

» Zur artikulatorischen Realisierung von Phonationstypen mittels eines selbstschwingenden Glottismodells

B.J. Kröger

Zusammenfassung: Es werden Ergebnisse zur artikulatorischen Synthese unterschiedlicher Phonationstypen (modal, behaucht, knarrend, Falsett, Flüstern) mit einem selbstschwingenden Glottismodell vorgestellt. Diese Simulationen geben Aufschluß über in vivo nur schwer meßbare Parameter zur artikulatorischen Voreinstellung der Stimmlippen (Stimmlippenspannung, Stimmlippenabstand etc.) und zur Aerodynamik (Luftdruck, Luftstrom etc.) im Bereich der glottalen Konstriktion und erlauben ein detailliertes Studium des resultierenden Schwingungsverhaltens der Stimmlippen einschließlich der auf die Stimmlippen einwirkenden Kräfte. Es wird gezeigt, daß drei artikulatorisch-laryngeale Parameter – i.e. Stimmlippenabstand im membranartigen und im knorpelartigen Teil der Glottis und Stimmlippenspannung – hinreichend sind, um die artikulatorisch-laryngealen Voreinstellungen für eine Vielzahl unterschiedlicher Phonationstypen zu definieren.

Schlüsselwörter: Phonationstypen – Stimmphysiologie – artikulatorische Sprachsynthese – selbstschwingende Glottismodelle

Different Types of Phonation and Appertaining Laryngeal Settings: Results of the articulatory speech synthesis of different phonation types (modal, breathy, creaky, falsetto, whisper) by means of a self-oscillating glottis model are presented. These simulations give information about the parameters describing the articulatory settings of the vocal folds (cord tension, vocal fold distance etc.) and the aerodynamics (pressure, volume flow etc.) in the area of glottal constriction which are difficult to measure in vivo and allow a detailed analysis of the resulting vocal fold oscillation including the forces acting on the vocal folds. It will be shown that three articulatory-laryngeal parameters – i.e. distance of the vocal folds within the membranous and within the cartilagenous part of the glottis and cord tension – are sufficient to define the articulatory-laryngeal setting for a great variety of phonation types.

Key words: Types of phonation – Physiology of the voice – Articulatory speech synthesis – Self-oscillating glottis models

Einleitung

Die artikulatorischen und akustisch-aerodynamischen Vorgänge im Kehlkopf sind sehr komplex und die Beobachtung dieser Vorgänge bei der Phonation ist nicht einfach durchführbar. Zumeist beschränken sich phonatorische Messungen auf die Geometrie und das Bewegungsverhalten der Stimmlippen (Laryngoskopie, Stroboskopie, Elektroglottographie) auf die Muskelaktivität (Elektromyographie) und auf aerodynamische Größen zumeist im oralen Bereich (Aerophone, Rothenbergmaske). In vielen Fällen stören diese Messungen aber den Sprachproduktionsprozeß und damit auch den Phonationsprozeß, da sie an erheblichen Meßaufwand gebunden sind. Somit bietet sich das Studium der bei der Phonation auftretenden Produktionsmechanismen mittels eines Modells geradezu an, da hier alle wichtigen geometrischen, mechanischen, dynamischen, aerodynamischen und akustischen Größen direkt beobachtet und die Resultate der phonatorischen Einstellungen anhand des generierten akustischen Signals perzeptiv überprüft werden können. Die Generierung unterschiedlicher Phonationstypen (modale Phonation, Hauchstimme, Knarrstimme, Falsett, Flüstern, siehe z.B. Laver, 1980; Wirth, 1987; S. 105–109) mittels unseres Modells ermöglicht die Untersuchung des Schwingungsverhaltens der Stimmlippen über die Gesamtheit möglicher phonatorischer Einstellungen (phonatory settings, Laver 1980; S. 93 ff). Es kann anhand dieser Arbeit auch abgeschätzt werden, welche und wie viele artikulatorisch-laryngeale Parameter zur Realisierung auftretender phonatorischer Einstellungen benötigt werden.

