Numerische Studie zum Einfluss laryngealer Areale auf individuelle und allgemeine akustischen Eigenschaften des menschlichen Vokaltrakts bei gehaltenen Vokalen

Mario Fleischer[1], Alexander Mainka[1,2], Dirk Mürbe[1,2]

[1]*Charité - Universitätsmedizin Berlin, Klinik für Audiologie und Phoniatrie, Berlin,*
[2]*Hochschule für Musik Carl Maria von Weber, Stimmstudio für Stimmforschung, Dresden*

Kurzfassung: Im Rahmen dieser Studie wurden die akustischen Eigenschaften des Vokaltrakts für drei Probanden analysiert. Diese phonierten sowohl singend als auch sprechend die Eckvokale /a/, /i/ und /u/ im MRT. Auf Basis der ermittelten 3D-Bildaufnahmen wurde der Vokaltrakt segmentiert und die abgeleiteten Geometriemodelle mittels der Finiten-Elemente-Methode einer numerischen Analyse im Frequenzbereich unterzogen. Erweitert dazu wurden zusätzlich gezielt geometrische Manipulationen der FE-Modelle vorgenommen: das gezielte Entfernen laryngealer Areale und der systematische Austausch des unteren Larynx zwischen sängerischer und sprecherischer Einstellung. Die insgesamt 90 Modelle wurden hinsichtlich der Formantfrequenzlage sowie diverser Verstärkungseigenschaften untersucht. Unter anderem zeigt sich, dass 1.) die Vokalformanten F1 und F2 auch bei größeren Manipulationen sehr stabil hinsichtlich ihrer Lage im Frequenzbereich sind, und 2.), größere Manipulationen hauptsächlich einen Verlust der Verstärkungseigenschaften im Bereich > 2 kHz nach sich ziehen aber auch die Anschlußimpedanzen des Vokaltrakts deutlich ändern. Diese Ergebnisse sind innerhalb der analysierten Kohorte unabhängig vom gewählten Vokal und der Stimmgattung und erlauben Rückschlüsse auf die Effizienz supraglottischer akustischer Quellen die auch im Rahmen einer klinischen Therapie eine Relevanz besitzen.

Einleitung

Als Vokaltrakt (VT) bezeichnet man die luftgefüllte Kavität zwischen den Stimmlippen und der Mundöffnung. Im Sinne einer postulierten Quelle-Filter-Theorie nach Fant [1] fungiert der VT als akustischer Resonator der eine frequenzabhängige Verstärkung des primären Stimmschalls bewirkt und für die menschliche Stimmgebung unabdingbar ist. Anderseits ist bekannt, dass die Anschlussimpedanz des VT die Schwingungsfähigkeit der Glottis durchaus signifikant beeinflusst, d.h. Quelle und Filter nicht unabhängig voneinander agieren [2].
Für diese Überlegungen sind dabei die individuellen globalen [3, 4] und lokalen geometrischen Eigenschaften des VT [5, 6, 7, 8], die Ankopplung mit der umgebenden Luft [9] und den umgebenden Weichteilstrukturen [10, 11] als auch die Rückwirkung auf die Glottis [12] von Relevanz.
Um ein möglichst vollständiges Verständnis über die akustischen Effekte sowie die Interaktionsprozesse zwischen der Luft und den umgebenden mechanischen Strukturen zu erhalten sind Kenntnisse über die morphometrische Komplexität des VT unabdingbar. Dafür stehen verschiedene bildgebende Verfahren wie Röntgenaufnahmen/Computertomografie (CT) [siehe z.B. 13, 14, 15] oder statische und dynamische Magnetresonanztomografieverfahren (MRT) [siehe

z.B. 16, 17, 18] zur Verfügung, die sich hinsichtlich Strahlenbelastung, Genauigkeit und Aufnahmedauer teilweise deutlich voneinander unterscheiden.
Die Analyse der resultierenden Bilddaten kann auf verschiedene Art und Weise geschehen. Eine Möglichkeit besteht in der Untersuchung von 1-dimensionalen oralen, velaren, pharyngealen und glottalen Maßen [14, 19, 20] z.B. im Midsaggitalschnitt und deren statistische Korrelation zu verschiedenen akustischen Kriterien wie z.B. Tonhöhe und Stimmgattung. Die Ermittlung von Flächenfunktionen als 2-dimensionales Maß zur Beschreibung der Änderung der Querschnittsfläche in Abhängigkeit zur Längskoordinate des VT läßt dagegen – unter der Annahme ebener Wellenausbreitung – eine Berechnung der akustischen Übertragungscharakteristik entlang des VT inkl. der Bestimmung von Resonanzen (Formantfrequenzen) und Bandbreiten näherungsweise zu [13, 21, 22]. Eine Erhöhung der Analysegenauigkeit für den kompletten Sprachbereich ist jedoch immer mit der 3-dimensionalen (3D) Betrachtung des VT verbunden [23, 6, 10, 24, 25]. Die akustische Analyse von 3D-Daten ist allerdings auch damit gekoppelt, dass 1.) – im Rahmen numerischer Modelle – die Rechenzeit von einer Echtzeitanalyse weit entfernt ist und 2.) ein einfacher Zusammenhang zwischen einer Länge, Fläche oder einem Volumen im 3D und dessen direkten Einfluß auf die (Strömungs-)Akustik immer nur ansatzweise angegeben werden kann [7].

