

# Analyse von MRT-Daten zur Entwicklung eines vokalischen Artikulationsmodells auf der Ebene der Areafunktion

Bernd J. Kröger

Institut für Deutsche Sprache und Linguistik der Humboldt-Universität zu Berlin und  
Zentrum für Allgemeine Sprachwissenschaft (ZAS) Berlin

kroeger@zas.gwz-berlin.de

## Zusammenfassung

Mittels Magnet-Resonanz-Tomographie (MRT) wurde die Formung des Ansatzrohres eines Sprechers für sechs Langvokale des Deutschen vermessen. Die resultierenden Areafunktionen wurden zunächst ohne Einbeziehung der Längenänderung des Ansatzrohres einer Faktorenanalyse unterzogen. Es zeigt sich, dass bereits wenige Faktoren einen hohen Prozentsatz der Varianz der Daten beschreiben. Darüber hinaus ist eine Interpretation der Faktoren innerhalb bekannter phonetisch-artikulatorischer Kategorien möglich. Auf der Basis dieses *vokalischen Artikulationsmodells* ist eine annähernd eindeutige Zuordnung der Zentralfrequenzen der ersten beiden Formanten zu den beiden artikulatorisch interpretierbaren Faktoren möglich. Somit kann dieses Modell auch als Basis zur Ermittlung physiologisch realistischer Areafunktionen anhand des Sprachsignals genutzt werden.

## 1 Einleitung

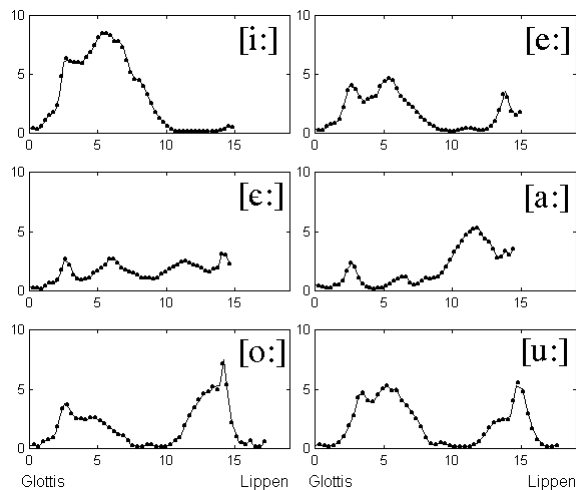
Zweidimensionale Artikulationsmodelle sind in Hinblick auf eine detaillierte Modellierung der Formung des Ansatzrohres nicht hinreichend. So kann insbesondere die für die Berechnung des akustischen Signals wichtige Areafunktion aus den von einem solchen zweidimensionalen Modell generierten mediosagittalen Schnittbild nicht befriedigend abgeleitet werden (Perrier et al. 1992). Eine Lösung hierfür stellt nur die direkte Messung der Areafunktion anhand dreidimensionaler Daten dar. Die so erhaltenen Areafunktionen können dann als Basis für die Entwicklung eines vokalischen Artikulationsmodells und zur Lösung des Inversionsproblems (d.h. Ermittlung artikulatorischer Parameter anhand des akustischen Signals) genutzt werden (Story und Titze 1998).

## 2 Die artikulatorische Datenbasis: MRT-Daten von sechs Langvokalen des Deutschen

Zunächst wurde die dreidimensionale Formung des Ansatzrohres von sechs Langvokalen des Deutschen ([i:], [e:], [ɛ:], [a:], [o:], und [u:]) eines Sprechers (JD, ZAS Berlin) mittels Magnet-Resonanz-Tomographie (MRT) vermessen (21 Sekunden Messzeit pro Vokal). Diese Vokale werden im folgenden auch als *Basisvokale* unseres Artikulationsmodells bezeichnet. Aufgrund unseres Messverfahrens (Kröger et al. 2000), das auf der Ermittlung der Form von Schnittflächen senkrecht zur Luftströmungsrichtung im Ansatzrohr basiert, kann direkt die Areafunktion  $A_p(x)$  (Querschnittsfläche des Ansatzrohres A eines Vokals p ( $p = 1...6$ ) als Funktion des aktuellen Abstandes x zur Glottis) ermittelt werden (durchgezogene Linien in Abb. 1). Mittels eines Interpolationsverfahrens (cubic spline interpolation) wird die Areafunktion für jeden Basisvokal in 50 äquidistante Sektionen ( $j = 1...50$ ) zerlegt (Punkte in Abb. 1 und Tab. 1).

