

An aeroacoustic approach to phonation

Richard S. McGowan

Haskins Laboratories, 270 Crown Street, New Haven, Connecticut 06511

(Received 6 April 1987; accepted for publication 6 October 1987)

A fluid mechanical, or aeroacoustic, point of view is followed to study possible sources of sound during phonation. Concentration is on two features of the vocal tract during phonation: abrupt area change from the glottis to the vocal tract and the finite length of the vocal tract. With these features, a source of sound distinct from the volume velocity source can be identified and a preliminary account of its effect on the acoustic field given. This source of sound is an oscillating force resulting from an interaction of rotational fluid motion with itself. Because of the schematic nature of the geometry of the model used here, this source may be considerably modified in actual phonation. It is concluded that specification of volume velocity is not enough to specify the source during phonation, even neglecting source-tract interaction.

PACS numbers: 43.70.Aj

INTRODUCTION

This article is part of an effort to characterize the voice source from what can be described as a fluid mechanical point of view. One of the possibilities seen immediately from such a point of view is that some of the fluid motion in the vocal tract may not be acoustic wave propagation. It is also known, however, that nonacoustic fluid motion can provide a source of sound. With these possibilities it is of interest to consider phonation using theoretical fluid mechanics. Fluid mechanics has started to receive more attention in the speech community, as shown by the works of Teager and Teager (1983a,b), Kaiser (1983), and others. We will pursue the connection between fluid motion and acoustics by using results of a specialty of fluid mechanics called aeroacoustics.

Theoretical aeroacoustics is the study of sound production and propagation in air using the equations of motion of air, including mass, momentum, and energy conservation. Both acoustic and nonacoustic types of fluid motion are considered from such a point of view. In other words, the aeroacoustic description is more general than the mechanical and circuit models currently in use. The reason for its greater generality is that theoretical aeroacoustics is a vector field theory, where the current mechanical and electrical models use a scalar field theory. In the fluid mechanical, or aeroacoustic, view, the dependent variables are fluid particle velocity, pressure, and density, although the latter two can be interchanged with any two thermodynamic quantities, such as pressure and entropy. In the aeroacoustic view, each point in space \mathbf{x} is associated with a three-dimensional vector \mathbf{v} , fluid particle velocity, as well as the two thermodynamic variables. All of these quantities are considered to be time varying. The three-dimensional velocity vector, $\mathbf{v}(\mathbf{x}, t)$, as a function of space and time is an example of a vector field.

In the current one-dimensional mechanical models of the vocal tract, the dependent variables are volume velocity, pressure, and density. Each plane perpendicular to the vocal tract axis is associated with values of these three scalars, which, again, are time variable. The electrical models' dependent variables are current instead of volume velocity and voltage instead of pressure. These are examples of scalar

field theories. In speech applications, one usually assumes that entropy is nearly constant so that only one thermodynamic variable is necessary in the description of acoustic motion, usually either pressure or density.

We will argue that the full three-dimensional character of the fluid velocity field should be part of the model of the voice source. In other words, it is not sufficient to specify the volume velocity at the glottis, which provides what is known as a monopole type source, or singularity, in the acoustic field at the glottal end of the vocal tract. We will show that this is not sufficient by showing that the nonacoustic fluid motion near the glottis will also provide, at the least, what is known as a dipole source, or singularity, in the acoustic field. This will become clear when we consider the three-dimensional character of the fluid motion near the glottis.

We will not be considering a geometry that is in detail like the vocal tract, but hope to include enough of the essential features to prompt a rethinking of voice source models. These essential features include an abrupt area change from the glottal opening into the vocal tract, where rotational motion of the fluid can be expected. This rotational motion is nonacoustic in that compressibility is not a necessary part of such motion, and the velocity field associated with such motion does not satisfy an acoustic wave equation. However, rotational motion will be seen to help provide a sound source through its pressure fluctuations. Another feature of the vocal tract is its finite length. This feature helps to determine the transfer function, or Green's function, which in turn determines how well different types of sources (monopole, dipole, etc.) radiate as a function of frequency.

The characterization of the voice source from a fluid mechanical point of view can be important if we wish to do voice synthesis given subglottal pressure and a model for vocal fold vibration. Also, since such a characterization will help to determine the strength of the sound sources, we will have a better idea as to correct approximations for sound propagation in the vocal tract.

I. DECOMPOSITION OF THE FLUID VELOCITY FIELD

In general, fluid mechanics uses the fluid particle velocity vector field instead of the scalar, fluid volume velocity

field. This means that the aeroacoustic models possess some important degrees of freedom that the one-dimensional models do not. Under unrestrictive conditions, a three-dimensional vector field can be decomposed into two different three-dimensional vector fields (Morse and Feshbach, 1953):

$$\mathbf{v} = \mathbf{v}_I + \mathbf{v}_S. \quad (1)$$

One field, \mathbf{v}_S , is called solenoidal and the other, \mathbf{v}_I , irrotational. The velocity vector at each point of space is written as the sum of these two types of vectors. The solenoidal part of the field can support rotational motion of the fluid, but not compression and expansion. The irrotational part of the field can support compression and expansion, but not rotation. Therefore, the irrotational field is a necessary part of acoustic motion. These properties can be stated mathematically with the following identities. The divergence of the solenoidal fields is zero and the curl of the irrotational field is zero:

$$\nabla \cdot \mathbf{v}_S = 0, \quad \nabla \times \mathbf{v}_I = 0. \quad (2)$$

The decomposition of Eq. (1) has been useful for describing the production and propagation of sound, and the interaction of sound with other modes of fluid motion. Chu and Kovaszany (1957) analyzed a general small disturbance in a fluid using this decomposition. Crow (1970) has analyzed the use of this decomposition in the context of aeroacoustics, where he reminds the reader that the decomposition is mathematical and that two velocity fields are not measured separately, even in the acoustic farfield. However, he does show that the pressure fluctuations associated with changes in the solenoidal field will propagate in the irrotational field as sound. In summary, the aeroacoustic point of view allows for two types of velocity fields: one irrotational for the compression and expansion, which supports acoustic motion, and the other, solenoidal, which will not. However, the latter field can provide a source of sound, even though sound cannot propagate in this field alone.