Unter Berücksichtigung der aufgeführten Vor- aber auch Nachteile wurden im Rahmen des vorliegenden Artikels 3D-Modelle des VT erstellt und numerisch berechnet und dabei keine explizit geometrischen Maße sondern Substrukturen im Sinne von definierten Arealen (speziell des unteren VT) variiert bzw. manipuliert und der Einfluß auf die akustischen Übertragungscharakteristiken und der Rückkopplungseigenschaften auf die Glottis untersucht. Am Beispiel gehaltener Vokale wird weiterhin der Unterschied zwischen sängerischer und sprechhafter Phonation direkt bzw. im Rahmen hybrider Modelle (Larynxaustausch) herausgestellt.

Material und Methoden

Datenerhebung und Modellerstellung

Für diese Studien wurden von drei Gesangsstudenten der Hochschule für Musik Carl Maria von Weber, Dresden, Deutschland (Stimmgattungen: Bariton, Bass und Tenor) MRT-Aufnahmen (3T-MRI; Magnetom Trio Trim, Siemens Medical Solutions, Erlangen, Germany) für die gehaltenen Eckvokale /a/, /u/ und /i/ sowohl in gesungener als auch sprechhafter Phonation (Phonationsdauer 9.2 s) erstellt [7]. Das MRT wurde mittels 12-Elemente-Kopf-Hals-Spule unter Nutzung einer 3D-volumeninterpolierten Sequenz mit 1.22 ms/4.01 ms Echo- bzw. Wiederholungszeit und einem Umkehrwinkel von 9° realisiert. Das Blickfeld betrug 300×300 mm^2 (288×288 Pixel^2) und 52 sagittale Schnitte mit einem Abstand von 1.8 mm wurden erstellt. Alle Untersuchungen waren im Einklang mit einem positiven Votum der zuständigen Ethikkommission an der Medizinischen Fakultät Carl Gustav Carus an der TU Dresden (EK153042011).

Die generierten Bilddaten wurden in ITK-SNAP (www.itksnap.org) segmentiert und als trianguliertes Gitter exportiert. Die Zähne wurden durch einen 3D-Scan (NextEngine, Santa Monica, California, USA) der individuellen Gipsabdrücke in ein Oberflächengittermodell überführt, durch nachträgliche Voxalisation mit den VT-Modellen verbunden und diese letztendlich in ein Volumengitter überführt (Abbildung 1). Details und weiterführende Informationen zur Modellerstellung sind in Fleischer et al. [25] zu finden. Im Ergebnis liegen 18 Modelle für die drei Probanden (je drei Vokale und zwei Phonationsmodi) vor. Neben den initialen Geometrien (Abbildung 2a) wurden des weiteren alle 18 Modelle dahingehend manipuliert, dass verschiedene supraglottische Areale abgeschnitten wurden. Im Detail wurden dabei der *Sinus Morgagni* (Abbildung 2b), der Endolarynx kaudal der Taschenfalten (Abbildung 2c) sowie der

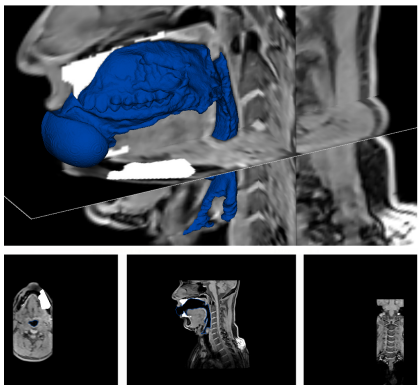
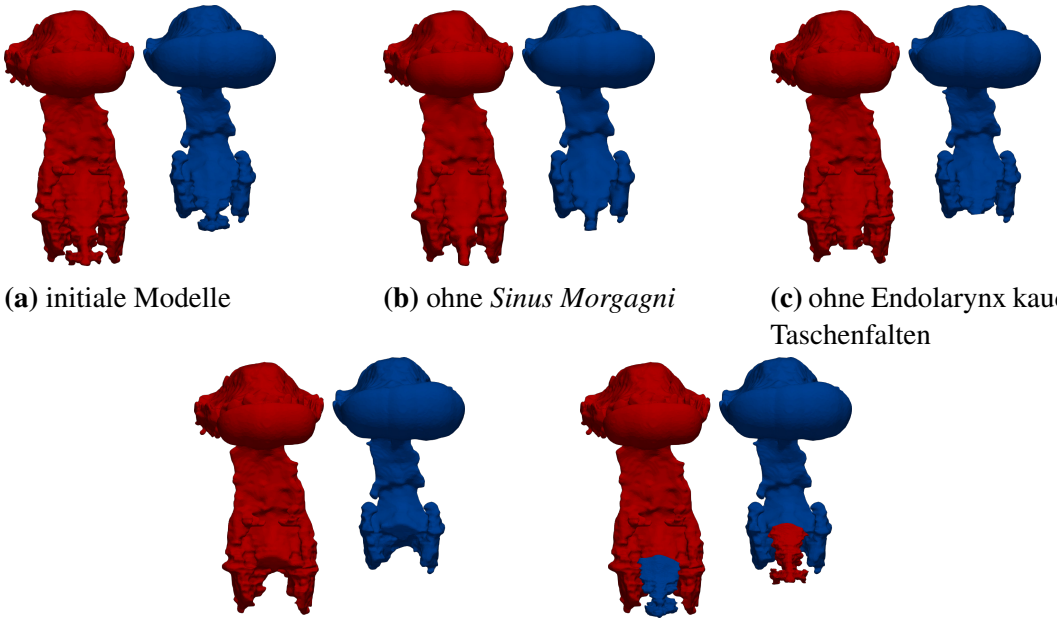


