgewählten Publikationen entsprechen denen, die in der untenstehenden gemeinsamen Erklärung mit dem/der Betreuer/in, angegeben sind. Sämtliche Publikationen, die aus dieser Dissertation hervorgegangen sind und bei denen ich Autor bin, entsprechen den URM (s.o.) und werden von mir verantwortet.

Die Bedeutung dieser eidesstattlichen Versicherung und die strafrechtlichen Folgen einer unwahren eidesstattlichen Versicherung (§156,161 des Strafgesetzbuches) sind mir bekannt und bewusst.“

Berlin, 21.02.2020

---

Unterschrift

## Anteilerklärung an den erfolgten Publikationen

Matthias Seipelt hatte folgenden Anteil an den eingereichten Publikationen:

Publikation 1: Caffier PP, Schmidt B, Gross M, Karnetzky K, Nawka T, Rotter A, **Seipelt M**, Sedlmaier B. A comparison of white light laryngostroboscopy versus autofluorescence endoscopy in the evaluation of vocal fold pathology. *Laryngoscope* 2013; 123(7):1729-1734.

Beitrag im Einzelnen: Patientenrekrutierung und -betreuung im Studienverlauf, Beteiligung an den Literaturrecherchen, Hauptanteil (neben Caffier PP) bei der Datenakquisition, Mitwirkung an der statistischen Auswertung der Primärdaten, Mitwirkung bei der Erstellung der Tabellen 1 und 2, Mitarbeit bei der Verfassung des Manuskripts.

Publikation 2: Salmen T, Ermakova T, Möller A, **Seipelt M**, Weikert S, Rummich J, Gross M, Nawka T, Caffier PP. The Value of Vocal Extent Measure (VEM) Assessing Phonomicrosurgical Outcomes in Vocal Fold Polyps. *J Voice* 2017; 31(1):114.e7-114.e15.

Beitrag im Einzelnen: Patientenrekrutierung und -betreuung im Studienverlauf, Beteiligung an der Durchführung der Untersuchungen; Mitwirkung bei der Erstellung der Tabelle 1, Mitarbeit bei der Verfassung des Manuskripts.

Publikation 3: **Seipelt M**, Möller A, Nawka T, Gonnermann U, Caffier F, Caffier PP. Monitoring the outcome of phonosurgery and vocal exercises with established and new diagnostic tools. *Biomed Res Int* 2020 Jan 23; 2020:4208189.

Beitrag im Einzelnen: Patientenrekrutierung und -betreuung im Studienverlauf, Durchführung von Literaturrecherchen (Datenbanken Medline, Embase, Cochrane Library, Web of Science), Hauptanteil bei der Datenakquisition, Erstellung der Tabellen 1 bis 3, Hauptanteil (neben Caffier PP) bei der Verfassung des Manuskripts.

---

Unterschrift, Datum und Stempel des betreuenden Hochschullehrers

---

Unterschrift des Doktoranden

## **Publikation 1**

Caffier PP, Schmidt B, Gross M, Karnetzky K, Nawka T, Rotter A, **Seipelt M**, Sedlmaier B. A comparison of white light laryngostroboscopy versus autofluorescence endoscopy in the evaluation of vocal fold pathology. *Laryngoscope* 2013; 123(7):1729-1734.

<https://doi.org/10.1002/lary.23931>













## **Publikation 2**

Salmen T, Ermakova T, Möller A, **Seipelt M**, Weikert S, Rummich J, Gross M, Nawka T, Caffier PP. The Value of Vocal Extent Measure (VEM) Assessing Phonomicrosurgical Outcomes in Vocal Fold Polyps. *J Voice* 2017; 31(1):114.e7-114.e15.

<https://doi.org/10.1016/j.jvoice.2016.03.016>



















### **Publikation 3**

**Seipelt M**, Möller A, Nawka T, Gonnermann U, Caffier F, Caffier PP. Monitoring the outcome of phonosurgery and vocal exercises with established and new diagnostic tools. *Biomed Res Int* 2020 Jan 23;2020:4208189.

<https://doi.org/10.1155/2020/4208189>

## Clinical Study

# Monitoring the Outcome of Phonosurgery and Vocal Exercises with Established and New Diagnostic Tools

Matthias Seipelt,<sup>1</sup> Andreas Möller,<sup>2</sup> Tadeus Nawka,<sup>1</sup> Ute Gonnermann,<sup>3</sup> Felix Caffier,<sup>1</sup> and Philipp P. Caffier<sup>1</sup> 

<sup>1</sup>Department of Audiology and Phoniatrics, Charité-University Medicine Berlin, Campus Charité Mitte, Chariteplatz 1, D-10117 Berlin, Germany

<sup>2</sup>Max-Planck Institute for Plasma Physics, Wendelsteinstraße 1, D-17491 Greifswald, Germany

<sup>3</sup>ENT Department, University of Greifswald, Fleischmannstraße 8, D-17475 Greifswald, Germany

Correspondence should be addressed to Philipp P. Caffier; philipp.caffier@charite.de

Received 14 August 2019; Accepted 20 December 2019; Published 23 January 2020

Academic Editor: Jan Plzak

Copyright © 2020 Matthias Seipelt et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Instrument-assisted measuring procedures expand the options within phoniatric diagnostics by quantifying the condition of the voice. The aim of this study was to examine objective treatment-associated changes of the recently developed vocal extent measure (VEM) and the established dysphonia severity index (DSI) in relation to subjective tools, i.e., self-evaluation via voice handicap index (VHI-12) and external evaluation via auditory-perceptual assessment of hoarseness (*H*). The findings for *H* (3 raters' group assessment), VHI-12, DSI, and VEM in 152 patients of both sexes (age range 16–75 years), taken before and 3 months after phonosurgery or vocal exercises, were compared and correlated. Posttherapeutically, all of the recorded parameters improved ( $p < 0.001$ ). The degree of *H* reduced on average by 0.5, the VHI-12 score sank by 5 points, while DSI and VEM rose by 1.5 and 19, respectively. The correlations of these changes were significant but showed gradual differences between *H* and VHI-12 ( $r = 0.3$ ), *H* and DSI ( $r = -0.3$ ), and *H* and VEM ( $r = -0.4$ ). We conclude that all investigated parameters are adequate to verify therapeutic outcomes but represent different dimensions of the voice. However, changes in the degree of *H* as gold standard were best recognized with the new VEM.

## 1. Introduction

The human voice is a very complex phenomenon that is difficult to quantify [1–3]. According to the basic protocol of the European Laryngological Society, a comprehensive assessment of vocal function can be gained using a multidimensional diagnostic approach [4]. Several measurements are recommended for voice evaluation, comprising subjective procedures such as self-assessment of the voice and external auditory-perceptual judgment, as well as objective procedures such as voice range profile (VRP) measurements, acoustic-aerodynamic analysis, and video-laryngostroboscopy (VLS).

