Akustische Charakterisierung eines gemittelten Vokaltraktmodells basierend auf den MRT-Daten professioneller Tenöre

Judith Probst[1], Alexander Lodermeyer[1], Matthias Echternach[2], Stefan Kniesburges[3], Stefan Becker[1]

[1] *Universität Erlangen-Nürnberg, Lehrstuhl für Prozessmaschinen und Anlagentechnik, 91058 Erlangen, Deutschland, Email: pro@ipat.uni-erlangen.de (J.P), lo@ipat.uni-erlangen.de (A.L), sb@ipat.uni-erlangen.de (S.B.)*

[2] *Klinikum der Universität München, Klinik für Hals-Nasen-Ohren-Heilkunde, 81377 München, Deutschland, Email: matthias.echternach@med.uni-muenchen.de*

[3] *Universitätsklinikum Erlangen, Abteilung Phoniatrie und Pädaudiologie, 91054 Erlangen, Deutschland, Email: stefan.kniesburges@uk-erlangen.de*

Einleitung

Die menschliche Stimme wird durch die strömungsinduzierte Schwingung der Stimmlippen erzeugt. Dieses akustische Primärsignal, das sich aus der Grundfrequenz der Stimmlippenschwingung, den Höherharmonischen und dem Breitbandschall zusammensetzt, wird durch den Vokaltrakt gefiltert und so zu den unterschiedlichen Lauten einer Sprache geformt [1].
Da eine Untersuchung der zugrundeliegenden physikalischen Prozesse der Phonation am Menschen nur sehr eingeschränkt möglich ist, wird meist auf Modelle zurückgegriffen. Hierbei wird zwischen computerbasierten, exzidierten und künstlichen Modellen unterschieden. Einen Überblick über bestehende Modelle bieten [2, 3, 4, 5].
Der menschliche Vokaltrakt entspricht einem Resonator, in dem sich stehende Wellen und damit Resonanzfrequenzen ausbilden. Diese Resonanzfrequenzen, die von der Geometrie des Vokaltrakts abhängen, werden als Formanten bezeichnet. Dabei dienen die ersten beiden Formanten zur Unterscheidung der verschiedenen Vokale. Höhere Formanten hingegen bestimmen die Klangfarbe einer Stimme [6].
Da die akustischen Eigenschaften des Vokaltrakts primär durch seine Geometrie bestimmt werden, werden für die Charakterisierung des Vokaltrakts meist geometrische Daten basierend auf MRT-Aufnahmen verwendet [7, 8, 9]. Zur Vereinfachung werden die resultierenden Geometrien häufig durch einen gerade gestreckten Vokaltrakt mit einzelnen Abschnitten unterschiedlichen Querschnitts umgesetzt [7]. Diese Vereinfachung ist aus akustischer Sicht zulässig, da aufgrund des geringen Querschnitts die Schallausbreitung bis zu einer Frequenz von etwa 5 kHz als eindimensional betrachtet werden kann [10].
Bisherige Studien basieren bedingt durch die Geometriegewinnung anhand von bildgebenden Verfahren nur auf den Daten einzelner Probanden. Um den Einfluss des Vokaltrakts auf die Fluid-Struktur-Akustik-Interaktion der Phonation und die Abstrahlung des Schalls zu untersuchen, ist jedoch ein allgemeineres Modell, das keine starken individuellen Einflüsse aufweist, notwendig. Ziel der Studie ist es daher, ein vereinfachtes Modell des Vokaltrakts mit reduzierten individuellen Eigenschaften zu erstellen, das trotzdem auf menschlichen Geometrien basiert. Dies wird durch eine Mittelung über sechs vereinfachte individuelle Vokaltraktgeometrien basierend auf den MRT-Aufnahmen professioneller Tenöre realisiert.
Zudem wurde der Einfluss einer Geometrievereinfachung des Vokaltrakts auf die akustischen Eigenschaften in Experimenten mit rein akustischer Anregung untersucht. Um auch den strömungsmechanischen Einfluss zu berücksichtigen, werden die Vokaltraktmodelle in einem experimentellen Aufbau mit künstlichen Stimmlippen untersucht.

