Speech Communication Seminar Stockholm. Aug. 1-3, 1974

A MATHEMATICAL MODEL OF SPEECH AERODYNAMICS

John J. Ohala

Phonology Laboratory, Department of Linguistics,

University of California, Berkeley, California 94720, USA

Abstract

A computer-implemented mathematical model of the aerodynamic processes in speech shows that many of the empirically observed variations in air pressure and air flow can be accounted for by assuming a constant pulmoric force but varying glottal and oral air resistance and, in the case of voiced stops, a varying oral cavity volume.

Introduction

A full understanding of the aerodynamic processes in speech would m the ability to accurately predict the DC variations in air pressure and flow in the vocal tract, including the subglottal cavities, given variantions in the pulmonic force applied to the lungs and in the glottal and supraglottal air resistance. One of the problems in reaching this goal is that it is easier to sample and measure the dependent variables, air pressure air flow, than it is the independent variables, pulmonic force and air resistance. In some cases indirect estimates of the air resistance can be obtained. Broad (1968), for example, derived effective mean glottal resistance, Rg, from simultaneous recordings of subglottal pressure, P_S, and transglottal air flow, Ug, via the relation, Rg = Ps/Ug.

Another approach to this problem is to construct a model of the aerodynamic system used in speech for which the time-varying values of pulmonic force and air resistance are guessed at and are used to derive the variations in air pressure and air flow (cf. Rothenberg 1968). We may have some confidence in the accuracy of our guesses if the derived pressure and flow values match those observed in real speech. I report here a preliminary attempt to devise such a model and to use it to explore certain controversial: issues in phonetics.

The issues

One of these issues is the relative contribution of the pulmonic and laryngeal systems in controlling fundamental frequency (F_0) of voice in speech. Recordings of P_s during speech frequently reveal it to be posit: by correlated with F_0 (Ladefoged 1963, Lieberman 1967, Vanderslice 1967, Ohala 1970, Atkinson 1973). Since P_s can vary as a function of both the pulmonic force and glottal (and supraglottal) resistance, it is possible attribute these P_s variations to either or both factors. Lieberman and Atkinson suggest that in certain circumstances the F_0 variations are caused by the PS variations which in turn are caused by variations in the pulmonic expiratory force. However, Isshiki (1969) and Ohala suggest that the subglottal pressure variation may be due in large part to variations in glottal resistance which would accompany the laryngeal muscles' action in varying Fo by changing the tension of the vocal cords.

Another issue surrounds the production of aspirated vs. unaspirated stops. Chomsky and Halle (1968), for reasons that are not entirely clear suggest that aspirated stops, e.g., [p^h] and [b^h], are produced with heightened Ps in contrast to unaspirated stops such as [p] and [b] which

would have normal Ps. It is implicit in their approach that they would regard this heightened Ps as a feature that is independent of (and thus not caused by) laryngeal features; therefore it could only be attributed to ' an increase in pulmonic force. One may guess that they thought the increased Ps necessary to account for the greater air flow accompanying aspirated stops. Ohala and Ohala (1972), however, sampled Ps during the speech of a Hindi speaker and found instances of heightened Ps during the closed portion of any stop, whether-aspirated or not, thus showing that the heightened Ps was not a distinguishing characteristic of aspirated stops. They attributed these Ps peaks to the effects of increased oral resistance and a continued lung volume decrement during the closure. They also found markedly decreased Ps immediately after the release of the aspirated stops, but not after the unaspirated stops. They explained this lowered Ps as being due to lowered glottal resistance immediately after the release of the aspirated stops and this, in turn, would explain the high rate of air flow characteristic of these stops. (Halle and Stevens (1971) present a new analysis of stops and make no reference to heightened Ps as the distinguishing feature of aspirated stops, presumably indicating they have abandoned this feature. However, they cite no new evidence in support of this move.)

A final issue to consider is what special action, if any, is necessary to maintain voicing during voiced obstruents. Halle and Stevens (1967) suggest that a change in the vibratory pattern of the vocal cords equivalent to a decrease in glottal resistance plus the enlargement of the oral cavity are necessary for the maintenance of voicing during obstruents. The aerodynamic model to be reported here may be able to shed light on this and the preceding issues.