II. ROTATIONAL MOTION AND PRESSURE HEAD LOSS

Is rotational motion a part of phonation? There is experimental evidence that it is, for example, the experiments of van den Berg *et al.* (1957), Scherer and Titze (1983), and Gauffin *et al.* (1983). In their static models, these investigators all note a loss of pressure head, or stagnation pressure, before and after their model glottises. Stagnation pressure for low Mach number flows is the sum of the usual static pressure p and kinetic pressure $(\rho_0/2)|\mathbf{v}|^2$:

$$p_{ST} = p + (\rho_0/2)|\mathbf{v}|^2, \quad (3)$$

where ρ is the density, and the subscript zero denotes the value in air at rest. In an inviscid, incompressible fluid, and in the absence of time variation and rotational fluid motion, the stagnation pressure p_{ST} is a constant. (The velocities are small enough that the fluid can be treated as incompressible just above the glottis.) The pressure head losses noted in these studies were too large to be solely due to viscous loss. Therefore, these investigators have shown, indirectly, the existence of rotational motion near their model glottises. In fact, van den Berg *et al.* (1957) attributed this head loss, or

nonlinear resistance, to turbulence, which is both rotational and random motion of fluid. The essential ingredient for the head loss is the exchange of energy from irrotational motion to rotational motion, so we will not be too concerned about the degree of randomness in the motion. While there may be a great deal of randomness in the static model configurations, there appears to be agreement that actual phonation in "normal" speaking voice does not allow enough time between cycles for the formation of much turbulence, and that the fluid motion is of a periodic nature (Teager, personal communication).

To apply the results of the static experiments to the dynamic voice source, we make the assumption that the acceleration terms in the equations of motion do not greatly affect the estimate of resistance. In other words, the resistance measured while the dynamic configuration is passing through a particular static configuration is the same as the resistance of that static configuration. We call this the quasisteady approximation, and it is a common approximation in speech acoustics (Flanagan, 1965). The limits of validity for this approximation are discussed in Sec. III.

We can get something of a worst case estimate of the head loss above the glottis using the quasisteady approximation. This is a well-known argument quoted by Batchelor (1970) and used by Ishizaka and Matsudaira (1972). Using momentum and mass conservation arguments, the change of pressure head is

$$\Delta p_{ST} = -(\rho_0/2)|\mathbf{v}_g|^2(1 - A_g/A_T)^2, \quad (4)$$

where \mathbf{v}_g is the fluid particle velocity in the glottis, A_g is the time-varying area of the glottis, and A_T is the vocal tract area (assuming a cylindrical shape for the vocal tract). Given the small ratio of maximum glottal area to the vocal tract area, the head loss can be quite large. The actual amount of head loss depends upon the exact geometry of the configuration. In Sec. III, the relationship between the pressure head loss and other quantities relevant to fluid motion will be discussed further.

III. FORCING ASSOCIATED WITH ROTATIONAL MOTION

In this section, we will make explicit the relationship between head loss and rotational motion using the quasisteady approximation. The loss of pressure head is the result of forcing between the solid boundaries and the air, which is realized in the air by what might be called the vorticity-velocity interaction force. Vorticity is a measure of the rotation of the fluid and is given by

$$\boldsymbol{\omega} = \nabla \times \mathbf{v} = \nabla \times \mathbf{v}_S. \quad (5)$$

The relation between loss of pressure head and this force can be understood by examining the momentum equation for an inviscid fluid, or Euler's equation:

$$\frac{\partial \mathbf{v}}{\partial t} + \mathbf{v} \cdot \nabla \mathbf{v} = -\frac{1}{\rho} \nabla p. \quad (6)$$

Using the vector identity $\mathbf{v} \cdot \nabla \mathbf{v} = \frac{1}{2} \nabla |\mathbf{v}|^2 - (\mathbf{v} \times \boldsymbol{\omega})$ and neglecting the acceleration term on the basis of the quasisteady approximation gives

$$(1/\rho)\nabla p + \frac{1}{2}\nabla|\mathbf{v}|^2 = \mathbf{v}\times\boldsymbol{\omega}. \quad (7)$$

In the case of an isentropic fluid (constant entropy),

$$c^2 dp = dp, \quad (8)$$

where c is the adiabatic speed of sound. (We will also be assuming homentropic conditions above the glottis, which means that the entropy is uniform in the air.) Consistent with the approximations already made in Eq. (7), density fluctuations can be neglected if the Mach number of the flow is assumed small. Equation (7) becomes

$$\nabla[p + (\rho_0/2)|\mathbf{v}|^2] = \rho_0(\mathbf{v}\times\boldsymbol{\omega}). \quad (9)$$

This equation relates the gradient of the stagnation pressure, or pressure head, to the term on the right, which we will call the vorticity-velocity interaction force (per unit volume).

We have used the quasisteady approximation, but have not yet set down the conditions under which it is valid. The quasisteady approximation can be made if the frequency of oscillation f , characteristic fluid particle velocity U , and length of region over which the head loss occurs Δx are related:

$$f\Delta x/U \ll 1. \quad (10)$$

Despite the fact that the head loss in static experiments may occur within 1 or 2 cm, this condition is very restrictive, even at $U = 5000$ cm/s (van den Berg *et al.*, 1957; Scherer and Titze, 1983; Scherer *et al.*, 1983).

Now, it is possible to sketch the essential parts of the real vocal tract that are pertinent to this article. We will idealize the vocal tract to be a straight, cylindrical tube of length l with a contraction for the glottis at $x_1 = 0$ (see Fig. 1). There is pressure head loss upstream and downstream of the glottis in static experiments. However, we are primarily interested in the vorticity-velocity force to the right (downstream) of the glottis, this region being the best coupled to the atmosphere. Therefore, we will assume that the contrac-

tion to the left (upstream) of the glottis is smooth in our idealization so that there is no head loss before the glottis. (Possible loss before the glottis could be taken into account with the addition of more detail, but doing so now would add little to the picture being presented here.) However, the expansion from the glottis to the vocal tract is abrupt so that there is separation and vorticity is unattached to the right of the glottis.

