

Aus der Klinik und Poliklinik für Hals-, Nasen- und Ohrenheilkunde
Klinik der Ludwig-Maximilians-Universität München
Vorstand: Prof. Dr. med. Martin Canis

**Automatisches Spracherkennungssystem zur Quantifizierung
von Veränderungen der Nase und Nasennebenhöhlen
auf die Sprachverständlichkeit**

Dissertation
zum Erwerb des Doktorgrades der Zahnmedizin
an der Medizinischen Fakultät der
Ludwig-Maximilians-Universität zu München

vorgelegt von

Katja Eckstein

aus

Oelsnitz (Vogtl.)

2019

Mit Genehmigung der Medizinischen Fakultät
der Universität München

Berichterstatlerin: Prof. Dr. med. Maria Schuster

Mitberichterstatler: PD Dr. med. Jesus Bujia

Mitbetreuung durch die
promovierte Mitarbeiterin: Dr. med. Susanne Mayr

Dekan: Prof. Dr. med. dent. Reinhard Hickel

Tag der mündlichen Prüfung: 09.05.2019

Inhaltsverzeichnis

	Seite
1 Zusammenfassungen	1
1.1 Zusammenfassung in deutsch	1
1.1.1 Hintergrund	1
1.1.2 Studienaufbau	1
1.1.3 Patienten und Methoden	1
1.1.4 Ergebnisse	1
1.1.5 Diskussion	1
1.1.6 Schlüsselwörter	2
1.2 Summary in english	3
1.2.1 Background	3
1.2.2 Design	3
1.2.3 Patients and methods	3
1.2.4 Results	3
1.2.5 Discussion	3
1.2.6 Keywords	3
2 Einleitung	4
3 Material und Methoden	7
3.1 Patienten	7
3.2 Testmaterial	8
3.3 Inspiratorische nasale Peak- Flow- Messung (PFM)	8
3.4 Spracherkennungssystem	9
3.5 Analyse und Auswertung der Daten	10
4 Ergebnisse	11
4.1 Worterkennungsrate WR	11
4.2 Peak Flow	16
4.3 Vergleich Worterkennungsrate und Peak Flow	18
5 Diskussion	19
6 Fazit	23
7 Anhang	24
8 Literaturverzeichnis	26
9 Danksagung	31
10 Eidesstattliche Versicherung	32

Zusammenfassung:

Hintergrund: Erkrankungen der Nase und Nasennebenhöhlen können zu Änderungen der Nasalität führen und damit Auswirkungen auf die Verständlichkeit von Gesprochenem haben. In dieser Arbeit wird die Veränderung der Sprachverständlichkeit nach nasenchirurgischen Eingriffen (Nasenseptumkorrektur, funktionelle endoskopische Nasennebenhöhlen-Eingriffe FESS) im Verlauf bestimmt.

Studienaufbau: prospektive Kohortenstudie

Patienten und Methoden: 20 Patienten (12 Männer, 8 Frauen, mittleres Alter 38 Jahre), die für eine Operation in der HNO-Klinik aufgenommen wurden, konnten nach deren Einwilligung in die Studie integriert werden. Nach klinischen Untersuchungen und nasalen Peak-Flow-Meter-Messungen wurde die Sprachverständlichkeit an Sprachaufnahmen eines gelesenen Standardtextes ermittelt. Mittels automatischer Spracherkennung wird der prozentuale Anteil korrekt gesprochener Wörter des Standardtextes als sogenannte Worterkennungsrate (WR) ermittelt. Die Untersuchungen erfolgten präoperativ (t1), am Tag nach der Operation mit Nasentamponade (t2) und drei Monate (t3) postoperativ. Zum Vergleich dient die Messung des maximalen inspiratorischen Nasendurchflusses mittels Peak-Flow-Meter.

Ergebnisse: Die Patienten wurden mit FESS, mit Conchotomien und Septorhinoplastiken, sowie verschiedenen Kombinationen aus diesen versorgt. Signifikante Veränderungen der WR zeigten sich von t1 zu t2 und von t2 zu t3. Insgesamt betrachtet änderte sich die WR prä- zu postoperativ (t1 zu t3) nur gering. Eine Korrelation der WR zu Alter und Geschlecht wurde nicht festgestellt. Die von t1 zu t3 leicht angestiegene Worterkennungsrate korrelierte mit den Ergebnissen der nasalen Peak-Flow-Metrie, welche ebenfalls die Tendenz zur Steigerung zeigten.

Diskussion: Die automatische Spracherkennung hat sich als valides Verfahren zur Quantifikation der Sprachverständlichkeit dargestellt. Eine vollständige nasale Obstruktion, hier durch Nasentamponaden, führte zu einer deutlichen Veränderung der Verständlichkeit. Eine signifikante Änderung der Sprachverständlichkeit durch Eingriffe an der Nase und damit an den

Resonanzräumen lässt sich nicht nachweisen. Dies ist insbesondere für Berufssprecher von großer Bedeutung.

Schlüsselwörter: Sprachverständlichkeit, automatische Spracherkennung, Worterkennungsrates, Sinusitis, Septumdeviation, FESS, nasaler Peak-Flow

Summary:

Background: A modified amnesic shows effects on the intelligibility of spoken words. The intelligibility of patients can be determined on a standardized text computer-aided by means of automatic speech recognition. In this work, the change in intelligibility is determined after nose surgery, (septum plastic surgery, functional endoscopic sinus surgery FESS) in the course.

Design: prospective cohort study

Patients and methods: 20 patients (12 men, 8 women, average age 38 years) taken for an operation in the ENT Clinic, after consent could be integrated into the study. Speech intelligibility on voice recordings of read standard text was determined according to clinical studies and nasal peak flow meters measurements. Using automatic speech recognition, the percentage of correctly spoken words of the standard text is calculated as so-called word recognition rate (WR). The investigations were performed preoperatively (t1), the day after surgery with nasal Tamponade (t2) and three months (t3) after surgery. The measurement of the peak flow is used for comparison.

Results: The patients were treated with functional endoscopic sinus surgery FESS, with conchotomy and septorhinoplasty, as well as various combinations of these. Significant changes of the WR are apparent from t1 to t2 and t2 to t3. Overall the WR changed pre-to postoperatively (t1 to t3) only slightly. A correlation of the WR to age and gender was not determined. The word recognition rate slightly increased from t1 to t3 correlated with the results of the nasal peak flow geometry, which also showed the trend to improvement.

Discussion: Represented ASR has presented himself as a valid method for the quantification of intelligibility. A complete nasal obstruction, through Nasal packing, resulted in a significant change of intelligibility. A significant change in the speech intelligibility through interventions on the nose and thus the resonant spaces cannot be proven. This is important in particular for professional speaker or singer.

Keywords: Speech intelligibility, automatic speech recognition, word recognition rate, sinusitis, nasal septum deviation, FESS, nasal peak-flow-measurement immediately

Einleitung:

Die menschliche Stimme stellt eine der wichtigsten Dimensionen unserer Persönlichkeit dar. Durch sie wird ein wesentlicher Teil der zwischenmenschlichen Kommunikation ermöglicht [12]. Die Sprache ist die übergeordnete Leistung des menschlichen Verhaltens bei der verbalen Kommunikation. Das Sprechen ist ihr als äußere Form untergeordnet. Es ist die Fähigkeit, Gedanken mittels der Sprech – und Stimmorgane auszudrücken [12]. Hierbei dient der Kehlkopf als Tongenerator. Dieses primäre Stimmsignal wird durch die Querschnitseigenschaften des Ansatzrohres (Resonanzraum) moduliert. Das Ansatzrohr umfasst den supraglottischen Raum sowie Rachen, Mundhöhle, Nase und Nebenhöhlen [12].

Die Sprachverständlichkeit wird durch viele Faktoren beeinflusst. Pathologische Veränderungen im Bereich der Nase, des Nasenrachens und der angrenzenden Nasennebenhöhlen können zu einer krankhaft veränderten Nasenresonanz führen, d.h. zu einer Rhinophonia clausa. Die veränderte Nasalität hat Auswirkungen auf die Verständlichkeit von Gesprochenem [35].

Es steht heute eine Reihe von Untersuchungsmethoden der Nasalität zur Verfügung. Noch vor kurzem bezeichnete man „das geschulte Ohr des Untersuchers“ als letztlich einziges relevantes „Messinstrument“, heute können wir an dieser Stelle computergestützte Spracherkennungssysteme nennen und verwenden [11].

Spracherkennungssysteme sind mittlerweile weit verbreitet (z.B. Telefonauskunftssysteme) und werden üblicherweise verwendet, um unbekannte Wörter oder Wortfolgen zu erkennen.

Die Technik basiert auf statistischen Modellen, welche die phonetischen und syntaktischen Besonderheiten der Sprache nutzen, z.B. welche Wörter am ehesten aufeinander folgen können. Für unseren diagnostischen Zweck wird das System nun umgekehrt genutzt: aus einer bekannten Wortfolge von Einzelwörtern (Text) wird nun die Qualität des Gesprochenen bewertet, in dem das System prozentual angibt, wie viele der Wörter einer gesprochenen Wortfolge richtig erkannt wurden [11,26].

