

Birkholz P, Kröger BJ, Neuschaefer-Rube C:

Stimmsynthese mit einem Zwei-Massen-Modell der Stimmlippen mit dreieckigem Öffnungsquerschnitt

Hintergrund

Für die Analyse und Synthese der Stimmfunktion werden häufig Zwei-Massen-Modelle (ZMM) nach dem Vorbild des Modells von Ishizaka und Flanagan [1] eingesetzt, da sie eine geringe Komplexität besitzen aber viele wesentliche Eigenschaften realer Phonation nachbilden. Die Stimmlippenelemente (Massen) aller bisherigen ZMM sind stets parallel zur dorso-ventralen Achse ausgerichtet und bewegen sich horizontal. Das Öffnen und Schließen der Stimmritze erfolgt dadurch immer gleichzeitig entlang der gesamten Länge der Stimmlippen (abrupt). Bei realer Phonation erfolgt das Öffnen und Schließen aber oft nicht in allen Stimmlippenabschnitten gleichzeitig. Insbesondere bei behauchter Phonation erfolgt das Öffnen und Schließen allmählich, und oft ist während der Verschlussphase ein glottaler Restspalt vorhanden. Daher ist insbesondere die Simulation behauchter Stimmen mit den bisherigen ZMM problematisch.

Um das gesamte Kontinuum der Stimmqualitäten von behaucht über normal bis gepresst simulieren zu können, haben wir eine Variante des klassischen ZMM entwickelt, bei der die Stimmritze in präphonatorischer Stellung die Form eines Dreiecks annimmt, das durch die Proc. Vocales und die vordere Kommissur aufgespannt wird (siehe Abb. 1). Die Stimmlippenelemente sind dabei schräg zur dorso-ventralen Achse ausgerichtet. Während der Phonation bewegen sich die Stimmlippenelemente wie im klassischen ZMM horizontal und behalten dabei ihren Winkel zur dorso-ventralen Achse bei. Abbildung 1 zeigt die daraus resultierende Geometrie der Glottis bei einer weit geöffneten und einer teilweise geschlossenen Glottis. Wenn die zwei oberen oder unteren Stimmlippenelemente in Kontakt sind, wirkt eine zusätzliche Rückstellkraft, die proportional zum Grad deren Überlappung ist. Im Gegensatz zum klassischen ZMM besteht die treibende Kraft auf die unteren Stimmlippenelemente nicht nur aus dem statischen Druck *zwischen* den Elementen, sondern auch aus dem Druck im konvergenten Einlaßbereich direkt unterhalb der Glottis.

Ob eine Stimme gepresst, behaucht oder normal klingt hängt in erster Linie von der lateralen Position der Proc. Vocales ab, und damit von der Ruheöffnungsfläche der

Glottis [2]. Je größer der präphonatorische Abstand der Proc. Vocales ist, desto größer ist die Ruheöffnungsfläche der Glottis und desto behauchter klingt die Stimme. Wenn dagegen die Stimmlippen durch die Stellknorpel fest aneinander gepresst werden, so ist die Glottis in der Ruhestellung fest verschlossen und es entsteht eine gepresste Stimme.

Material und Methoden

Wir haben für das klassische und das modifizierte ZMM untersucht, inwiefern sich bei einer Variation des Ruheabstandes der Stimmlippenelemente unterschiedliche Stimmqualitäten ergeben, wie sie in der Realität zu erwarten sind. Zuerst haben wir für jedes Modell ermittelt, für welchen Bereich von Ruheabständen eine selbsterhaltende Schwingung möglich war. Diese Bereiche wurden dann in zehn äquidistante Abstände unterteilt. Beim klassischen Modell wurde die Auslenkung der Stimmlippenelemente von der Mittellinie zwischen -0.15 und 0.35 mm in Schritten von 0.05 mm variiert. Beim neuen Modell wurde die Auslenkung der Stimmlippen auf Höhe der Proc. Vocales zwischen -0.2 und 0.7 mm in Schritten von 0.1 mm variiert. Durch die dreieckige Form der Glottis im neuen Modell ist ihr Flächeninhalt bei einer Auslenkung von 0.7 mm vergleichbar mit dem des klassischen rechteckigen Modells bei einer Auslenkung von 0.35 mm. Das untere und obere Massenelement jeder Stimmlippe wurden jeweils gleich weit ausgelenkt. Für jede präphonatorische Einstellung der Modelle wurde der glottale Volumenstrom berechnet und durch ein Vokaltraktmodell für den Vokal [a:] gefiltert. Dabei wurde die Stimmlippenspannung jeweils für eine Grundfrequenz von 120 Hz eingestellt und ein subglottischer Druck von 1 kPa verwendet.

Die insgesamt 20 Stimuli (zehn Stimuli je Modell) wurden in zufälliger Reihenfolge zehn Testhörern über Kopfhörer vorgespielt, die die Stimmqualität jedes Stimulus auf einer Skala von 1 (sehr gepresst) bis 5 (sehr behaucht) beurteilen sollten.

