AN AEROACOUSTICS APPROACH TO PHONATION: SOME EXPERIMENTAL AND THEORETICAL OBSERVATIONS

Richard S. McGowan

Abstract. We examine the sources of sound during phonation using an aeroacoustic formulation. Some sources of sound during phonation are found to have a dipole character. The most important of these in the low frequency limit is the result of vorticity-velocity interaction. This picture is in contrast to the usual picture of the voice source as a monopole that can be modeled as a piston in a tube. A fluid mechanical approach to voice modeling is promoted in this note.

Introduction

The characterization of the voice source in terms of fluid dynamic variables is a part of the subject called aeroacoustics. Aeroacoustic theories are formulated from the conservation equations of fluid mechanics, so that such questions as the amount of total fluid energy that is converted into the fluid energy of acoustic motion during phonation can be answered. To date, no satisfactory partition of energy has been proposed, as noted by Teager and Teager (1983) and Kaiser (1983). They have suggested a fluid mechanical approach to phonation, and it is hoped that this note will contribute in that direction.

During phonation, there is modulated fluid movement in the region near the glottis. It is known that fluid motion can be decomposed into two kinds: solenoidal and irrotational, where the latter can support acoustic oscillation, and the former cannot (Batchelor, 1970). In the standard model of the voice source, the entire oscillatory field in the glottal region is treated as acoustic: the volume velocity at the glottis is the input to the one-dimensional analog circuit of the vocal tract. This picture of the voice source treats the glottal source as a piston in a tube, which can be classified as a monopole source. Here we will argue that this is not the correct model of the source of phonation.

Some of the general results from the aeroacoustics literature will be considered along with a discussion of their application to the voice source. The main results of this discussion can be summarized as follows. The solenoidal field, which contains the rotational motion of the fluid, and hence vorticity, is important for creating sound. The solenoidal field creates the

Acknowledgment. The author thanks Vin Gulisano for his photography, and Ed Wiley, Dick Sharkany, and Don Hailey for help with experimental hardware. Thanks goes to Professor K. R. Sreenivasan for his help with flow visualization.

[HASKINS LABORATORIES: Status Report on Speech Research SR-86/87 (1986)]
necessary potential energy for acoustic motion through dynamic pressure fluctuations near the folds. This type of source can be classified as a dipole type source. As a result, the fraction of total oscillatory fluid kinetic energy that is converted into acoustic energy is small, but of course, not insignificant in acoustic terms.

Along with the theoretical discussion given the voice source, we have a few measurements to support some of the ideas presented here. The measurements are taken from an oscillatory jet that mimics the air flow of the glottal region, with two exceptions: there are no moving surfaces near the jet, and the jet exits into a nearly unbounded region. However, this situation allows us to view some important aspects of sound production, like vortex formation. Further, it allows us to observe the sound produced by vortex formation, abstracted from that produced by moving boundaries. These measurements are incidental to the main point of urging an aeroacoustics approach to vocal tract acoustics in the future.

Sound Production

The production of sound by the interaction of fluid with itself and solid surfaces falls within the field of aeroacoustics, whose modern beginnings came with the work of M. J. Lighthill in 1952. Lighthill derived a nonlinear wave equation from the equations of motion for a Newtonian fluid. He wrote it so the familiar linear wave operator on the left when applied to density is set equal to nonlinear terms on the right. The right-hand side terms are to be identified with "sources" for the acoustic propagation of the left-hand side, density perturbation. This identification is known as Lighthill's acoustic analogy. Largely motivated by engineering problems, much work has been done on understanding this and other similar equations by solving them in various geometries and by rewriting the source terms. The analogy is difficult to test directly because the source terms usually have not been measured.

The work reported here relies on the simplification of the source terms from the work of Powell (1964). For low Mach number flows without entropy spottiness, Powell shows that the dominant source of sound involves the nonlinear interaction of vorticity and velocity. In fact, the wave equation appears as:

\[
\frac{\partial^2 \rho}{\partial t^2} - c_s^2 \nabla^2 \rho = \rho_o \text{div} (J \wedge v) + \rho_o \nabla^2 |v|^2
\]

where:

- \(v\) = fluid particle velocity
- \(J\) = vorticity = \(\nabla \wedge v\)
- \(\rho\) = perturbation fluid density
- \(\rho_o\) = ambient fluid density
- \(c_s\) = ambient speed of sound

Since vorticity is a quantity that appears in the source term, we attempt to determine whether vorticity is a part of the vocal tract flow in the next section. Later, a formal solution to this equation will be exhibited.
First, it is argued that vorticity occurs as part of the flow from the glottis. The geometry of the vocal tract in the region of the glottis can be idealized as that of one cylindrical pipe of relatively small diameter emptying into a cylindrical pipe of relatively large diameter. The ratio of the areas is taken as approximately ten. Air flowing from the small pipe into the larger pipe, or even into an unbounded region, forms a jet. A jet is a region of shear flow, which is necessary to meet the boundary conditions at the walls of the larger pipe, or at infinity in the case of an unbounded region.