Abbildung 1 – MRT-Aufnahme (Graustufen) inkl. nachträglich implementierter Zahnmodelle (weiß) überlagert mit dem detaillierten Oberflächengittermodell (blau) des VT am Beispiel /a/ (Bass, sprechhaft).



(a) initiale Modelle

(b) ohne *Sinus Morgagni*

(c) ohne Endolarynx kaudal der Taschenfalten

(d) ohne Larynx (komplett)

(e) Hybridisierung des kompletten Larynx

Abbildung 2 – Modellhierarchie am Beispiel /a/ (Bass) für den gesungenen (rot) und den sprechhaften Phonationsmodus (blau)

komplette Larynx im Modell entfernt (Abbildung 2d). Weiterhin wurde zwischen den Modellen mit sängerischer und sprechhafter Einstellung der komplette Larynx kreuzweise vertauscht, so dass hybride Modelle vorlagen (Abbildung 2e). In Summe bilden somit 90 Modelle die zu analysierende Kohorte.

Numerische Modellierung

Im Rahmen dieser Studie erfolgt ausschließlich eine numerische Berechnung der akustischen Übertragungscharakteristik aller Modelle. Dabei wird die quellfreie, skalare Helmholtz-Gleichung

$$-(\kappa^2 + \nabla^2)\, P(\vec{\mathbf{x}}, \omega) = 0 \tag{1}$$

für den komplexwertigen akustischen Druck P in Abhängigkeit vom Ort $\vec{\mathbf{x}}$ und der Kreisfrequenz ω gelöst. $\kappa = \omega/c$ ist die Wellenzahl, $c = 343$ m/s die Schallgeschwindigkeit für eine Temperatur von 20° Celsius. Weiterhin wurden die Randbedingungen am geöffneten Mund, an

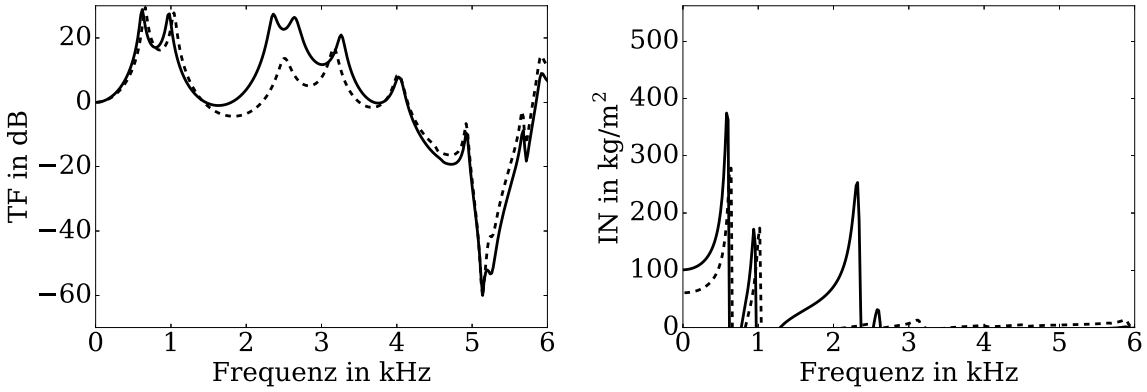


Abbildung 3 – Transferfunktionen (links) und Inertanzen (rechts) am Beispiel /a/, Bariton, sprechhaft für das initiale Modell (Volllinien) und das zugehörige Modell komplett ohne Larynx (Strichlinien).