In the case of phonation, we can imagine an oscillating jet with associated shear fluid motion above the glottis. Vorticity associated with the shear layer of the jet is directed azimuthally, so that the vectors representing the vorticity, as a set, wrap themselves around the x_1 axis in a ring. There may also be a secondary flow between the jet shear and the walls in the form of an oscillating ring (donut) of rotating fluid with an axis of symmetry coincident with the x_1 axis. The vorticity associated with such motion is, again, azimuthally directed. The fluid velocity in the region of the jet will have components along the x_1 axis and in the radial direction from this axis. Thus the vectors representing the vorticity-velocity interaction force, $\rho_0(\mathbf{v}\times\boldsymbol{\omega})$, have components in both the axial and radial directions. By the symmetry of the situation, the radial components tend to cancel one another, while the axial components add. A spatial average of these vectors near the glottis produces a vector directed toward the glottis and oscillating in magnitude. These statements are consistent with the observed pressure head loss in the axial direction [see Eq. (9)].

It should be noted that the vorticity-velocity interaction force is a nonlinear function of velocity because vorticity is a function of velocity by Eq. (5). However, in phonation, the fluid velocity is a periodically time-varying function with a steady component. Despite the fact that the vorticity-velocity interaction force is a nonlinear function of fluid velocity, the fundamental frequency, as well as a steady component and harmonics, appears in the forcing term.

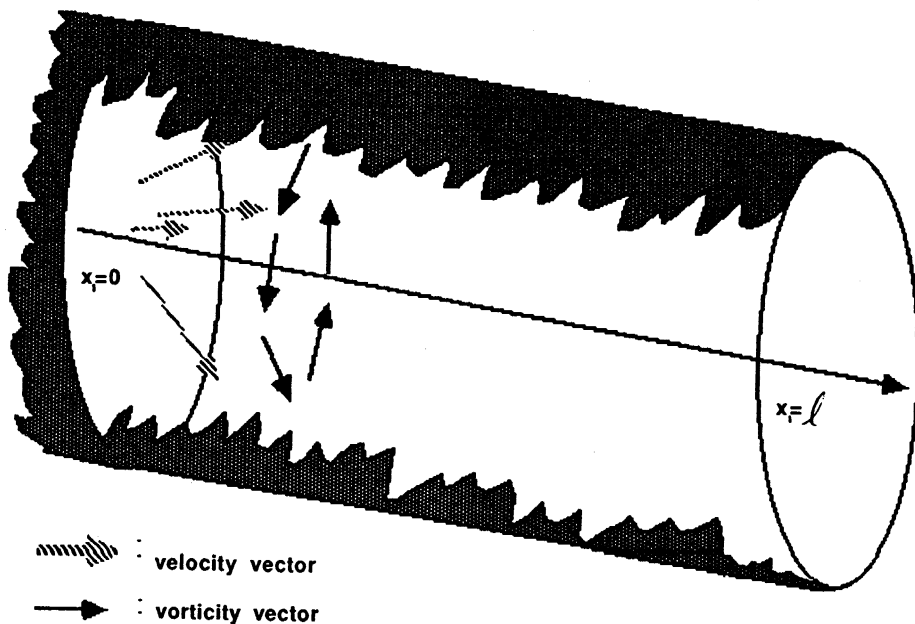


FIG. 1. Geometry of tube.

We conclude that the net effect of the vectors $\rho_0(\mathbf{v} \times \boldsymbol{\omega})$ is a force vector that is oscillating in magnitude and directed upstream, toward the glottis. In Sec. IV, we discuss the relationship of this force to sound production.

IV. SOUND SOURCES

To have sound, it is necessary to have the potential energy of compression and expansion, as well as the kinetic energy of fluid motion. One way to create sound is to change the density of fluid by adding mass into a fixed volume in an oscillatory way. If the region where this addition occurs is small compared to the sound speed divided by the frequency, or wavelength, such a source is called a monopole.

If two monopoles of equal strength, but 180 deg out of phase, are put near one another, that is, at a small distance compared to wavelength, the resulting source is called a dipole. This is an inefficient way to produce sound because the two monopoles nearly cancel at small separations. The monopole and dipole described above are known as acoustically compact sources because their sizes are small compared to wavelength (Lighthill, 1978).

When there is an oscillating force, fluid will accelerate from one region to another and back again, without a net change in density. If the spatial domain where this forcing is effective is small compared to the wavelength of sound, then this domain behaves as a dipole source. There is a force between the vocal folds and the fluid, realized in the fluid by the vorticity-velocity interaction force. We need to specify this force to help characterize the voice source. Specifying the volume velocity is not enough.

Before going on to the formalism of aeroacoustic theory, we will summarize the argument thus far. From the loss of pressure head we are able to infer an oscillating vorticity-velocity interaction force under the quasisteady approximation ($f\Delta x/U \ll 1$). This force supplies a dipole type of singularity in the sound field, assuming that the region of head loss is small compared to wavelength of sound, that is, assuming acoustic compactness ($f\Delta x/c \ll 1$). The heuristic arguments leading to this conclusion have been sketched out without the mathematical theory of aeroacoustics that relates forces to sound production. It will be helpful to go on and use this theory to help clarify the argument. It should be noted that only our estimate of the size of the vorticity-velocity interaction force, using Eqs. (4) and (9), depends on the correctness of the quasisteady approximation: Its existence is independent of this approximation. In fact, there is experimental evidence of pressure head loss, or nonlinear resistance, in dynamic situations (Sivian, 1935; Ingard and Ising, 1967). However, whether we can use some common approximations in aeroacoustic theory will be seen to depend on the less restrictive acoustic compactness assumption.

V. AEROACOUSTICS

Aeroacoustics concerns itself with the production of sound by the interaction of fluid with itself and solid surfaces. A classic work in this area is that of Lighthill in 1952. Lighthill derived an exact nonlinear wave equation from the

equations of motion for air, where he wrote the familiar linear wave operator on the left applied to the perturbation density to be equal to the nonlinear right-hand side (Lighthill, 1952).