Nr.	[i:]	[e:]	[ɛ:]	[a:]	[o:]	[u:]
1	0.4085	0.2338	0.2735	0.3885	0.2954	0.2956
2	0.3730	0.2233	0.2848	0.3267	0.2086	0.2209
3	0.5684	0.5560	0.1438	0.2709	0.5689	0.1631
4	1.1396	0.7798	0.3962	0.2747	0.7700	0.2654
5	1.4868	0.8797	0.6357	0.5075	0.9100	0.5322
6	1.7493	1.2128	0.6487	0.5378	1.8399	1.0442
7	2.3956	2.1880	0.9221	0.8423	3.3930	1.9427
8	4.8401	3.6751	1.7532	1.7217	3.6682	2.7774
9	6.3264	4.1054	2.6556	2.3126	2.9166	4.2695
10	6.0858	3.6884	2.1800	1.9968	2.5131	4.6810
11	6.0557	3.0923	1.3705	1.1293	2.5045	4.0651
12	6.0291	2.5964	0.9709	0.5703	2.4138	3.9110
13	5.9402	2.8399	0.9498	0.3497	2.6057	4.5118
14	6.4417	3.0154	1.0337	0.2253	2.5865	5.0446
15	6.8952	3.1691	1.0819	0.1653	2.3553	5.2741
16	7.4155	4.0142	1.4852	0.2277	2.0773	4.8391
17	8.1160	4.4009	1.7210	0.2738	1.7556	4.8945
18	8.4465	4.6464	1.9350	0.4060	1.4212	4.0462
19	8.4529	4.4720	2.1825	0.6115	1.2144	3.5821
20	8.2737	3.7834	2.7157	0.8347	1.0809	3.0335
21	7.8378	3.1601	2.7293	0.9936	0.6569	2.5238
22	7.7645	2.7732	2.3711	1.1513	0.2320	1.8452
23	7.3094	2.4617	1.9708	1.2047	0.1327	0.9595
24	6.2111	2.1667	1.7680	0.6419	0.1821	0.3813
25	5.1531	1.8063	1.6376	0.5017	0.3143	0.3365
26	4.5529	1.3858	1.5383	0.5698	0.3049	0.4722
27	4.4703	1.1261	1.3473	0.8756	0.2008	0.4058
28	3.9733	0.8042	1.1108	1.0822	0.2037	0.2330
29	3.3063	0.5805	1.1064	1.0354	0.1879	0.1409
30	2.5190	0.3808	1.0911	1.0550	0.3435	0.1535
31	1.7735	0.2841	0.9795	1.2009	0.6822	0.1442
32	1.2771	0.2280	1.1123	1.5220	0.9506	0.1768
33	0.8979	0.1913	1.4147	1.9774	1.9309	0.3084
34	0.6023	0.1997	1.5852	2.6048	2.7446	0.6083
35	0.3413	0.2278	1.8253	3.2627	3.4691	1.1404
36	0.1627	0.3160	2.0588	3.6964	4.1240	1.7408
37	0.1589	0.4398	2.2217	4.2340	4.5977	2.1917
38	0.1503	0.4293	2.4359	4.6751	4.7536	2.3253
39	0.1489	0.3769	2.5003	4.9838	5.1928	2.4489
40	0.1553	0.2809	2.3721	5.2178	4.9173	2.4139
41	0.1530	0.2534	2.2064	5.3086	7.1313	4.0098
42	0.1494	0.3598	2.0815	4.8002	5.4129	5.5472
43	0.1498	0.5794	1.9505	4.3794	2.1605	4.7689
44	0.1511	1.0332	1.6831	4.1368	0.9795	2.9259
45	0.1476	1.9571	1.6167	3.4908	0.4987	1.6455
46	0.1468	3.3473	1.8342	2.7862	0.3740	0.7442
47	0.2188	3.0913	1.9583	2.8892	0.6480	0.3060
48	0.3675	1.8347	3.0800	3.3913	0.1644	0.1455
49	0.6296	1.5191	3.0580	2.9951	0.1457	0.1531
50	0.5004	1.7465	2.2988	3.5710	0.6078	0.2467
dl	0.2965	0.2979	0.2919	0.2884	0.3436	0.3517

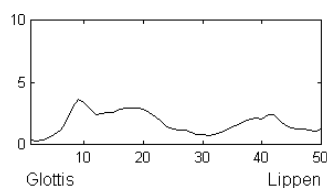
**Tabelle 1.** Querschnittsflächenwerte  $A_{pj}$  [cm<sup>2</sup>] und Längen  $dl_p$  [cm] der Sektionen für alle sechs Basisvokale nach Interpolation der Rohdaten von Kröger et al. (2000) in äquidistante Sektionen. Die Gesamtlänge des Ansatzrohres  $L_p$  ergibt sich für jeden Vokal als  $L_p = 50 dl_p$ .