der Modellwand und der Glottis mit

$$\nabla P_{\text{Mund}} \cdot \vec{n} = -j\kappa \frac{\rho c}{Z_{\text{Mund}}} P_{\text{Mund}};\ \nabla P_{\text{Wand}} \cdot \vec{n} = -j\kappa \frac{\rho c}{Z_{\text{Wand}}} P_{\text{Wand}};\ \nabla P_{\text{Glottis}} \cdot \vec{n} = -j\omega\rho V_{\text{Glottis}} \quad (2)$$

vorgegeben. Z_{Mund} ist abhängig von der Mundöffnungsfläche und entspricht in guter Näherung der Abstrahlimpedanz einer steifen Zylinderfläche gegen einen unendlichen großen Raum [26]. Um einen alleinigen Einfluß der Geometrie des VT zu quantifizieren wurde die Wandimpedanz $Z_{\text{Wand}} = 500 \cdot \rho c$ und die Dichte zu $\rho = 1.2$ kg/m^3 gesetzt. Als Eingangsgröße wurde die Normalenkomponente der akustischen Volumenschnelle V_{Glottis} an der Glottis vorgegeben. Dabei muss beachtet werden, dass für die Modelle ohne den Endolarynx kaudal der Taschenfalten (Abbildung 2c) und ohne den kompletten Larynx (Abbildung 2d) die Glottisfläche der neuen Schnittfläche entspricht. Gleichungen 1 und 2 wurden in der freien Finite-Elemente-Bibliothek FEniCS (www.fenicsproject.org) gelöst. Weiterführende und ergänzende Informationen zur numerischen Modellierung können Fleischer et al. [25] entnommen werden.
Im Postprocessing wurde die Volumenstrom-Transferfunktion

$$\text{TF} = \vec{V}_{\text{Mund}} \cdot \vec{n} / V_{\text{Glottis}} \quad (3)$$

inkl. der ersten fünf Formantfrequenzen F1-F5 (Resonanzfrequenzen in TF) und die Inertanz

$$\text{IN} = \Re\{P_{glottis}/(j\omega V_{\text{Glottis}})\} > 0 \quad (4)$$

des VT berechnet (Abbildung 3). Um den Einfluß der Modellmanipulationen zu quantifizieren, wurden diverse Differenzfunktionen der Transfer- und Inertanzfunktionen untersucht. Dabei wurde jeweils die TF eines manipulierten Modells von dem initialen Modell (z.B. $\Delta\text{TF}^{\text{ohne Larynx}} = \text{TF}^{\text{initial}} - \text{TF}^{\text{ohne Larynx}}$), das hybridisierte Modell von dem initialen Modell ($\Delta\text{TF}^{\text{hybrid}} = \text{TF}^{\text{initial}} - \text{TF}^{\text{hybrid}}$) sowie die TF der sprechhaften Einstellung von der TF des gesungenen Konfiguration ($\Delta\text{TF}^{\text{gesungen/sprechhaft}} = \text{TF}^{\text{gesungen}} - \text{TF}^{\text{sprechhaft}}$) subtrahiert (IN analog). Diese Berechnungen wurden für alle Vokale und Probanden getrennt durchgeführt, sodass insgesamt 2×81 Differenzfunktionen vorliegen. Um eine Datenreduktion zu erreichen wurden von diesen Differenzfunktionen ausschließlich die positiven Anteile extrahiert (entspricht dem Verstärkungsanteil) und deren Flächen (z.B. $\Delta\text{TF}_{\text{Gain}}^{\text{ohne Larynx}}$ bzw. $\Delta\text{IN}_{\text{Gain}}^{\text{ohne Larynx}}$ und zugehörigen geometrischen Schwerpunkte auf der Frequenzachse f_{Gain} ermittelt (siehe Abbildung 4 zur Illustration).

Ergebnisse und Diskussion

Formantanalyse

Die Ergebnisse der Formantanalyse können Abbildung 5 entnommen werden. Unabhängig von Proband (Stimmgattung), der Phonationsart aber auch der Art der Modellmanipulation stimmen