The form in which Lighthill originally wrote the equation is

$$\frac{\partial^2 \rho}{\partial t^2} - c^2 \nabla^2 \rho = \frac{\partial^2 T_{ij}}{\partial x_i \partial x_j}, \quad (11)$$

where

$$T_{ij} = \text{stress tensor} = \rho v_i v_j + p_{ij} - c^2 \rho \delta_{ij},$$

$$p_{ij} = p \delta_{ij} + \mu \left(-\frac{\partial v_i}{\partial x_j} - \frac{\partial v_j}{\partial x_i} + \frac{2}{3} \frac{\partial v_k}{\partial x_k} \delta_{ij} \right),$$

$$\delta_{ij} = \begin{cases} 1, & \text{if } i=j, \\ 0, & \text{otherwise,} \end{cases}$$

and μ is the coefficient of viscosity, $\mathbf{v} = (v_1, v_2, v_3)$ is the fluid particle velocity, t is time, and $\mathbf{x} = (x_1, x_2, x_3)$ is the spatial coordinate. The left-hand side describes the propagation of an acoustic wave in a uniform fluid at rest. If the quantity on the right-hand side has bounded support (is nonzero only in a bounded region), we may think of it as providing the stress differences necessary for sound propagation into the remaining space, outside the support of the right-hand side (Lighthill's acoustic analogy).

Largely motivated by engineering problems, much work has been done on understanding the acoustic analogy by rewriting Eq. (11), solving it in various geometries, and simplifying the source terms. Lighthill neglected viscous stress, and he argued that pressure and density are isentropically related and there are no entropy inhomogeneities to a good approximation in low Mach number flows without heat sources. In this case,

$$T_{ij} \approx \rho_0 v_i v_j. \quad (12)$$

We make the same assumptions and approximations in this article. A vector identity used above can be invoked again to write

$$\frac{\partial T_{ij}}{\partial x_i \partial x_j} \approx \rho_0 \operatorname{div}(\boldsymbol{\omega} \times \mathbf{v}) + \frac{\rho_0}{2} \nabla^2 |\mathbf{v}|^2, \quad (13)$$

where a term involving compressibility has been neglected.

In 1964, Powell argued that the source term exhibited in Eq. (13) can be further simplified to a good approximation, in certain cases, by dropping the second term on the right-hand side (Powell, 1964). This is intuitively plausible because regions of large vorticity are associated with large amounts of sound, while the second term seems to have little to do with sound production. It is known that vorticity is associated with sound production because of such phenomena as the sound produced by vorticity shed from a wire (Aeolian tones). This was made more precise by Howe, who treated the specific stagnation enthalpy as the acoustic quantity that is propagated instead of density or pressure (Howe, 1975). This has the effect of moving the second term on the right-hand side of Eq. (13) to the left-hand side of the wave equation, Eq. (11). In the absence of entropy inhomogeneities, and for small Mach number flow, Howe's equation can be written

$$\left(\frac{1}{c^2} \frac{D^2}{Dt^2} - \nabla^2\right)B = \text{div}(\omega \times \mathbf{v}), \quad (14)$$

where $B = \int (dp/\rho) + \frac{1}{2}|\mathbf{v}|^2$ is the stagnation enthalpy of the isentropic air, $D/Dt = \partial/\partial t + \bar{U}(\partial/\partial x_1)$, and \bar{U} is the mean flow particle in the x_1 direction (mean flow in the sense of time average) (Howe, 1975).

The propagated quantity, or dependent variable, is now stagnation enthalpy. We have not written the most general form for the stagnation enthalpy, but have written a special form based on the assumption of no entropy variations in time or space. The relationship between the usual acoustic perturbation pressure, $p' = p - p_0$, to the time-varying portion of stagnation enthalpy B' is that of a limiting case of a more general model. In the case that second-order time-variable flow quantities can be neglected, B' can be linearized. Specifically, in the case of plane-wave propagation, this linearization can be written $B' = (p'/\rho_0)(1 + M)$, where $M = \bar{U}/c$ is the mean flow Mach number. Howe's formulation is general because it allows for wave propagation, even in the case the second-order, time-variable quantities cannot be neglected. Also, the wave operator is slightly modified to take account of mean (time average) convection of the sound wave.

It is interesting to examine Eq. (14) in the region where the right-hand side is nonzero. Again, if this region is acoustically compact, we can let $c \rightarrow \infty$ and write, approximately,

$$-\nabla^2 B \approx \text{div}(\omega \times \mathbf{v}), \quad (15)$$

where $\nabla B = \nabla(p/\rho_0 + \frac{1}{2}|\mathbf{v}|^2)$. [By Eq. (8), $\rho \approx \rho_0$, a constant to the approximation used here.] Equation (15) is just the divergence of Eq. (9), and shows how the Laplacian of pressure head relates to the divergence of the vorticity-velocity interaction force in acoustically compact regions. Unlike Eq. (9), this relationship has been written without the restrictive quasisteady assumption. Only acoustic compactness needs to be invoked.

It should be noted that, in the case that $(\omega \times \mathbf{v})$ is zero,

$$\nabla B = -\frac{\partial \mathbf{v}}{\partial t}. \quad (16)$$

This follows from Euler's equation, Eq. (6), without any compactness assumptions. We will use Howe's form of the acoustic analogy, Eq. (14), in our discussion.

VI. INTEGRAL SOLUTION

We will take the idealized cylindrical tube in Fig. 1 and formally solve Eq. (14) to the right of the glottis using the method of Green's functions. A Green's function, being the space-time analog to the impulse response function, produces an integral solution to the partial differential equation (14) based on the principle of superposition. First, we need to specify a control volume (see Fig. 1). The control volume V consists of the tube to the right of the glottis and to the left of the mouth, that is, from $x_1 = 0$ to $x_1 = l$. The tube will be assumed hard walled, and only plane-wave propagation will be considered for now. This reduces the problem to solving Eq. (14) in one spatial variable, x_1 , so that the boundary conditions at the ends of V can be considered as averages over the appropriate disk. [Note that, in order to derive the

source on the right-hand side of Eq. (14), the three-dimensional character of the fluid field still needs to be considered.]