In order to quantify the self-experienced extent of a vocal problem, the subjective impairment can be assessed using

standardized questionnaires [5, 6]. The original voice handicap index (VHI) consists of 30 questions addressing functional, physical, or emotional aspects in the context of dysphonia [7]. Shorter VHI versions were designed because many patients perceive the answering of 30 questions tedious and partly redundant [8]. The 9-item VHI-9i and the 12-item VHI-12 had been created after original item reduction based on factor analysis and test-retest validation. They represent reliable, commonly used questionnaires with improved acceptance and practicability in clinical routine [9]. Regardless of the subjective self-evaluation, the examiner's auditory perception of the patient's voice is considered in many medical studies as the gold standard for voice assessment [10–12]. Different evaluation systems were developed, assessing various parameters including the

Common pillars in voice diagnostics:  
 DSI: dysphonia severity index  
 MPT: maximum phonation time  
 RBH: roughness (*R*), breathiness (*B*), and (overall grade of) hoarseness (*H*)  
 VEM: vocal extent measure  
 VHI: voice handicap index  
 VLS: videolaryngostroboscopy  
 VRP: voice range profile

FIGURE 1: List of common abbreviations in voice diagnostics.

perceived roughness (*R*), breathiness (*B*), and the overall grade of hoarseness (*H*). The application of the RBH scale is considered to be reliable, particularly when group assessments are used for further analysis [13–15].

The inclusion of instrument-assisted measurement procedures can support and usefully expand the diagnostic investigation by objectively quantifying the current condition of the voice [16, 17]. The established dysphonia severity index (DSI) is calculated as a weighted combination of the highest possible fundamental frequency ( $F0_{max}$ ), the lowest phonation intensity ( $I_{min}$ ), maximum phonation time (MPT), and jitter [18]. Since the DSI quantifies dysphonia as a negative criterion and involves the risk of inaccurate results due to its multidimensional acquisition, we recently developed the one-dimensional vocal extent measure (VEM) for objective VRP evaluation [19]. The VEM quantifies the subject's dynamic performance and frequency range and is calculated as a relation of area and perimeter of the VRP. The VEM describes the vocal abilities and enables a classification of voice performance as a positive criterion [20]. A list of common abbreviations in voice diagnostics is given in Figure 1.

For comprehensive documentation of vocal status and treatment, it is necessary to ensure that changes in voice quality can be adequately identified. For this purpose, the measurement data used must be sensitive to the slightest changes in voice quality, and the registration equipment must be able to detect them. In order to investigate the suitability of objective and subjective parameters for the assessment of vocal improvement after phonosurgery and vocal exercises, the changes in DSI and VEM values should be monitored and compared with those of the subjective auditory perception via RBH and VHI-12.

## 2. Materials and Methods

A total of 152 patients with various voice problems underwent therapy in a clinical prospective study. The trial was conducted in accordance with the Declaration of Helsinki and on approval by the local ethical review board. All data were collected at the pretherapeutic visit and three months after the intervention. According to the diagnosed clinical pathology and the previous course of the disease, the patients received either surgical treatment or conservative vocal exercises. Logopedic voice therapy was conducted by qualified speech therapists and included 20 sessions (2 times per week, for 45 minutes). Phonomicrosurgery was performed by 3 experienced senior phonosurgeons via direct microlaryngoscopy in general anesthesia (TIVA with propofol/remifentanyl).

Various established examination instruments were applied to evaluate the treatment outcome. Digital VLS was performed using a high-resolution rigid video laryngoscope (10 mm; 70°) with an integrated microphone (XION medical, Berlin, Germany) [21]. Laryngoscopy served to discriminate between organic dysphonia and functional dysphonia. Stroboscopy visualized the vocal fold vibrations during phonation and indicated impairment by showing reduced/absent mucosal wave propagation or reduced/eliminated phonatory vibration.

The LingWAVES program (WEVOSYS, Forchheim, Germany) was used for standardized registration of the VRP and acoustic-aerodynamic analysis. Several acoustic and aerodynamic parameters were recorded, such as  $I_{min}$ ,  $F0_{max}$ , MPT, and jitter. Based on the defined combination of these parameters, the DSI was calculated to classify the voice into nondysphonic ( $\geq 4.2$ ) versus mildly ( $< 4.2$  to  $\geq 1.8$ ), moderately ( $< 1.8$  to  $\geq -1.2$ ), or severely ( $< -1.2$ ) dysphonic [22]. In addition, the VEM as a recently introduced new diagnostic tool for the objective assessment of vocal capacity was computed [19, 23]. VEM calculation was done after VRP measurement by a proprietary software program (AVA) which can extract various other parameters from the VRP, thereby enabling VRP comparisons [24]. The VEM multiplies the VRP area by the quotient of the theoretical perimeter of a circle with the VRP surface area and the actual VRP circumference. The mathematical derivation of the equation of this measure is explained elsewhere [19]. The VEM quantifies the dynamic performance and frequency range of the voice by a one-dimensional, interval-scaled value without unit, typically between 0 and 120. These limits may be exceeded at both ends ( $VEM_{min} \geq -150$ ;  $VEM_{max} \leq 150$ ), describing a small vocal capacity by a small VEM and a large VRP by a high VEM.

The VHI-12 was applied for the patient's subjective self-assessment of the own voice [9]. Study participants rated all 12 questions on a scale from 0 to 4 (0: never, 1: almost never, 2: sometimes, 3: almost always, 4: always), followed by one question concerning the overall voice impairment at the present time (VHIs) on a scale from 0 to 3 (0: normal, 1: mild, 2: moderate, 3: severe). The VHI-12 total score allowed an impairment-related severity classification (0–7: no dysphonia, 8–14: mild dysphonia, 15–22: moderate dysphonia, 23–48: severe dysphonia).

External auditory-perceptual voice evaluation was assessed using the RBH system when the patients were reading the standardized text "The north wind and the sun" (German version). The perceived roughness (*R*), breathiness (*B*), and overall grade of hoarseness (*H*) of the patients' voices were scored on a scale from 0 to 3 (0: not existing, 1: mild, 2: moderate, 3: severe) by three experienced examiners (one phoniatric physician, one clinical linguist, and one biomedical engineer). To enhance the evaluation objectivity, all audio recordings were rated independently in one session after being shuffled and blinded regarding the patient assignment and pre-/posttherapeutic status. The degree of *H* served as a gold standard to provide an indication of the therapy success.

The outcome analysis was based on pre- and post-therapeutic voice function diagnostics and VLS. The

parameters  $H$ , VHI-12, VHIs, DSI, and VEM were compared with each other before and three months after the intervention, as well as their changes. It was tested whether the therapy resulted in a significant difference in the parameters measured. In addition, the measurement data were correlated with each other before and after therapy, as were the respective changes. Statistical methods applied were the calculation of Spearman's rank-order correlation coefficients ( $r$ ), as well as the paired  $t$ -test. The level of significance was set at  $\alpha = 0.05$ .