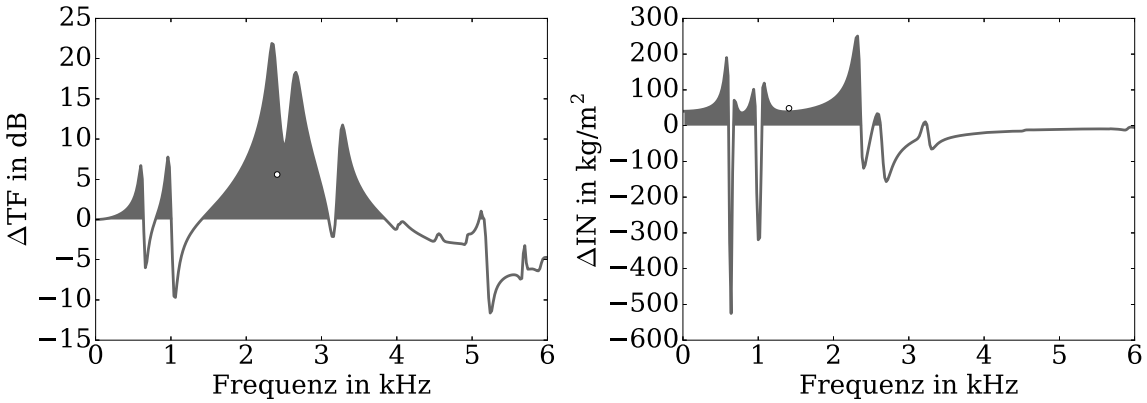


Abbildung 4 – Differenzen (Volllinien) der Transferfunktionen (ΔTF, links) und Inertanzen (ΔIN, rechts) am Beispiel /a/, Bariton, sprechhaft (initiales Modell minus das zugehörige Modell ohne Larynx, siehe auch Abbildung 3). Eine Untermenge bilden dabei alle Werte die größer als 0 sind. Die jeweiligen Summen der Einzelflächen (grau) weisen eine spezifische Größe $\Delta TF_{Gain}^{ohne\,Larynx}$ bzw. $\Delta IN_{Gain}^{ohne\,Larynx}$ und einen spezifischen geometrischen Schwerpunkt auf der Frequenzachse f_{Gain} auf (Kreis) und können als Maß einer geometriebedingten Verstärkung der Übertragung bzw. der Glottisschwingung interpretiert werden.

die ersten beiden Formanten mit der zu erwartenden Lage im Formantdreieck überein. Die beste Übereinstimmung ist dabei beim /u/ zu beobachten wo sowohl F1 als auch F2 sehr gut mit publizierten Werten übereinstimmen. Tendenziell sind aber auch für /a/ und /i/ ähnliche Schwankungsbreiten wie in den publizierten Daten zu verzeichnen. Prinzipiell ist zu beobachten, dass F1 und F2 kaum von der Vielzahl der betrachtenden Freiwerte beeinflußt wird. Das zeigen auch die Streuungen ΔF in Abbildung 5. Um eine Vergleichbarkeit über die Kohorte zu errreichen wurden für ΔF jeweils die Formantfrequenzen für einen Proband und Vokal bestimmt (je 10) und die zugehörigen Mediane subtrahiert. Damit werden die jeweiligen Veränderungen in den Formantfrequenzlagen durch Modelländerung ersichtlich. Es zeigt sich, analog zu den bereits im Vokaldreieck gemachten Beobachtungen, dass F1 und F2 relativ gering variieren (160 Hz bzw. 280 Hz). Bei den höheren Formanten F1 bis F5 steigt die Streuung dagegen auf 480, 1020 bzw. 1730 Hz deutlich an.

Diese Ergebnisse sind insofern beachtenswert, dass z.B. eine deutliche Längenänderung in der Größenordnung von 10^{-2} m, wie sie z.B. beim Entfernen des Larynx auftritt, die großen stehenden Wellen >0.3 m bei Frequenzen <1 kHz kaum beeinflußt. Das hängt u.a. damit zusammen, dass aufgrund des hohen Kalibersprungs am Larynxeingang die stehenden Wellen eine Knotenfläche aufweisen. Dagegen zeigen sich deutliche Auswirkungen der Modellmanipulationen bei den höheren Formantlagen. Das wiederum zeigt, dass der, vor allem für die klassische Gesangsstrategie wichtige und aber auch das Timbre kennzeichnende, hohe Frequenzbereich >2 kHz durch den unteren Larynx determiniert wird. Tatsächlich bestehen Hinweise, dass klassische Sänger Einstellungsvorgänge gerade in diesem Bereich zur Ausbildung des Sänger-Formant-Clusters nutzen [1, 7, 27].

Verstärkung

Im Gegenteil zur Formantanalyse stellen die eingeführten Verstärkungsmaße ein globales, den vollständigen untersuchten Frequenzbereich berücksichtigendes Maß dar. Hierbei muss deutlich unterschieden werden zwischen der Auswertung der Differenzen der TF und der IN. Erstere sind ein Maß dafür, welche Erhöhung der Verstärkung des als vom VT unabhängig angenommenen glottalen Quellspektrums resultiert. Letztere sind dagegen ein Anhalt dafür inwieweit der angekoppelte Vokaltrakt die Schwingungsfähigkeit der Glottis (und damit des Energieeintrages in den VT) unterstützt. Die Ergebnisse für beide Analysen können Abbildung 6 entnommen wer-