Methodik

Vokaltraktmodelle

Die verwendeten Vokaltraktgeometrien basieren auf den MRT-Aufnahmen professioneller Tenöre, die am Institut für Musikermedizin am Universitätsklinikum Freiburg von Echternach et al. durchgeführt wurden [8, 11]. Ausgehend von diesen Daten wird für jeden der sechs Tenöre jeweils ein Vokaltraktmodell mit realistischer Geometrie und mit vereinfachter Form erstellt. Die realistischen Geometrien werden durch eine manuelle Segmentierung der MRT-Daten in der Open-Source-Software *3D Slicer* erzeugt. Die vereinfachten Modelle basieren auf den Flächenfunktionen der Tenöre von Echternach et al. [11] und bilden gerade Modelle mit kreisförmigen Querschnitten.
Um den Einfluss des Vokaltrakts auf die Phonation und die zugrundeliegenden physikalischen Vorgänge zu untersuchen, wird neben den individuellen Vokaltraktmodellen ein allgemeineres, vereinfachtes Modell des Vokaltrakts erstellt. Diese gemittelte Vokaltraktgeometrie wird durch eine Mittelung über die Flächenfunktionen der sechs Tenorvokaltrakte berechnet. Dabei werden, wie in Abbildung 1 dargestellt, die Vokaltrakte auf eine Länge normiert und die Querschnittsflächen gemittelt.
Für die Untersuchungen der akustischen Eigenschaften werden die Vokaltraktmodelle additiv gefertigt.

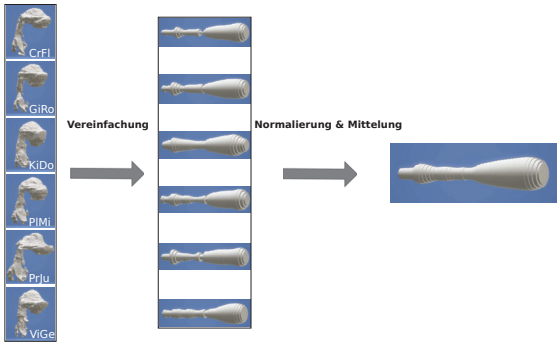

Abbildung 1: Vorgehensweise für die Erstellung des gemittelten Vokaltrakts

Experimenteller Aufbau

Die Untersuchung der akustischen Eigenschaften des Vokaltrakts erfolgt in zwei unterschiedlichen Aufbauten.
In einem Aufbau werden die Vokaltraktmodelle mit einem Druckkammerlautsprecher rein akustisch mit einem Stimmlippensignal angeregt.
Für die Untersuchung der akustischen Eigenschaften unter Berücksichtigung der Stimmlippenschwingung und des resultierenden Strömungsfelds im Vokaltrakt wird ein Versuchsaufbau basierend auf dem Aufbau von Kniesburges et al. [12, 13] und Lodermeyer et al. [14, 15] verwendet. Er setzt sich aus einer Einheit zur Bereitstellung des definierten Massenstroms, einem Silencer zur Dämpfung akustischer Fluktuationen in der Zuströmung und dem künstlichen Larynxaufbau zusammen. Der künstliche Larynx besteht aus dem subglottalen Kanal mit einer Querschnittsfläche von 15 mm x 18 mm und einem künstlichen Stimmlippenpaar aus Silikon. Diese werden, basierend auf dem M5-Modell von Scherer et al. [16] und Thomson et al. [17], aus dem Silikon *Ecoflex 30* (Smooth-On, USA) gegossen und haben einen E-Modul von E = 4,4 kPa. Die einzelnen Vokaltraktmodelle werden stromabwärts des künstlichen Larynx befestigt, siehe Abbildung 2.