In the solution of this problem, periodicity in time will be assumed, so that all dependent variables can be Fourier decomposed, the problem solved in each frequency f , and the results superposed for the solution. We will consider the solution for an arbitrary frequency component. The Fourier transform of Eq. (14) gives

$$\left(\frac{d^2}{dx_1^2} + \frac{2iM}{1-M^2}k \frac{d}{dx_1} + \frac{k^2}{1-M^2}\right)\tilde{B} = -\frac{\tilde{F}}{1-M^2}, \quad (17)$$

where $k = \Omega/c$, $\Omega = 2\pi f$, $F = \text{div}(\omega \times \mathbf{u})$, and a tilde denotes Fourier transform. The boundary conditions are

$$\frac{d\tilde{B}}{dx_1}\Big|_{x_1=0} = \tilde{Q}, \quad \tilde{B}|_{x_1=l} + Z_l \frac{d\tilde{B}}{dx_1}\Big|_{x_1=l} = 0,$$

where $Q = -(1/A_g)(\partial A_g/\partial t)$. Recall that \mathbf{v}_g denotes the particle velocity at the glottis, A_T denotes the area of the cylindrical tube, and A_g denotes the area of the glottis. The boundary condition at $x_1 = 0$ is inhomogeneous because of the periodic mass addition at the glottis [see Eq. (16)]. We have approximated this situation with no source-tract interaction, that is, as a constant volume velocity source. This is an approximation that has been used in speech acoustics, despite the fact that tract impedance is comparable to source impedance at resonant frequencies (Flanagan, 1965).

Here, Z_l denotes the ratio of pressure head, divided by density, to fluid particle acceleration at the mouth in the absence of vorticity-velocity forcing [see Eq. (16)]. Further, if fluctuating vorticity is assumed to decay before reaching the mouth by convection, we may linearize B' , and the boundary condition at $x_1 = l$ is the radiation boundary condition, with Z_l related to the acoustic impedance by the expression $Z_l = (-i\Omega\rho_0)^{-1} \times (\text{acoustic impedance})$. More experimental facts are needed before this boundary condition can be set with great confidence.

The form of Eq. (15) is not the most conducive for solution via the method of Green's functions. We would like to transform it into a self-adjoint problem (Carrier and Pearson, 1968). Let

$$\tilde{w} = \tilde{B} \exp\{[iM/(1-M^2)]kx_1\}. \quad (18)$$

With this change of variable, the differential equation (17) becomes

$$\left[\frac{d^2}{dx_1^2} + \left(\frac{k}{1-M^2}\right)^2\right]\tilde{w} = \frac{-\tilde{F}}{1-M^2} \exp\left(\frac{ik}{1-M^2}x_1\right), \quad (19)$$

with boundary conditions

$$\frac{-iM}{1-M^2}k\tilde{w}\Big|_{x_1=0} + \frac{d\tilde{w}}{dx_1}\Big|_{x_1=0} = \tilde{Q}$$

and

$$\left(1 - Z_l \frac{iM}{1-M^2}\right)\tilde{w}\Big|_{x_1=l} + Z_l \frac{d\tilde{w}}{dx_1}\Big|_{x_1=l} = 0.$$

To solve the transformed problem, we find the Green's function, $\tilde{G} = \tilde{G}(x_1, y_1; f)$, which satisfies

$$\left[\frac{d^2}{dx_1^2} + \left(\frac{k}{1-M^2} \right)^2 \right] \tilde{G} = \delta(x_1 - y_1), \quad (20)$$

where y_1 is the source position and δ is the Dirac delta function. Here, \tilde{G} satisfies the homogeneous boundary conditions

$$\frac{-iM}{1-M^2} k \tilde{G} \Big|_{x_1=0} + \frac{d\tilde{G}}{dx_1} \Big|_{x_1=0} = 0$$

and

$$\left(1 - Z_l \frac{iM}{1-M^2} k \right) \tilde{G} \Big|_{x_1=l} + Z_l \frac{d\tilde{G}}{dx_1} \Big|_{x_1=l} = 0.$$

For $l > x_1 > y_1 > 0$, the Green's function has the form (see Appendix A)

$$\tilde{G}(x_1, y_1; f) = A^+ \cos\left(\frac{k}{1-M^2} x_1\right) + D^+ \sin\left(\frac{k}{1-M^2} x_1\right), \quad (21)$$

where

$$D^+ = \frac{1-M^2}{k} \left(\frac{\cos\{[k/(1-M^2)]y_1\} + iM \sin\{[k/(1-M^2)]y_1\}}{1-iM\eta} \right),$$

$$A^+ = \eta D^+,$$

and

$$\eta = \frac{\tan\{[1/(1-M^2)]kl\} + ([1/(1-M^2)]kZ_l/[1 - [iM/(1-M^2)]kZ_l])}{([1/(1-M^2)]kZ_l/[1 - [iM/(1-M^2)]kZ_l])\tan[1/(1-M^2)kl] - 1}.$$

After transforming back to the dependent variable \tilde{B} , and using the definitions of \tilde{F} and \tilde{Q} , the integral solution can be written (see Appendix B)

$$\begin{aligned} \tilde{B}(x_1; f) &\equiv \tilde{B}_d(x_1; f) + \tilde{B}_m(x_1; f) \\ &= \left(\frac{1}{1-M^2} \right) \int_0^l (\omega \times \mathbf{v})_y \\ &\quad \times \frac{d}{dy_1} \left[\tilde{G} \exp\left(\frac{-iM}{1-M^2} k(x_1 - y_1)\right) \right] dy_1 \\ &\quad - \tilde{G} \Big|_{y_1=0} \exp\left(\frac{-iM}{1-M^2} kx_1\right) \frac{1}{A_T} \left(\frac{\partial A_s |\mathbf{v}_s|}{\partial t} \right) \end{aligned} \quad (22)$$

for $l > x > 0$, where $(\omega \times \mathbf{v})_y$ denotes the spatial average of the x_1 component of $(\omega \times \mathbf{v})$ over the area of the tube at y_1 from the glottis. We have assumed that $(\omega \times \mathbf{v})$ is zero at $y_1 = 0$ and $y_1 = l$ to perform an integration by parts. The second term on the right represents the standard volume velocity source, which can be considered to be a monopole source. The first term represents the source due to the vorticity-velocity interaction force and is considered to be of a dipole type. This can be seen by noting that the Green's function for the dipole source is the spatial derivative of the Green's function for the monopole source. Under the assumption of acoustic compactness, that is, that the pressure head loss occurs in a distance Δx small compared to wavelength, this integral is nonzero, and hence provides a dipole source.