Beschreibung des selbstschwingenden Glottismodells

Das von uns realisierte selbstschwingende Glottismodell nach Ishizaka und Flanagan (1972) ist weit verbreitet in der Simulation der Sprachproduktion, da es einen gelungenen Kompromiß aus Einfachheit und Vollständigkeit darstellt. Es besteht aus zwei Feder-Masse-Paaren, die den unteren und oberen Teil der Stimmlippen repräsentieren (Abb. 1a). Die Schwingungsanregung erfolgt aufgrund der aerodynamischen Vorgaben: Im Fall des Verschlusses der Stimmlippen wird durch den subglottalen Druck p_s eine Kraft auf die Stimmlippen ausgeübt, die zu ihrer Öffnung führt. Im Fall der geöffneten Stimmlippen entsteht ein Luftstrom (glottaler Volumenstrom u_g), der einen Unterdruck bewirkt (Bernoulli-Effekt) und damit wiederum zur Schließung der Stimmlippen führt. Die Schwingung der beiden Massepaare ist phasenverschoben (Massepaar 1 läuft dabei zeitlich voraus) und modelliert die von unten nach oben fortschreitende Öffnungs- und Schließbewegung der Stimmlippen. Diese Phasenverschiebung ist von entscheidender Bedeutung für die Energieübertragung und damit

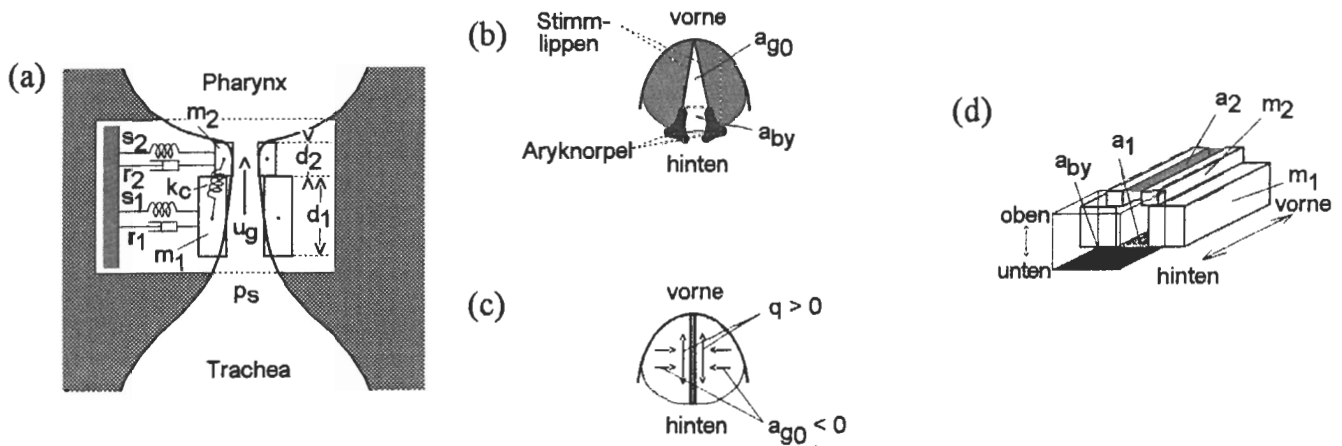


Abb. 1 Zur Geometrie des Glottismodells: (a) Ansicht der glottalen Konstriktion von vorn. Das Modell umfaßt zwei gekoppelte gedämpfte Feder-Massepaare (Masse m_i , Höhe d_i , Federkonstante r_i , Dämpfungskonstante s_i , $i = 1,2$; Kopplungskonstante k_c). Die Stimmlippen-schwingung entsteht aufgrund aerodynamischer Vorgaben (subglottaler Druck p_s und glottaler Volumenstrom u_g). (b) und (c) Ansicht der glottalen Konstriktion von oben zur Verdeutlichung der Unterscheidung von membranartigem Teil (Stimmlippen mit Öffnungsfläche a_{g0}) und knorpelartigem Teil (Aryknorpel mit Öffnungsfläche a_{by}) der Glottis und mit Eintragung der Wirkungsrichtung der Stimmlippen-spannung q und der seitlichen Kompression im Falle $a_{g0} < 0$. (d) Dreidimensionale Ansicht des Glottismodells mit Angabe der momentanen Öffnungsfläche zwischen den beiden Massepaaren a_1 und a_2 und der Öffnungsfläche des knorpelartigen Teils der Stimmlippen a_{by} . Hier sind auch die beiden Massepaare m_1 und m_2 dargestellt.

für die Aufrechterhaltung der Stimmlippen-schwingung. Bei der Benutzung nur eines Massepaares (Flanagan und Landgraf, 1968) hingegen ist die Übertragung der Energie des aerodynamischen Systems in das mechanische Schwingungssystem der Stimmlippen nicht immer möglich und damit das einwandfreie Schwingen der Stimmlippen in bestimmten Fällen nicht gegeben.