Messaufbau und Auswertungsmethoden

Die Schalldruckmessungen werden in einem schallarmen Raum mit einem 1/2"-Freifeldmikron *Type 4189* (Brüel & Kjaer, Dänemark), das in einer Entfernung von 90 cm und unter einem Winkel von 45° vom Ausgang des Vokaltrakts platziert ist, durchgeführt. Das aufgenommene Signal wird mit dem *Nexus Conditioning Amplifier* (Brüel & Kjaer, Dänemark) verstärkt und mit dem Multifunktionsmodel *PXIe-6356* (National Instruments, USA) mit 44,1 kHz abgetastet.
Die Auswertung der Messdaten und die Berechnung der spektrale Leistungsdichte des akustischen Drucks erfolgt in *Matlab* (Mathworks, USA). Basierend auf den resultierenden Spektren werden die Formantfrequenz der einzelnen Vokaltrakte bestimmt.

Ergebnisse und Diskussion

Akustische Anregung

In Abbildung 3 sind die Schalldruckspektren für die Vokaltraktmodelle der sechs Probanden jeweils in realistischer, MRT-basierter und in vereinfachter Ausführung für eine rein akustische Anregung zu sehen. Jedes der Spektren weist die Spitzen der Grundfrequenz und ihrer Harmonischen aus. Darüber hinaus sind die Formanten zu erkennen, die durch gestrichelten Linien markiert sind. Als Referenzwerte sind außerdem die physiologischen Werte der Formanten, die parallel zu den Messungen von Echternach et al. [11] bestimmt wurden, eingezeichnet.
Der Vergleich zwischen den Spektren der MRT-basierten und vereinfachten Modelle zeigt, dass diese über den gesamten Frequenzbereich sehr gut übereinstimmen. Dies bestätigt die Ergebnisse von Story et al. [7] und Arnela et al. [10], dass trotz der Vereinfachung des Vokaltrakts durch ein gestrecktes Modell mit abschnittsweise variierendem Durchmesser die akustischen Eigenschaften des Vokaltrakts gut abgebildet werden können.

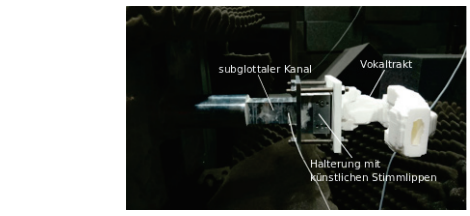

Abbildung 2: Versuchsaufbau mit subglottalem Kanal, künstlichen Stimmlippen und MRT-basiertem Vokaltrakt

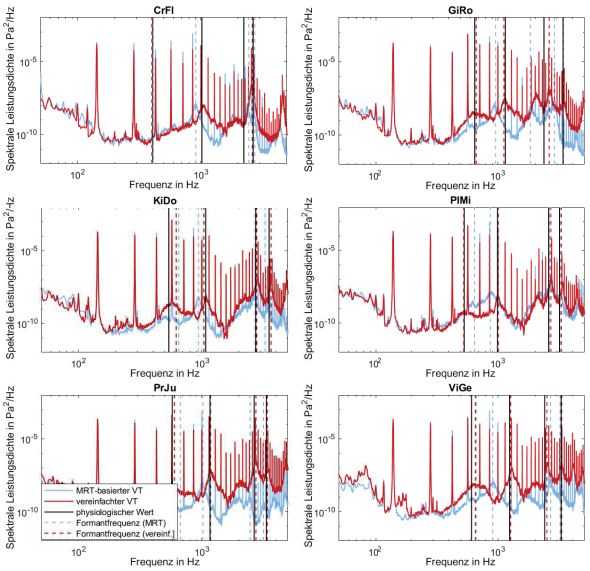

Abbildung 3: Spektrale Leistungsdichte der MRT-basierten (blau) und vereinfachten (rot) Modelle für alle Probanden bei einer akustischen Anregung mit einem Stimmlippensignal

Anregung mit oszillierenden Stimmlippen

In Abbildung 4 sind die Schalldruckspektren für die MRT-basierten und vereinfachten Modelle der sechs Probanden für eine Anregung mit oszillierenden Stimmlippen abgebildet. Vergleichbar zu Abbildung 3 sind auch hier die Grundfrequenz, die Harmonischen und die Formanten zu erkennen.
Der Vergleich zwischen den Spektren der MRT-basierten und vereinfachten Modelle zeigt für diese Anregung je-

doch, dass sich der Breitbandanteil des Schalldrucks deutlich unterscheidet. Diese Unterschiede deuten darauf hin, dass das Strömungsfeld und der daraus resultierende Strömungsschall im Vokaltrakt durch die Vereinfachung verändert werden.