Die Nase und ihre angrenzenden anatomischen Strukturen, insbesondere die Nasennebenhöhlen, können einen entscheidenden Einfluss auf die Stimmqualität nehmen [23]. Eine Obstruktion, vor allem der Nasenhaupthöhle, führt in der Regel zu einer Hyponasalität. Hierbei sind hauptsächlich die im Deutschen enthaltenen Nasale /m/, /n/ und /ng/ verändert oder durch pränasalisierte Plosive ersetzt. Es zeigt sich zu wenig Nasenresonanz mit dumpfer, verstopft klingender Stimme. Veränderungen lassen sich im Stimmsignal vor allem in tieferen Frequenzbändern messen [23].

Im Gegensatz dazu liegt bei insuffizientem Abschluss des Nasenrachens während der Artikulation ein offenes Näseln oder Hypernasalität vor. Es entweicht Luft durch die Nase, der intraorale Luftdruck wird vermindert und es entstehen zusätzliche Geräusche bei der Bildung von oralen Konsonanten [32]. In dieser Studie wurden ausschließlich Patienten mit postoperativer Nasentamponade gewählt, um somit den Einfluss der maximalen Hyponasalität auf die Sprachverständlichkeit zu verdeutlichen.

Für den Stimmklang ist die Nase von hoher Bedeutung. Obgleich es im Deutschen nur drei echte Nasale gibt (/m/, /n/, /ng/), führt eine Veränderung der nasalen Resonanz sofort zu hörbaren Einflüssen auf das Sprechen. So kann z. B. eine einfache Rhinitis acuta meist leicht auch vom Unbedarften allein an der Stimme bemerkt werden [32].

Auch die Nasennebenhöhlen nehmen Einfluss auf den Stimmklang. Die Datenlage ist dazu allerdings übersichtlich und weist zum großen Teil methodische Schwächen auf. Einige Meinungen gehen davon aus, dass die Pneumatisation der Gewichtsersparnis des Schädels diene, andere Meinungen vermuten eine Resonanzfunktion der Stimme [5]. Die thermische Isolation des ZNS, der Schutz der Schädelkalotte, die Funktion zur Anwärmung und Befeuchtung der Atemluft und die Nasennebenhöhlen als Bestandteil des Riechorgans werden ebenso diskutiert [22,23]. Die meisten Hypothesen konnten durch anatomische Studien widerlegt werden. Belegt ist hingegen eine aktive Schleimsekretion an der Mukosaauskleidung der Nasennebenhöhlen kombiniert mit einer aktiven Konvektion in den Sinus, vermutlich mit dem Ziel einer Wasserdampfsättigung der inspiratorischen Atemluft [43]. Durch die endonasale Nasennebenhöhlenchirurgie wird die Belüftung der betroffenen Strukturen wiederhergestellt und somit eine Ausheilung dieser ermöglicht. Die

krankhaft veränderte Schleimhaut wird abgetragen, die gesunde wird weitestgehend geschont [2].

Neuere Untersuchungen von Havel et al. konnten akustische Veränderungen des Stimmsignals durch isolierte Eingriffe an den Nasennebenhöhlen darstellen. Die Arbeitsgruppe hatte in mehreren Ansätzen (in vivo, am Modell und ex vivo) aufzeigen können, dass ein isolierter Verschluss der Nasennebenhöhlen durch Einbringen einer dichten Masse in die Ostien eine Veränderung der Resonanz erbringt, sowie auch – gegenläufig am akustischen Signal – die Öffnung der Nasennebenhöhlen. Insbesondere weite Öffnungen zu den Ostien der großen Nasennebenhöhlen zeigten signifikante Auswirkungen auf die Resonanzfunktion der Nase [18,19].

Für viele Berufsgruppen ist es wichtig zu erfahren, ob eine chirurgische Therapie der Nasenhaupthöhle und der Nasennebenhöhlen Einfluss auf die Sprachverständlichkeit nimmt. Es sollte geklärt werden, ob langfristig mit Verbesserungen oder Einschränkungen der Verständlichkeit nach einer Operation am Naseninneren durch die Veränderungen der nasalen Resonanzräume gerechnet werden kann bzw. muss. Dies ist insbesondere auch für das Aufklärungsgespräch vor einem Eingriff an der Nase und den Nasennebenhöhlen von Bedeutung.

In dieser prospektiven Studie wurden nun erstmals objektive Daten mittels automatischer Spracherkennungstechnik zur Verständlichkeit von Patienten mit Erkrankungen im Bereich der Nase und den angrenzenden Nasennebenhöhlen mit Hilfe einer computergestützten stochastischen Analyse gesprochener Sprache ermittelt. Es werden drei Zeitpunkte verglichen – vor , direkt nach einer Operation und im Verlauf. Die Daten vor der Operation und im Verlauf werden verglichen zu Messungen der Nasenluftpassage.

Material und Methoden:

Patienten:

Zwanzig Patienten (8 Frauen, 12 Männer) im Alter von 18 bis 93 Jahren (mittleres Alter 38 Jahre) (Anhang Tabelle 1), die aufgrund einer chirurgischen Maßnahme an der Nase in der Hals- Nasen- Ohrenklinik der Universität Erlangen-Nürnberg im Zeitraum von April 2007 bis Juli 2007 stationär aufgenommen wurden, konnten in diese Studie integriert werden. Alle Patienten willigten nach genauer Information über den Studienablauf in die von der Ethikkommission der Friedrich-Alexander-Universität geprüfte Untersuchung ein. Bestandteil der Studie waren klinische HNO-ärztliche Untersuchungen mittels vorderer Rhinoskopie und starrer Endoskopie, Sprachaufnahmen mit unten näher beschriebenen Testmaterial zu drei Zeitpunkten, einen Tag vor der Operation (t1), einen Tag danach (t2) und ein weiteres Mal nach drei Monaten (t3). Zum Zeitpunkt t1 und t3 erfolgte darüber hinaus noch die Messung des maximalen inspiratorischen Nasendurchflusses mittels einem nasalen Peak-Flow-Meter.

Sieben Patienten erhielten eine endonasale Nasennebenhöhlen-Operation (functional endoscopic sinus surgery FESS), bei zwei der Patienten wurde eine FESS mit einer Septorhinoplastik kombiniert und bei zwei Patienten mit einer Conchotomie. Ein weiterer Patient erhielt eine FESS zusammen mit einer Septorhinoplastik und einer Conchotomie. Sechs Patienten erhielten nur eine Septorhinoplastik einschließlich einer Conchotomie. Ausschließlich eine Septorhinoplastik und ausschließlich eine Conchotomie wurden bei jeweils einem Patienten vorgenommen.

Bei der ersten postoperativen Untersuchung (t2) waren die Patienten mit Nasentamponaden versorgt, so dass zu diesem Zeitpunkt kein nasaler Peak Flow bestimmt werden konnte.

Die Patienten gaben alle deutsch als ihre Muttersprache an und hatten keine Sprecherausbildung. Unter den Testpersonen befand sich ein Raucher. Die klinische Untersuchung des Larynx ergab in keinem Fall einen pathologischen Befund, der zu einer Stimmstörung hätte führen können. Es fand sich kein Hinweis auf eine neurogene Sprechstörung bei den Patienten.

Testmaterial:

Die Erstellung der Sprachaufnahmen erfolgte standardisiert. Die Patienten lasen an jedem der Untersuchungszeitpunkte den Text „Der Nordwind und die Sonne“ (Anhang). Der Text wurde mit PEAKS, einem Programm zur Bewertung aller Arten von Sprachstörungen, am Bildschirm in 10 syntaktisch sinnvoll aufgeteilten Untereinheiten angezeigt. Die Aufzeichnung erfolgte digital mit einem dnt Call 4U Comfort-Headset mit einer Auflösung von 16 kHz und 16 bit. Der gelesene Text wurde digital aufgezeichnet.

Inspiratorische nasale Peak-Flow-Messung (PFM).

Die Messung des Luftdurchflusses ist eine gängige Methode, um die respiratorische Funktion der Nase zu bestimmen. Für diese Untersuchung wurde das Luftvolumen bestimmt, das bei schneller Einatmung durch die Nasenhaupthöhle binnen einer Sekunde fließt. Die Messung des maximalen inspiratorischen Nasendurchflusses wurde mit einem Peak-Flow-Meter durchgeführt, bei welchem das Mundansatzstück durch eine Maske zum luftdichten Verschluss des naso-oralen Bereichs ersetzt wurde. Mit einem inspiratorischen Peak-Flow-Meter erhält man innerhalb von Sekunden aufschlussreiche Information über Inspirationsfähigkeit und Inhalationstechnik. Die Ergebnisse können sofort auf der Skala des Gerätes (bis 350 l/min) abgelesen werden.