We follow Batchelor (1970) in arguing that vorticity is formed in the special case of a periodically modulated jet, which occurs above the glottis. While Batchelor considers the steady case, we will make the quasi-steady assumption.

Initially, the jet is in a region where the flow from the smaller pipe is mixing with the fluid in the larger pipe. The initial mixing region extends a distance $\Delta x$ in the direction of the pipe axis, after which the jet fills the entire pipe. We use $V$ to denote particle velocity in the axial direction, $p$ to denote pressure, and $A$ to denote cross-sectional area. The subscript 1 denotes the smaller pipe, and the subscript 2 the larger pipe. If $f$ is the frequency of motion under consideration, the quasi-steady assumption can be made if:

$$f \ll \frac{\text{unsteady part of } V}{\Delta x}$$

This assumption can be seen to be approximately valid above the glottis for $f < 1000 \text{ Hz}$, the unsteady part of $V_1 = 4,000 \text{ cm/sec}$ and $\Delta x < 2 \text{ cm}$. (From van den Berg's experiments [van den Berg, Zantema, & Doornenbal, 1957] the assumption that $\Delta x < 2 \text{ cm}$ appears well founded, because all the loss of pressure head appears to occur before their final transducer. From their figure 1, the distance from the jet exit to the final transducer is apparently under 2 cm.) We do not argue the validity of the common quasi-steady assumption in real voice, but use it knowing the limitations.

![Figure 1. Jet exiting into cylinder.](image-url)
The equations for mass and momentum conservation can be written in integral form, using control planes at the exit and after the mixing region (see Figure 1). Irrotational motion before and after the mixing region is assumed. Mass conservation gives:

\[ V_1 A_1 = V_2 A_2 \]

Momentum conservation gives:

\[ p_1 A_2 + \rho_0 V_1^2 A_1 = p_2 A_2 + \rho_0 V_2^2 A_2 \]

Solving for pressure:

\[ p_2 = p_1 + \rho_0 V_1^2 \left( \frac{A_1}{A_2} \right) (1 - \frac{A_1}{A_2}) \]

If we were to assume irrotational flow in the mixing region, Bernoulli's relation under the quasi-steady assumption gives:

\[ p_2^* = p_1 + \rho_0 V_1^2 / 2 \left( 1 - \left( \frac{A_1}{A_2} \right)^2 \right) \]

The difference is:

\[ p_2^* - p_2 = \rho_0 V_1^2 / 2 \left( 1 - \frac{A_1}{A_2} \right)^2 \]

The difference in pressure implies that energy must be going into rotational motion, and then into heat, and, possibly sound.

The rotational motion discussed in the above paragraph may be in the form of a vortex ring. As the jet enters the larger tube, fluid is pulled along side of the jet by viscous action. Farther from the jet, fluid must return because of mass conservation. Taking a cross-section in plane containing the cylinders' axis, the following picture of the velocity field may be made (see Figure 1). It can be seen that because the jet has cylindrical symmetry, the rotational flow field is toroidal, that is, a vortex ring. In a real vocal tract, the vortex will no longer be toroidal, but it may be topologically equivalent to a toroid.

In the experimental situation to be described, we verified the existence of vorticity in an indirect way. Instead of the larger tube, a nearly unbounded region (a large box) was used, where the boundaries can be considered to be at infinity. If vorticity is found in our experiment, then it will be found in the case of two tubes. The effect of the cylindrical wall of the larger tube is to increase shear, and hence vorticity. So this experiment is less favorable to the formation of vorticity than the case of two tubes.

Our apparatus was primarily a flow visualization apparatus using smoke. The modulated air flow was produced using an air compressor from an electric tire pump. The volume displacement of the compressor piston was 2.57 cm³, and it was equipped with a valve to block flow into the piston cylinder during the downstroke. This resulted in a series of pulses of air with the mean level of flow on the same order of magnitude as the peak flow. The speed of the compressor was varied using the voltage control on a D.C. power supply. The output of the compressor was fed by rubber hose to a cylindrical brass nozzle with inner diameter .9 cm. The brass nozzle exhausted air into a cardboard box lined with flat-black paper, with a slot for photography cut in the side.
Smoke was supplied using a cigarette in a plastic tube, with one end connected to a ducted fan and the other connected through an intervenous needle to the hose between the compressor and the nozzle. Lighting was provided at the floor of the box using a General Radio 1531-A Strobatac, which was wired through a General Radio 1531-P2 delay. Using an electric eye, the strobe was triggered from the drive shaft of the electric motor for synchrony.

