

ANSÄTZE EINES ZEITABHÄNGIGEN ARTIKULATIONSMODELLS

WILHELM H. VIEREGGE

1. VORBEMERKUNG

Auf dem Gebiet der akustischen Artikulation versuchen Vertreter von Netzwerktheorien (Fant 1960, Flanagan 1965) das Ansatzrohr durch elektrische Analogieschaltungen zu simulieren. Eine zweite Gruppe von Forschern (Ungeheuer 1962, Heinz 1962, Schroeder 1967, Mermelstein 1967, Vieregge 1970) beschreiben die bei der akustischen Artikulation ablaufenden Vorgänge durch eine spezielle Differentialgleichung, die sog. Webstersche Horngleichung, in welche die akustischen Größen wie Schalldruck und Schallschnelle direkt als Variable eingehen. Diese Gleichung liefert in Abhängigkeit eines sich räumlich verändernden Ansatzrohres die diesem entsprechenden Formantfrequenzen.

Eine zweite von Ungeheuer (1962, S.113) angegebene verallgemeinerte Horngleichung läßt sich dazu heranziehen, das in der Zeit sich ändernde Ansatzrohr zum Untersuchungsgegenstand zu machen. Die Lösung dieser Gleichung steht noch aus; erste Ansätze zur Lösung liefern jedoch schon aufschlußreiche Ergebnisse.

2. UNTERSUCHUNGEN MITHILFE DER WEBSTERSCHEN HORNGLEICHUNG FÜR EIN RÄUMLICH EINDIMENSIONALES ANSATZROHR $\sigma(z)$

Ungeheuer (1962) benutzt die Webstersche Horngleichung als Grundlage seiner Vokalartikulationstheorie; diese Gleichung hat die Form:

$$(1) \quad \frac{\partial^2 p}{\partial z^2} + \frac{\partial(\ln \sigma)}{\partial z} \frac{\partial p}{\partial z} = \frac{1}{c^2} \frac{\partial^2 p}{\partial t^2}.$$

wobei $p(z, t)$ den Schalldruck, $\sigma(z)$ den Querschnittsverlauf des Ansatzrohres, z die Raumachse, t die Zeit und c die Schallgeschwindigkeit bedeuten. Er führt diese Gleichung auf ein Eigenwertproblem zurück; die zu berechnenden Eigenwerte liefern die Formantfrequenzen des artikulierten Sprachschalls.

Vieregge (1969, 1970) simulierte die Webstersche Horngleichung auf einem elektronischen Rechner, so daß für beliebig vorgegebene Ansatzrohrkonfigurationen die

numerischen Werte der Formantfrequenzen bestimmt werden konnten, was bei Ungeheuer nur für einige Spezialfälle möglich war.

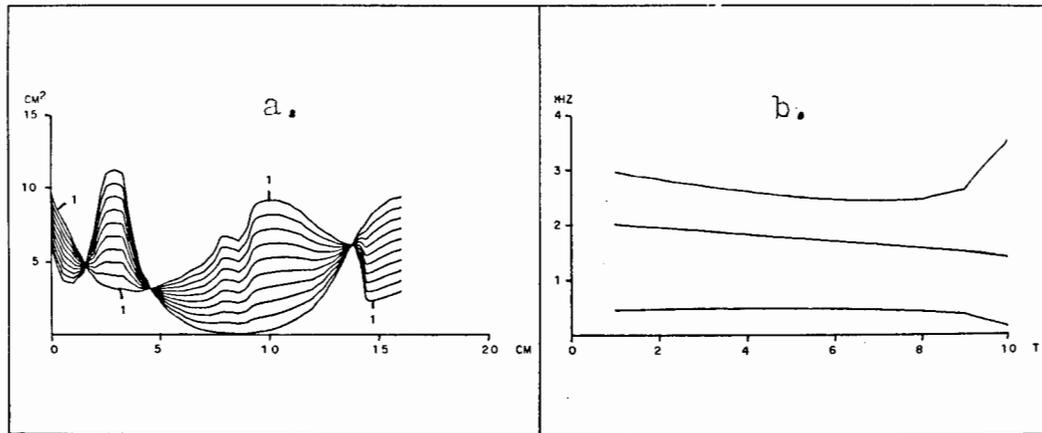


Abb. 1 a,b

Ein erster Schritt in Richtung zeitabhängiges Artikulationsmodell konnte durch sog. "Pseudozeitverläufe" (siehe Abbildung 1) in Vieregge (1969, 1970) verwirklicht werden. Pseudozeitverläufe stellen diskrete Auswahlen von Ansatzrohrkonfigurationen dar, d.h. es wird die kontinuierliche Aufeinanderfolge von Querschnittsverläufen beim Sprechen durch eine zeitliche Quantelung ersetzt.