### 3. Results

Altogether, 152 patients were examined before and after therapeutic treatment: 102 females (17–70 years, median 48) and 50 males (16–75 years, median 42). A total of 304 data sets were collected. Subjects of both sexes were comparable in terms of age, sociodemographic characteristics, hoarseness level ( $H$ ), and underlying pathologies. Sixty-six individuals (43%) used their voice in a nonprofessional manner (e.g., clerks, IT-specialists, and laborers), whereas 86 patients (57%) had a high vocal strain in their profession (e.g., teachers, actors, and singers). Pretherapeutically, the patients exhibited various clinical disorders. VLS revealed in 101 subjects (66%) organic diseases at vocal fold level. Classification of the resulting organic dysphonia according to the underlying pathology showed in 41 patients (27%) diseases of the lamina propria (e.g., nodules, polyps, cysts, and edema), in 24 patients (16%) movement disorders (vocal fold paralysis, spasmodic dysphonia), in 19 patients (12%) diseases of the epithelium (e.g., leukoplakia, hyperkeratosis, carcinoma, and papillomatosis), and in 5 patients (3%) arytenoid pathologies (granuloma). Fifty-one participants (34%) had normal laryngeal anatomy but suffered from a vocal load-induced functional dysphonia. Altogether, 46 subjects (30%) had initially no hoarseness (H0), including 29 patients with functional dysphonia, 13 patients with small glottal findings (marginal edema, nodules, and leukoplakia), and 4 patients with pathologies distant from the vocal fold level (arytenoid granuloma). A summary of relevant pretherapeutic patient characteristics, pathology classification according to Rosen and Murry [25], and a listing of all diagnoses are shown in Table 1.

Posttherapeutically, all investigated vocal parameters had improved. Regarding subjective evaluation, the mean RBH status exposed less roughness, breathiness, and overall grade of hoarseness ( $p < 0.001$ ).  $H$  decreased from 1.2 to 0.7, changing in most patients from H1 to H0 ( $n = 29$ ; i.e., 19%) and from H2 to H1 ( $n = 22$ ; i.e., 14%). The VHI-12 reduced on average from 15 to 10, corresponding to a self-assessed improvement from moderately to mildly impaired ( $p < 0.001$ ). Respectively, the overall VHIs score sank from 1.4 to 0.7, changing most often from mild voice impairment to normal ( $n = 44$ ; i.e., 29%) and from moderate to mild voice impairment ( $n = 35$ ; i.e., 23%). Figure 2 summarizes the mean pre- and posttherapeutic data in all patients for the investigated subjective vocal parameters using column diagrams.

A comparison of objective parameters revealed for the DSI a mean increase from 2.2 to 3.7, showing significant

improvement ( $p < 0.001$ ) that remained at the level of mild dysphonia. The VEM rose from 60 to 79, reflecting voice improvement with significantly enhanced vocal capacity ( $p < 0.001$ ). Figure 3 illustrates the pre- and posttherapeutic data for both objective voice parameters using boxplots. It indicates additionally the different distribution of DSI and VEM in relation to the degree of  $H$ .

A comparison of treatment groups revealed that phonosurgery had the largest impact on voice function with higher numerical improvement of subjective and objective parameters. A patient example demonstrating phonosurgery-induced changes of laryngeal and vocal findings is presented in Figure 4.

To evaluate the extent of treatment-related benefits, Table 2 shows the mean differences between pre- and posttherapeutic values and the 95% confidence intervals for them. The numeric outcome of the values after conservative logopedic therapy was much smaller, but the vocal capabilities improved in most patients, too. Furthermore, Table 2 displays the pre- and posttherapeutic comparison concerning both dysphonia groups (functional/organic) and all pathology classification subgroups. In general, age and gender had no significant influence on the treatment outcome.

The correlation of the parameters  $H$ , VHI-12, VHIs, DSI, and VEM with each other showed a significant ( $p < 0.01$ ) but weak to moderate linear relationship. The strength of the relationship changed only slightly due to the therapy. However, the DSI proved an exception in this regard. While the DSI data before treatment showed a moderate negative relationship with  $H$  ( $r = -0.4$ ) and moderate positive relationship with VEM ( $r = 0.6$ ), these correlations decreased considerably after therapy, revealing weaker relationships for  $H$  ( $r = -0.3$ ) and VEM ( $r = 0.3$ ). The weak relationship between DSI and VHI-12 as well as DSI and VHIs did not show relevant changes posttherapeutically. In contrast, the VEM correlated with the VHI-12 at  $r = -0.4$  and with  $H$  at  $r = -0.7$ , revealing moderate and strong negative relationships, irrespective of the therapy status. Furthermore,  $H$  and VHI-12 correlated after therapy at  $r = 0.4$ , and  $H$  and VHIs at  $r = 0.5$ .

The investigation of therapy-induced changes ( $\Delta$ ) in the individual measurement data indicated that correlations of these changes resulted in rather small coefficients for all parameters. The relationship between  $\Delta$ DSI and  $\Delta$  $H$  was  $r = -0.3$  ( $p < 0.01$ ). The  $\Delta$ DSI showed no significant relationship to  $\Delta$ VHI-12 ( $r = -0.04$ ) and  $\Delta$ VHIs ( $r = -0.09$ ). The relationship between  $\Delta$ VEM and  $\Delta$  $H$  was  $r = -0.4$  ( $p < 0.001$ ). In contrast to DSI,  $\Delta$ VEM revealed also a significant relationship to  $\Delta$ VHI-12 and  $\Delta$ VHIs ( $r = -0.2$  each,  $p < 0.01$ ). Besides,  $\Delta$  $H$  and  $\Delta$ VHI-12 correlated at  $r = 0.3$ , and  $\Delta$  $H$  and  $\Delta$ VHIs at  $r = 0.4$  ( $p < 0.01$ ). Finally, the relationship between  $\Delta$ DSI and  $\Delta$ VEM was moderate at  $r = 0.5$  ( $p < 0.01$ ). A summary of all correlation results can be found in Table 3.

### 4. Discussion

It was possible to show that all parameters under investigation reacted to the therapy and improved on average, thus

TABLE 1: Pretherapeutic patient characteristics.

Characteristics	No. of all patients	% of total group (n = 152)	No. of male patients	% of male group (n = 50)	No. of female patients	% of female group (n = 102)
<i>Gender</i>						
Male	50	33	—	—	—	—
Female	102	67	—	—	—	—
<i>Age</i>						
Years (mean ± SD)	45 ± 16	—	43 ± 17	—	45 ± 15	—
<i>Main voice use</i>						
Nonprofessional	66	43	22	44	44	43
Professional	86	57	28	56	58	57
<i>Sociodemographic</i>						
Scholar	10	7	4	8	6	6
Student/apprentice	14	9	7	14	7	7
Employed	97	64	29	58	68	66
Unemployed	8	5	2	4	6	6
Pensioner	23	15	8	16	15	15
<i>Overall hoarseness level (H)</i>						
H0 (not existing)	46	30	12	24	34	33
H1 (mild)	6	40	21	42	39	38
H2 (moderate)	27	18	10	20	17	17
H3 (severe)	19	12	7	14	12	12
<i>Pathology classification*</i>						
Functional dysphonia	51	34	15	30	36	35
Organic dysphonia	101	66	35	70	66	65
Rosen I (epithelium)	19	12	10	20	9	9
Rosen II (lamina propria)	41	27	7	14	34	33
Rosen III (arytenoid)	5	3	3	6	2	2
Rosen IV (other)	36	24	15	30	21	21
<i>Organic diagnosis</i>						
Vocal fold paralysis	18	11.8	9	18.0	9	8.8
Vocal fold nodules	13	8.6	0	—	13	12.7
Vocal fold polyp	9	5.9	4	8.0	5	4.9
Reinke's edema	9	5.9	0	—	9	8.8
Laryngeal papillomatosis	8	5.3	3	6.0	5	4.9
Marginal edema	6	3.9	2	4.0	4	3.9
Spasmodic dysphonia	6	3.9	0	—	6	5.9
Contact granuloma	5	3.2	3	6.0	2	2.0
Vocal fold atrophy	5	3.2	4	8.0	1	1.0
Hyperkeratosis	4	2.6	3	6.0	1	1.0
Leukoplakia	4	2.6	3	6.0	1	1.0
Sulcus vocalis	3	2.0	2	4.0	1	1.0
Glottal carcinoma (pT1a)	3	2.0	1	2.0	2	2.0
Vocal fold cyst	3	2.0	1	2.0	2	2.0
Varix cordis	1	0.7	0	—	1	1.0
Laryngotracheal stenosis	1	0.7	0	—	1	1.0
Glottal web	1	0.7	0	—	1	1.0
Traumatic laryngeal fracture	1	0.7	0	—	1	1.0
Bamboo nodes	1	0.7	0	—	1	1.0