The model

The aerodynamic processes in speech were modeled mathematically with the model being implemented on a small general-purpose digital computer. : The basic elements of the model are shown in figure 1.

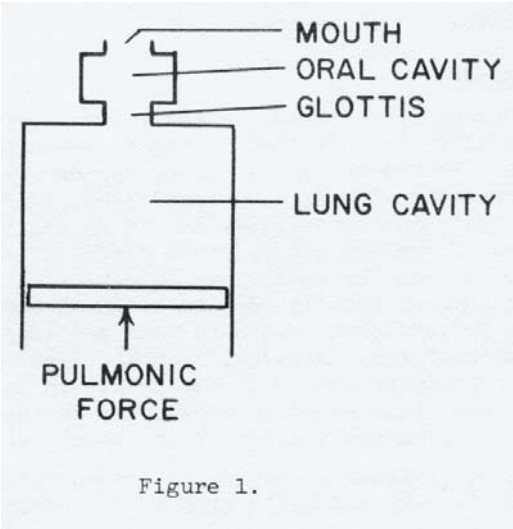


Figure 1.

Two connected air cavities, the lung cavity and the oral cavity, are defined by their respective volumes and air masses. Between the oral cavity and the "outside" there is an aperture, the mouth. Between the lung cavity and the oral cavity there is another aperture, the glottis. Both of these apertures are defined by their respective resistances. The volume of the lung cavity may decrease as the pulmonic force moves the chest wall and causes a lung volume decrement. The volume of the oral cavity is allowed to increase during voiced stop closures. The pressure inside

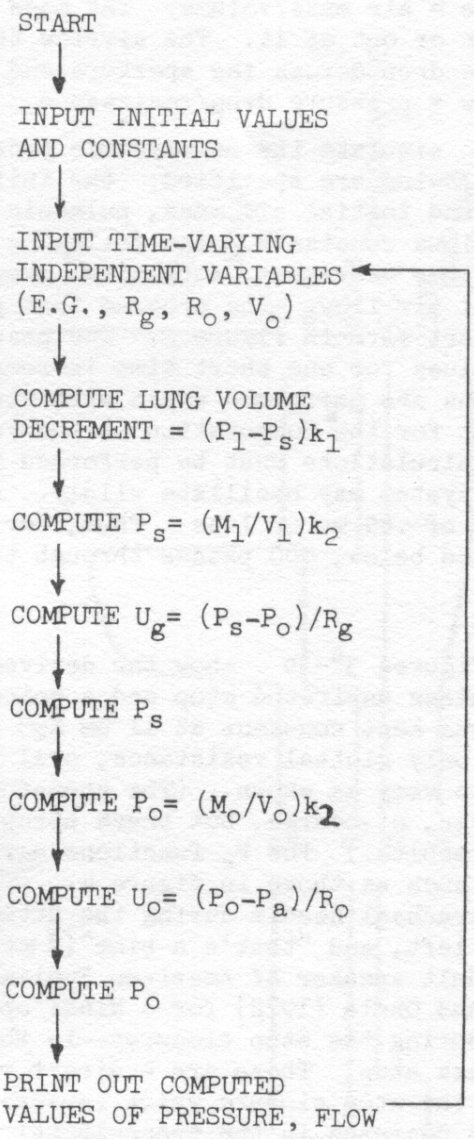
a cavity is derived by the relation: pressure = air mass/volume. The mass of air inside a cavity varies as air flows in or out of it. The airflow through an aperture is a function of the pressure drop across the aperture and the resistance of the aperture: air flow = pressure drop/resistance.

To simulate the aerodynamic processes during a given sample of speech, the following are specified: the initial lung volume and air mass, the oral volume and initial air mass, pulmonic force, glottal and oral resistance, and various constants. The following are computed for each time increment: lung volume decrement, subglottal pressure, oral pressure, glottal air flow, and oral air flow. The program that performs these computations is given in flow chart form in figure 2. One pass through the program derives the relevant values for one short time increment. Then, on the next pass, the calculations are performed again with the most recently derived values serving as input for the computation of the values for the next time increment. These calculations must be performed for a sufficiently small time increment or the system may oscillate wildly. I found it necessary to use a time increment of .45 ms or less. Thus, for the 400 ms samples of speech to be discussed below, 880 passes through the program were required.