In dieses Modell geht die echte Zeitinformation nicht ein, d.h. es ist nicht möglich, eine Aussage darüber zu machen, ob die Zeit einen Einfluß auf die Formantfrequenzen hat oder nicht. Dennoch ergab sich folgendes interessante Ergebnis: relativ kleine Veränderungen der Geometrie des Ansatzrohres können relativ große Veränderungen der akustischen Artikulationsdaten bewirken und umgekehrt. Weiterhin sind Größe der Veränderung und Lage der Formantfrequenzen abhängig von der geometrischen Position der artikulatorischen Veränderung im Ansatzrohr; d.h. es findet eine Art Ort-Frequenz-Transformation von artikulatorischen in akustische Signaldaten statt.

Dieses Modell stellt einen ersten Schritt dar, das zeitabhängige Verhalten des Ansatzrohres theoretisch zu erfassen.

3. UNTERSUCHUNGEN MITHILFE DER HORNGLEICHUNG FÜR EIN ZEITLICH VERÄNDERLICHES ANSATZROHR $\sigma(z, t)$

Die Horngleichung für ein zeitlich variables Ansatzrohr hat die Form:

$$(2) \quad \frac{\partial}{\partial z} \left[\frac{\partial p}{\partial z} \sigma \right] = \frac{1}{c^2} \frac{\partial^2 (p \sigma)}{\partial t^2}$$

Es soll nun versucht werden, für einen speziellen Ansatz der Querschnittsfunktion Gl. (2) auf ein Eigenwertproblem zurückzuführen, um damit die zeitliche Abhängigkeit des Ansatzrohres auf das Ergebnis zu demonstrieren. Mit dem speziellen Ansatz für die Sigmafunktion

$$(3) \quad \sigma(z, t) = s(z) v(t) = s(z) \hat{v} e^{-\delta t}$$

geht Gl. (2) in folgendes Eigenwertproblem über:

$$(4) \quad \varphi'' + \varphi' \frac{s'}{s} + \varphi \frac{(2\pi f)^2 + \frac{\delta^2}{4}}{c^2} = 0 \quad \lambda = \frac{(2\pi f)^2 + \frac{\delta^2}{4}}{c^2}$$

$$\varphi(z) \Big|_{z=0} = 0, \quad \varphi'(z) \Big|_{z=a} = 0 \quad \text{mit} \quad p(z, t) = \varphi(z) g(t)$$

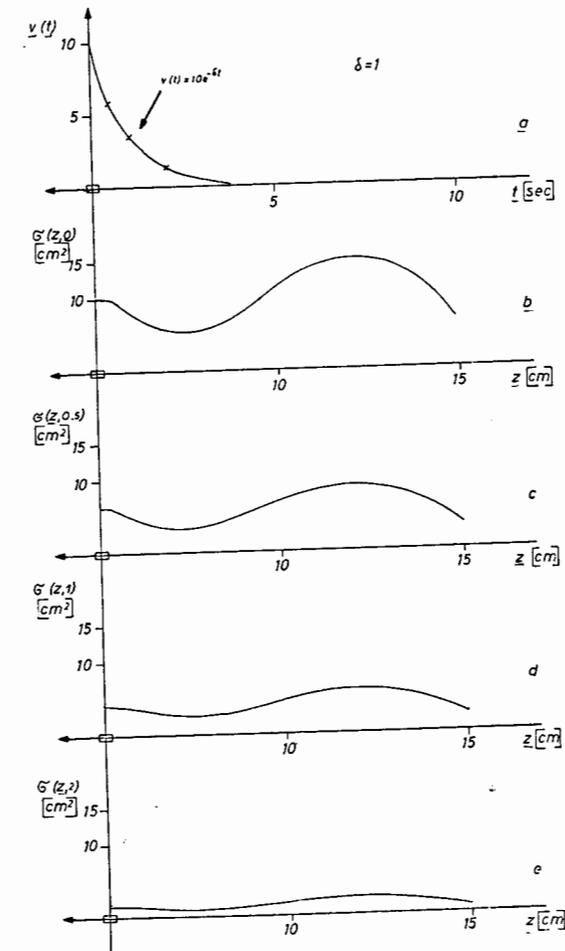


Abb. 2.