Unless otherwise specified, data expressed as number of patients and percentage of group. \* Pathology classification according to Rosen and Murry [25], i.e., I: epithelium, e.g., leukoplakia, hyperkeratosis, CIS (=carcinoma in situ), carcinoma, and papillomatosis. II: lamina propria, e.g., Reinke's edema, polyps, cysts, scars, and vascular malformation. III: arytenoid, e.g., granuloma and infection. IV: other, including movement disorders, hypo-/atrophy, and malformation as e.g., sulcus or glottal web.

presenting their general suitability for documentation of the therapeutic process. However, due to the individual construction and intention, each objective and subjective parameter performed differently. The established DSI represents a weighted sum of I\_min, F0\_max, MPT, and jitter [18] and therefore integrates parameters of VRP, aerodynamic, and acoustic measures. As assumed, most of

our patients showed DSI values ranging from -5 to 5, whereby -5 corresponded to a very dysphonic voice and 5 to a perceptual normal voice. Due to the special structure of our patient cohort including a considerable number of elite vocal performers and subjects with extremely dysphonic voices, more study participants than expected (34%) had initial DSI values which exceeded these boundaries at both ends. After

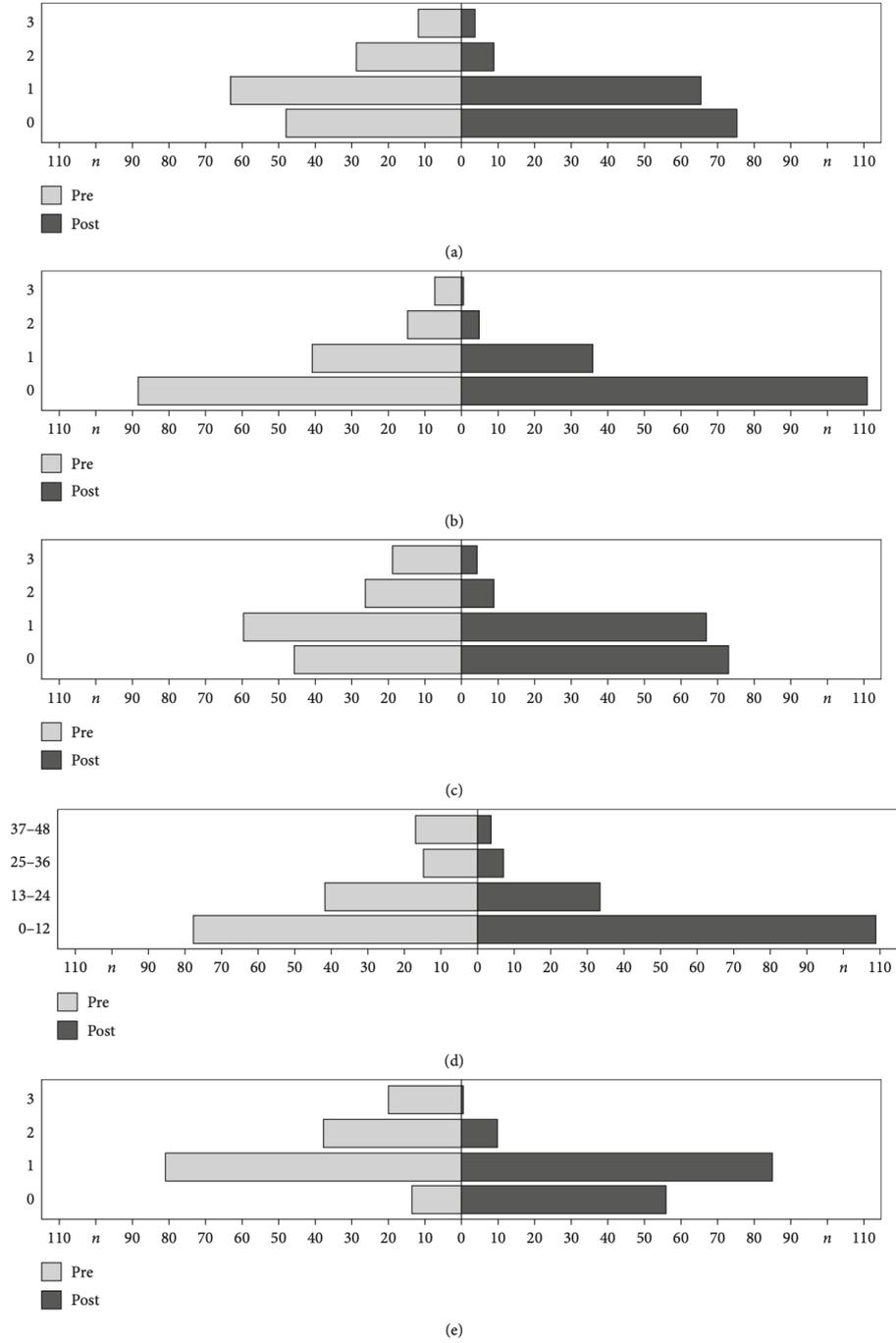


FIGURE 2: Subjective vocal parameters: (a) *R*, (b) *B*, (c) *H*, (d) VHI-12, and (e) VHIs before treatment (light grey, left columns) and after treatment (dark grey, right columns). The abscissae show the number of patients (*n*), and the ordinates represent the scales of the RBH system and VHIs system (0-3) as well as the VHI-12 score.

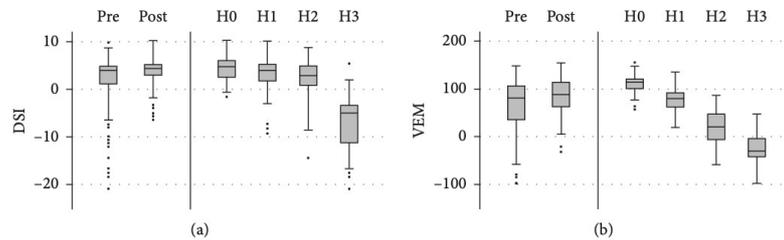


FIGURE 3: Objective vocal parameters: (a) DSI and (b) VEM before and after treatment, as well as their distribution according to the degree of *H*. The boxplots display the median, quartiles, range of values covered by the data, and any outliers (single spots).