Results

Figures 3a-3b show the derived pressure and air flow functions for a voiceless aspirated stop and a voiced stop, respectively. The pulmonic force was kept constant at 11 cm H20 (over atmospheric pressure) in both cases; only glottal resistance, oral resistance, and oral volume were allowed to vary as shown. (The step-function changes in resistance are unrealistic, of course, but these abrupt variations do not seem to give unusual results.) The P_s functions agree well with those obtained for real speech such as those in figure 4. (The P_s curves in figure 4 were sampled via a tracheal needle during the utterances "that's a pine" [ðætsəp^{h}ɑjn], on the left, and "that's a bine" [ðætsəbɑjn], on the right, as spoken by a male adult speaker of American English. See Ohala 1970.) As was noted by Ohala and Ohala (1972) for a Hindi speaker, there are momentary increases in P_s during the stop closures—in this case the rise is greater for the voiceless stop. These are a direct result of the increased oral resistance during the stop closure which causes an increase in oral pressure and a consequent decrease in the transglottal pressure drop which in turn causes diminished glottal flow. The P_s then approaches the pulmonic force asymptotically. For 50 ms after the release of the voiceless aspirated stop the glottal resistance remains low. Consequently the air flow out of the lung cavity is very high, with the result that the subglottal pressure is momentarily lowered. Again, this agrees well with the real speech data (cf. figure 4 and the findings of Ohala and Ohala 1972).

It is clear from many other studies that the oral pressure for voiced stops is significantly lower than that for voiceless stops (Fischer-Jørgensen 1972 and references therein). This is necessary in order that a positive transglottal pressure drop be maintained so that there will be a continuing glottal air flow and thus voicing. To achieve this with this model one or both of the following would be necessary: a) an increase in glottal resistance during the stop closure, or b) an increase in the volume of the oral cavity during the stop. Halle and Stevens' (1967) suggestion that glottal resistance be lowered during stop closures would make the problem worse: oral pressure would reach that of subglottal pressure even more rapidly and voicing would cease. As there is no evidence (that I know of) for {a}, but there is evidence for (b) (Ewan and Krones 1972), I allowed the oral cavity to gradually increase by 2 cm3 during the 100 ms stop closure. This allowed oral pressure to be less than subglottal pressure and thus yielded continued air flow and voicing throughout the stop closure.



P_a = atmospheric pressure

P_l = pulmonic force

P_s = subglottal pressure

P_o = oral pressure

R_g = glottal resistance

R_o = oral resistance

U_g = glottal air flow

U_o = oral air flow

V_l = lung volume

V_o = oral volume

M_l = lung air mass

M_o = oral air mass

k_1, k_2 = constants

Figure 2. Flow chart of computer program simulating speech aerodynamics.