Abbildung 2a. zeigt gemäß Gl. (3) für $\hat{v} = 10$ und $\delta = 1$ einen Verlauf $v(t)$. Die daraus resultierenden Querschnittsverläufe sind zu verschiedenen Zeitpunkten in Abb. 2b, c, d, e dargestellt. Der Verlauf der e-Funktion läßt eine Aussage darüber zu, wie schnell das sich zeitlich verändernde Ansatzrohr gegen Null geht. Aus Gl. (4) für λ geht hervor, daß nur für sehr große δ -Werte der λ -Wert wesentlich verändert wird. Das bedeutet, daß $\sigma(z, t)$ mit großer Geschwindigkeit verändert werden muß, um eine merkliche Veränderung der Resonanzfrequenz zu erzielen. Immerhin kann so gezeigt werden, daß die Änderungsgeschwindigkeit des Ansatzrohres einen Einfluß auf die Resonanzfrequenzen nimmt.

Die Zurückführung der Gl. (2) auf das Eigenwertproblem (4) stellt erst einen kleinen Schritt in Richtung zeitabhängiges Artikulationsmodell dar. Man wird bei einer nicht so speziellen Lösung von Gl. (2) mit dem Ansatz für p in Gl. (4) — d.h. einer Analyse nach stehenden Wellen — nicht mehr auskommen, da ein solches Artikulationsmodell ein in der Zeit variables berechenbares Signal liefern soll. Dies bedeutet weiterhin, daß eine Generatorschwingung als Simulierung der Glottisimpulse berücksichtigt werden muß. Auch die Sigmafunktion $\sigma(z, t)$ wird anhand von Röntgenfilmen vorgegeben werden müssen, um ein berechenbares Sprachsignal zu erhalten, welches mit natürlichen Sprachsignalen verglichen werden kann, um eventuell Korrekturen am mathematischen Modell vornehmen zu können. An diesen Problemen wird z.Zt. gearbeitet.

*Institut für Kommunikationsforschung
und Phonetik
Universität Bonn*

LITERATUR

- Fant, G.
1960 *Acoustic Theory of Speech Production* (The Hague).
- Flanagan, J.L.
1965 *Speech Analysis, Synthesis and Perception* (Berlin, Springer Verlag).
- Heinz, J.M.
1962 "Reduction of Speech Spectra to Descriptions in Terms of Vocal Tract Area Functions", ScD. Thesis, (M.I.T.).
- Mermelstein, P.
1967 "Determination of the Vocal Tract Shape from Measured Formant-Frequencies", *Journal of the Acoustical Society of America* 41:1283-1294.
- Schroeder, M.R.
1967 "Determination of the Geometry of the Human Vocal Tract by Acoustic Measurements", *Journal of the Acoustical Society of America* 41:1002-1010.
- Ungeheuer, G.
1962 *Elemente einer akustischen Theorie der Vokalartikulation* (Berlin, Springer Verlag).
- Vieregge, W.H.
1970 *Untersuchungen zur akustischen Artikulation der Plosivlaute* (= *Bibliotheca Phonetica* 10) (Basel, S. Karger Verlag).

- 1969 "Untersuchungen zur akustischen Struktur der Plosivlaute, elementare Darstellungen zur Signalphonetik mithilfe elektronischer Rechenmethoden", Phil.Diss., (Bonn).
- 1971 "Untersuchungen zur akustischen Artikulation", in *DAWID-II, Beiträge zur automatischen Spracherkennung*, (= *IKP-Forschungsberichte*) (Hamburg, Helmut Buske Verlag).

DISCUSSION

TILLMANN (Bonn)

Kann man schon etwas darüber sagen, welchen speziellen Einfluß die Berücksichtigung der Zeit bei einer sich ändernden Querschnittsfunktion $\sigma(x, t)$ auf das durch die 2. Gleichung beschriebene akustische Sprachsignal ausübt?

VIIEGGE

Bezüglich des Einflusses der Zeit läßt sich sagen, daß nach der 2. Gleichung ein sich in der Zeit veränderndes Sprachsignal resultiert. Nach der 1. Gleichung wurden lediglich Parameter des Sprachsignals (Formant-Frequenzen) berechnet, in die die Zeit nicht eingehen.