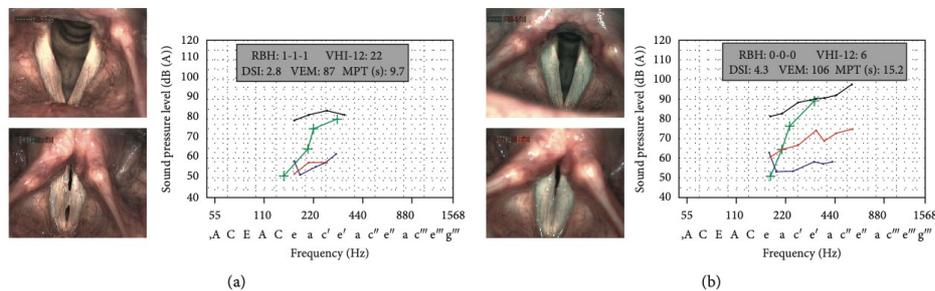


FIGURE 4: Example of phonosurgery-induced changes of laryngeal and vocal findings in a 49-year-old female dental assistant suffering from persisting dysphonia and dysodia. (a) Preoperative VLS shows a marginal edema of the right vocal fold with a glottal gap during phonation, asynchronous oscillations, and impaired mucosal wave propagation. The preoperative VRP pattern displays envelope curves for the loudest (black lines) and softest (blue lines) singing voice and for the speaking voice at different vocal intensity levels (green lines) with little dynamic and frequency ranges. The singer's formant levels (red lines) are low, characterizing the impaired concentration of acoustic energy by resonator amplification of certain frequency ranges in the vocal tract. The values of all objective voice parameters (DSI: dysphonia severity index; VEM: vocal extent measure; MPT: maximum phonation time) and subjective voice parameters (RBH: roughness, breathiness, overall grade of hoarseness; VHI-12: twelve-item voice handicap index) are reduced. (b) Three months after phonomicrosurgical removal of the edema, the treated vocal fold shows a straight margin. The glottal closure is complete, and the oscillations have normalized (mucosal wave propagation regular and symmetric). The patient reveals higher dynamic and frequency ranges of speaking and singing voice with considerably improved objective and subjective parameters.

TABLE 2: Changes in vocal measures after treatment for all patients and separated for both intervention groups (logopedics/phonosurgery), both dysphonia groups (functional/organic), and all pathology classification subgroups according to Rosen and Murry [25].

	<i>H</i>	VHI-12	VHIs	DSI	VEM
Total group of patients ( <i>n</i> = 152)	-0.5 (-0.6; -0.4)	-4.8 (-6.2; -3.3)	-0.7 (-0.8; -0.6)	1.5 (1.0; 2.0)	18.7 (13.4; 24.1)
Logopedic treatment group ( <i>n</i> = 79)	-0.3 (-0.4; -0.2)	-1.4 (-2.6; -0.1)	-0.5 (-0.7; -0.4)	0.2 (-0.1; 0.5)	6.5 (1.6; 11.4)
Phonosurgery group ( <i>n</i> = 73)	-0.8 (-1.0; -0.6)	-8.5 (-10.9; -6.0)	-0.9 (-1.1; -0.7)	2.8 (1.9; 3.7)	32.0 (22.9; 41.0)
Functional dysphonia group ( <i>n</i> = 51)	-0.3 (-0.4; -0.2)	-0.3 (-1.4; 0.7)	-0.5 (-0.7; -0.4)	0.1 (-0.2; 0.4)	3.6 (0.1; 7.1)
Organic dysphonia group ( <i>n</i> = 101)	-0.7 (-0.8; -0.5)	-7.0 (-9.0; -5.0)	-0.8 (-1.0; -0.7)	2.2 (1.5; 2.9)	26.4 (18.9; 33.9)
Rosen I subgroup (epithelium) ( <i>n</i> = 19)	-0.7 (-1.2; -0.3)	-5.1 (-9.6; -0.6)	-1.0 (-1.3; -0.7)	2.0 (0.5; 3.6)	30.2 (10.1; 50.3)
Rosen II subgroup (lamina propria) ( <i>n</i> = 41)	-0.6 (-0.7; -0.4)	-4.6 (-6.8; -2.4)	-0.7 (-0.9; -0.5)	1.2 (0.2; 2.1)	15.7 (7.3; 24.1)
Rosen III subgroup (arytenoid) ( <i>n</i> = 5)	-0.1 (-0.4; 0.3)	-1.6 (-6.0; 2.8)	-0.3 (-0.9; 0.3)	0.2 (-0.8; 0.9)	2.5 (-11.8; 16.7)
Rosen IV subgroup (other) ( <i>n</i> = 36)	-0.8 (-1.1; -0.5)	-11.5 (-15.7; -7.4)	-1.0 (-1.2; -0.7)	3.9 (2.6; 5.3)	39.9 (24.7; 55.1)

Data expressed as mean differences of preoperative and postoperative values (upper line), with 95% confidence intervals (lower line, in brackets). VEM = vocal extent measure; DSI = dysphonia severity index; VHI-12 = voice handicap index; VHIs = self-perceived impairment of voice at the present time.

therapy, the mean DSI measurement data did not improve by an average of one degree of severity [22] but still showed a significant increase. This confirms previous studies, which describe the DSI as a useful parameter to measure the severity of dysphonia and the improvement after therapy

[26–29]. However, various studies could show that the DSI is influenced by differences in measurements of the registration programs as well as by age or gender [19, 30–32]. Therefore, we developed and investigated the VEM as a new objective vocal parameter unimpacted by these interacting factors.

TABLE 3: Results of correlation analysis.

		DSI	VEM	<i>H</i>	VHI-12	VHIs
Pretherapeutic	DSI	1	0.6	-0.4	-0.2	-0.3
	VEM	0.6	1	-0.7	-0.4	-0.3
	<i>H</i>	-0.4	-0.7	1	0.4	0.4
	VHI-12	-0.2	-0.4	0.4	1	0.6
	VHIs	-0.3	-0.3	0.4	0.6	1
Posttherapeutic	DSI	1	0.3	-0.3	-0.3	-0.3
	VEM	0.3	1	-0.7	-0.3	-0.4
	<i>H</i>	-0.3	-0.7	1	0.4	0.5
	VHI-12	-0.3	-0.3	0.4	1	0.7
	VHIs	-0.3	-0.4	0.5	0.7	1
Therapy-induced changes ( $\Delta$ )	DSI	1	0.5	-0.3	-0.04 (ns)	-0.09 (ns)
	VEM	0.5	1	-0.4	-0.2	-0.2
	<i>H</i>	-0.3	-0.4	1	0.3	0.4
	VHI-12	-0.04 (ns)	-0.2	0.3	1	0.6
	VHIs	-0.09 (ns)	-0.2	0.4	0.6	1

All correlation coefficients were significant ( $p < 0.01$ ), unless otherwise specified (ns = not significant).