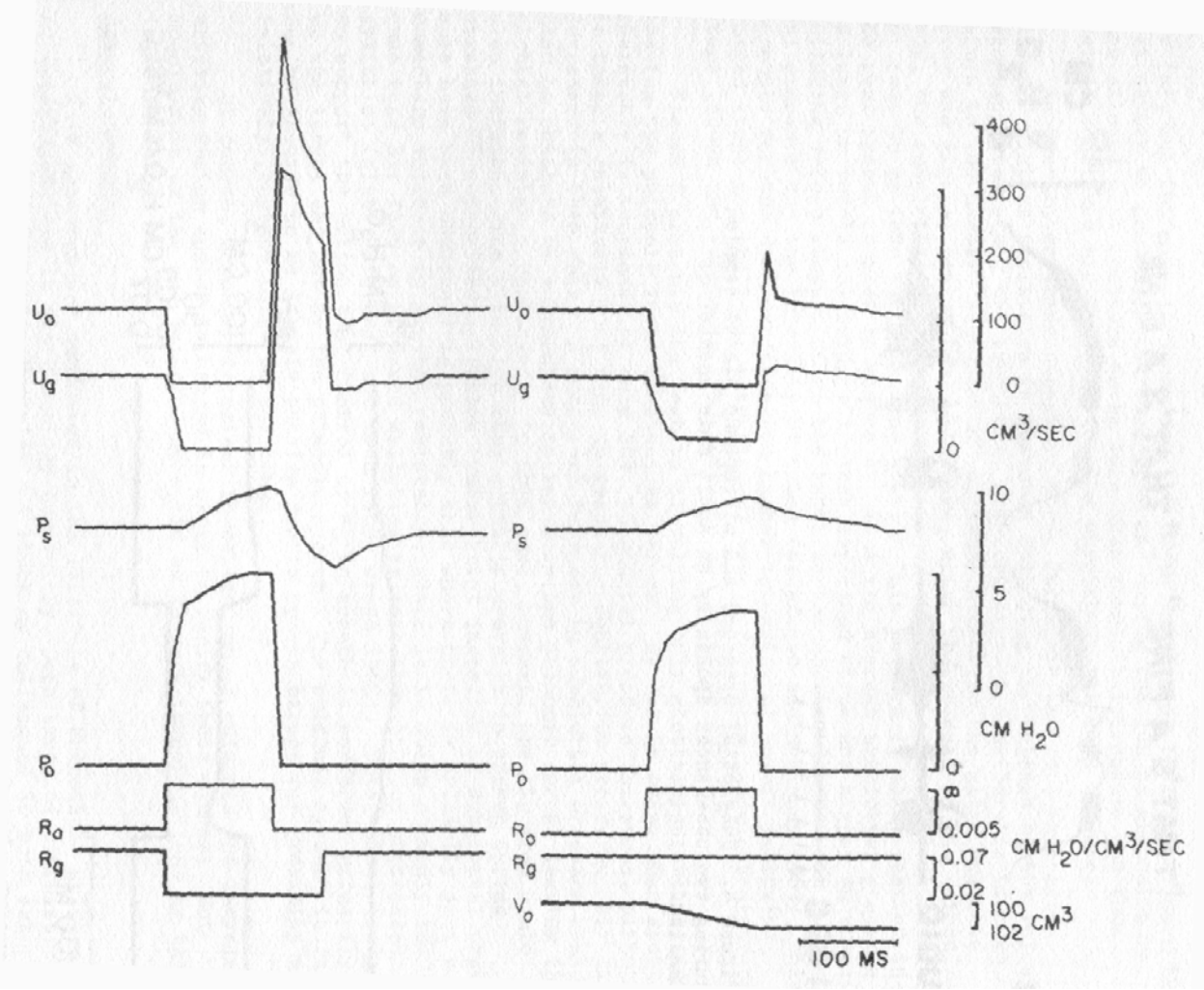


Figure 3. Output of aerodynamic model. A: intervocalic voiceless aspirated stop. B: intervocalic voiced stop. Parameters, from top: oral air flow, glottal air flow, subglottal pressure, oral pressure, oral resistance, glottal resistance, oral cavity volume.