Der Patient wird aufgefordert, ganz auszuatmen, das Gerät waagrecht zu halten und dafür zu sorgen, dass die Gesichtsmaske luftdicht um die Nase abschließt. Anschließend wird der Patient angehalten, kräftig durch die Nase einzuatmen. Beim Einatmen durch die Nase zum Messen des maximalen Einatemstroms muss der Patient etwa eine Sekunde lang kurz und kräftig einatmen. Der Einatemstrom wird dokumentiert, und das Gerät wird in seine Grundstellung zurückversetzt.

Zur validen Bestimmung werden in der Regel drei Messungen des maximalen Einatemstroms durch die Nase vorgenommen, von denen das höchste Ergebnis zur Berechnung herangezogen wird.

Spracherkennungssystem:

Zur objektiven Messung der Sprachqualität bzw. der funktionellen Einschränkung des Sprechens aufgrund von Erkrankungen der Nase und Nasennebenhöhlen kam in dieser Studie ein Spracherkennungssystem zur Anwendung, das bereits an anderen Kollektiven geprüft wurde [7]. Die automatische Spracherkennung kann man als ein Problem der Mustererkennung auffassen. Die Aufgabe besteht darin, den Inhalt einer Äußerung durch Vergleich mit vorhandenen Sprachmustern zu bestimmen [8]. Sprachdaten werden dabei zuerst in kleine zeitliche Einheiten unterteilt (16 ms), die überlappend analysiert werden hinsichtlich derer temporalen und spektralen Merkmale und dann kategorisiert und mit vorliegenden Wortmodellen verglichen werden.

Ein Spracherkennungssystem wertet zwei unterschiedliche Informationskanäle aus: Das Wissen über das akustische Signal wird mit stochastischen Wortmodellen (Hidden- Markov- Modelle) modelliert, das Wissen über die Abfolge von Wörtern einer Sprache wird mit stochastischen Sprachmodellen repräsentiert. Für unseren Zweck wurde nur ersteres einbezogen, um die akustischen Eigenschaften des Gesprochenen stärker zu bewerten (Unigramm-Sprachmodell) [9]. Das eingesetzte Erkennungssystem arbeitet auf akustischer Ebene polyphon-basiert, d.h., die akustischen Merkmale eines Phonems werden mit koartikulatorischen Modulationen des Phonems dargestellt, so dass auch unterschiedliche phonetische Realisationen eines Lauts adäquat erfasst werden können. So wird z. B. der Laut /e/ in „Welle“ anders ausgesprochen als in „Weg“.

Die Test-Perplexität (durchschnittliche Anzahl von Wörtern, die auf ein bestimmtes Wort innerhalb einer Sequenz folgen können) des Unigramm-Sprachmodells beträgt 94 entsprechend der Anzahl der Wörter im Text, was ausreichend ist, um eine schlechte Verständlichkeit als niedrige Worterkennungsrate (WR, siehe unten) abzubilden und umgekehrt. Eine weitere Verbesserung des Spracherkennungsergebnisses wurde durch sprechabhängige Adaption (maximum likelihood linear regression, MLLR [10]) erreicht.

Die Worterkennungsrate WR ist der prozentuale Anteil korrekt erkannter Wörter des gelesenen Textes.

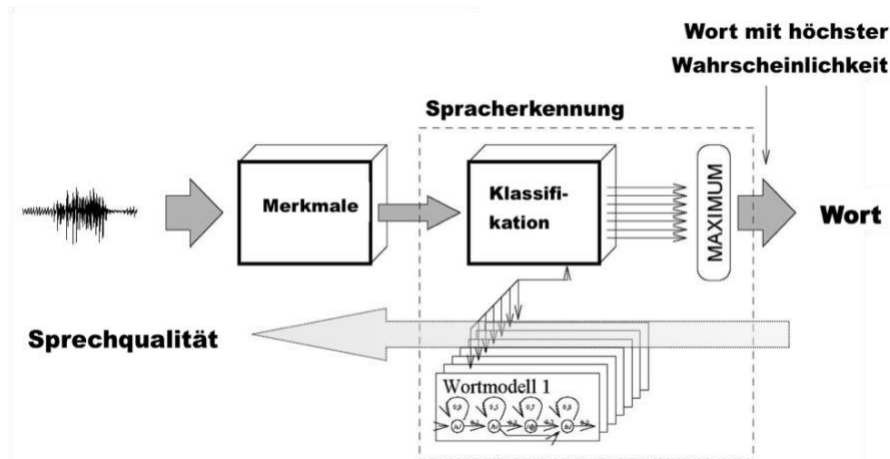


Abbildung 1: schematische Darstellung der Funktionsweise der automatischen Spracherkennung. Die akustischen Eigenschaften von Sprache werden in kleinen zeitlichen Einheiten hinsichtlich ihrer Merkmale analysiert und mit Wortmodellen verglichen. Das Wort, das mit der höchsten Wahrscheinlichkeit mit dem akustischen Signal übereinstimmt, wird als das erkannte Wort gewählt.

Analyse und Auswertung der Daten:

Die statistische Analyse und Auswertung erfolgte mittels *Microsoft Excel*, *SPSS* und *PEAKS*, einem Programm zur Bewertung aller Arten von Sprachstörungen, das von Java unterstützt wird. Im Letzteren wurde die vollautomatische Bewertungsvariante verwendet. Diese nutzt die Information, die beim Aufnehmen am PC entsteht. Es lässt sich eine Referenz ermitteln, die zur Berechnung der Worterkennungsrates herangezogen werden kann. Die automatische WR kann nahezu mit der Leistung transkriptionsbasierter Wortakkuratheit gleichgesetzt werden.

Des Weiteren wurden die Anamnese der Patienten, klinische Untersuchungs- und Fragebögen bezüglich ihres Allgemeinbefindens, der Nasennebenhöhlenbeschwerden, der Stimme und deren Einschränkung in die Studie einbezogen. Die Daten wurden mittels *SPSS* und *Excel* analysiert und berechnet. Die Korrelationsanalysen erfolgten in *Excel*.

Ergebnisse:

Die verschiedenen Untergruppen der Testpersonen basieren auf deren Anamnese und der Art des operativen Eingriffs. Bei fortgeschrittener chronischer Rhinosinusitis wurde eine endonasale, endoskopische Nasennebenhöhlenchirurgie (FESS) durchgeführt, bei funktionell bedeutsamer Veränderung der Nasenform mit nasaler Obstruktion erhielten die Testpersonen eine Septorhinoplastik. Alle Patienten klagten präoperativ über Einschränkungen der Nasenatmung.

Art der Operation (n=20)	n	chronische Sinusitiden	polypöse Schleim- hautveränderungen	Schiefnase
NNH-OP	9	7	2	
NNH-OP mit Conchotomie	1	1	1	
NNH-OP mit Septorhinoplastik	1	1		
NNH-OP mit SRP und Conchotomie	1	1		
Septorhinoplastik	1			1
Septorhinoplastik mit Conchotomie	6		6	
Conchotomie	1		1	

Abbildung 2: Hauptsymptomatik bei den verschiedenen Operationsformen

Worterkennungsrate WR:

Am Tag vor der Operation (t1) lag die Worterkennungsrate (WR) zwischen 47,2% und 78,7% (65,7% \pm 16,5%) (Abbildung 3;5).

Es zeigte sich kein signifikanter Zusammenhang von Verständlichkeit und Geschlecht (Abbildung 3) bzw. Alter (Abbildung 9) der Patienten ($p > 0,05$), wobei die weiblichen Testpersonen bei t1 im Mittelwert mit 71,2 % höhere

Werte der WR aufwiesen als die männlichen Testpersonen mit einem Mittelwert von 67,2% (Abbildung 3). Die Streuung der Ergebnisse ist bei den weiblichen Patienten deutlich niedriger (Abbildung 3).