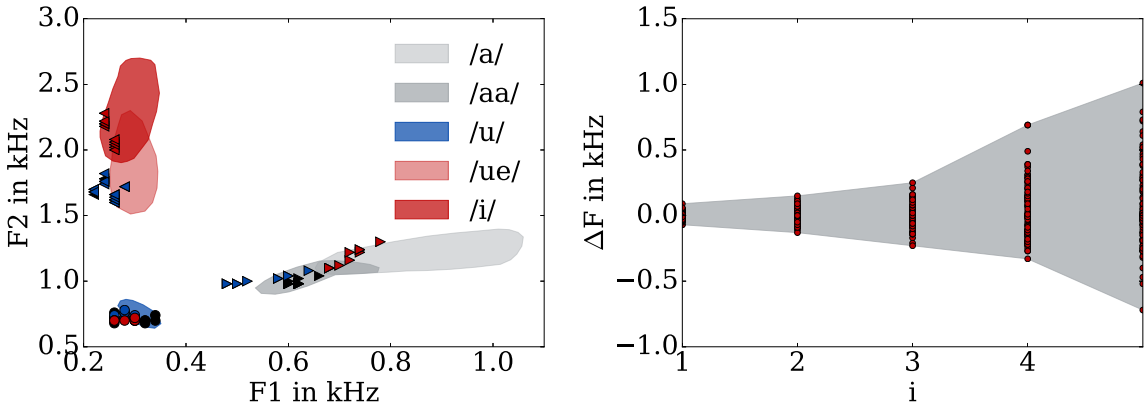


Abbildung 5 – Formantfrequenzdreieck (links) für initiale als auch alle manipulierten Modelle (Symbolfarben stehen für Stimmgattung; schwarz–Bariton, blau–Bass, rot–Tenor) im Vergleich zu Standardwerten (farbig laut Legende, aus [28]). Ergänzend ist die normalisierte Schwankung ΔF (rote Kreise) von Fi mit $i \in [1:5]$ innerhalb der gesamten Kohorte inkl. einer Einhüllenden (grau) dargestellt (rechts).

den.

Die Auswertung der ΔTF_{Gain} zeigt, dass der Schwerpunkt der Verstärkung im Bereich zwischen ca. 2 und 4.8 kHz und damit im Timbrebereich liegt. Weiterhin ist zu sehen, dass die Höhe der Schnittmanipulation systematisch mit der Höhe der Verstärkung zusammenhängt und bei Abwesenheit des vollständigen Larynx maximal wird [6]. Indirekt kann damit abgeleitet werden, dass akustische Quellen, die z.B. bei supraglottischer Stimmgebung (Taschenfaltenstimme) evoziiert werden oder aber auch nach Kehlkopfentfernung (z.B. bei Tumorbefall) durch Kehlkopfprothesen entstehen nicht in der Lage sind das ursprüngliche, natürliche Timbre zu aktivieren ohne es mit einer signifikanten Erhöhung der eingetragenen Energie in diesem Frequenzbereich zu kompensieren. Im Hinblick auf den Einfluß von Konfigurationsänderungen des Larynx auf die Verstärkung während Phonation im klassischen Gesangsmodus im Vergleich zur sprechhaften Phonation kann festgehalten werden, dass die Verstärkung in ähnlichen Größenordnungen wie bei der Entfernung diverser Areale erfolgen kann. Das ist nicht unbedingt zu erwarten, da die geometrischen Änderungen weitaus geringer sind als beim partiellen Abschneiden der Modelle. Kleinere Unterschiede zwischen den Manipulationsarten sind in den Frequenzbereichen zu sehen. Während Abschneiden diverser Areale sich primär in etwas tieferen Frequenzbereichen zwischen ca. 2.2 und 3.8 kHz auswirkt, liegen die Unterschiede zwischen den Phonationsmodi im Bereich von etwa 2.1 bis 4.8 kHz.

Die Auswertung ΔIN_{Gain} zeigt dagegen ein etwas geändertes Bild. Zwar sind auch hier die Auswirkungen primär im Bereich zwischen 2 und 4.5 kHz wirksam, allerdings zeigen die Modelle wo eine Resektion vorgenommen wurde keine eindeutige Staffelung hinsichtlich der Höhe der Verstärkung, jedoch können sie ebenfalls deutlich voneinander separiert werden. Weiterhin ist ersichtlich, dass der sängerische Phonationmodus hinsichtlich einer verbesserten Schwingungsfähigkeit der Glottis zu einer Verstärkung in einer ähnlichen Größenordnung führt wie bei einer Abstinenz des *Sinus Morgagni*.

Die Analysen sowohl für ΔTF_{Gain} als auch ΔIN_{Gain} in Abbildung 6 sind weder von der Stimmgattung noch vom phonierten Vokal abhängig (nicht dargestellt). Die Daten legen außerdem nahe, dass bestimmte anatomische Kehlkopfkonfigurationen bzw. die Fähigkeit diese Einstellungsvorgänge zu erzeugen günstig für das potentielle stimmliche Leistungsvermögen sind.