**Abbildung 1.** Areefunktionen der Basisvokale. Abszisse: Abstand von Glottis in cm, Ordinate: Querschnittsfläche in  $\text{cm}^2$ . Rohdaten aus Kröger et al. 2000 (durchgezogene Kurven) und Interpolation in 50 äquidistante Sektionen (Punkte).

### 3 Hauptkomponentenanalyse ohne Einbeziehung der Länge des Ansatzrohres

Die Quadratwurzelwerte der Areefunktionen  $A_{pj}$  wurden zunächst ohne Einbeziehung der Länge des Ansatzrohres einer Faktorenanalyse (PCA, Principal Component Analysis, Hauptkomponentenanalyse) unterzogen (Story und Titze 1998). Die Wurzelfunktion wurde gewählt, weil die aus einer direkten Faktorenanalyse der Querschnittsflächenwerte  $A_{pj}$  entstehenden Rekonstruktionen der Areefunktionen in bestimmten Fällen zu (unerwünschten) negativen Flächenwerten führen (ibid., S. 229). Die Wurzelfunktion hat überdies die Eigenschaft, dass kleine Querschnittsflächenwerte (also Konstriktionen) expandiert und große Werte komprimiert werden. Dies ist von Vorteil, da die Übertragungsfunktion des Ansatzrohres von der Areefunktion im Bereich von Konstriktionen stärker beeinflusst wird als in Bereichen großer Hohlräume (Wood 1979). Bei der Faktorenanalyse wird nun zunächst eine *neutrale Areefunktion*  $A_0(x)$  berechnet (Abb. 2). Ein Querschnittsflächenwert  $A_{0j}$  dieser Areefunktion ergibt sich hierbei als arithmetisches Mittel über die entsprechenden Querschnittsflächenwerte aller  $p$  Basisvokale für jede Sektion  $j$ . Die zur Faktorenanalyse benutzten Eingangsdaten  $\alpha_{pj}$  sind dann jeweils die *Flächendifferenzen* zur neutralen Areefunktion  $\alpha_{pj} = A^{1/2}_{pj} - A^{1/2}_{0j}$ .



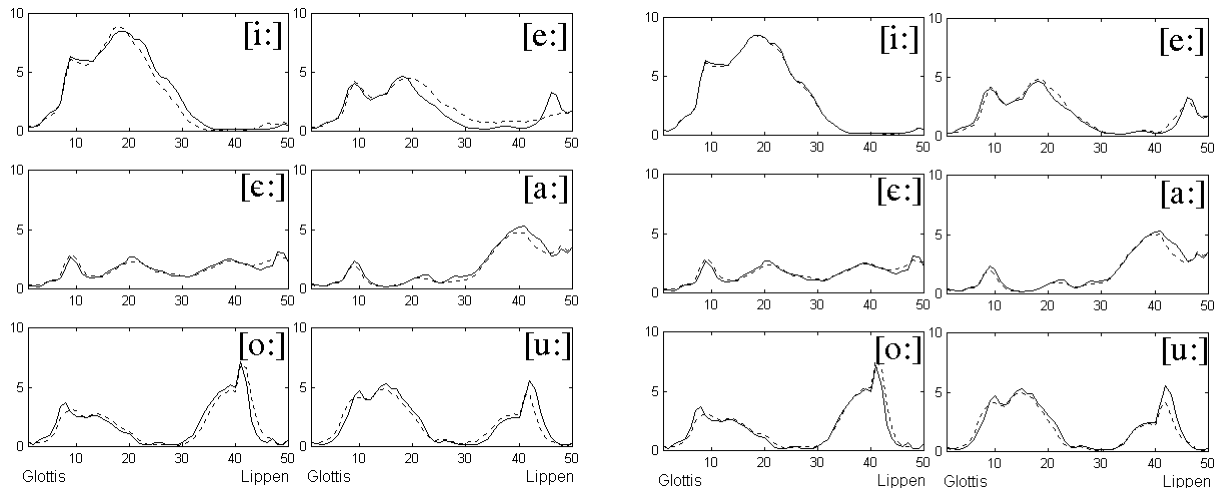
**Abbildung 2.** Neutrale Areefunktion  $A_0(x)$  (Mittelwert aus allen Basisvokalen, Abb. 1) als Basis der Faktorenanalyse. Abszisse: Nummer der Sektion, Ordinate: Querschnittsfläche in  $\text{cm}^2$ .