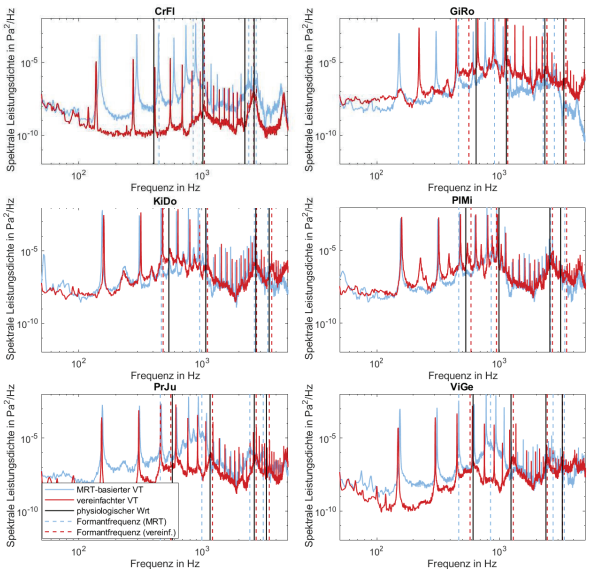

Abbildung 4: Spektrale Leistungsdichte der MRT-basierten (blau) und vereinfachten (rot) Modelle für alle Probanden bei einer Anregung mit oszillierenden Stimmlippen

Analyse der Formantfrequenzen

Die ermittelten Formantfrequenzen aller Vokaltraktmodelle im Vergleich zu den physiologischen Formanten sind in Abbildung 5 dargestellt. Ebenso wie für die Aufnahmen der Tenöre können auch für alle Modelle vier Formantfrequenzen im Bereich zwischen 50 Hz und 5 kHz bestimmt werden. Die mittleren Abweichungen der Formantfrequenzen zu den Referenzwerten liegen dabei für die MRT-basierten Modelle unter 15 % bei rein akustischer Anregung und unter 18 % bei einer Anregung mit Stimmlippen. Bei den vereinfachten Modellen ist die Abweichung kleiner 5 % bei rein akustischer Anregung und kleiner 8 % bei einer Anregung mit Stimmlippen. Der dritte und vierte Formant weisen dabei jeweils deutlich kleinere Abweichungen auf als der erste und zweite Formant.
Ein Vokal wird jedoch nicht durch eine diskrete Formantfrequenz charakterisiert, sondern unterliegt je nach Individuum einer gewissen Streuung [18]. Die vokalspezifischen Filtereigenschaften des Vokaltrakts werden daher ausreichend genau reproduziert.

Analyse des gemittelten Vokaltrakts

Der Einfluss der Mittelung auf die akustischen Eigenschaften des Vokaltrakts wird anhand von Abbildung 6 untersucht. Diese stellt die Schalldruckspektren des gemittelten Vokaltrakts und der vereinfachten Vokaltraktmodelle bei einer Anregung mit künstlichen Stimmlippen dar. Sie zeigt, dass die akustischen Eigenschaften sowohl im Bereich der Formanten, als auch im Breitbandanteil sehr gut durch den gemittelten Vokaltrakt dargestellt werden. Zur Überprüfung der Vokalverständlichkeit sind die ersten beiden Formantfrequenzen des gemittelten Vokaltrakts gemeinsam mit den physiologischen Werten in ein Formantendiagramm [18], siehe in Abbildung 7, eingetragen. Der Vergleich zeigt, dass die Formanten des gemittelten Vokaltrakts um 1 % vom Mittelwert der ersten Formanten und um 6 % vom Mittelwert der zweiten Formanten abweichen. Das zeigt, dass die Mittelung über die Querschnittsflächen entlang des gestreckten Vokaltrakts ein geeignetes Mittel ist, um die individuellen Einflüsse zu reduzieren und dabei die wesentlichen akustischen Eigenschaften des Vokaltrakts zu erhalten.