The VEM calculation is based on the size and shape of the VRP, instead of gathering data from a combination of different objective parameters [19, 24]. Most of our patients showed VEM values between 0 and 120. As expected, these limits were exceeded: (1) at the upper end in functionally impaired singers with professionally trained “great” voices and very large VRP, and (2) at the lower end in extremely dysphonic and nearly aphonic voices with very small VRP. These findings support the underlying idea, during the construction of this vocal parameter, that the ideal VRP should not show abrupt differences in the dynamic range of notes produced by the patients along with their frequency range [24]. A well-balanced dynamic extent approximated the optimal VRP shape to a circle where the area is the biggest for a given perimeter compared to other geometric figures [19]. Our results confirmed that larger and “smoother” VRP without relevant “jumps” in intensity achieved higher values. Vocal capacity quantified in this way, i.e., as a relation of area and perimeter of the VRP, showed a very distinct increase during the course of therapy in our study participants. This is in line with the results of the very few phonosurgical studies assessing VEM values in patients with Reinke’s edema [20], vocal fold polyps [23], and nodules [33]. All of these investigations observed significantly increased VEM values after treatment.

Concerning subjective parameters, the VHI-12 successfully quantified the self-experienced extent of the vocal problem in our patients. Due to therapy, most of them rated an improvement from moderately to mildly impaired. These results correspond to other studies using VHI questionnaires for the investigation of surgical and conservative treatment success in organic dysphonia and functional dysphonia [20, 23, 33–36]. Our overall impression supports the general acknowledgment that short-form VHI versions represent reliable instruments with excellent acceptance and practicability in clinical routine [8, 9]. Compared to previous investigations, the examiners’ auditory perception was the main indicator in our study for the assessment of therapy success [10–13]. We consider the RBH system to be reliable, particularly in case of evaluation by group assessments

[14, 15, 23]. The mean RBH status of our patients’ voices revealed significantly less roughness, breathiness, and overall grade of hoarseness. These results also confirmed the outcomes of former studies [23, 33–35]. Additionally, our analysis of the degree of *H* in relation to the objective parameters DSI and VEM revealed a better representation and graphical distinction of the auditory-perceptual assessment via the VEM. This is a new and important study finding which was confirmed in our investigations of correlation.

Correlations between the changes in individual parameters are able to show how well the improvement in one measurement value is reproduced by another measuring procedure. The generally weak correlations in our results can be explained by the different approaches to the individual parameters and are likewise a manifestation of the additional information content of the respective measurement methods. A relatively high correlation between parameters confirms the success of the therapy from the different aspects of these parameters. Regarding DSI, only a weak negative correlation with *H* and no significant relationship to VHI-12 could be found. Overall, this implies that although the DSI seems suitable for indicating the success of therapy, the increase of DSI has very little to do with the improvement in the degree of hoarseness and with the patient’s perception of the vocal problem. This is also seen in the decreasing correlation of the values for DSI and *H* after therapy. The changes in the VHI-12, on the other hand, had a weak relationship to the changes of the VEM. Moreover, changes in the VEM demonstrated a moderately negative relationship to the changes of *H*, which means that the VEM increases as hoarseness decreases. Thus, in addition to the quantification of vocal capacity, the VEM is validated by the auditory assessment. The novel numeric description of the VRP by means of the interval-scaled VEM provides the researcher with a diagnostic parameter which is suitable for monitoring the course of treatment.

While interpreting these results, some limitations of our study should be considered. First, the number of patients was too small and the cohort was too heterogeneous to examine comparably sized groups of *H* levels, pathology

classification, or diagnosis-related subgroups. Therefore, there could be participation bias. Second, individual treatment recommendations depended on phoniatic indication and were based on comprehensive counseling related to clinical signs, symptoms, individual vocal requirements, abilities, and medical history. Nevertheless, at the end, the patients decided about the kind of intervention; thus, there may be selection bias. Third, our posttherapeutic follow-up of three months was too short to allow statements about the long-term outcome. Fourth, the investigated treatment modalities are often used in a combined mode to accelerate and optimize vocal improvement. We were not able to control whether patients after phonosurgery received hidden other therapies. Additional logopedic treatment or singing lessons are easily accessible and could influence the results especially in the recovery of operated patients. Therefore, there may be also performance bias. Finally, some well-known factors influencing the VRP registration have to be taken into account, such as the routine of the examiner, musicality and motivation of the patients, and the absence of generally accepted specifications regarding the number of registered tones. However, all VRPs were recorded by one experienced examiner under practically equal conditions, so that most of the mentioned factors can be ignored in this study.

Overall, our specific therapeutic outcomes confirmed the results of other studies investigating treatment effects in patients with various voice problems [33–39]. As expected, phonosurgery had the largest numeric impact on the improvement of vocal function. Conservative therapy provided smaller quantitative enhancements but often also qualitative vocal restoration with recovered artistic capabilities, particularly in singers with functional dysphonia. Logopedic training goals typically included reducing extrinsic laryngeal tension, using a relaxed laryngeal posture, and effective abdominal-diaphragmatic support for all phonation events [40]. Specific attention was given to the balance of respiratory forces, laryngeal coordination, and optimal filtering of the source signal via resonance and articulatory awareness [41, 42]. With this approach, also some of our patients with organic findings gained substantial voice improvement. As known from the literature, mainly younger patients with short duration of dysphonia and small benign pathologies of the lamina propria (e.g., vocal fold polyps, marginal edema) due to overuse benefitted from voice therapy [43–45].

## 5. Conclusions

The investigated parameters DSI, VEM, VHI, and RBH are all suitable for monitoring the course of voice treatment and adequate to quantify the outcomes of phonosurgery and logopedic vocal exercises. Correlation analysis confirms the clinical impression that DSI, VEM, VHI, and RBH represent different dimensions of the voice and are complementing objective or subjective measurements either for the evaluation of voice quality, vocal performance, or perceived vocal handicap. The VEM proves to be a comprehensible and easy-to-use parameter for objective VRP evaluation. Changes in the degree of hoarseness as gold standard were best

recognized with the new VEM. Thus, in addition to the quantification of vocal capacity, the VEM is supported and validated by the auditory findings and provides an interval-scaled parameter for documentation.

## Data Availability

The datasets generated and analyzed for this study are not publicly available, as they were obtained from a proprietary database via a licensing agreement.

## Conflicts of Interest

The authors have no funding, financial relationships, or conflicts of interest to disclose.

## Acknowledgments

The authors wish to thank Professor Eleanor Forbes for manuscript proofreading and language editing.