Mittels der Kovarianzmatrix  $\mathbf{R}$  über  $\alpha_{pj}$  können die normalisierten Eigenvektoren der Faktoren  $\phi_i$  ( $i = 1 \dots 50$ ) berechnet werden:  $\mathbf{R}\phi = \phi\mathbf{I}\lambda$ , wobei  $\mathbf{I}$  die Einheitsmatrix und  $\lambda$  den Vektor der Eigenwerte darstellt. Es zeigt sich, dass bereits wenige Eigenvektoren einen hohen Prozentsatz der Varianz der Daten (d.h. der sechs Basisvokale) beschreiben (Tab. 2). Die Rekonstruktion der Basisvokale mittels zweier und auch dreier Faktoren ist in Abb. 3 dargestellt. Um die Ergebnisse dieser Faktorenanalyse - im folgenden auch als *vokalisches Artikulationsmodell* bezeichnet - zu interpretieren, sollen die Eigenvektoren der vier wichtigsten Faktoren betrachtet werden (Abb. 4). Diese Querschnittsflächenwerte geben für jeden Eigenvektor jeweils die Richtung und Stärke der Änderung einer mittels des zugehörigen Faktors rekonstruierten Areefunktion  $A_j$  (in Wurzel aus Fläche) an. Der Ausgangspunkt ist dabei die neutrale Areefunktion. Es kann bereits erkannt werden, dass der erste Faktor zu einer Erhöhung der Querschnittsfläche im vorderen und zu einer Verringerung der Querschnittsfläche im hinteren

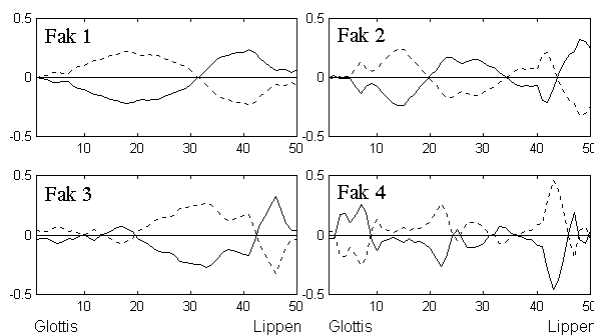
Bereich des Ansatzrohres führt. Faktor 2 hingegen führt zu einer Erhöhung der Querschnittsfläche im mittleren Bereich des Ansatzrohres und im Bereich der Lippen sowie zu einer Verringerung in allen anderen Bereichen.

Faktorenzahl	kumulative Varianz
1	68.91
2	88.64
3	96.19
4	99.43

**Tabelle 2.** Prozentsatz der kumulativen Varianz der ersten vier Faktoren über die Daten.



**Abbildung 3.** Originaldaten (durchgezogene Kurven) und ihre Modellierung (gestrichelte Kurven) mittels zweier Faktoren (links) und mittels dreier Faktoren (rechts).

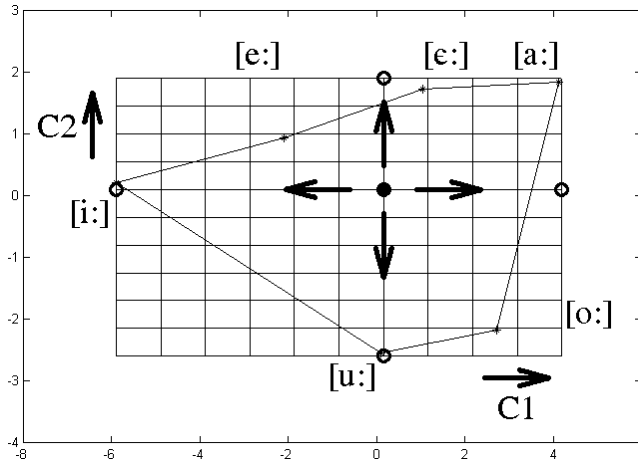


**Abbildung 4.** Eigenfunktion  $\phi_i$  ( $i$ : Faktornummer: 1...4). Jede Eigenfunktion  $i$  gibt die Änderung der neutralen Arefunktion aufgrund des Faktors  $i$  an. Abszisse: Nummer der Sektion, Ordinate: Wurzel aus Fläche (in cm). Durchgezogene Kurve: Amplitudenkoeffizient +1, gestrichelte Kurve: Amplitudenkoeffizient -1.