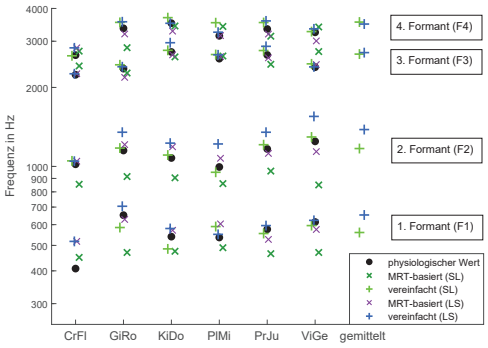

Abbildung 5: Formantfrequenzen aller Vokaltraktmodelle für beide Anregungsformen

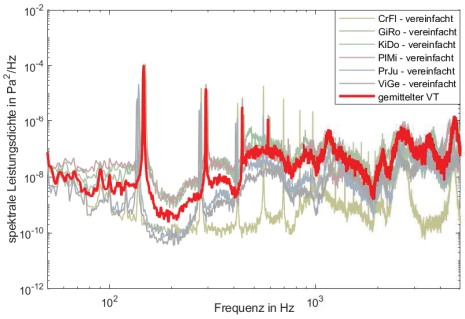

Abbildung 6: Spektrale Leistungsdichte der vereinfachten Vokaltraktmodelle im Vergleich zu dem gemittelten Vokaltraktmodell bei einer Anregung mit oszillierenden Stimmlippen

Zusammenfassung

In der vorliegenden Studie werden die akustischen Eigenschaften des Vokaltrakts in einem experimentellen Aufbau für realistische und vereinfachte Vokaltraktmodelle untersucht. Dabei wird zwischen einer rein akustischen Anregung und einer Anregung durch eine strömungsinduzierte Stimmlippenschwingung unterschieden. Um die individuelle Einflüsse zu reduzieren, wird darüber hinaus ein vereinfachter gemittelter Vokaltrakt erzeugt und untersucht.
Die Ergebnisse zeigen, dass die Vokalcharakteristika mit allen Vokaltraktmodellen reproduziert werden können. Allerdings wird entgegen der aktuellen Literatur eine Veränderung des Schalldruckspektrums durch die gängige Geometrievereinfachung erkannt, wenn die Schwingung der Stimmlippen und die daraus resultieren-

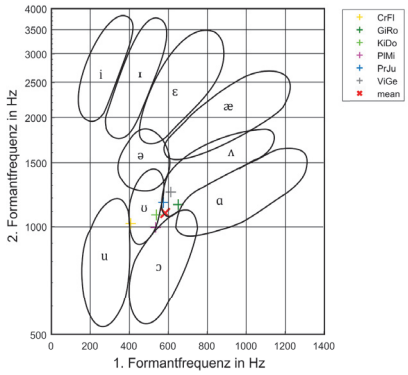

Abbildung 7: Das Formantendiagramm nach Peterson und Barney [18] zeigt die Formantfrequenzen des gemittelten Vokaltrakts im Vergleich zu den physiologischen Formanten der Tenöre.

de Strömung berücksichtigt wird. Dies deutet darauf hin, dass die Geometrieveränderug des Vokaltrakts zu einer Veränderung des Strömungsfeldes im Vokaltrakt und damit der strömungsinduzierten Schalls führt.
Die Validierung des gemittelten Modells anhand der physiologischen Referenzwerte zeigt, dass die akustischen Eigenschaften und der vorgegebene Vokalklang reproduziert werden. Der vorgestellte gemittelte Vokaltrakt ist also dazu geeignet, in zukünftigen Untersuchungen den Einfluss des Vokaltrakts auf die Phonation zu ermitteln.

Danksagung: Die Arbeit wurde durch die Else-Kröner-Fresenius Stiftung (Projektnr.: 2016_A78) im Rahmen des Erstantragsstellerprogramms finanziell gefördert.