A 35mm Konica single lens reflex camera with a close-up lens and tri-X, 400 ASA film was used to photograph this nozzle region. The image plane was approximately 9 inches from the jet. The compressor was operated at a frequency of about 20 Hz, in order to use the high intensity setting on the Strobatac. The best exposures occurred for shutter speeds of between 1/8 and 1/4 sec and an f-stop of 2.0. (We did not try lower f-stops for high intensity exposure.) The resulting photographs, one of which is shown below, showed signs of an oscillating vortex ring. There are bands of smoke perpendicular to the jet, which indicates rotational motion in a vortex ring. More detail could be seen if we had stroboscopically luminated the jet in a cross-section parallel to the jet axis, and used a higher density smoke.

Figure 2. Vorticity in an oscillating jet.

Although we did not obtain a quantitative measure of vorticity, we do have good reason to suggest that vorticity of appreciable strength may be generated near the glottis. Also, this vorticity may be modeled as a vortex ring.
Having discussed some aspects of the acoustic source, we now exhibit a solution for the acoustic quantities in the far-field (i.e., at distances large compared to the wavelength). The integral solution to the wave equation shown above was proposed by Powell (1964). We will discuss Powell's solution applied to the vocal tract and to our experimental situation, described below. Powell assumes that the Mach number of the flow is small and the product of the Mach number with the source compactness parameter is also small. The source compactness parameter is the product of the source length scale and the typical wavenumber of the acoustic wave. Because we are considering only low frequencies, the mixing region is presumed short (i.e., $\Delta x < 2$ cm), and the Mach numbers small, Powell's assumptions appear to be valid for the vocal tract. If we draw a control volume that contains the great majority of the vorticity, the observer position $x$, and with surfaces along solid boundaries or within the fluid where acoustic relations are valid, then the solution in the far-field can be written (Powell, 1964):

$$
\rho(x) = \frac{-\rho_0}{4\pi xc_0} \int \frac{3}{|x|} \left( J \cdot \vec{v} \right) dV(y) - \frac{\rho_0}{4\pi x c_0} \int \left( p + \frac{1}{2} \rho_0 |\vec{v}|^2 \right) n \cdot \frac{\vec{x}}{|x|} dS(y) - \frac{\rho_0}{4\pi x c_0} \int \frac{3}{|x|} \left( J \cdot \vec{v} \right) n \cdot dS(y)
$$

* denotes acoustic time delay
$x$ = far-field coordinate
$y$ = source coordinate
$n$ = normal to surface pointing away from control volume

where $S_o$ denotes the part of the surface of the control volume $V_o$, which coincides with a solid surface. (The surface integrals appear because we used the free-space fundamental solution. Future research should include finding the Green's function suitable for vocal tract geometry. Loosely, Green's functions are to boundary value problems what impulse responses are to initial value problems.)

For the vocal tract, we take the control volume bounded by the lungs, trachea, glottis, pharynx, mouth, and a large sphere outside the mouth. The first and second integrals apparently provide dipole sources, and the third, apparently, a monopole source.

Because there are solid boundaries present, there can be a net energy exchange between the fluid and the solid, and because $\rho_0 J \cdot \vec{v}$ is proportional to the time rate of change in the momentum in the fluid, the first integral can be approximated without considering acoustic time delays. Therefore, the first integral truly provides a dipole source. This integral will be nonzero in the region just above the glottis, where we assert the existence of a strong oscillating vortex ring with an axis of symmetry coincident with that of the dipole. This term is associated with the loss of pressure head
discussed in the section on vortex formation, and we call it the vorticity-velocity interaction term. It was seen that the loss of head was an order one multiple of \( \frac{P_0}{2} (V_g)^2 \), where \( V_g \) is the glottal fluid particle velocity. This can be taken as the order of magnitude of the acoustic pressure provided by such a source.

The second integral also provides for dipole sources of sound with axes normal to the interior surfaces of the vocal tract. Because of the direction of the axes, this term should contribute little to the propagation of sound, except perhaps in the region of the vocal folds. The quantity \( p + \left( \frac{P_0}{2} \right) |\mathbf{V}|^2 \) is equal to \( -\rho_0 \int \mathbf{V} \cdot \mathbf{d}y \) on the surface of the folds, so that an order-of-magnitude comparison between the first and second integrals can be carried out. The ratio of the second to the first is on the order of magnitude: \( \frac{(r \cdot f)}{V_g} \) where \( r \) is the radius of the vocal tract and \( f \) is the frequency of sound under consideration. In a low-frequency approximation, consistent with the quasi-steady approximation used earlier, the first integral dominates the second.