## References

- [1] B. Liu, E. Polce, H. Raj, and J. Jiang, "Quantification of voice type components present in human phonation using a modified diffusive chaos technique," *Annals of Otolaryngology & Laryngology*, vol. 128, no. 10, pp. 921–931, 2019.
- [2] B. Barsties and M. De Bodt, "Assessment of voice quality: current state-of-the-art," *Auris Nasus Larynx*, vol. 42, no. 3, pp. 183–188, 2015.
- [3] N. Roy, J. Barkmeier-Kraemer, T. Eadie et al., "Evidence-based clinical voice assessment: a systematic review," *American Journal of Speech-Language Pathology*, vol. 22, no. 2, pp. 212–226, 2013.
- [4] P. H. Dejonckere, P. Bradley, P. Clemente et al., "A basic protocol for functional assessment of voice pathology, especially for investigating the efficacy of (phonosurgical) treatments and evaluating new assessment techniques," *European Archives of Oto-Rhino-Laryngology*, vol. 258, no. 2, pp. 77–82, 2001.
- [5] M. Behlau, G. Madazio, F. Moreti et al., "Efficiency and cutoff values of self-assessment instruments on the impact of a voice problem," *Journal of Voice*, vol. 30, no. 4, pp. 506.e9–506.e18, 2016.
- [6] M. Behlau, F. Zambon, F. Moreti, G. Oliveira, and E. de Barros Couto Jr., "Voice self-assessment protocols: different trends among organic and behavioral dysphonias," *Journal of Voice*, vol. 31, no. 1, pp. 112.e13–112.e27, 2017.
- [7] B. H. Jacobson, A. Johnson, C. Grywalski et al., "The voice handicap index (VHI)," *American Journal of Speech-Language Pathology*, vol. 6, no. 3, pp. 66–70, 1997.
- [8] C. A. Rosen, A. S. Lee, J. Osborne, T. Zullo, and T. Murry, "Development and validation of the voice handicap index-10," *The Laryngoscope*, vol. 114, no. 9, pp. 1549–1556, 2004.
- [9] T. Nawka, I. M. Verdonck-de Leeuw, M. de Bodt et al., "Item reduction of the voice handicap index based on the original version and on European translations," *Folia Phoniatica et Logopaedica*, vol. 61, no. 1, pp. 37–48, 2009.
- [10] Y. Lee, G. Kim, and S. Kwon, "The usefulness of auditory perceptual assessment and acoustic analysis for classifying the voice severity," *Journal of Voice*, no. 19, pp. 30087–30096, 2019, In press.

- [11] G. B. Kempster, B. R. Gerratt, K. Verdolini Abbott, J. Barkmeier-Kraemer, and R. E. Hillman, "Consensus auditory-perceptual evaluation of voice: development of a standardized clinical protocol," *American Journal of Speech-Language Pathology*, vol. 18, no. 2, pp. 124–132, 2009.
- [12] R. Schönweiler, P. Wübbelt, M. Hess, and M. Ptok, "Psychoakustische skalierung akustischer stimmparameter durch multizentrisch validierte RBH-bewertung1," *Laryngo-Rhino-Otologie*, vol. 80, no. 3, pp. 117–122, 2001.
- [13] M. Ptok, C. Schwemmler, C. Iven, M. Jessen, and T. Nawka, "Zur auditiven bewertung der stimmqualität," *HNO*, vol. 54, no. 10, pp. 793–802, 2006.
- [14] H. Hanschmann and R. Berger, "Perceptual and acoustic evaluation of hoarseness," *Laryngorhinotologie*, vol. 90, no. 2, pp. 68–70, 2011.
- [15] R. Schönweiler, M. Hess, P. Wübbelt, and M. Ptok, "Novel approach to acoustical voice analysis using artificial neural networks," *Journal of the Association for Research in Otolaryngology: JARO*, vol. 1, no. 4, pp. 270–282, 2000.
- [16] P. H. Dejonckere, "Assessment of voice and respiratory function," in *Surgery of Larynx and Trachea*, M. Remacle and H. E. Eckel, Eds., Springer, Berlin, Germany, pp. 11–26, 2010.
- [17] D. D. Mehta and R. E. Hillman, "Voice assessment: updates on perceptual, acoustic, aerodynamic, and endoscopic imaging methods," *Current Opinion in Otolaryngology & Head and Neck Surgery*, vol. 16, no. 3, pp. 211–215, 2008.
- [18] F. L. Wuyts, M. S. D. Bodt, G. Molenberghs et al., "The dysphonia severity index," *Journal of Speech, Language, and Hearing Research*, vol. 43, no. 3, pp. 796–809, 2000.
- [19] P. P. Caffier, A. Möller, E. Forbes, C. Müller, M. L. Freymann, and T. Nawka, "The vocal extent measure: development of a novel parameter in voice diagnostics and initial clinical experience," *BioMed Research International*, vol. 2018, Article ID 3836714, 10 pages, 2018.
- [20] T. Salmen, T. Ermakova, A. Schindler et al., "Efficacy of microsurgery in Reinke's oedema evaluated by traditional voice assessment integrated with the vocal extent measure (VEM)," *Acta Otorinolaryngologica Italica: Organo Ufficiale Della Societa Italiana di Otorinolaringologia E Chirurgia Cervico-Facciale*, vol. 38, no. 38, pp. 194–203, 2018.
- [21] P. P. Caffier, B. Schmidt, M. Gross et al., "A comparison of white light laryngostroboscopy versus autofluorescence endoscopy in the evaluation of vocal fold pathology," *The Laryngoscope*, vol. 123, no. 7, pp. 1729–1734, 2013.
- [22] U. Gonnermann, *Quantifizierbare Verfahren Zur Bewertung Von Dysphonien*, Peter Lang, Frankfurt/Main, Germany, 2007.
- [23] T. Salmen, T. Ermakova, A. Möller et al., "The value of vocal extent measure (VEM) assessing phonosurgical outcomes in vocal fold polyps," *Journal of Voice*, vol. 31, no. 1, pp. 114.e7–114.e15, 2017.
- [24] P. Caffier and A. Möller, "Das stimmumfangsmaß SUM als neuer parameter in der objektiven stimmdiagnostik," *Sprache · Stimme · Gehör*, vol. 40, no. 4, pp. 183–187, 2016.
- [25] C. A. Rosen and T. Murry, "Nomenclature of voice disorders and vocal pathology," *Otolaryngologic Clinics of North America*, vol. 33, no. 5, pp. 1035–1045, 2000.
- [26] M. M. Hakkesteeft, M. P. Brocaar, and M. H. Wieringa, "The applicability of the dysphonia severity index and the voice handicap index in evaluating effects of voice therapy and phonosurgery," *Journal of Voice*, vol. 24, no. 2, pp. 199–205, 2010.
- [27] K. Nemr, M. Simões-Zenari, G. S. de Souza, A. Hachiya, and D. H. Tsuji, "Correlation of the dysphonia severity index (DSI), consensus auditory-perceptual evaluation of voice (CAPE-V), and gender in Brazilians with and without voice disorders," *Journal of Voice*, vol. 30, no. 6, pp. 765.e7–765.e11, 2016.
- [28] V. Uloza, B. B. V. Latoszek, N. Ulozaitė-Staniene, T. Petrauskas, and Y. Maryn, "A comparison of dysphonia severity index and acoustic voice quality index measures in differentiating normal and dysphonic voices," *European Archives of Oto-Rhino-Laryngology*, vol. 275, no. 4, pp. 949–958, 2018.
- [29] M. Brockmann-Bauser, B. Balandat, and J. E. Bohlender, "Immediate lip trill effects on the standard diagnostic measures voice range profile, jitter, maximum phonation time, and dysphonia severity index," *Journal of Voice*, no. 18, p. 30573, 2019, In press.
- [30] M. M. Hakkesteeft, M. P. Brocaar, M. H. Wieringa, and L. Feenstra, "Influence of age and gender on the dysphonia severity index," *Folia Phoniatrica et Logopaedica*, vol. 58, no. 4, pp. 264–273, 2006.
- [31] P. Aichinger, F. Feichter, B. Aichstill, W. Bigenzahn, and B. Schneider-Stickler, "Inter-device reliability of DSI measurement," *Logopedics Phoniatrics Vocology*, vol. 37, no. 4, pp. 167–173, 2012.
- [32] B. B. V. Latoszek, N. Ulozaitė-Staniene, Y. Maryn, T. Petrauskas, and V. Uloza, "The influence of gender and age on the acoustic voice quality index and dysphonia severity index: a normative study," *Journal of Voice*, vol. 33, no. 3, pp. 340–345, 2019.
- [33] P. Caffier, T. Salmen, T. Ermakova et al., "Phonomicrosurgery in vocal fold nodules: quantification of outcomes in professional and non-professional voice users," *Medical Problems of Performing Artists*, vol. 32, no. 4, pp. 187–194, 2017.
- [34] M. D. M. Ropero Rendón, T. Ermakova, M.-L. Freymann, A. Ruschin, T. Nawka, and P. P. Caffier, "Efficacy of phonosurgery, logopedic voice treatment and vocal pedagogy in common voice problems of singers," *Advances in Therapy*, vol. 35, no. 7, pp. 1069–1086, 2018.
- [35] S. Reetz, J. E. Bohlender, and M. Brockmann-Bauser, "Do standard instrumental acoustic, perceptual, and subjective voice outcomes indicate therapy success in patients with functional dysphonia?," *Journal of Voice*, vol. 33, no. 3, pp. 317–324, 2019.
- [36] T. Yilmaz, "Surgical treatment of glottic web using butterfly mucosal flap technique: experience on 12 patients," *The Laryngoscope*, vol. 129, no. 6, pp. 1423–1427, 2019.
- [37] J. Ruotsalainen, J. Sellman, L. Lehto, and J. Verbeek, "Systematic review of the treatment of functional dysphonia and prevention of voice disorders," *Otolaryngology—head and Neck Surgery*, vol. 138, no. 138, pp. 557–565, 2008.
- [38] E. Sielska-Badurek, E. Osuch-Wójcikiewicz, M. Sobol, E. Kazanecka, A. Rzepakowska, and K. Niemczyk, "Combined functional voice therapy in singers with muscle tension dysphonia in singing," *Journal of Voice*, vol. 31, no. 4, pp. 509.e23–509.e31, 2017.
- [39] S. M. Zeitels, R. E. Hillman, M. Mauri, R. Desloge, and P. B. Doyle, "Phonomicrosurgery in singers and performing artists: treatment outcomes, management theories, and future directions," *Annals of Otolaryngology, Rhinology & Laryngology*, vol. 111, no. 12, pp. 21–40, 2002.
- [40] J. C. Goffi-Fynn and L. M. Carroll, "Collaboration and conquest: MTD as viewed by voice teacher (singing voice specialist) and speech-language pathologist," *Journal of Voice*, vol. 27, no. 3, pp. 391.e9–391.e14, 2013.