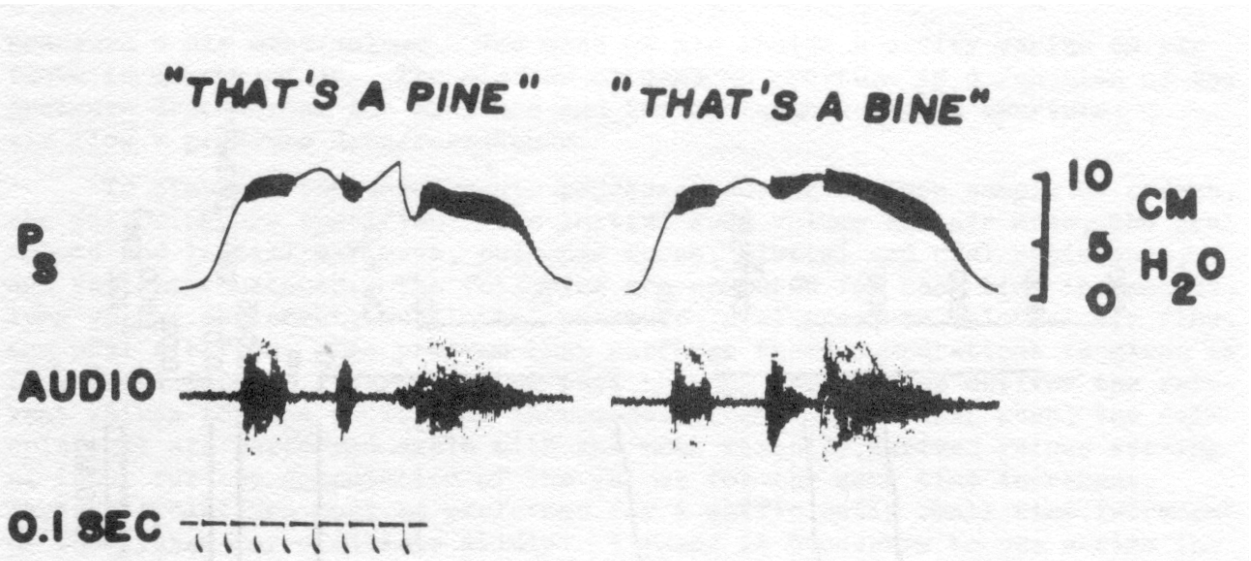


Figure 4. Subglottal pressure and microphone signal sampled during two utterances spoken by an adult male speaker of English.

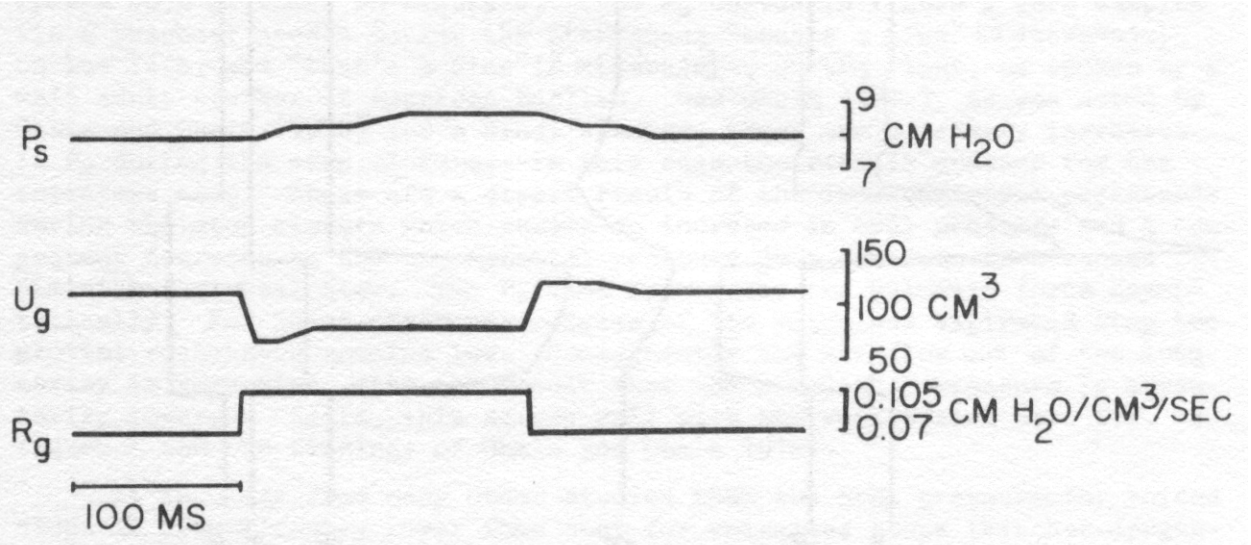


Figure 5. Output of aerodynamic model showing effects of increased glottal resistance on subglottal pressure (top) and glottal air flow (second line).

Another interesting aspect of these curves is the fact that after "normal" glottal resistance is restored following the release of the voiceless aspirated stop, the subglottal pressure takes a considerable time to return to the normal "equilibrium" pressure proper to the given pulmonic force and glottal resistance. Likewise, after the release of the voiced stop, the subglottal pressure is maintained at a higher-than-normal level for some 90 ms into the following vowel. This pattern is also observed in the real speech samples in figure 4. Thus the average subglottal pressure is lower on vowels following voiceless aspirated stops and higher on vowels following voiced stops. Given the known causal correlation between subglottal pressure and the intensity

of voice (Ladefoged and McKinney 1963), this accounts for the commonly observed higher intensity of vowels following voiced stops and the lower intensity of vowels following voiceless aspirated stops (House and Fairbanks 1953, Lehiste and Peterson 1959).