Tabelle 3 gibt nun die Amplitudenkoeffizienten der ersten vier Faktoren  $C_i(p)$  ( $i = 1...4$ ) zur Rekonstruktion eines Basisvokals  $p$  ( $p = 1...6$ ) an. Die Lage der ersten zwei Amplitudenkoeffizienten wurde für die sechs Basisvokale auch graphisch dargestellt (Abb. 5). Es zeigt sich, dass Faktor 1 insbesondere die Vorderzungenvokale [i:], [e:] und [ε:] bis hin zum Vokal [a:] differenzieren. Faktor 1 kann somit als Hebung - Senkung des Zungenrückens interpretiert werden. Faktor 2 separiert im wesentlichen [ε:] (und [a:]) von [u:] (und [o:]). Der Faktor 2 kann somit als Änderung der Kontur des Zungenrückens von annähernd neutral bis (/u/-förmig) gebogen interpretiert werden.

Faktor	[i:]	[e:]	[ɛ:]	[a:]	[o:]	[u:]
C1	-5.8928	-2.1034	1.0471	4.1302	2.7294	0.0895
C2	0.2146	0.9369	1.7328	1.8406	-2.1814	-2.5434
C3	-1.3491	1.8522	-0.1856	-0.3671	-0.8894	0.9391
C4	-0.1777	0.7861	-0.1174	-0.4830	1.0321	-1.0401

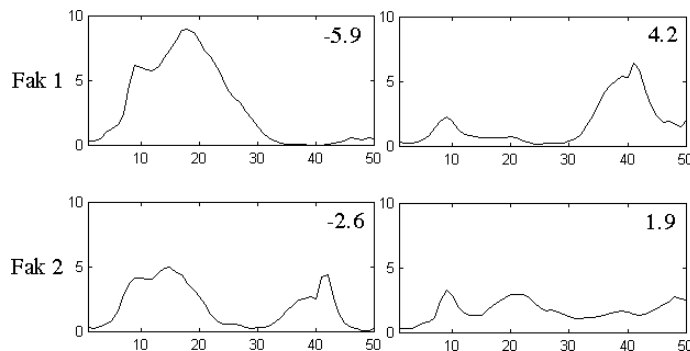
**Tabelle 3.** Amplitudenkoeffizienten  $C_i$  der ersten vier Faktoren zur Realisierung der sechs Basisvokale.



**Abbildung 5.** Das Gitternetz repräsentiert den artikulatorischen Vokalraum (C1-C2-Ebene mit:  $-5.9 < C_1 < 4.2$  und  $-2.6 < C_2 < 1.9$ ). Die kleinen Sterne und die ihre Verbindungslinien kennzeichnen die Lage der Basisvokale (siehe auch Tab. 3). Großer Punkt: Lage der neutralen Areafunktion; Pfeile: Richtung der Änderung der neutralen Areafunktion durch jeweils einen Faktor; Kreise: Lage der vom Modell generierbaren peripheren Areafunktionen.

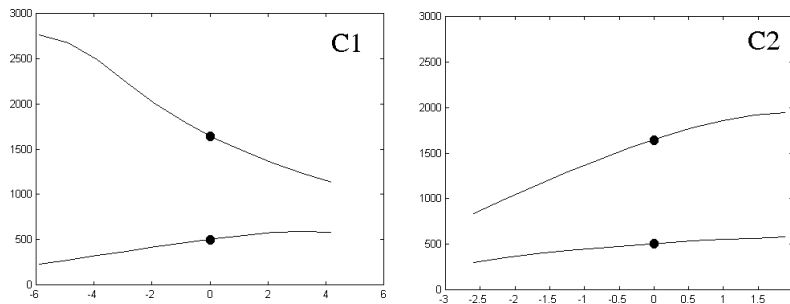
#### 4 Modellierung der akustisch-artikulatorischen Relationen des Vokalraumes

Da bereits zwei Faktoren die Varianz der Daten zu 89% modellieren, können die errechneten ersten beiden Eigenfunktionen zusammen mit der neutralen Areafunktion als Basis für ein (vorläufiges) vokalisches Artikulationsmodells (Zwei-Faktoren-Modell) angesehen werden. Es ist ausreichend, die Amplitudenkoeffizienten der beiden Faktoren im Intervall  $-5.9 \leq C_1 \leq 4.2$  und  $-2.6 \leq C_2 \leq 1.9$  zu variieren, da innerhalb dieser Intervalle insbesondere auch die Eckvokale [i:], [a:] und [u:] realisierbar sind (siehe Tab. 3). Somit definieren diese Intervalle der Amplitudenkoeffizienten den *artikulatorischen Vokalraum* dieses Artikulationsmodells (Abb. 5). In Abb. 6 sind nun die zugehörigen „peripheren“ Areafunktionen (d.h. die vom Modell generierte Areafunktionen bei den jeweiligen Extremwerten der Amplitudenkoeffizienten eines Faktors) für jeden Faktor gegeben. Die zugehörigen Orte im artikulatorischen Vokalraum sind in Abb. 5 durch Kreise dargestellt. Insbesondere bestätigt die Form dieser Areafunktionen nochmals die in § 3 gegebene phonetisch-artikulatorische Interpretation der beiden Faktoren des Modells.