The overall picture presented here is similar to that of Howe (1980) and Bechert (1980), who describe the attenuation of sound exiting from a pipe with mean flow. Some of the energy from the irrotational flow at the glottis is transferred to rotational energy because of the abrupt area change; the remainder is transmitted as sound. The forcing associated with this transfer of energy is the vorticity-velocity

interaction force, which, in turn, acts as a dipole source of sound. The partition of energy is shown schematically in Fig. 2.

VII. SOME ESTIMATES

The ratio of the sound field due to the dipole, forcing source to that due to the monopole, volume velocity source depends on certain flow parameters. This dependence can be estimated using Eq. (22), and an attempt to set a numerical upper bound on the size of the ratio of fields will be made, assuming an infinite glottal impedance. Because we must use very crude estimates of the parameters, this upper bound cannot be considered a very close estimate of the ratio. This estimate should be made experimentally, but we will be able to indicate some trends using this bound.

An upper bound on the vorticity-velocity forcing is provided by Eq. (4). Using acoustic compactness, and for ob-

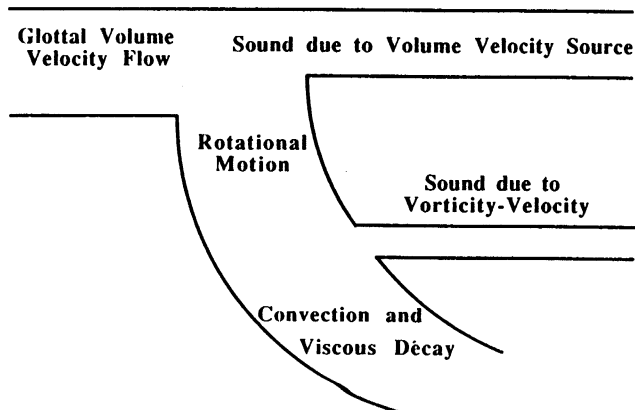


FIG. 2. Partition of energy.

ervation point x_1 away from the source at y_1 , we can approximate the dipole as a point dipole at an axial position $y_1 = K\Delta x$, where K is a constant of order 1. With these simplifications, Eq. (22) can be written

$$\frac{|\tilde{B}_d|}{|\tilde{B}_m|} = \left| \frac{A_T}{2c} \frac{[(1 - A_g/A_T)^2 |\tilde{v}_g|^2]}{(\tilde{A}_g |\tilde{v}_g|)} \right| \times \left[\left(K \frac{(1 + M^2)}{(1 - M^2)} k\Delta x \right)^2 + 4M^2 \right]^{1/2} + O\left(\frac{k\Delta x}{1 - M^2}\right)^2, \quad (23)$$

where the O term denotes the asymptotic dependence of the error on $k\Delta x/(1 - M^2)$, which is smaller under the compactness assumption and for small Mach numbers. The first factor is the ratio of the dipole source strength (divided by Δx) to the monopole source strength. The second factor is the ratio of the dipole transfer function (multiplied by Δx) to the monopole transfer function.

To illustrate a possible frequency dependence for the first factor, we will take a simple model for the time waveform of the glottal fluid particle velocity and the glottal area function. The glottal fluid particle velocity is assumed to be a periodic rectangular wave, while the glottal area function is assumed to be triangular (Flanagan, 1965). If U_g is the amplitude of the glottal particle velocity, \mathcal{A}_g is the maximum glottal area, β is the duty cycle, and f_0 is the fundamental frequency, then the Fourier transforms for the waveforms, normalized to the period $1/f_0$, can be written

$$(\tilde{A}_g |\tilde{v}_g|) = 4U_g \mathcal{A}_g / \beta \pi \{ \sin^2[(n/2)\beta\pi] / n^2 \} \quad (24)$$

and

$$\left(1 - \frac{\mathcal{A}_g}{A_T}\right)^2 |\tilde{v}_g|^2 = 2U_g^2 \left(1 - \frac{\mathcal{A}_g}{A_T}\right)^2 \times \frac{\sin(n\beta\pi)}{n} + \frac{1}{n} O\left(\frac{\mathcal{A}_g}{A_T}\right),$$

where n is the harmonic number. Combining these results, Eq. (23) becomes

$$\frac{|\tilde{B}_d|}{|\tilde{B}_m|} = \frac{U_g}{c} \frac{A_T}{\mathcal{A}_g} n \left| \frac{\beta\pi}{4} \left(1 - \frac{\mathcal{A}_g}{A_T}\right)^2 \right| \times \frac{\sin(n\beta\pi)}{\sin^2[(n/2)\beta\pi]} + O\left(\frac{\mathcal{A}_g}{A_T}\right) \times \left[\left(K \frac{(1 + M^2)}{(1 - M^2)} k\Delta x \right)^2 + 4M^2 \right]^{1/2} + O\left(\frac{k\Delta x}{1 - M^2}\right)^2. \quad (25)$$

Here, U_g/c is the peak Mach number of the glottal pulse, which is on the same order of magnitude as the mean flow Mach number M . It is reasonable that the ratio of dipole strength to monopole strength increases with peak Mach number because the dipole strength is a nonlinear function of glottal fluid particle velocity, while the monopole is a linear function of this same quantity. This nonlinear versus linear dependence also accounts for the source strength ratio's de-

pendence on frequency. From the first factor in Eq. (25), it is seen that the ratio of the magnitudes of the spectral envelopes of source strengths is an increasing function of frequency, after the fundamental. (Here, we are assuming that $\beta > \frac{1}{2}$ and $\mathcal{A}_g/A_T < \frac{1}{2}$.)