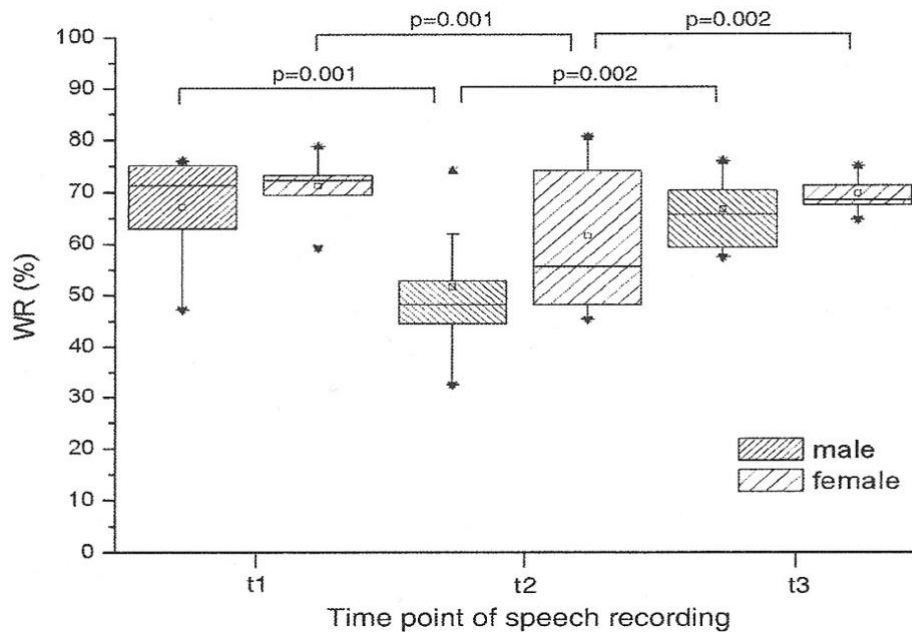


Abbildung 3: Die Worterkennungsrate WR in Abhängigkeit vom Geschlecht zum Zeitpunkt t1 (präoperativ), t2 (postoperativ mit Nasentamponade) und t3 (drei Monate postoperativ)

Die WR zeigte postoperativ (t2) Werte zwischen 32,4% und 80,6% ($55,7\% \pm 12,8\%$) und ist damit signifikant geringer ($p = 0,033$) als zu t1 (Abbildung 4,5,8). Zum Alter (Abbildung 9) und Geschlecht (Abbildung 3) konnten auch zu diesem Zeitpunkt keine relevanten Auffälligkeiten abgeleitet werden.

Postoperativ nach drei Monaten (t3) zeigten sich Werte der WR zwischen 57,4% und 75,9% ($68,0\% \pm 5,3\%$), die zu t2 signifikant erhöht sind ($p = 0,019$) (Abbildung 4,5,8). Alter (Abbildung 9) und Geschlecht (Abbildung 3) nahmen wiederum keinen relevanten Einfluss auf die Ergebnisse, es zeigt sich postoperativ in Bezug auf den präoperativen Zustand erneut eine geringere Breite bei den Werten der weiblichen Testpersonen.

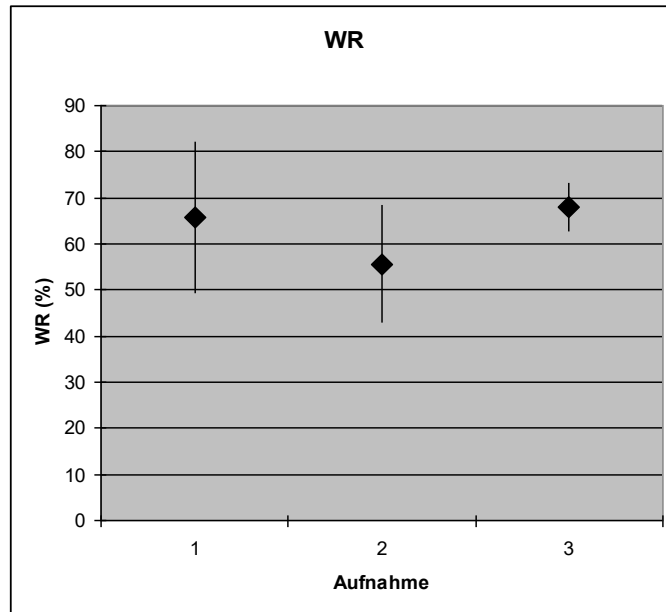


Abbildung 4: Die Abbildung zeigt die Mittelwerte der Worterkennungsraten (y-Achse) zu den Aufnahmezeiten t1 (1), t2 (2) und t3 (3) graphisch dargestellt

Stellt man die Ergebnisse der ersten und der letzten Sprachaufnahme gegenüber, so ist mit einem Wert von ($p = 1,000$) zwischen diesen Zeiten kein signifikanter Unterschied gegeben (Abbildung 5;8).

Es gibt somit eine signifikante Änderung von t1 zu t2, ebenso eine signifikante Änderung von t2 zu t3. Zwischen t1 und t3 konnte kein signifikanter Zusammenhang ermittelt werden (Abbildung 5,8).

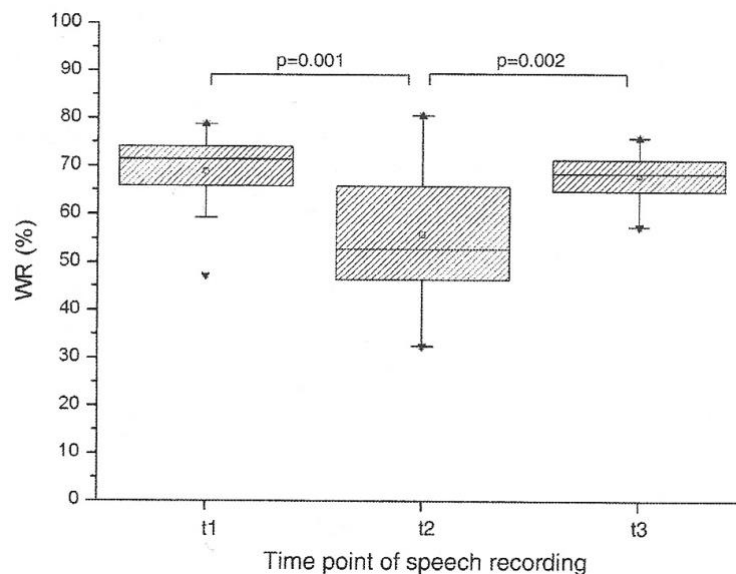


Abbildung 5: Die Worterkennungsrate WR der gesamten Testpersonen ($n=20$) zum Zeitpunkt t1 (präoperativ), t2 (postoperativ mit Nasentamponaden und t3 (postoperativ nach drei Monaten)

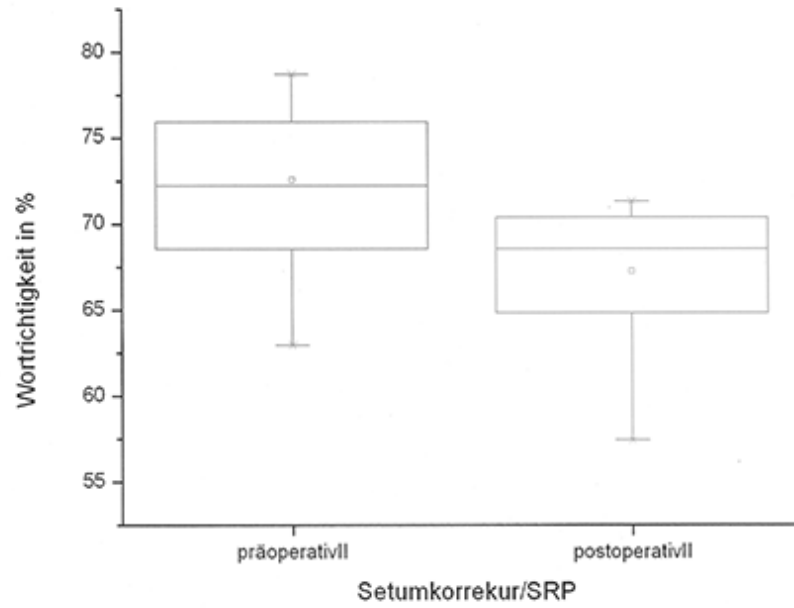


Abbildung 6: Die WR der Testpersonen mit einer Septumkorrektur/SRP zum Zeitpunkt t1 (präoperativ) und t3 (postoperativ nach drei Monaten)

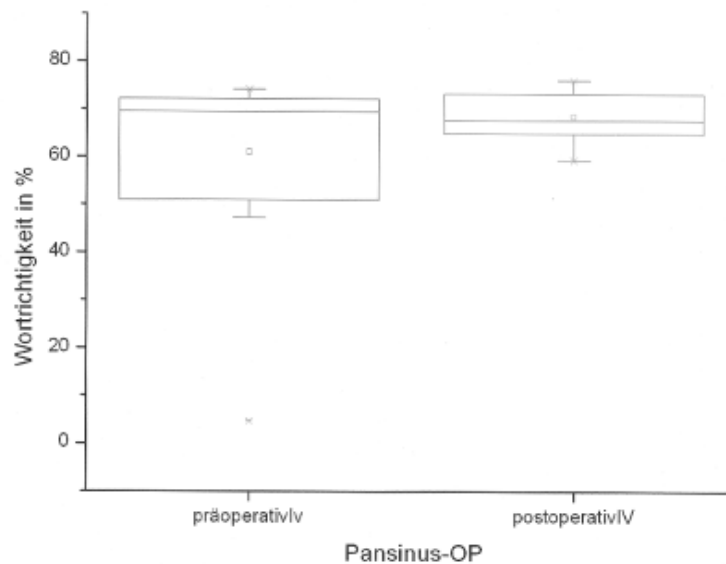


Abbildung 7: Die WR der Testpersonen mit einer Pansinus- OP zum Zeitpunkt t1 (präoperativ) und t3 (postoperativ nach drei Monaten)

Vergleich der Zeitpunkte	p
t1 zu t2	0,033
t2 zu t3	0,019
t1 zu t3	1,000

Abbildung 8: Berechnung des Signifikanzniveaus nach Bonferroni

- t1 Sprachaufnahme einen Tag vor der Operation,
- t2 Sprachaufnahme einen Tag nach der Operation,
- t3 Sprachaufnahme drei Monate nach der Operation

Es gibt keinen signifikanten Bezug zu Alter (Abbildung 9) und Geschlecht (Abbildung 3) der Patienten. Die Dauer der Aufnahmen nahm von t1 zu t3 ab.