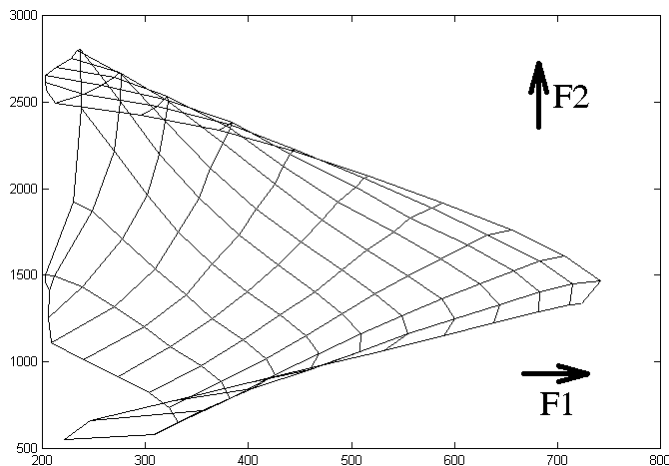
**Abbildung 6.** Periphere Areafunktionen für die ersten zwei Faktoren C1 und C2 bei Extremwerten der Amplitudenkoeffizienten. Faktor 1:  $C_1 = -5.9, 4.2$  und  $C_2 = 0$ ; Faktor 2:  $C_2 = -2.6, 1.9$  und  $C_1 = 0$ .

Zur Spezifikation des *akustischen Vokalraumes* wurden mittels des akustischen Modells von Sondhi und Schroeter (1987) die Übertragungsfunktion des Ansatzrohres (128 Punkte  $f_s = 10$  kHz) und mittels Peak-Picking und parabolischer Interpolation die Zentralfrequenzen der ersten vier Formanten berechnet. Als Länge des Ansatzrohres wurde der Mittelwert der Ansatzrohrlänge über alle sechs Basisvokale angesetzt ( $L = 15.58$  cm). Der aus den Areafunktionen des Modells resultierende Verlauf der Zentralfrequenzen der ersten beiden Formanten (F1 und F2) ergibt für Faktor 1 eine starke Absenkung für F2 und eine Erhöhung für F1 (Abb. 7). Dies entspricht einem [i:]-[a:]-ähnlichen Übergang. Der Verlauf der Formantfrequenzen kann für Faktor 2 einem [u:]-[ε:]-ähnlichen Übergang zugesprochen werden. Zusätzlich wurden in Abb. 7 die Formantwerte der neutralen Areafunktion gekennzeichnet.



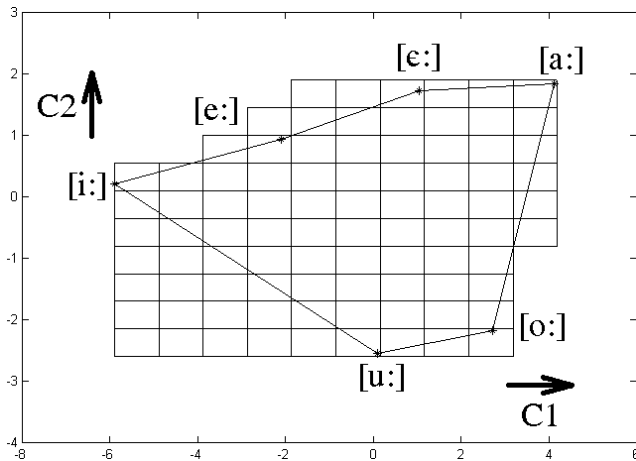
**Abbildung 7.** F1-/F2-Verläufe (in Hz) bei kontinuierlicher Variation von Amplitudenkoeffizient C1 bzw. C2, wobei der jeweils andere Koeffizient 0 beträgt (Pfeile in Abb. 5). Punkte: F1-F2-Werte der neutralen Areafunktion.

Zur genaueren Charakterisierung des akustischen Vokalraumes wurden die Areafunktionen für jede Kombination der Amplitudenkoeffizienten beider Faktoren generiert und die zugehörigen Formantfrequenzen berechnet. Dazu wurde der artikulatorische Vokalraum für jeden Faktor in 10 äquidistanten Abstände geteilt (Gitternetz in Abb. 5). Die Darstellung des Gitternetzes des akustischen Vokalraumes ist in Abb. 8 gegeben.