The second factor in Eq. (25), representing the ratio of the different Green's, or transfer, functions is small for small mean flow Mach numbers and small $k\Delta x/(1 - M^2)$. However, it is a strictly increasing function of $k\Delta x/(1 - M^2)$ and, therefore, of frequency f . This adds another power to the frequency dependence already noted for the ratio of source strengths. The Green's function dependence on frequency is easily given a physical interpretation. The monopole source is a velocity source located at a pressure maximum, and therefore supplies energy to the acoustic field efficiently. On the other hand, the dipole is a pressure source near the same pressure maximum. This source and the pressure maximum get closer, in terms of wavelength, as the frequency decreases, and so it is a particularly inefficient source at low frequencies.

While the assertion that the ratio of the magnitudes of the spectral envelopes in the first factor of Eq. (25) is an increasing function of frequency is bounded below by the fundamental frequency, the estimate on the ratio of transfer functions is taken in the low-frequency limit. That is, Eq. (25) should be valid above the fundamental, while for $K\Delta x \sim 1$ cm, the dipole source region can be considered to be compact up to about 3000 Hz, which provides the upper frequency limit for Eq. (25).

At a fundamental frequency of 200 Hz, the upper bound on the ratio $|\tilde{B}_d|/|\tilde{B}_m|$ provided by Eq. (25), neglecting the trigonometric functions, is 0.015 or -36 dB for $A_T/\mathcal{A}_g = 10$, $U_g/c = 0.1$, $M = 0.05$, $K\Delta x = 1$ cm, and $\beta = \frac{1}{2}$. At $f = 3000$ the upper bound is 3.4 or 10 dB.

It should be clear now that we have had to make some crude estimates of some of the parameters involved in Eq. (25) to obtain these bounds. As stated above, this upper bound provides a poor estimate, particularly as the frequency increases. The use of Eq. (4), which provides an estimate of maximum possible head loss, becomes suspicious as frequency increases. This estimate not only is made under the quasisteady approximation, but it is assumed for the integral in Eq. (22) that the maximum vorticity-velocity force applies over the entire cross-sectional area of the duct, which requires acoustic compactness in the radial direction. These approximations become more suspect with increasing frequency.

The use of a rectangular wave for the glottal particle velocity overestimates the importance of the high frequencies for the dipole, while the triangular wave for the volume velocity underestimates the high frequencies for the monopole source. A more refined analysis would show a slower growth in the ratio of dipole field to monopole field with frequency than what was shown here. In general, however, we can expect the ratio of the magnitudes of the spectral envelopes to be an increasing function of frequency, for large enough frequencies, because the quadratic dependence of the dipole strength on velocity increases the importance of any discontinuities in the particle velocity waveform.

We are particularly unsure of the estimates of $K\Delta x$ and the glottal impedance. While the glottal end was assumed to have infinite impedance, thus favoring the monopole source, $K\Delta x$ was probably overestimated to compensate. All numerical estimates are very sensitive to changes in these parameters.

VIII. CONCLUSION

The sources of sound in phonation cannot be completely characterized by volume velocity. In considering a schematic version of the vocal tract, we have indentified a dipole source along with the monopole volume velocity source. The dipole source is the result of forcing between the solid surfaces and the air, which is realized in the air as the vorticity-velocity interaction force, and which also results in resistance to fluid flow. In the circuit models, this resistance is not considered as a source or sink of sound, but rather as a means of deriving the relationship between subglottal pressure and volume velocity. Indeed, the resistance has a role in determining the time evolution of the glottal pulse given the subglottal pressure, but, beyond this role, it has acoustic consequences. Although we were unable to estimate the absolute magnitude of the field due to the dipole source, the acoustic consequences can be small, especially at low frequencies. As the folds are abducted for breathy voice, the dipole component should get stronger, and, in fact, is recognized as the main source of random noise. In general, we can expect that the magnitude of the spectral envelope of the ratio of the dipole source strength to the monopole source strength increases with frequency.

In the work presented here we have used some aeroacoustic theory to explain how certain fluid mechanical phenomena can be related to sound. In doing so we have kept the linear source-filter picture of speech production by referring to nonlinear terms as "source" terms for linear propagation. There will be some controversy as to whether this is the best approach. Further, some source-tract interaction will have

to be included in the future. The volume velocity is not a given, but, rather, is a part of the entire flow field from the lungs to the atmosphere. Possible small sources provided by the movement of the folds themselves should also be considered.

Because of the schematic nature of the geometrical model used here, we cannot expect that the source named here can be transferred directly to the real vocal tract. However, the vorticity-velocity interaction force must be considered in more realistic geometries, because it provides part of the inhomogeneous term in the acoustic analogies [e.g., Eq. (14)]. The action of this term in a more realistic geometry depends on the interaction of the rotational motion of the air with the epiglottis, palate, and even the lips, as well as the actual shape of the jet from the glottis. Put in more physical terms, the forces between solid structures and the air must be considered before a complete picture of the voice sources during phonation can be known. The tools that the aeroacoustics literature provides allow us to fill the conceptual gap between the fluid mechanics of air in its generality and acoustics.

ACKNOWLEDGMENTS

The work on this article has been supported by National Institute of Health Grants NS-13617 and HD-01994 to Haskins Laboratories. The author appreciates the comments of Dr. Ronald Scherer, as well as those of the reviewers. The author extends special thanks to Dr. Thomas Baer for many helpful discussions.

APPENDIX A

The Green's function satisfies Eq. (20) with the homogeneous boundary conditions. The delta function singularity is generated by the second derivative of the Green's function in this equation at $x_1 = y_1$. Otherwise, the Green's function satisfies the homogeneous equation. This can be summed up in the following:

$$\tilde{G} = \begin{cases} A^+ \cos\{[k/(1-M^2)]x_1\} + D^+ \sin\{[k/(1-M^2)]x_1\}, & \text{for } l > x_1 > y_1 > 0, \\ A^- \cos\{[k/(1-M^2)]x_1\} + D^- \sin\{[k/(1-M^2)]x_1\}, & \text{for } 0 < x_1 < y_1 < l, \end{cases}$$

$$\lim_{x_1 \rightarrow y_1^-} \tilde{G} - \lim_{x_1 \rightarrow y_1^+} \tilde{G} = 0,$$

$$\lim_{x_1 \rightarrow y_1^-} \frac{d\tilde{G}}{dx_1} - \lim_{x_1 \rightarrow y_1^+} \frac{d\tilde{G}}{dx_1} = -1.$$

The latter two conditions, along with the homogeneous boundary conditions, serve to determine A^+ , D^+ , A^- , and D^- . The solution for $x_1 > y_1$ is shown in Eq. (21).