Alter in Jahren	n (n=20)	t1 (%)	t2 (%)	t3 (%)
≤ 20	2	75.9 (± 2.6)	65.8 (± 10.2)	72.2 (± 2.8)
20-30	10	70.6 (± 5.3)	51.9 (± 11.5)	67.8 (± 4.9)
30-50	3	61.6 (± 10.7)	56.5 (± 12.9)	64.8 (± 4.0)
60-70	4	65.5 (± 10.7)	58.7 (± 11.5)	68.5 (± 3.2)
93	1	65.7 (± 0)	48.2 (± 0)	59.3 (± 0)

Abbildung 9: WR in Abhängigkeit vom Alter der Patienten

Die WR der Patienten mit einer endonasalen Nasennebenhöhlen-Operation FESS unterscheidet sich postoperativ unwesentlich von den präoperativen Ausgangsbefunden (Abbildung 7). Bei den Patienten mit einer Septumkorrektur oder Septorhinoplastik ist bei der WR ein leichter Abfall zwischen dem präoperativem Ausgangszustand und dem postoperativ ermittelten Wert festzustellen (Abbildung 6). Eine Signifikanz ist nicht nachweisbar.

Peak – Flow- Metrie PFM:

Testpersonen mit einer endonasalen Nasennebenhöhlen-Operation (FESS) tendierten bei der inspiratorischen nasalen Peak-Flow-Metrie (PFM) zu einer Steigerung der Werte zwischen der Messung zu den Zeitpunkten t1 und t3 (Abbildung 11), Testpersonen mit einer Septumkorrektur oder Septorhinoplastik zeigten einen deutlichen Anstieg der Werte zwischen t1 und t3 (Abbildung 12). Gesamt betrachtet ist das nasale Atemzugvolumen der Testgruppe postoperativ nach drei Monaten gegenüber dem präoperativen Ausgangszustand angestiegen, eine Signifikanz ist aber nicht nachweisbar (Abbildung 10).

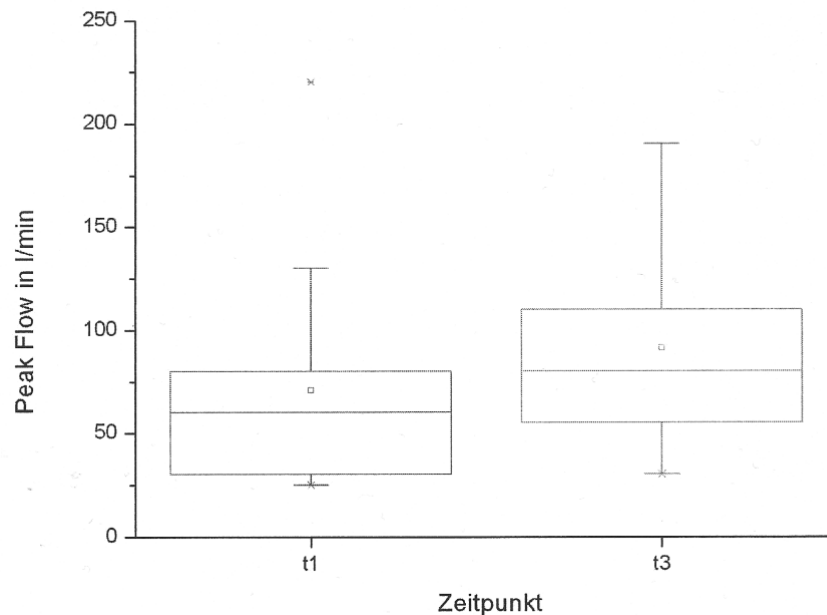


Abbildung 10: Die Abbildung zeigt die Werte der Peak-Flow- Metrie der gesamten Testpersonen (n=20) prä- und postoperativ nach drei Monaten. Eine Zunahme des Volumens ist zu erkennen.

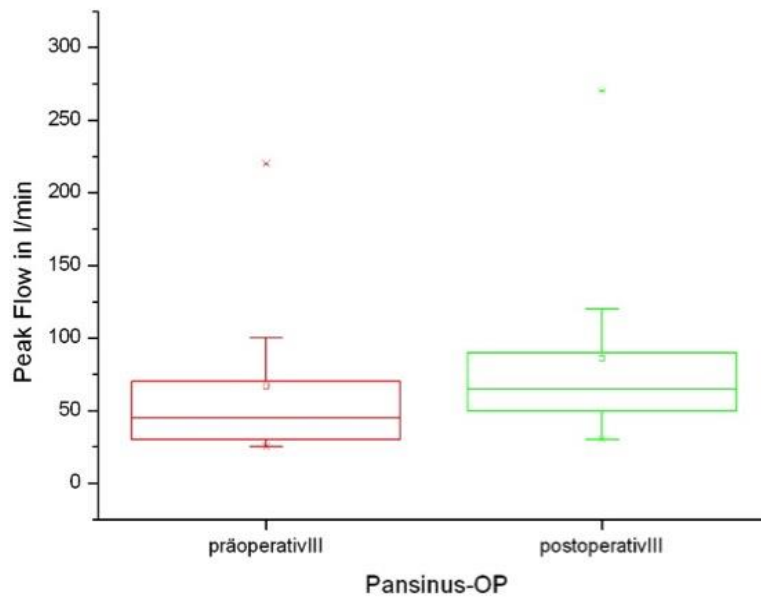


Abbildung 11: Die Abbildung zeigt die Werte der Peak-Flow- Metrie prä- und postoperativ nach drei Monaten bei den Testpersonen mit einer Pansinus- OP. Eine Zunahme des Volumens ist zu erkennen.

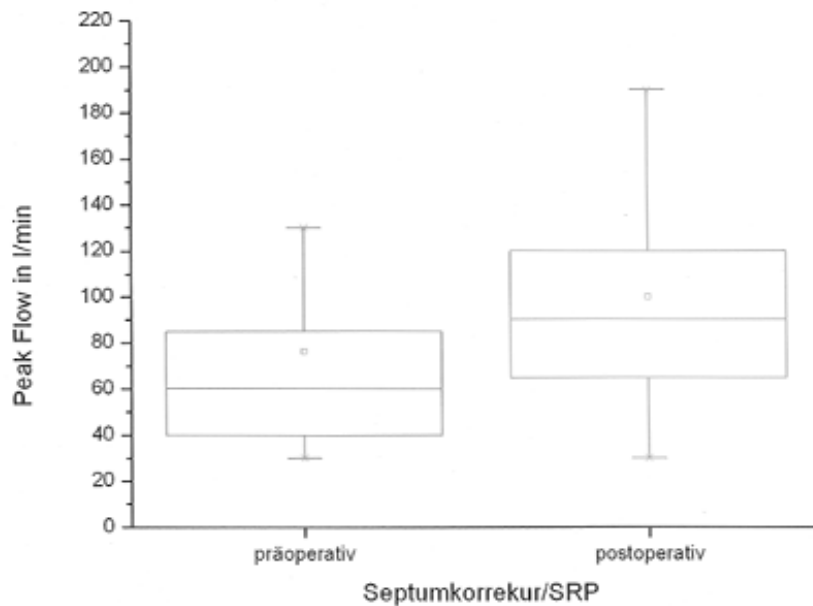


Abbildung 12: Die Abbildung zeigt die Werte der Peak-Flow- Metrie prä- und postoperativ nach drei Monaten bei den Testpersonen mit einer Septumkorrektur/SRP. Eine Zunahme des Volumens ist deutlich zu erkennen.

Vergleich Worterkennungsrate und Peak Flow- Metrie:

Die Korrelationsanalyse zwischen dem Mittelwert der prä- und postoperativen WR und der inspiratorischen nasalen Peak-Flow-Metrie zeigt eine hohe Korrelation ($r=1$) (Abbildung 14). Dies verdeutlicht, dass beide Parameter eine positive Korrelation zueinander haben, stark voneinander abhängig sind und jeweils eine Wertsteigerung aufweisen, eine Signifikanz in den Wertsteigerungen selbst ist jedoch nicht nachweisbar.

Die Korrelation beider Parameter zum Zeitpunkt t1 (präoperativ) liegt bei $r=0,23$, zum Zeitpunkt t3 (postoperativ nach drei Monaten) ist $r= 0,24$ - beide Analysen zeigen eine geringe, positive Korrelation.