Der Ansatz nach Ishizaka und Flanagan (1972) realisiert aber nur Unterschiede des glottalen Öffnungsgrades in horizontal-lateraler und in vertikaler Richtung, nicht aber in der horizontal-mediosagittalen Richtung. Da die Glottis in einen schwingungsfähigen membranartigen und einen nicht schwingungsfähigen knorpelartigen Teil unterteilt werden kann (Abb. 1b), ist die Modellierung eines unterschiedlichen Öffnungsabstandes und damit die Modellierung des unterschiedlichen Schwingungsverhaltens in dieser Richtung (parallel zur Glottis) wichtig. Dies kann durch eine Erweiterung des Modells um einen Bypass (Cranen und Boves, 1987; Kröger, 1994; Cranen u. Schroeter, 1995) mit konstanter, aber voreinstellbarer Öffnungsfläche a_{by} realisiert werden (Abb. 1d).

Die artikulatorisch-laryngealen Steuerparameter

Es ergeben sich damit im wesentlichen drei laryngeale Steuerparameter (Tab. 1). 1. Die *glottale Ruheöffnungsfläche* a_{g0} betrifft den membranartigen Teil der Stimmlippen und definiert den artikulatorisch voreinstellbaren Stimmlippenabstand. Im Fall positiver Werte und ohne Vorliegen einer Stimmlippen-schwingung gibt dieser Wert die resultierende horizontale Glottisöffnungsfläche wieder (Abb. 1b). Bei Vorliegen einer Stimmlippen-schwingung wird durch diesen Parameter die Ruhelage der schwingenden Massepaare definiert. Im Fall negativer Werte gibt der Betrag dieses Wertes die Höhe der seitlichen Kompression (medial compression, Laver, 1980; S. 108) an (siehe auch Abb. 1c). 2. Die *Bypassöffnungsfläche* a_{by} betrifft nur den knorpelartigen Teil der Stimmlippen und gibt die horizontale Öffnungsfläche zwischen den Aryknorpeln an (Abb. 1b). 3. Die *Stimmlippen-spannung* (Longitudinalspannung der

Stimmlippen, Abb. 1c) betrifft den membranartigen Teil der Glottis und reguliert in erster Linie die Schwingungsfrequenz der Stimmlippen. Der Parameter wirkt in definierter Weise auf untergeordnete Parameter des selbstschwingenden Glottismodells: Masse m_i , Federsteife s_i , Höhe d_i der einzelnen Massepaare ($i = 1,2$), Kopplungskonstante beider Massepaare k_c (siehe Ishizaka u. Flanagan, 1972, S. 1259f; Sondhi u. Schroeter, 1987; S. 958, siehe auch Abb. 1a).