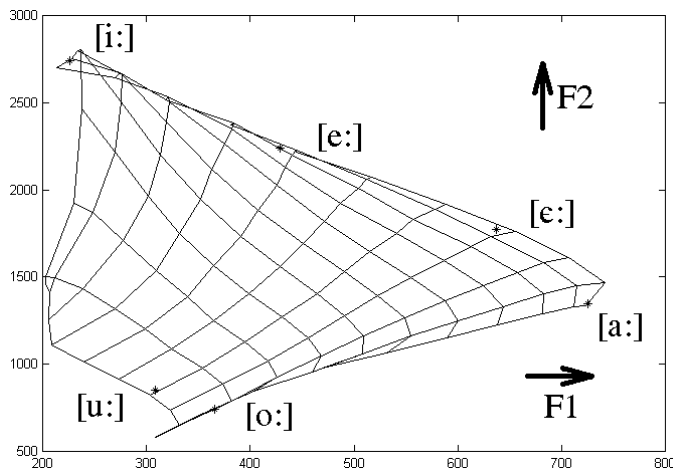


**Abbildung 8.** Akustischer Vokalraum (F1-F2-Ebene in Hz) als Gegenstück zum artikulatorischen Vokalraum (Abb. 5). Das Gitternetz entspricht genau dem Gitternetz des artikulatorischen Vokalraumes (Abb. 5), wobei die Gitternetzpunkte hier die zugehörigen Formantfrequenzen wiedergeben.

Die Formung dieses Netzes stellt somit die artikulatorisch-akustischen Relation dar. In weiten Bereichen des Vokalraumes ist diese Relation umkehrbar eindeutig und es kann direkt von den akustischen Daten (F1 und F2) auf das zugehörige C1-C2-Wertepaar und damit auf die zugehörige Areafunktion zurückgeschlossen werden. Nur bei niedrigem F1 zusammen mit sehr hohem oder sehr niedrigem F2 ergeben sich Mehrdeutigkeiten durch das Übereinanderfallen von Gitternetzlinien. Das Netz zeigt hier starke Verbiegungen. Zur Eliminierung dieser Mehrdeutigkeiten wurden zunächst einige Gitternetzpunkte eliminiert, die nicht unmittelbar im durch die sechs Basisvokale aufgespannten artikulatorischen Vokalraum liegen (modifizierter Vokalraum, Abb. 9). Die Abbildung des zugehörigen akustischen Vokalraumes zeigt bereits eine deutliche Verringerung der Mehrdeutigkeiten (Abb. 10).

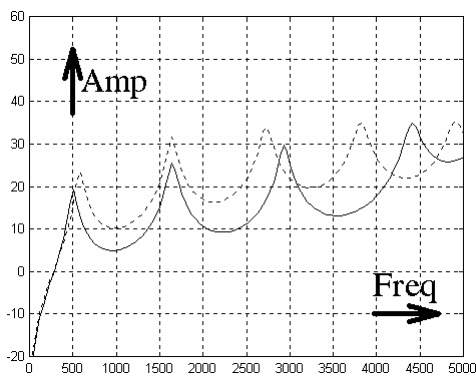


**Abbildung 9.** Modifizierter artikulatorischer Vokalraum des Modells (C1-C2-Ebene) unter Eliminierung von Gitterpunkten, die außerhalb des durch die sechs Basisvokale aufgespannten Vokalraumes liegen (vgl. Abb. 5).



**Abbildung 10.** Akustischer Vokalraum (F1-F2-Ebene in Hz) als Gegenstück zum modifizierten artikulatorischen Vokalraum des Modells (Abb. 9). Zusätzlich wurde die Lage der Basisvokale angegeben. Die Berechnung der Formantfrequenzen der Basisvokale erfolgte mit den in Tab. 1 gegebenen Daten aber bei normierter Länge des Ansatzrohres ( $L = 15.58$  cm).

Abschließend sollen noch die Übertragungsfunktion des Ansatzrohres und die resultierenden Zentralfrequenzen der ersten vier Formanten der neutralen Arealfunktion und eines gleichlangen querschnittskonstanten Ansatzrohres ( $A = 1 \text{ cm}^2$ ) angegeben werden (Abb. 11 und Tab. 4). Für F1 und F4 ergeben sich Abweichungen von über 10%. Somit kann die neutrale Arealfunktion nur bedingt mit einem querschnittskonstanten Ansatzrohr gleichgesetzt werden. Es ist aber anzunehmen, dass alle in Tab. 4 angegebenen Formantfrequenzen im Realisierungsbereich des Schwa-Lautes des Sprechers liegen.