APPENDIX B

The integral solution can be derived by considering the following identity (Green's theorem):

$$\int_0^l \left(\tilde{G} \frac{d^2 \tilde{w}}{dy_1^2} - \tilde{w} \frac{d^2 \tilde{G}}{dy_1^2} \right) dy_1 = \left(\tilde{G} \frac{d\tilde{w}}{dy_1} - \tilde{w} \frac{d\tilde{G}}{dy_1} \right) \Big|_0^l.$$

Using the boundary conditions associated with the differential Eqs. (19) and (20), we obtain

$$\left(\tilde{G} \frac{d\tilde{w}}{dy_1} - \tilde{w} \frac{d\tilde{G}}{dy_1} \right) \Big|_0 = -\tilde{G} \Big|_{y_1=0} \tilde{Q}.$$

The differential equations (19) and (20) provide the following identity:

$$\begin{aligned} \int_0^l \left(\tilde{G} \frac{d^2 \tilde{w}}{dy_1^2} - \tilde{w} \frac{d^2 \tilde{G}}{dy_1^2} \right) dy_1 \\ = -\frac{1}{1-M^2} \int_0^l \tilde{F} \tilde{G} \exp\left(\frac{-iM}{1-M^2} ky_1\right) dy_1 - \tilde{w}(x_1). \end{aligned}$$

Combining and transforming back to the dependent variable $\tilde{B}(x_1)$, we obtain Eq. (22).

- Batchelor, G. K. (1970). *Introduction to Fluid Dynamics* (Cambridge U. P., Cambridge, England), pp. 373–375.
- Bechert, D. W. (1980). "Sound absorption caused by vorticity shedding, demonstrated with a jet flow," *J. Sound Vib.* **70**, 389–405.
- Carrier, G. F., and Pearson, C. E. (1968). *Ordinary Differential Equations* (Ginn, Waltham, MA), pp. 20–21.
- Chu, B.-T., and Kovaszany, L. S. G. (1957). "Non-linear interactions in a viscous heat-conducting compressible gas," *J. Fluid Mech.* **3**, 494–514.
- Crow, S. C. (1970). "Aerodynamic sound emission as a singular perturbation problem," *Stud. Appl. Math.* **44**, 21–44.
- Flanagan, J. L. (1965). *Speech Analysis, Synthesis, and Perception* (Springer, Berlin), pp. 39–47.
- Gauffin, J., Binh, N., Ananthapadmanabha, T. V., and Fant, G. (1983). "Glottal geometry and volume velocity waveform," in *Vocal Fold Physiology: Contemporary Research and Clinical Issues*, edited by D. M. Bless and J. H. Abbs (College-Hill, San Diego, CA), pp. 194–201.
- Howe, M. S. (1975). "Contributions to the theory of aerodynamic sound, with applications to excess jet noise and theory of the flute," *J. Fluid Mech.* **71**, 625–673.
- Howe, M. S. (1980). "The dissipation of sound at an edge," *J. Sound Vib.* **70**, 407–411.
- Ingard, U., and Ising, H. (1967). "Acoustic nonlinearity of an orifice," *J. Acoust. Soc. Am.* **42**, 6–17.
- Ishizaka, K., and Matsudaira, M. (1972). *Fluid Mechanical Considerations of Vocal Cord Vibration* (Speech Communications Research Laboratory, Santa Barbara, CA), pp. 12–17.
- Kaiser, J. (1983). "Some observations on vocal tract operation from a fluid flow point of view," in *Vocal Fold Physiology: Biomechanics, Acoustics, and Phonatory Control*, edited by I. R. Titze and R. H. Scherer (Denver Center for Performing Arts, Denver), pp. 358–386.
- Lighthill, M. J. (1952). "On sound generated aerodynamically, I. General theory," *Proc. R. Soc. A* **211**, 564–587.
- Lighthill, M. J. (1978). *Waves in Fluids* (Cambridge U. P., Cambridge, England), pp. 1–41.
- Morse, P. M., and Feshbach, H. (1953). *Methods of Theoretical Physics* (McGraw-Hill, New York), pp. 52–53.
- Powell, A. (1964). "Theory of vortex sound," *J. Acoust. Soc. Am.* **36**, 177–195.
- Scherer, R., and Titze, I. (1983). "Pressure-flow relationships in a model of the laryngeal airway with diverging glottis," in *Vocal Fold Physiology: Contemporary Research and Clinical Issues*, edited by D. M. Bless and J. H. Abbs (College-Hill, San Diego, CA), pp. 179–193.
- Scherer, R. C., Titze, I. R., and Curtis, J. F. (1983). "Pressure-flow relationships in two models of the larynx having rectangular glottal shapes," *J. Acoust. Soc. Am.* **73**, 668–676.
- Sivian, L. J. (1935). "Acoustic impedance of small orifices," *J. Acoust. Soc. Am.* **7**, 94–101.
- Teager, H. (1987). Personal communication.
- Teager, H., and Teager, S. (1983a). "The effects of separated air flow on vocalization," in *Vocal Fold Physiology: Contemporary Research and Clinical Issues*, edited by D. M. Bless and J. H. Abbs (College-Hill, San Diego, CA), pp. 124–143.
- Teager, H., and Teager, S. (1983b). "Active fluid dynamic voice production models, or is there a unicorn in the garden?" in *Vocal Fold Physiology: Biomechanics, Acoustics, and Phonatory Control*, edited by I. R. Titze and R. C. Scherer (Denver Center for Performing Arts, Denver), pp. 387–401.
- van den Berg, Jw., Zantema, J. T., and Doornenbal, P., Jr. (1957). "On the air resistance and the Bernoulli effect of the human larynx," *J. Acoust. Soc. Am.* **29**, 626–631.