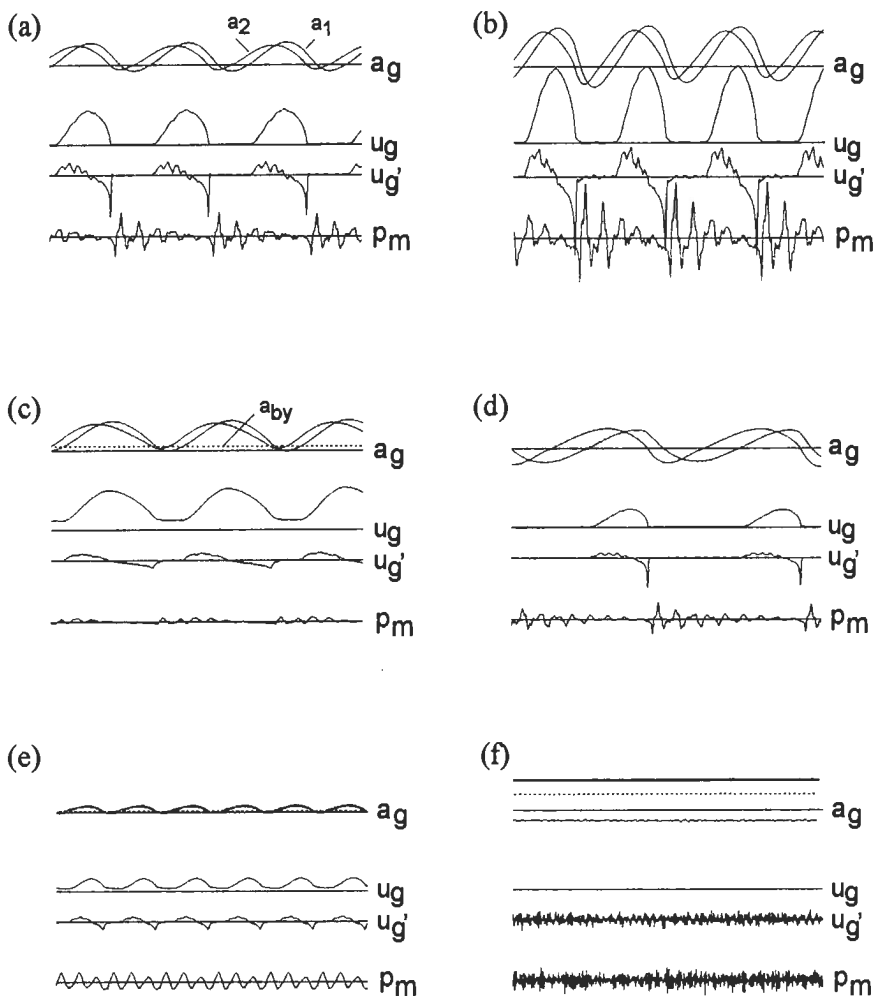
Das von uns implementierte Glottismodell ist Teil eines Gesamtsystems (phonetisches Produktionsmodell, siehe Kröger, 1993), so daß auch die akustische Rückkopplung des Sprechtraktes auf die Stimmlippen Berücksichtigung findet (Kröger, 1991). Das Modell erlaubt neben der Berechnung einer Vielzahl von laryngealen Parametern (Stimmlippen-schwingungsform, Druck und Volumenstrom im Bereich der glottalen Konstriktion, auf die Stimmlippen einwirkende Kräfte, etc.) die Generierung akustischer Sprachsignale und damit die perzeptuelle Evaluierung unterschiedlicher Stimmklänge.

Zur Generierung der Phonationstypen

Die Generierung unterschiedlicher Phonationstypen wurde unter Voreinstellung des Sprechtraktes auf den Vokal /a:/ durchgeführt. Die generierten Stimuli unterschieden sich damit nur durch die Einstellung der artikulatorisch-laryngealen Steuerparameter (phonatorische Einstellung). Die akustischen Signale wurden durch einen trainierten Phonetiker den Phonationstypen modal (normal, modal voice), behaucht (breathy voice), knarrend (creaky voice, vocal fry), Falsett (falsetto) und Flüstern (whisper) zugeordnet. Für jeden Phonationstyp sind die vom Modell generierte Stimmlippen-schwingung, die Schwingungsform des glottalen Volumenstroms und seiner Zeitableitung und das Ozillogramm des zugehörigen akustischen Signals dargestellt (Abb. 2). Die aus den Simulationen resultierenden Einstellungen der laryngealen Steuerparameter werden im folgenden beschrieben. Die hierzu angegebenen Werte sind allerdings nur als ungefähre Mittelwerte zur

Tab. 1 Bedeutung der drei artikulatorisch-laryngealen Steuerparameter innerhalb bestimmter Wertebereiche.