- [41] K. Pietsch, T. Lyon, and V. K. Dhillon, "Speech language pathology rehabilitation," *Medical Clinics of North America*, vol. 102, no. 6, pp. 1121-1134, 2018.
- [42] D. E. Hazlett, O. M. Duffy, and S. A. Moorhead, "Review of the impact of voice training on the vocal quality of professional voice users: implications for vocal health and recommendations for further research," *Journal of Voice*, vol. 25, no. 2, pp. 181-191, 2011.
- [43] W. D. LeBorgne and E. N. Donahue, "Voice therapy as primary treatment of vocal fold pathology," *Otolaryngologic Clinics of North America*, vol. 52, no. 4, pp. 649-656, 2019.
- [44] M. Sahin, S. Gode, M. Dogan, T. Kirazli, and F. Ogut, "Effect of voice therapy on vocal fold polyp treatment," *European Archives of Oto-Rhino-Laryngology*, vol. 275, no. 6, pp. 1533-1540, 2018.
- [45] J. Agarwal, A. Wong, W. Karle, M. Naunheim, M. Mori, and M. Courey, "Comparing short-term outcomes of surgery and voice therapy for patients with vocal fold polyps," *The Laryngoscope*, vol. 129, no. 5, pp. 1067-1070, 2019.

Mein Lebenslauf wird aus datenschutzrechtlichen Gründen  
in der elektronischen Version meiner Arbeit nicht veröffentlicht.

## Publikationsliste

**Seipelt M**, Möller A, Nawka T, Gonnermann U, Caffier F, Caffier PP. Monitoring the outcome of phonosurgery and vocal exercises with established and new diagnostic tools. *Biomed Res Int* 2020 Jan 23;2020:4208189. doi: 10.1155/2020/4208189.

Salmen T, Ermakova T, Möller A, **Seipelt M**, Weikert S, Rummich J, Gross M, Nawka T, Caffier PP. The Value of Vocal Extent Measure (VEM) Assessing Phonomicrosurgical Outcomes in Vocal Fold Polyps. *J Voice* 2017; 31(1):114.e7-114.e15. doi: 10.1016/j.jvoice.2016.03.016.

Caffier PP, Schmidt B, Gross M, Karnetzky K, Nawka T, Rotter A, **Seipelt M**, Sedlmaier B. A comparison of white light laryngostroboscopy versus autofluorescence endoscopy in the evaluation of vocal fold pathology. *Laryngoscope* 2013; 123(7):1729-1734. doi: 10.1002/lary.23931.

## **Danksagung**

Vor allen Anderen möchte ich meinem Doktorvater PD Dr. med. Philipp Caffier, Oberarzt der Klinik für Audiologie und Phoniatrie der Charité, danken. Seine unermüdliche fachliche und menschliche Unterstützung, sein Gehör und die vielen hilfreichen Hinweise trugen maßgeblich zur Anfertigung dieser Arbeit bei.

Weiterhin danke ich auch dem gesamten Kollektiv der Klinik für Audiologie und Phoniatrie der Charité – Universitätsmedizin Berlin für die fruchtbare und ereignisreiche Zeit, insbesondere Prof. Dr. med. Tadeus Nawka, der mit seiner Begeisterung für die Medizin rund um die Stimme alle Mitarbeiter inspirierte und inspiriert.

Auch dem verstorbenen Prof. Dr. med. Manfred Gross, dem ehemaligen Direktor der Klinik für Audiologie und Phoniatrie, gebührt mein Dank, da ich nach langer Tätigkeit in der HNO als Weiterbildungsassistent in der Phoniatrie und Pädaudiologie unter seiner Leitung und Führung strukturiert die Qualifikation zum Phoniater und Pädaudiologen erwerben durfte. Er war es unter meinen phoniatischen Lehrern, von dessen ärztlichem Vorgehen ich das Meiste in mein eigenes Vorgehen und Handeln übernommen habe.

Und, letztgenannt, aber keineswegs letztrangig, sei meine Familie, der ich zu Dank verpflichtet bin, genannt: Meinen Eltern für die endlose Geduld, die motivierende Ermutigung und ihr menschliches Vorbild, meiner Frau und meinen Kindern Charlotte und Ferdinand für ihre liebende Unterstützung.