Literatur

[1] I.R. Titze: Principles of voice production. Prentice Hall, 1994

[2] Kniesburges, S.; L Thomson, S.; Barney, A.; Triep, M.; Sidlof, P.; Horacek, J.; Brucker, C.; Becker, S. In vitro experimental investigation of voice production. Current bioinformatics 2011, 6, 305–322.

[3] Alipour, F.; Brucker, C.; D Cook, D.; Gommel, A.; Kaltenbacher, M.; Mattheus, W.; Mongeau, L.; Nauman, E.; Schwarze, R.; Tokuda, I.; others. Mathematical models and numerical schemes for the simulation of human phonation. Current Bioinformatics 2011, 6, 323–343.

[4] Mittal, R.; Erath, B.D.; Plesniak, M.W. Fluid dynamics of human phonation and speech. Annual Review of Fluid Mechanics 2013, 45.

[5] Döllinger,M.; Kaltenbacher,M. Preface: Recent Advances in Understanding the Human Phonatory Process. Acta Acustica united with Acustica 2016, 102, 195–208.

[6] Johan Sundberg. Vocal tract resonance in singing. The NATS Journal, 44:11-19, 1988.

[7] Story, B.H.; Titze, I.R.; Hoffman, E.A. Vocal tract area functions from magnetic resonance imaging. The Journal of the Acoustical Society of America 1996, 100, 537–554.

[8] Echternach, M.; Sundberg, J.; Markl, M.; Richter, B. Professional opera tenors' vocal tract configurations in registers. Folia Phoniatrica et Logopaedica 2010, 62, 278–287.

[9] Aalto, D.; Aaltonen, O.; Happonen, R.P.; Jääsaari, P.; Kivelä, A.; Kuortti, J.; Luukinen, J.M.; Malinen, J.; Murtola, T.; Parkkola, R.; others. Large scale data acquisition of simultaneous MRI and speech. Applied Acoustics 2014, 83, 64–75.

[10] Arnela, M.; Dabbaghchian, S.; Blandin, R.; Guasch, O.; Engwall, O.; Van Hirtum, A.; Pelorson, X. Influence of vocal tract geometry simplifications on the numerical simulation of vowel sounds. The Journal of the Acoustical Society of America 2016, 140, 1707–1718.

[11] Echternach, M.; Sundberg, J.; Baumann, T.; Markl, M.; Richter, B. Vocal tract area functions and formant frequencies in opera tenors' modal and falsetto registers. The Journal of the Acoustical Society of America 2011, 129, 3955–3963.

[12] Kniesburges, S.; Hesselmann, C.; Becker, S.; Schlücker, E.; Döllinger,M. Influence of vortical flow structures on the glottal jet location in the supraglottal region. Journal of Voice 2013, 27, 531–544.

[13] Kniesburges, S.; Lodermeyer, A.; Becker, S.; Traxdorf, M.; Döllinger, M. The mechanisms of subharmonic tone generation in a synthetic larynx model. The Journal of the Acoustical Society of America 2016, 139, 3182–3192.

[14] Lodermeyer, A.; Becker, S.; Döllinger, M.; Kniesburges, S. Phase-locked flow field analysis in a synthetic human larynx model. Experiments in Fluids 2015, 56, 77.

[15] Lodermeyer, A.; Tautz, M.; Becker, S.; Döllinger, M.; Birk, V.; Kniesburges, S. Aeroacoustic analysis of the human phonation process based on a hybrid acoustic PIV approach. Experiments in Fluids 2018, 59, 13.

[16] Scherer, R.C.; Shinwari, D.; De Witt, K.J.; Zhang, C.; Kucinschi, B.R.; Afjeh, A.A. Intraglottal pressure profiles for a symmetric and oblique glottis with a divergence angle of 10 degrees. The Journal of the Acoustical Society of America 2001, 109, 1616–1630.

[17] Thomson, S.L.; Mongeau, L.; Frankel, S.H. Aerodynamic transfer of energy to the vocal folds. The Journal of the Acoustical Society of America 2005, 118, 1689–1700.

[18] Peterson, G.E.; Barney, H.L. Control methods used in a study of the vowels. The Journal of the acoustical society of America 1952, 24, 175–184.