Modellparameter	Bedeutung
$a_{g0} < 0$	Verschuß des membranartigen Teils der Glottis mit seitlicher Kompression; der Betrag von a_{g0} beschreibt die Stärke der seitlichen Kompression
$0 < a_{g0} \leq 0,2 \text{ cm}^2$	Verschuß des membranartigen Teils der Glottis ohne seitliche Kompression; die (geringen) positiven Werte reflektieren die Ruhelage der Stimmlippen (Stimmlippenschwingung möglich)
$a_{g0} > 0,2 \text{ cm}^2$	Mittlere und weite Öffnung der Glottis; der Betrag von a_{g0} ist die Stimmlippenöffnungsfläche des membranartigen Teils der Glottis und gibt den artikulatorisch aktiv voreingestellten Stimmlippenabstand an
$a_{by} = 0$	Verschuß des knorpelartigen Teils der Glottis
$0 < a_{by}$	Öffnung des knorpelartigen Teils der Glottis (z. B. als Zusatzöffnung zur Stimmlippenschwingung oder als Flüsterdreieck)
$0 < q < 0,5$	geringe Längsspannung der Stimmlippen
$0,5 < q < 1,5$	mittlere Längsspannung der Stimmlippen (üblicher Bereich beim Sprechen in normaler Stimmlage)
$q > 1,5$	hohe Längsspannung der Stimmlippen

**Abb. 2** Glottisöffnungsfläche a_g (zusammengesetzt aus a_1 , a_2 und a_{by} , siehe Abb. 1 d), glottaler Volumenstrom u_g , zeitliche Ableitung des glottalen Volumenstroms u_g' und am Mund abgestrahltes Schallsignal p_m als Funktion der Zeit für unterschiedliche Phonationstypen. (a) Modale Phonation bei normaler Sprechstärke, (b) modale Phonation bei lauter Sprechstärke, (c) Hauchstimme, (d) Knarrstimme, (e) Falsett und (f) Flüsterstimme.

optimalen Generierung des jeweiligen Phonationstyps zu interpretieren.

Ausgehend vom *normalen Phonationsmodus* ($a_{g0} = a_{by} = 0$, $p_s = 6,6 \text{ cmH}_2\text{O}$, $q = 1$ für normale Sprechstärke, Abb. 2a; gleiche Parametereinstellung bis auf $p_s = 13,3 \text{ cmH}_2\text{O}$ für laute Sprechstärke, Abb. 2b) kann die *Hauchstimme* erzeugt werden, in-

dem sowohl die glottale Ruheöffnungsfläche als auch die Bypassöffnungsfläche erhöhte wird ($a_{g0} = 0,05 \text{ cm}^2$, $a_{by} = 0,02 \text{ cm}^2$, Abb. 2c). Zur Realisierung dieses Phonationstyps ist der Bypass von großer Wichtigkeit, da bei Fehlen des Bypass durch die alleinige Erhöhung der Ruheöffnungsfläche bei $a_{g0} \approx 0,2 \text{ cm}^2$ ein abrupter Übergang von modaler Phonation in einen nicht schwingungsfähigen Zustand auftritt (siehe auch Ishiza-

ka u. Flanagan, 1972; S. 1254). Die Hauchstimme ist dann nicht realisierbar. Die *Knarrstimme* entsteht im Vergleich zur modalen Phonation durch Einstellung von negativen Werten der glottalen Ruheöffnungsfläche (seitliche Kompression) und durch Verringerung der Longitudinalspannung ($a_{g0} = 0,2 \text{ cm}^2$, $q = 0,5$, Abb. 2d). Insbesondere ist auch eine Reduktion des subglottalen Drucks ($p_s = 3,3 \text{ cmH}_2\text{O}$) vorteilhaft, damit die Amplituden der Stimmlippenauslenkung nicht unrealistisch groß werden. Allerdings konnten keine aperiodisch auftretenden einzelnen Glottisimpulse realisiert werden. Eine weitere Erhöhung der seitlichen Kompression ($a_{g0} \leq -0,4 \text{ cm}^2$, $p_s = 6,6 \text{ cmH}_2\text{O}$) führt zum festen Glottisverschluß. Der Modus des *Falsett* kann in diesem Modell nur andeutungsweise ausgehend von der modalen Phonation durch Erhöhung der Stimmlippenspannung und durch geringe Vergrößerung der Ruheöffnungsfläche und der Bypassöffnungsfläche erreicht werden ($q = 2$, $a_{g0} = 0,01 \text{ cm}^2$, $a_{by} = 0,005 \text{ cm}^2$, Abb. 2e). Insbesondere erhöht sich mit dem Faktor q auch die Kopplungskonstante k_c zwischen beiden Massepaaren. Damit nimmt der Phasenunterschied in der Schwingung beider Massepaare ab und es entsteht als typisches Merkmal des Falsetts eine Verringerung der Auslenkungsunterschiede der Glottis in vertikaler Richtung. *Flüstern* wird durch Verschluß und hohe seitliche Kompression des membranartigen Teils der Glottis zur Verhinderung von Stimmlippen-schwingungen bei gleichzeitiger Öffnung des knorpelartigen Teils der Glottis realisiert ($a_{g0} = -0,5 \text{ cm}^2$, $a_{by} = 0,1 \text{ cm}^2$, Abb. 2f).