Sowohl der Gesundheitszustand im Allgemeinen, als auch der Zustand der Stimme, beide Parameter anhand der klinischen Fragebögen ermittelt, hat sich nach dem Eingriff und der Heilungsphase verbessert oder blieb unverändert, Verschlechterungen wurden nicht angegeben (Abbildung 13).

Subjective score	General health		Nasal resonance	
	Preoperative	Postoperative	Preoperative	Postoperative
1-2	–	1	3	8
3-4	4	11	7	5
5-6	12	6	8	7
7-8	3	2	2	–
9-10	1	–	–	–

Abbildung 13: Prä- und postoperativer Vergleich der subjektiven Angaben zum Gesundheitszustand im Allgemeinen und dem Zustand der Stimme (Score von 1-10 mit 1 für exzellent)

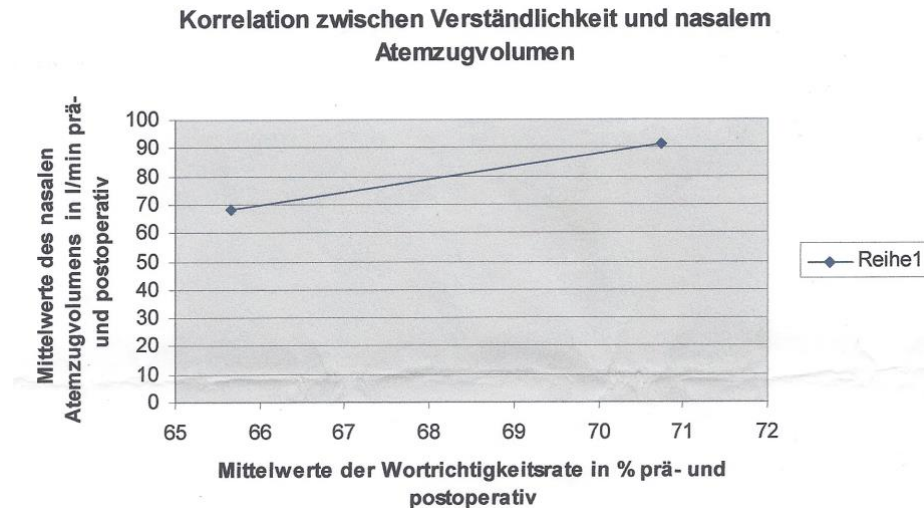


Abbildung 14: Korrelation zwischen Verständlichkeit und nasalem Atemzugvolumen. Die Abbildung zeigt den Mittelwert von WR und Atemzugvolumen prä- und postoperativ korreliert ($r=1$)

Diskussion:

Dass die luftgefüllten Nasennebenhöhlen des Gesichtsschädels einen entscheidenden akustischen Effekt zur menschlichen Sprache beitragen, ist eine oft diskutierte und immer wiederkehrende These. Bereits im 17. Jahrhundert wurden derartige Effekte von Bartholinus beschrieben. Er stellte fest, dass Individuen mit einer „defekten Sprache“ keine Nasennebenhöhlen aufwiesen.

Howell studierte die Maori auf Neuseeland und beobachtete eine „eigenartige, tote Sprache“ bei den Menschen, bei denen eine Unterentwicklung der Nebenhöhlen vorlag. Es gibt aber auch eine Reihe von Autoren, die keinen Zusammenhang zwischen der Sprachqualität des Menschen und den Nasennebenhöhlen in ihren Studien feststellen konnten.

Vom chirurgischen Standpunkt aus folgt der Sinuschirurgie eine Sprachänderung. Dies wird aber immer wieder kontrovers diskutiert, da einige Autoren keine postoperative Änderung der Stimmqualität nach endonasaler Chirurgie feststellen konnten, andere wiederum messbare akustische Unterschiede gefunden haben [9,26].

Es wird deutlich, dass die sehr unterschiedlichen Ergebnisse ein Problem der In- Vivo- Studien sind.

Mittlerweile konnte nachgewiesen werden, dass die Nasennebenhöhlen sehr wohl einen Einfluss auf den Stimmklang haben [17]. Ob sich dies tatsächlich auf die Verständlichkeit auswirkt, wurde bislang aber noch nicht nachgewiesen.

In dieser Studie wurde nun dafür ein objektives Verfahren angewendet, das bereits zuvor für andere Fragestellungen zur Verständlichkeit im Einsatz war.

Schuster et al verglichen die Mittelwerte der WR eines Teams aus erfahrenen Sprechern, der Sprache von Kindern mit Lippen-, Kiefer- und Gaumenspalten, der Sprache von Erwachsenen nach einer Laryngektomie und Erwachsenen mit malignen Tumoren im Mundbereich. Die Autoren fanden in dieser Studie eine signifikante Korrelation von -0.90 , $p > 0.01$ bei einem Multi-rater-Kappa drei erfahrener Logopäden von 0.48 [33], bzw. -0.84 , $p < 0.01$, mit einer Korrelation zwischen den drei bzw. vier erfahrenen Bewertern von 0.68 und 0.85 [35]. Bei Sprechstörungen nach der Behandlung von Mundhöhlenkarzinomen zeigten sich ähnliche Werte zur Übereinstimmung der perzeptiven Bewertung mittels dem neuen objektiven Verfahren mit einer Korrelation von -0.93 ; $p = 0.01$, bei einer Übereinstimmung der drei erfahrenen Sprecher von $0,58$ (kappa) [34].

Vorangegangene Arbeiten nutzten die Methode der Wahrnehmungsanalyse im Vergleich mit automatischer Spracherkennung. Die Studien ergaben eine sehr hohe Korrelation von $r=0.88$ (Pearson) oder $r=0.93$ (Spearman) zwischen der WR und dem Mittelwert bei ungeübten Hörern, die das Gehörte transkribierten [17].

Die Qualität der phonematischen Ausformung von Wörtern und somit die Ermittlung der Verständlichkeit dient als globale Bewertung der Sprechqualität [40]. In dieser Studie quantifizierten wir die Sprachverständlichkeit bei Patienten mit Beschwerden im Bereich der Nase und der Nasennebenhöhlen und der entsprechenden chirurgischen Therapie. Die Sprachverständlichkeit wurde in der vorliegenden Studie erstmals bei einer solchen Patientengruppe mit einer objektiven und apparativen Methode, der automatischen Spracherkennungstechnik, bestimmt.

Der Sprecher, das Sprachregister (Sprache beim Lesen, Spontansprache), das Vokabular, die grammatische Komplexität oder Perplexität (durchschnittliche Anzahl von Wörtern, die auf ein bestimmtes Wort innerhalb einer Sequenz folgen können) und die akustischen Eigenschaften des Aufnahmeegerätes stellen relevante Einflussfaktoren für die Qualität der Messung mittels automatischer Spracherkennung dar. Durch die Anwendung eines Standardtextes wie in dieser Studie „Der Nordwind und die Sonne“ (Anhang), und stabile Aufnahmebedingungen kann der Einfluss der meisten dieser Faktoren minimiert werden. Der Sprecher bzw. die Qualität des Gesprochenen bleiben als hauptsächlicher Faktor der Messung bestehen. Der Prozentsatz korrekt erkannter Wörter einer Wortfolge dient dabei als Maß für den Verständlichkeitsgrad [11].

Die ermittelten Worterkennungsraten sind natürlich nicht gleichbedeutend mit der Verständlichkeit. Der subjektive Faktor, also die Interpretation und Weiterverarbeitung von gehörter Sprache in den entsprechenden Hirnarealen ist ebenso ein entscheidendes Kriterium zur Erkennung wie die Stimmqualität, die phonetische und morpho-syntaktische Struktur, die Hintergrundgeräusche, die Lautstärke und die Sprechgeschwindigkeit von Gesprochenem selbst. Deutlich wird dies durch das Erschließen von gesprochener Sprache mithilfe des Kontextes und pragmatischer und prosodischer Elemente [11]. Allerdings ist eine rein subjektive Bewertung besonders bei zu erwartenden nicht-großen Unterschieden wie in dieser Studie nicht ausreichend reliabel. Dies zeigen auch die oben angeführten Werte zur Multirater-Varianz in den Studien von Paal [31]. Schuster [33;34;35]. und Windrich [44].