Ergebnisse

Abbildung 2 zeigt die gemittelte Bewertung der Testhörern für die verschiedenen Ruheauslenkungen der Proc. Vocales. Die Stimmqualitäten der Stimuli des klassischen Modells wurden trotz verschiedener Abduktionsgrade stets relativ undifferenziert als normal bis leicht gepresst bewertet. Die Stimuli des neuen Modells decken dagegen das gesamte Kontinuum von Stimmqualitäten ab. Die

Wahrnehmung der Stimmqualität korreliert dabei wie in der Realität mit dem Abduktionsgrad.

Diskussion

Das modifizierte Modell ermöglicht gegenüber dem klassischen Modell die Stimm synthese unterschiedlicher Stimmqualitäten ohne die Modellkomplexität wesentlich zu erhöhen.

Literatur

1. Ishizaka K, Flanagan JL. Synthesis of voiced sounds from a two-mass model of the vocal cords. The Bell System Technical Journal 1972;51(6):1233-1268
2. Klatt DH, Klatt LC. Analysis, synthesis, and perception of voice quality variations among female and male talkers. Journal of the Acoustical Society of America 1990; 87(2):820-857

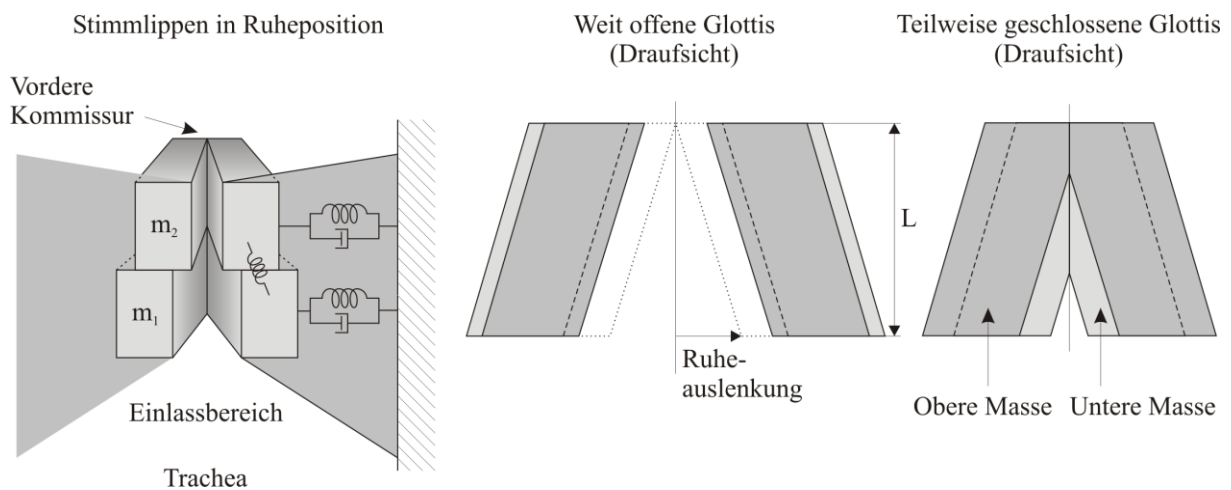


Abbildung 1: Modifiziertes Zwei-Massen-Modell der Stimmlippen.

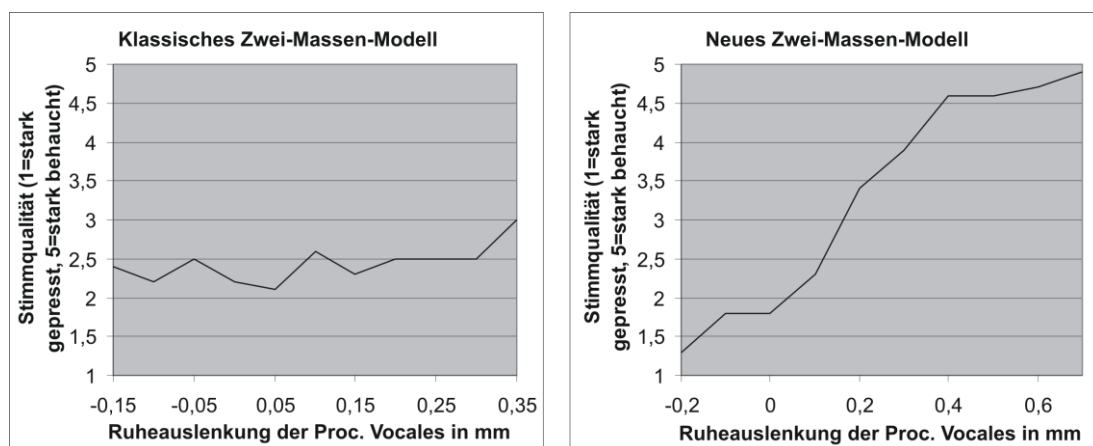


Abbildung 2: Wahrnehmung der Stimmqualität für verschiedene Abduktionsgrade der Stimmlippen im klassischen und modifizierten Zwei-Massen-Modell.