Aufgrund des integralen Charakters der Auswertungen (Flächenmaße anstatt frequenzspezifischer Einzelmaße, wie z.B. Maximalwerte in Abbildung 4) kann zusätzlich davon ausgegangen werden, dass diese ein hohes Maß an Robustheit bzw. Insensitivität gegenüber Variationen und Streuungen in den Modelldaten haben.

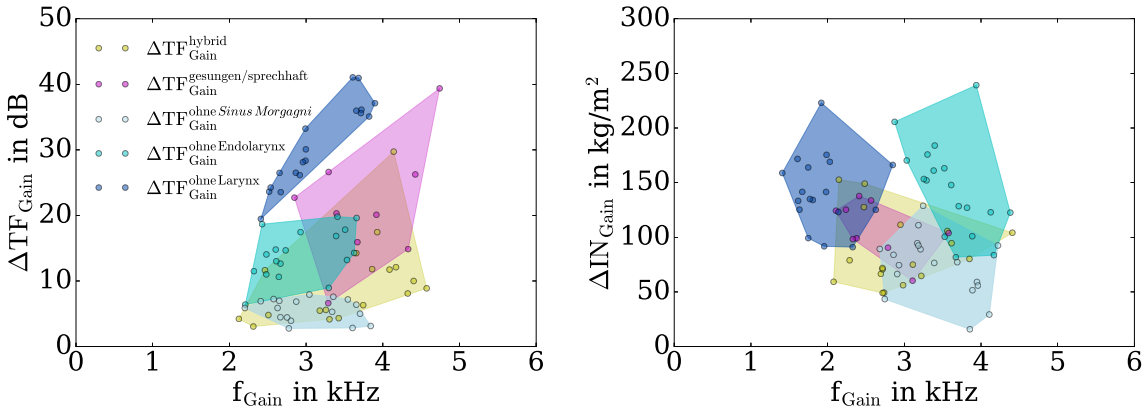


Abbildung 6 – Bereich aller TF_{Gain} (links) und IN_{Gain} (rechts) über den jeweiligen Flächenschwerpunkten f_{Gain} getrennt nach der Art der Manipulation laut Legende.

Zusammenfassung

Die Auswertung dieser umfangreichen Modellstudie zeigt, dass Geometrieänderungen im unteren VT die akustischen Eigenschaften primär höherfrequent beeinflussen. Da die beobachtbaren Verstärkungen im Sänger- bzw. Sprecherformantbereich liegen ist die anatomische sowie die aktiv einstellbare Konfiguration des Kehlkopfes als unabhängiger Einflußfaktor auf das individuelle stimmliche Leistungsvermögen zu sehen. Neben der etablierten Formantanalyse wurden in dieser Studie durch die Auswertung von Differenzfunktionen neue skalare Maße eingeführt, die sowohl Aussagen über Verstärkungseigenschaften diverser Regionen im VT als auch Aussagen über die Kopplungseigenschaften des VT mit der Glottis zulassen und aufgrund ihres integralen Charakters robust gegenüber Detailänderungen bzw. kleineren Schwankungen sind.

Literatur

[1] FANT, G.: *Acoustic theory of speech production*. The Hague, Mouton & Co. N.V., 1960.

[2] TITZE, I. R. und B. H. STORY: *Acoustic interactions of the voice source with the lower vocal tract*. *J Acoust Soc Am*, 101, S. 2234–2243, 1997.

[3] YAN, N., M. L. NG, M. M. K., und T. H. TO: *Vocal tract dimensional characteristics of professional male and female singers with different types of singing voices*. *Int J Speech-Lang Pathol*, 2013.

[4] LAMMERT, A. C. und S. S. NARAYANAN: *On short-time estimation of vocal tract length from formant frequencies*. *PLOS ONE*, 10(7), S. 1–23, 2015.

[5] DANG, J. und K. HONDA: *Acoustic characteristics of the piriform fossa in models and humans*. *J Acoust Soc Am*, 101, S. 456–465, 1997.

[6] TAKEMOTO, H., S. ADACHI, T. KITAMURA, P. MOKHTARI, und K. HONDA: *Acoustic roles of the laryngeal cavity in vocal tract resonance*. *J Acoust Soc Am*, 120, S. 2228–2238, 2006.

[7] MAINKA, A., A. A. POZNYAKOVSKIY, I. PLATZEK, M. FLEISCHER, J. SUNDBERG, und D. MÜRBE: *Lower vocal tract morphologic adjustments are relevant for voice timbre in singing*. *PLoS One*, 10, S. e0132241, 2015.

[8] TRASER, L., P. BIRKHOLZ, T. FLÜGGE, R. KAMBERGER, M. BURDUMY, B. RICHTER, J. G. KORVINK, und M. ECHTERNACH: *Relevance of the implementation of teeth in three-dimensional vocal tract models*. *J Speech Lang Hear Res*, 60, S. 2379–2393, 2017.