Figure 5 presents the results of varying only glottal resistance and leaving the pulmonic force constant as before. As can be seen, when glottal resistance is increased by only 50%, subglottal pressure increases, although it takes a relatively long time to reach the equilibrium pressure. Air flow decreases in this case. A momentary increase in subglottal pressure could also be obtained by a momentary increase in the pulmonic force, leaving the glottal resistance unchanged. In this case, however, the air flow would also increase. The situation that actually prevails in speech during stressed or emphasized syllables (where brief increases of subglottal pressure have been observed) is probably that where there is primarily just a momentary increase in glottal resistance, since it is quite commonly the case that air flow on stressed syllables is less than that on unstressed syllables (Klatt, Stevens, and Mead 1968, Broad 1968). This, then, tends to support the notion that control of F_o in speech is performed primarily by the larynx and not by the pulmonic system. The pulmonic system, in fact, can be assumed to be largely passive during speech except for providing a relatively constant force to the lungs.

Of course, more physiological investigation of pulmonic and laryngeal activity during speech is needed in order to verify these claims. But models such as the one reported here aid us in such investigations by telling what things to look for.

Acknowledgements

This research was conducted in the Institute of Phonetics, University of Copenhagen, and in the Phonology Laboratory, Department of Linguistics, University of California, Berkeley. I thank my colleagues in both laboratories, especially Robert Krones, for helpful comments. This work was supported in part by the National Science Foundation.

References

Atkinson, J. E. (1973) Aspects of intonation in speech: implications from an experimental study of fundamental frequency. University of Connecticut dissertation.

Broad, D. J. (1968) Some physiological parameters in prosodic description. Speech Communication Research Laboratories Monograph No. 3. Santa Barbara.

Chomsky, N. and Halle, M. (1968) The sound pattern of English. Harper and Row.

Ewan , W. G. and Krones, R. (1972) A study of larynx height in speech using the thyro-umbrometer. JASA 53*:345.*

Fischer-Jørgensen, E. (1972) PTK et BDG français en position intervocalique accentuée. In: A. Valdman, (ed) Papers in linguistics and phonetics to the memory of Pierre Delattre. Mouton. Pp. 143-200.

Halle, M. and Stevens, K. N. (1967) On the mechanism of glottal vibration for vowels and consonants. M.I.T. Q.P.R. 85:267-70.

Halle, M. and Stevens, K. N. (1971) A note on laryngeal features. M.I.T. Q.P.R. 101:198-213.

House, A. S. and Fairbanks, G. (1953) The influence of consonant environment upon the secondary acoustical characteristics of vowels. JASA 25:105-13.

Isshiki, N. (1969) Remarks on mechanisms for vocal intensity variations. JSHR 12:669-72.

Klatt, D. H., Stevens, K. N., and Mead, J. (1968) Studies of articulatory activity and airflow during speech. In: A. Bouhuys (ed.) Sound Production in man. Annals of the New York Academy of Sciences 155:42-54.

Ladefoged, P. (1963) Some physiological parameters in speech. Language and Speech 6:109-19.

Ladefoged, P. and McKinney, N. (1963) Loudness, sound pressure and sub-glottal pressure in speech. JASA 35:453-60.

Lehiste, I. and Peterson, G. E. (1959) Vowel amplitude and phonemic stress in American English. JASA 31:428-35.

Lieberman, P. (1967) Intonation, perception, and language. M.I.T. Press.

Ohala, J. (1970) Aspects of the control and production of speech. Working papers in phonetics, no. 15. UCLA.

Ohala, M. and Ohala, J. (1972) The problem of aspiration in Hindi phonetics. Annual Bulletin, Research Institute of Logopedics and Phoniatrics, University of Tokyo 6:39-46, and: Project on Linguistics Analysis Reports, Berkeley 16:63-70.

Rothenberg, M. (1968) The breath-stream dynamics of simple-released-plosive production. Bibliotheca Phonetica 6.

Vanderslice, R. (1967) Larynx vs. lungs: cricothyrometer data refuting some recent claims concerning intonation and archetypality. Working papers in phonetics 7:69-79. UCLA.