The final integral involves the movement of the folds themselves. This integral appears to provide for a monopole source of sound. However, the integral is identically zero when acoustic time delays are neglected, because the folds do not change volume as they oscillate. Further, this integral is negligible in relation to the first, especially at low frequencies, because the peak velocity of the folds is so much less than that of the particle fluid velocity at the glottis.

This is only one possible form for an integral solution in the aeroacoustics literature. There are others that show the boundary forcing more explicitly, but without explicit reference to vorticity-velocity interaction (Goldstein, 1976). Also, a formulation by Howe (1975) combines the effects of the first and second integral and exhibits the vorticity-velocity interaction explicitly.

We have argued that the three integrals above provide dipole sources in the region of the glottis. The first integral, which is the result of energy transfer between the solid surface and the fluid, is arguably the largest of the three in the low frequency limit. This is not the whole story, because there is time varying motion of the fluid above the vorticity producing mixing region, which is required by mass conservation. This may provide an acoustic signal beyond what we have discussed so far, perhaps obeying nonlinear propagation laws. We are not prepared to compare the amplitude of this wave with that produced by the terms already discussed. The aeroacoustic formulation is not complete as we have discussed it here, but we have identified sources that have not been considered previously.

It should be noted that since the acoustic pressure fluctuations provided by the first integral are on the order of \( \frac{P_0}{2} (V_g)^2 \), this source is inefficient. The ratio of the acoustic intensity radiated by this term to the flux of fluid kinetic energy density in the glottal region is on the order of the square of the peak Mach number of the oscillatory part of the glottal flow.

We performed an experiment with the modulated jet to determine the importance of the vorticity-velocity interaction as a sound source. Using the same compressor described in the section on vorticity, we attached the hose
from the compressor to a lamp post wrapped in packing foam to minimize its effect on the field. The nozzle thus was oriented horizontally, about 2 ft. above the floor. We used a B+K Sound Level Meter, with a wind shield, at 1 ft. 10 in. from the nozzle and in the horizontal plane of the nozzle. We ran the compressor at 80 Hz, and used no band-pass filter for measuring the intensity. Measurements were taken at 45° intervals from -90° to 90° to the centerline of the nozzle.

We model this situation with a control volume consisting of the interior of the tube down to the piston connected with a large sphere outside the tube. The sphere has a cylindrical section removed, which contains the tube (see Figure 3). Thus, part of the bounding surface of the control volume contains the piston of the compressor. Because there is an oscillatory change in volume we have a monopole source, which is represented by the third integral in the integral solution. However, because there is oscillatory vortex shedding from the tube exit, there is some dipole component of sound seen in the far field.

Figure 3. Control volume
The experimental results show a directivity pattern that is not omnidirectional and shows a large dipole component. These results are summarized in the table below.

<table>
<thead>
<tr>
<th>angle</th>
<th>SPL</th>
</tr>
</thead>
<tbody>
<tr>
<td>90°</td>
<td>58 dB</td>
</tr>
<tr>
<td>45°</td>
<td>65 dB</td>
</tr>
<tr>
<td>0°</td>
<td>67 dB</td>
</tr>
<tr>
<td>-45°</td>
<td>64 dB</td>
</tr>
<tr>
<td>-90°</td>
<td>55 dB</td>
</tr>
</tbody>
</table>

Indeed, if we take 58 dB to be the intensity of the monopole source at the distance the measurements took place, then the main lobe of the dipole field adds 9 dB. This indicates that a great deal of fluid energy is in the form of vorticity, or the result of boundary forcing, which goes into making the dipole source. Only a small fraction of acoustic energy can be associated with the monopole source. (Covective amplification, where vorticity is convected in a mean flow, is known to alter the directivity of the sound field. However, since the mean flow Mach number is small, this effect cannot account for the present deviation from an omni-directional pattern.)

Conclusion

In this note we have found good reason for supposing the existence of vorticity above the glottis. This nonacoustic motion can be shown to produce sound via the mechanism of a fluctuating pressure head near the folds. This fluctuating pressure head is the result of an exchange of energy between the solid and fluid, which is realized in the fluid as vorticity-velocity interaction. This source is in addition to any oscillating fluid motion that may be considered to be acoustic above the region of strong vorticity-velocity interaction. We have not accounted for this latter wave in this presentation, so that our application of the aeroacoustic formulation may yet be incomplete.

The picture presented here stands in contrast to the standard picture of a piston source. In the picture presented here, a large amount of fluid energy goes into rotational motion near the folds, with only a small fraction of this energy being converted to sound. This process cannot be accounted for by a piston in a tube.

Future research should include finding the Green's function appropriate to the vocal tract, and experimental measurements and theoretical predictions of the oscillatory fluid field just above the glottis and the forcing of the fluid by the vocal folds. These ingredients, difficult to obtain, will completely characterize the acoustic field produced during phonation.
References