[9] ARNELA, M., R. BLANDIN, S. DAPPAGHCHIAN, O. GUASCH, F. ALÍAS, X. PELORSON, A. VAN HIRTUM, und O. ENGWALL: *Influence of lips on the production of vowels based on finite element simulations and experiments*. *J Acoust Soc Am*, 139(5), S. 2852–2859, 2016.

[10] FLEISCHER, M., S. PINKERT, W. MATTHEUS, A. MAINKA, und D. MÜRBE: *Formant frequencies and bandwidths of the vocal tract transfer function are affected by the mechanical impedance of the vocal tract wall*. *Biomech Mod Mechanobiol*, 14, S. 719–733, 2015.

[11] HANNA, N., J. SMITH, und J. WOLFE: *Frequencies, bandwidths and magnitudes of vocal tract and surrounding tissue resonances, measured through the lips during phonation*. *J Acoust Soc Am*, 139(5), S. 2924–2936, 2016.

[12] TITZE, I. R.: *Acoustic interpretation of resonant voice*. *J Voice*, 15(4), S. 519–528, 2001.

[13] PERRIER, P., L.-J. BOË, und R. SOCK: *Vocal tract area function estimation from midsagittal dimensions with ct scans and a vocal tract cast: modeling the transition with two sets of coefficients. Journal of speech and hearing research*, 35(1), S. 53–67, 1992.

[14] ROERS, F., D. MÜRBE, und J. SUNDBERG: *Voice classification and vocal tract of singers: A study of x-ray images and morphology*. *J Acoust Soc Am*, 125, S. 503–512, 2009.

[15] VAMPOLA, T., J. HORÁČEK, A.-M. LAUKKANEN, und J. G. ŠVEC: *Human vocal tract resonances and the corresponding mode shapes investigated by three-dimensional finite-element modelling based on ct measurement. Logoped Phoniat Vocol*, 40(1), S. 14–23, 2015.

[16] CRARY, M. A., I. M. KOTZUR, J. GAUGER, M. GORHAM, und S. BURTON: *Dynamic magnetic resonance imaging in the study of vocal tract configuration*. *Journal of Voice*, 10(4), S. 378–388, 1996.

[17] NIEBERGALL, A., S. ZHANG, E. KUNAY, G. KEYDANA, M. JOB, M. UECKER, und J. FRAHM: *Real-time mri of speaking at a resolution of 33 ms: Undersampled radial flash with nonlinear inverse reconstruction*. *Magn Reson Med*, 69(2), S. 477–485, 2013.

[18] MAINKA, A., I. PLATZEK, W. MATTHEUS, M. FLEISCHER, A.-S. MÜLLER, und D. MÜRBE: *Three-dimensional vocal tract morphology based on multiple magnetic resonance images is highly reproducible during sustained phonation*. *J Voice*, 31(4), S. 504.e11–504.e20, 2017.

[19] ROERS, F., D. MÜRBE, und J. SUNDBERG: *Predicted singers' vocal fold lengths and voice classification – a study of x-ray morphological measures*. *J Voice*, 23, S. 408–413, 2009.

[20] VOS, R. R., D. T. MURPHY, D. M. HOWARD, und H. DAFFERN: *Determining the relevant criteria for three-dimensional vocal tract characterization*. *Journal of Voice*, 32(2), S. 130 – 142, 2018.

[21] STORY, B. H., I. R. TITZE, und E. A. HOFFMAN: *Vocal tract area functions from magnetic resonance imaging*. *J Acoust Soc Am*, 100, S. 537–554, 1996.

[22] ECHTERNACH, M., J. SUNDBERG, T. BAUMANN, M. MARKL, und B. RICHTER: *Vocal tract area functions and formant frequencies in opera tenors' modal and falsetto registers*. *J Acoust Soc Am*, 129, S. 3955–3963, 2011.

[23] MOTOKI, K.: *Three-dimensional acoustic field in vocal-tract*. *Acoust Sci Tech*, 23, S. 207–212, 2002.

[24] VAMPOLA, T., J. HORÁČEK, und J. G. ŠVEC: *Modeling the influence of piriform sinuses and valleculae on the vocal tract resonances and antiresonances*. *Acta Acust united Ac*, 101, S. 594–602, 2015.

[25] FLEISCHER, M., A. MAINKA, S. KÜRBIS, und P. BIRKHOLZ: *How to precisely measure the volume velocity transfer function of physical vocal tract models by external excitation*. *PLoS One*, 13(3), S. 1–16, 2018.

[26] MORSE, P. M. und K. U. INGARD: *Theoretical acoustics*. McGraw-Hill, New York, 1968.

[27] SUNDBERG, J.: *Articulary interpretation of the "singing formant"*. *J Acoust Soc Am*, 55, S. 838–844, 1974.

[28] SUNDBERG, J.: *The Science of the Singing Voice*. Northern Illinois Univ, 1989.