Zur Erzeugung stimmloser Laute tritt im Vergleich zur modalen Phonation insbesondere eine weite Öffnung des membranartigen Teils der Glottis ($a_{g0} > 0,2 \text{ cm}^2$) auf, die ein Schwingen der Stimmlippen auch bei hohem Luftstrom verhindert. Eine weitere Erhöhung der Glottisöffnung führt zur Atemstellung.

Zusammenfassung

Das hier beschriebene selbstschwingende Glottismodell ist in der Lage, unterschiedliche Phonationstypen in perzeptiv befriedigender Weise zu realisieren. Es ergeben sich drei laryngeale Parameter, die eine hinreichende funktionale Beschreibung der laryngealen Artikulation für alle wichtigen phonatorischen Einstellungen der Stimmlippen erlauben. Diese drei Parameter (Stimmlippenöffnungsfläche für den membranartigen und für den knorpelartigen Teil der Glottis und Stimmlippenspannung, Tab. 1) sind mit den von Laver (1980, S. 108 ff) gegebenen Parametern vergleichbar.

In weiteren Untersuchungen sollen mit Hilfe des Glottismodells die artikulatorisch-physiologischen und die physikalischen (i.e. aerodynamischen und physikalisch-dynamischen) Ursachen der Produktion unterschiedlicher Phonationstypen untersucht werden, da bei vorgegebenen artikulatorischen Parameterwerten das Verhalten aller wichtigen Größen wie Glottisöffnungsfläche (ergibt im Verlauf der Zeit die Stimmlippen-schwingung, siehe Abb. 2), glottaler Luftstrom (ergibt im Verlauf der Zeit das akustische Anregungssignal des Sprechtraktes, siehe Abb. 2), Luftdruck unterhalb, innerhalb und oberhalb der glottalen Konstriktion und die auf die Stimmlippen wirkenden Kräfte berechnet und damit in direkter Weise beobachtet werden können. Eine Erweiterung des Modells zum Studium pathologischer Phonation ist wünschenswert und wird angestrebt.

Literatur

- Cranen, B., L. Boves: On subglottal formant analysis. *Journal of the Acoustical Society of America* 81 (1987) 734–746
- Cranen, B., J. Schroeter: Modelling a leaky glottis. *Journal of Phonetics* 23 (1995) 165–177
- Flanagan, J. L., L. L. Landgraf: Self-oscillating source for vocal-tract synthesizers. *IEEE Transactions in Audio and Electroacoustics AU-16* (1968) 57–64
- Ishizaka, K., J. L. Flanagan: Synthesis of voiced sounds from a two-mass model of the vocal cords. *The Bell System Technical Journal* 51 (1972) 1233–1268
- Kröger, B. J.: Zur Auswirkung der Glottis-Sprechtrakt-Kopplung auf die Stimmreinheit. *Sprache – Stimme – Gehör* 15 (1991) 139–142
- Kröger, B. J.: A gestural production model and its application to reduction in German. *Phonetica* 50 (1993) 213–233
- Kröger, B. J.: Simulation of vocal fold oscillation behaviour by a self-oscillating glottis model. *Journal de Physique* 4 (Colloque 5, 3ème Congrès français d'acoustique, Toulouse) (1994) 457–460
- Laver, J.: *The Phonetic Description of Voice Quality*. Cambridge University Press, Cambridge (1980)
- Sondhi, M. M., J. Schroeter: A hybrid time-frequency domain articulatory speech synthesizer. *IEEE Transactions on Acoustics, Speech, and Signal Processing ASSP-35* (1987) 955–967
- Wirth, G.: *Stimmstörungen*. Deutscher Ärzte Verlag, Köln, 2. Auflage (1987)

B. J. Kröger

Institut für Phonetik der Universität zu Köln
Greinstraße 2
50939 Köln