**Abbildung 11.** Übertragungsfunktion des Ansatzrohres (akustisches Modell nach Sondhi und Schroeter 1987) für die neutrale Arealfunktion des vokalischen Artikulationsmodells (durchgezogene Linie) und eines querschnittskonstanten Rohres ( $A = 1 \text{ cm}^2$ ) der gleichen Länge (gestrichelte Linie). Abszisse: Frequenz in Hz; Ordinate: Amplitude in dB. Die zugehörigen Formantfrequenzwerte sind in Tab. 4 zusammengestellt.

Ansatzrohrform	F1	F2	F3	F4
querschn.konst. [Hz]	582	1642	2728	3819
neutr. Areafunk. [Hz]	506	1646	2934	4415
Differenz [%]	13.1	0.2	7.6	15.6

**Tabelle 4.** Zentralfrequenzen der ersten 4 Formanten der in Abb. 11 gezeigten Übertragungsfunktion des Ansatzrohres und prozentuale Differenz zwischen beiden Fällen.

## 5 Diskussion der Ergebnisse und Ausblick

Es konnte gezeigt werden, dass bereits die Ansatzrohrgeometrien der hier definierten Basisvokale - also der vorderen ungerundeten und der hinteren (gerundeten) Langvokale des Deutschen - zur Erstellung eines einfachen vokalischen Artikulationsmodell (Zwei-Faktoren-Modell) hinreichen. Story und Titze (1998) - die ein solches Modell anhand von Daten eines Sprechers des Amerikanischen Englisch erstellten - extrahierten qualitativ ähnlich interpretierbare Eigenwerte. Ein Nachteil dieser Artikulationsmodelle ist aber die Vernachlässigung der Längenänderung des Ansatzrohres. Story und Titze (1998) erweitern ihr Modell durch Einbeziehung der Länge als „Korrekturfaktor“. Es ist aber auch möglich, die Ansatzrohrlänge der Vokale einfach als zusätzlichen Parameter in den Vektor der Eingabedaten (vgl. § 3) aufzunehmen und damit direkt ins Modell zu integrieren. Dies soll im nächsten Schritt der Modellerweiterung realisiert werden.

In weiteren Schritten sollen auch die Ansatzrohrgeometrien der gerundeten Vorderzungenvokale und einiger Konsonanten (labialer, apikaler und dorsaler Verschluß) mit einbezogen werden. Es ist zu erwarten, dass diese Geometrien in dem hier realisierten Artikulationsmodell nicht mehr befriedigend realisiert werden können. Wahrscheinlich wird eine Zerlegung nach Artikulatoren (Lippen, Zungenspitze, Zungenrücken) bzw. eine Hinzunahme weiterer Faktoren nötig. Darüber hinaus soll auch die neutrale Areafunktion des Modells zur Realisierung unterschiedlicher „settings“ (Laver 1980) variiert werden (Story und Titze 2000).

**Danksagung.** Die vorgestellte Untersuchung wurde im Rahmen des DFG-Projektes GWZ 4/4-1 P1.1 gefördert. Mein Dank gilt Bernd Pompino-Marschall und Christine Mooshammer für ihre tatkräftige Unterstützung zur Realisierung dieser Arbeit.

## Literatur

- Kröger, B.J., Winkler, R., Mooshammer, C., Pompino-Marschall, B. (2000): “Estimation of vocal tract area function from magnetic resonance imaging: Preliminary results“, *Proceedings of 5<sup>th</sup> Seminar on Speech Production: Models and Data*, Kloster Seeon, Bavaria, p. 333-336.
- Laver, J. (1980): *The Phonetic Description of Voice Quality* (Cambridge University Press, Cambridge).
- Perrier, P., Boë, L.J., Sock, R. (1992): “Vocal tract area function estimation from midsagittal dimensions with CT scans and a vocal tract cast: modelling the transition with two sets of coefficients“, *Journal of Speech and Hearing Research* **35**, 53-67.
- Sondhi, M.M., Schroeter, J. (1987): “A hybrid time-frequency domain articulatory speech synthesizer“, *IEEE Transactions on Acoustics, Speech, and Signal Processing* **ASSP-35**, 955-967.
- Story, B.H., Titze, I.R. (1998): “Parameterization of vocal tract area functions by empirical orthogonal modes“, *Journal of Phonetics* **26**, 223-260.
- Story, B.H., Titze, I.R. (2000): “An investigation of voice quality based on modifications of the neutral vocal tract shape“, *Proceedings of 5th Seminar on Speech Production: Models and Data*, Kloster Seeon, Bavaria, p. 349-352.
- Wood, S. (1979): “A radiographic analysis of constriction location for vowels“, *Journal of Phonetics* **7**, 25-43.