Die automatische Spracherkennung basiert auf der statistischen Modellierung von akustisch-phonetischem und linguistischem Wissen, der Schwerpunkt lag bei dieser Studie auf dem akustisch-phonetischen Hintergrund [26]. Das Programm PEAKS vermeidet durch Wertung jedes einzelnen Wortes eine unsystematische Verfälschung einer subjektiven Bewertung und belegt in einer Vorarbeit [14] die Übereinstimmung dessen mit der perzeptiven Beurteilung bei Einzelwörtern. Folglich würden die Ergebnisse einer Wiederholungsmessung am selben Sprachmaterial nicht voneinander abweichen. Die automatische

Spracherkennung kann daher als genaues objektives Maß für den Sprachverständlichkeitsgrad gelten. [11,14]

Die Verständlichkeit von Patienten mit Erkrankungen im Bereich der Nase und der Nasennebenhöhlen und deren chirurgischer Therapie wurde nun erstmals in dieser Studie mittels eines standardisierten Textes untersucht. Das Alter und das Geschlecht zeigten keinen signifikanten Einfluss auf die vorliegenden Messergebnisse. Eine Verfälschung der Ergebnisse durch veränderte Schalldruckpegel konnte ausgeschlossen werden. Die Dauer der Aufnahme wurde kontinuierlich geringer als Hinweis auf einen Lernprozess nach mehrmaligem Lesen des Textes. Dieser Effekt sollte in Anbetracht der Ergebnisse nicht wesentlich für die Unterschiede zwischen den einzelnen Untersuchungszeitpunkten sein. Zum Zeitpunkt t2 mit höchstem Trainingseffekt, zwei Tage nach dem ersten Lesen, ist die WR am geringsten. Dies hängt damit zusammen, dass ein vollständiger Verschluss der Nasenhaupthöhlen mit Ausschaltung der nasalen Resonanz doch zu relevanten Veränderungen der Sprechqualität und damit der WR führt. Mit der eingesetzten Methode ist das Phänomen der veränderten Nasalität also prüfbar.

In dieser Verlaufsstudie wurde die WR zu verschiedenen Zeitpunkten verglichen. Bei der ersten Sprachaufnahme einen Tag vor dem Eingriff wurde ein Wert ermittelt, welcher dann mit dem zweiten Wert, ermittelt am ersten postoperativen Tag, und dem Dritten, ermittelt nach einer ausreichenden Heilungsphase von drei Monaten, gemittelt verglichen wurde.

Wie erwartet, zeigte sich unmittelbar postoperativ eine signifikante Verminderung der Verständlichkeit, da die Patienten zu diesem Zeitpunkt mit Nasentamponaden versorgt waren. Der Vergleich der ermittelten Verständlichkeit vom präoperativen zum ausgeheilten postoperativen Zustand zeigte keine signifikanten Unterschiede. Von einer nasennebenhöhlenchirurgischen Therapie und auch von chirurgischen Therapieformen der Nase wie Septumkorrektur oder Conchotomie ist also keine signifikante Änderung der Sprachverständlichkeit zu erwarten. Dies zeigt auch die Betrachtung der Einzeldaten im Anhang. Die Veränderung der nasalen

Resonanzräume hat also einen wesentlichen Einfluss auf die Verständlichkeit. Diese ist aber nach der Heilungsphase nicht mehr nachweisbar.

Die nun vorliegende Aussage, belegt am Signifikanzniveau, ist von großer Bedeutung für Patienten, die einen Sprechberuf ausüben und auf eine gute Sprech- und Sprachverständlichkeit angewiesen sind. Nicht zuletzt ist das Sprechen nicht nur von beruflicher Relevanz, sondern stellt, wie bereits im einleitenden Teil näher erläutert, einen wesentlichen Teil der Kommunikation zwischen den Menschen und ihrer psychosozialen Integrität dar.

Fazit:

Die Spracherkennungstechnik konnte in dieser Studie, wie auch schon in vorangegangenen Arbeiten, zur objektiven Bewertung der Verständlichkeit bei Patienten mit Erkrankungen der Nase und der Nasennebenhöhlen sowie deren chirurgischer Therapie herangezogen werden.

Die postoperativ verminderte Worterkennungsrates bei iatrogen obstruierten Nasenhaupthöhlen zeigt direkt postoperativ die Angemessenheit der Methode für das beschriebene Problem auf.

Es besteht kein signifikanter Unterschied des Verständlichkeitsgrades zwischen den präoperativen und den postoperativen Ergebnissen der Patienten nach mehrmonatiger Heilungsphase. Von einem relevanten Einfluss auf die Sprechqualität ist durch Eingriffe an den Nasenhaupt- und Nebenhöhlen also nicht auszugehen. Details sollten an einem größeren Kollektiv überprüft werden.

Anhang :

Text im Wortlaut, der von den Patienten vorgelesen wurde:

Der Nordwind und die Sonne

Einst stritten sich Nordwind und Sonne, wer von ihnen beiden wohl der Stärkere wäre, als ein Wanderer, der in einen warmen Mantel gehüllt war, des Weges daherkam. Sie wurden einig, dass derjenige für den Stärkeren gelten sollte, der den Wanderer zwingen würde, seinen Mantel auszuziehen. Der Nordwind blies mit aller Macht, aber je mehr er blies, desto fester hüllte sich der Wanderer in seinen Mantel ein. Endlich gab der Nordwind den Kampf auf. Nun erwärmte die Sonne die Luft mit ihren freundlichen Strahlen, und schon nach wenigen Augenblicken zog der Wanderer seinen Mantel aus. Da musste der Nordwind zugeben, dass die Sonne von ihnen beiden der Stärkere war.

Tabelle 1: Worterkennungsrate WR bei 20 Patienten vor der Operation (t1), einen Tag nach der Operation (t2) und drei Monate nach der Operation(t3).

Patient	Geschlecht	Alter bei OP in Jahren	WR t1 in %	WR t2 in %	WR t3 in %
L.A.	weiblich	18	73,2	55,6	75,0
K.P.	weiblich	18	78,7	75,9	69,4
A.H.	männlich	21	71,3	62,0	64,8
D.F.	männlich	22	75,9	52,8	70,4
T.S.	männlich	22	75,0	65,7	64,8
Ö.Y.	männlich	22	75,9	44,4	71,3
C.F.	weiblich	24	71,3	45,4	73,2
G.M.	weiblich	24	59,3	48,2	64,8
M.G.	männlich	25	63,0	32,4	57,4
W.I.	männlich	25	74,1	74,1	75,0
B.B.	männlich	26	68,5	46,3	65,7
F.B.	männlich	28	71,3	47,2	70,4
C.W.	weiblich	35	72,2	74,1	68,5
J.D.	männlich	41	04,6	51,9	66,7
K.W.	männlich	44	50,9	43,5	59,3
G.B.	männlich	66	47,2	50,9	75,9
H.B.	weiblich	67	72,2	59,3	68,5
G.A.	weiblich	68	73,2	80,6	67,6
E.W.	weiblich	70	69,4	54,6	71,3
M.L.	männlich	93	65,7	48,2	59,3
MW ± Standardabweichung		38 ± 22	65,7 ± 16,5	55,7 ± 12,8	68,0 ± 5,3

Literaturverzeichnis:

1. Blomgren K, Simola M, Hytonen M, Pitkaranta A. Peak nasal inspiratory and expiratory flow measurements—practical tools in primary care? *Rhinology* 41 (4): 206-10, 2003.
2. Chen MY, Metson R. Effects of sinus surgery on speech. *Arch Otolaryngol Head Neck Surg* 123 (8): 845-52, 1997.
3. Clement PA. Committee report on standardization of rhinomanometry. *Rhinology* 22 (3): 151-5, 1984.
4. Clement PA, Gordts F. Consensus report on acoustic rhinometry and rhinomanometry. *Rhinology* 43 (3): 169-79, 2005.
5. Clement PA, van Dishoeck A, van de Wal J, Stoop P, Hoek T, van Strick R. Nasal provocation and passive anterior rhinomanometry (PAR). *Clin Allergy* 11 (3): 293-301, 1981.
6. Dames F, Maier A, Schützenberger A, Stelzle F, Holst A, Noth E, Eysholdt U, Schuster M. Intelligibility of children with bilateral and unilateral cleft lip and palate. *Laryngorhinootologie* 88 (11): 723-8, 2009.
7. Dang J, Honda K, Suzuki H. Morphological and acoustical analysis of the nasal and the paranasal cavities. *J Acoust Soc Am* 96 (4): 2088-100, 1994.
8. Deitmer T. [Modern functional diagnosis of the nose and paranasal sinuses]. *Eur Arch Otorhinolaryngol Suppl* 1: 1-71, 1996.
9. Dworkin JP, Marunick MT, Krouse JH. Velopharyngeal dysfunction: speech characteristics, variable etiologies, evaluation techniques, and differential treatments. *Lang Speech Hear Serv Sch* 35 (4): 333-52, 2004.
10. Ellegard E. Practical aspects on rhinostereometry. *Rhinology* 40 (3): 115-7, 2002.
11. Euler E, *Grundkurs Spracherkennung*, Vieweg-Verlag, Wiesbaden, 2006.
12. Friedrich G, Bigenzahn W, *Phoniatrie- Einführung in die medizinischen, psychologischen und linguistischen Grundlagen von Stimme und Sprache*, Verlag Hans Huber, Bern, 1995

13. Gales M, Pye D, Woodland P. Variance compensation within the MLLR framework for robust speech recognition and speaker adaptation. In Proc. ICSLP 1996; 3: 1832-1835, Philadelphia, USA, 1996.
14. Gallwitz F, Niemann H, Nöth E. Spracherkennung-Stand der Technik, Einsatzmöglichkeiten und Perspektiven. *Wirtschaftsinformatik* 41(6):538-547, 1999.
15. Grutzenmacher S, Mlynski G, Mlynski B, Lang C. [Objectivation of nasal swelling--a comparison of four methods]. *Laryngorhinootologie* 82 (9): 645-9, 2003.
16. Haderlein T, Eysholdt U, Riedhammer K, Nöth E, Rosanowski F. Automatisierung des Post-Laryngektomie-Tests PLTT. Aktuelle phoniatische Aspekte 2007
17. Haderlein T, Riedhammer K, Noth E, Toy H, Schuster M, Eysholdt U, Hornegger J, Rosanowski F. Application of automatic speech recognition to quantitative assessment of tracheoesophageal speech with different signal quality. *Folia Phoniatr Logop* 61 (1): 12-7, 2009.
18. Havel M, Becker S, Schuster M, Johnson T, Maier A, Sundberg J. Effects of functional endoscopic sinus surgery on the acoustics of the sinonasal tract. *Rhinology*. 2017 Mar 1;55(1):81-89.
19. Havel M, Ertl L, Bauer D, Schuster M, Stelter K, Sundberg J. Resonator properties of paranasal sinuses: preliminary results of an anatomical study. *Rhinology* 52: 178-182, 2014
20. Henningsson G, Kuehn DP, Sell D, Sweeney T, Trost-Cardamone JE, Whitehill TL. Universal parameters for reporting speech outcomes in individuals with cleft palate. *Cleft Palate Craniofac J* 45 (1): 1-17, 2008.
21. Hirschberg A. Rhinomanometry: an update. *ORL J Otorhinolaryngol Relat Spec* 64 (4): 263-7, 2002.
22. Hong KH, Kwon SH, Jung SS. The assessment of nasality with a nasometer and sound spectrography in patients with nasal polyposis. *Otolaryngol Head Neck Surg* 117 (4): 343-8, 1997.
23. Hosemann W, Gode U, Dunker JE, Eysholdt U. Influence of endoscopic sinus surgery on voice quality. *Eur Arch Otorhinolaryngol* 255 (10): 499-503, 1998.

24. Lam DJ, James KT, Weaver EM. Comparison of anatomic, physiological, and subjective measures of the nasal airway. *Am J Rhinol* 20 (5): 463-70, 2006.
25. Lang C, Grutzenmacher S, Mlynski B, Plontke S, Mlynski G. Investigating the nasal cycle using endoscopy, rhinoresistometry, and acoustic rhinometry. *Laryngoscope* 113 (2): 284-9, 2003.
26. Maier A, Honig F, Bocklet T, Noth E, Stelzle F, Nkenke E, Schuster M. Automatic detection of articulation disorders in children with cleft lip and palate. *J Acoust Soc Am* 126 (5): 2589-602, 2009.
27. Malm L, Gerth van Wijk R, Bachert C. Guidelines for nasal provocations with aspects on nasal patency, airflow, and airflow resistance. *International Committee on Objective Assessment of the Nasal Airways, International Rhinologic Society. Rhinology* 38 (1): 1-6, 2000.
28. Mathelitsch L, Friedrich G, *Die Stimme-Instrument für Sprache, Gesang und Gefühl*, Springer-Verlag Berlin, Heidelberg, 1995
29. Mlynski G, Grutzenmacher S, Plontke S, Mlynski B, Lang C. Correlation of nasal morphology and respiratory function. *Rhinology* 39 (4): 197-201, 2001.
30. Numminen J, Dastidar P, Rautiainen M. Influence of sinus surgery in rhinometric measurements. *J Otolaryngol* 33 (2): 98-103, 2004.
31. Paal S, Reulbach U, Strobel-Schwarthoff K, Nkenke E, Schuster M. Evaluation of speech disorders in children with cleft lip and palate. *J Orofac Orthop* 66 (4): 270-8, 2005.
32. Petursson M, Neppert J, *Elementarbuch der Phonetik*, Buske Verlag Hamburg, 1996.
33. Schuster M, Haderlein T, Noth E, Lohscheller J, Eysholdt U, Rosanowski F. Intelligibility of laryngectomees' substitute speech: automatic speech recognition and subjective rating. *Eur Arch Otorhinolaryngol* 263 (2): 188-93, 2006.
34. Schuster M, Lohscheller J, Hoppe U, Kummer P, Eysholdt U, Rosanowski F. Voice handicap of laryngectomees with tracheoesophageal speech. *Folia Phoniater Logop* 56 (1): 62-7, 2004.
35. Schuster M, Maier A, Haderlein T, Nkenke E, Wohlleben U, Rosanowski F, Eysholdt U, Noth E. Evaluation of speech intelligibility for children with

- cleft lip and palate by means of automatic speech recognition. *Int J Pediatr Otorhinolaryngol* 70 (10): 1741-7, 2006.
36. Simola M, Malmberg H. Sensation of nasal airflow compared with nasal airway resistance in patients with rhinitis. *Clin Otolaryngol Allied Sci* 22 (3): 260-2, 1997.
 37. Singh A, Patel N, Kenyon G, Donaldson G. Is there objective evidence that septal surgery improves nasal airflow? *J Laryngol Otol* 120 (11): 916-20, 2006.
 38. Stelzle F, Ugrinovic B, Knipfer C, Bocklet T, Noth E, Schuster M, Eitner S, Seiss M, Nkenke E. Automatic, computer-based speech assessment on edentulous patients with and without complete dentures - preliminary results. *J Oral Rehabil*.
 39. Szucs E, Clement PA. Acoustic rhinometry and rhinomanometry in the evaluation of nasal patency of patients with nasal septal deviation. *Am J Rhinol* 12 (5): 345-52, 1998.
 40. Vogt B, Maier A, Batliner A, Noth E, Nkenke E, Eysholdt U, Schuster M. [Numerische Quantifizierung der Verständlichkeit bei Schulkindern mit isolierten und kombinierten Gaumenspalten]. *HNO* 55 (11): 891-8, 2007.
 41. Watelet JB, Bachert C, Gevaert P, Van Cauwenberge P. Wound healing of the nasal and paranasal mucosa: a review. *Am J Rhinol* 16 (2): 77-84, 2002.
 42. Weinhold I, Mlynski G. Numerical simulation of airflow in the human nose. *Eur Arch Otorhinolaryngol* 261 (8): 452-5, 2004.
 43. Winck JC, Delgado JL, Almeida JM, Marques JA. Heated humidification during nasal continuous positive airway pressure for obstructive sleep apnea syndrome: objective evaluation of efficacy with nasal peak inspiratory flow measurements. *Am J Rhinol* 16 (3): 175-7, 2002.
 44. Windrich M, Maier A, Kohler R, Noth E, Nkenke E, Eysholdt U, Schuster M. Automatic quantification of speech intelligibility of adults with oral squamous cell carcinoma. *Folia Phoniatr Logop* 60 (3): 151-6, 2008.
 45. Wustenberg EG, Huttenbrink KB, Hauswald B, Hampel U, Schleicher E. [Optische Rhinometrie. Kontinuierliche und direkte Messung der Schwellung der Nasenschleimhaut mit allergener Provokation.

Echtzeitüberwachung der nasalen Provokationstests mit optischer Rhinometrie]. HNO 52 (9): 798-806, 2004.

Danksagung:

An erster Stelle gilt mein Dank meiner Doktormutter Frau Prof. Dr. med. Maria Schuster für ihre wissenschaftliche und persönliche Unterstützung während der gesamten Bearbeitungsphase meiner Dissertation.

Außerdem gilt mein Dank Frau Dr. med. Susanne Mayr, die mich in der Anfangsphase aktiv unterstützte.

Herrn Andreas Maier danke ich für die Bereitstellung und fachliche Beratung zu dem Spracherkennungsprogramm PEAKS, Frau Annika Ströbele und Julia von Ochsenstein für ihre Unterstützung in der Datenauswertung.

Ich danke Frau Katrin Rogler für ihr Mitwirken und ihr Engagement.

Besonders möchte ich an dieser Stelle auch meiner Familie für die unermüdliche Stärkung und Motivation danken.

Eidesstattliche Versicherung:

Name, Vorname: Eckstein, Katja

Ich erkläre hiermit an Eides statt,
dass ich die vorliegende Dissertation mit dem Thema

**Automatisches Spracherkennungssystem zur Quantifizierung von
Veränderungen der Nase und Nasennebenhöhlen auf die
Sprachverständlichkeit**

Selbständig verfasst, mich außer der angegebenen keiner weiteren Hilfsmittel bedient und alle Erkenntnisse, die aus dem Schrifttum ganz oder annähernd übernommen sind, als solche kenntlich gemacht und nach ihrer Herkunft unter Bezeichnung der Fundstelle einzeln nachgewiesen habe.

Ich erkläre des Weiteren, dass die hier vorgelegte Dissertation nicht in gleicher oder in ähnlicher Form bei einer anderen Stelle zur Erlangung eines akademischen Grades eingereicht wurde.

Katja Eckstein

Plauen, 04.11.2018

Unterschrift Doktorandin