ACOUSTIC TUBE ANALYSIS OF FORMANT BANDWIDTHS AND FREQUENCIES IN HELIUM SPEECH

M. A. Richards
Engineering Experiment Station

R. W. Schafer
School of Electrical Engineering

Georgia Institute of Technology
Atlanta, Georgia 30332

ABSTRACT

We use the lossy electric transmission-line analog model of the vocal tract to study the acoustics of speech produced in a hyperbaric helium-oxygen atmosphere. The analysis extends previous work by including more completely the effects of the wall vibration, glottal, and radiation impedances, and by analyzing the formant bandwidths and amplitudes in addition to the formant frequencies. It shows that (1) the classic Fant and Lindquist formula somewhat overstates the formant frequency shift when glottal and radiation effects are included; (2) the lower formant bandwidths increase by much more than commonly assumed; and (3) the upper formant amplitudes are higher relative to the lower formants in helium speech than in normal speech. These results are useful in developing advanced helium speech enhancement algorithms.

INTRODUCTION

Deep sea divers working at depths of about 60 meters or more must, for physiological reasons related to the high ambient pressure, breathe a helium-oxygen ("heliox") atmosphere in place of air (1). Unfortunately, the resulting "helium speech" is nearly unintelligible. Various devices known as helium speech unscramblers have been designed to enhance the intelligibility of helium speech.

Current unscrambler designs are based on a simplified model of helium speech acoustics. Although they aid communication significantly, improved performance might be obtained with better compensation of formant frequency shifts and bandwidth effects. New digital signal processing techniques offer the possibility of more refined analysis and enhancement algorithms (2). However, a better model of helium speech acoustics is needed to guide their design.

The human vocal tract can be modelled to a first approximation as a single lossy acoustic tube, which can be analyzed with an electrical transmission line analog of the form shown in Figure 1. Flanagan (3) develops this model in detail, relating all of the impedances to physical properties of the atmosphere and vocal tract. Fant and Lindquist (4) first used a similar approach to study helium speech acoustics. They derived the now widely accepted relationship between the frequency F_a of a formant in air and the frequency F_h of that same formant in hyperbaric heliox:

$$F_h^2 = \alpha^2 F_a^2 + F_o^2 \qquad (1)$$

where α is the ratio of the speed of sound in the heliox environment to its speed in air, and F_o is an atmosphere and geometry-dependent constant. Typically, α ranges from 2.5 to 2.8 and F_o is about 200 to 400 Hz. Clearly, the shift in formant frequencies and the increase in speech bandwidth are dramatic. The form of the model of Equation (1) has received substantial empirical support, although the values of α and especially F_o which give the best fit to real data are often not exactly as predicted.

Fant and Lindquist's analysis did not account for either the glottal or radiation impedances, nor did it consider the formant bandwidths and amplitudes. It also neglected the effect of cavity wall vibration on the characteristic impedance of the vocal tract. In this paper, we give the results of a theoretical study of the behavior of the frequencies, bandwidths, and amplitudes of speech in a hyperbaric heliox atmosphere. Our method is to use Flanagan's model of the vocal system as a single, uniform cross-section, lossy acoustic tube; exercise it with atmospheric parameters (density, sound velocity, etc.) appropriate to hyperbaric heliox; and compare the formant characteristics to those obtained from the same model with

CH1945-5/84/0000-0398 $1.00

parameters appropriate to a normal atmosphere. Our analysis overcomes the shortcomings of Fant and Lindquist's, and also extends Flanagan's work by treating all of the model impedances simultaneously.

THE VOCAL TRACT MODEL

Figure 1 shows the transmission line analog of the vocal tract model. Z_g and Z_r are the glottal and radiation impedances. Z_1 and Z_2 are functions of the characteristic impedance Z_0 and propagation constant ζ_{tw} of the vocal tract, which in turn depend on the per-unit-length series impedance Z_s and shunt admittance Y_p. Z_s accounts for viscous friction and the inertance of the air mass. Y_p is the sum of two admittances, one accounting for heat conduction and air compressibility, and the other for vocal tract wall vibration. As we shall see, this last term has a major influence on helium speech acoustics.

The complete definition of the impedances of Figure 1 in terms of physical properties of the atmosphere and vocal tract requires too much space to give here. The reader is referred to reference (3) or (5) for details. We note that most of the components are functions of frequency. Furthermore, they depend on the atmospheric properties of density (ρ), sound velocity (c), viscosity (μ), thermal conductivity (κ), specific heat at constant pressure (c_p), and adiabatic constant (α). The last enters as (α-1).

HELIUM-OXYGEN ATMOSPHERES

In analyzing the acoustics of helium speech, it is necessary to know the above six gas parameters for hyperbaric heliox. These can be readily predicted if the volume fractions of helium and oxygen and the ambient pressure are known (5). To restrict the possibilities somewhat, consider an extended saturation dive at a depth of about 75 meters or more. The ambient pressure at a depth of d meters is

$$P = 1 + 0.0968d \text{ ATA} \tag{2}$$

An oxygen partial pressure of PO_2 = 0.4 ATA is suitable for this case (1) and is assumed throughout. The helium fraction is then

$$p_h = \frac{P - 0.4}{P} \tag{3}$$

It follows from these comments that useful heliox mixtures range from about 95% to 99.5% helium.

Figure 2 shows the ratio of the six relevant gas parameters in heliox to their values in air, as a function of mixture with PO_2= 0.4 ATA. Viscosity and the adiabatic constant are not greatly changed. Thermal conductivity and specific heat, on the other hand, are increased by five times or more. However, it is the density and sound velocity which are most important. Sound velocity is 2.5 to nearly 3 times greater in hyperbaric heliox than in air. Density, which is proportional to pressure, increases dramatically as depth increases. Also, since all other parameters are changing slowly with changes in the mixture, the difference between one heliox mixture and another is essentially one of density. Finally, note that the products ρc and ρc^2, which appear several times in the detailed expressions for the vocal tract impedances (3,5), can increase by factors of about 20 to over 60.

ANALYSIS OF THE VOCAL TRACT MODEL

The model of Figure 1 can be analyzed to calculate the approximate pole locations for the vocal "network," from which formant frequencies, bandwidths, and amplitudes readily follow. It is shown in reference (5) that the poles are located at

$$s_n = -\sigma_n + j\Omega_n \tag{4}$$
$$= \frac{1}{k(\Omega_n)} \left(-(\sigma_{tw}+\sigma_g+\sigma_r) \pm j\Omega_n^*\right)$$

where Ω_n^* is the frequency of the n-th formant in the lossless acoustic tube and the various σ subscripts refer to "tract" (friction and heat) and wall vibration, glottal, and radiation losses, respectively. The factor $k(\Omega)$ involves wall vibration, glottal, and radiation effects. It represents the perturbation of the actual pole frequency from its lossless counterpart.

The formant frequency is that value of Ω_n such that $\Omega_n k(\Omega_n) = \Omega_n^*$. This value is most easily found iteratively, especially at low frequencies and great depths where the perturbation is large. Once Ω_n is known, the various σs are evaluated at Ω_n to obtain σ_n. The formant bandwidths are then approximately $2\sigma_n$ rad/s, and the amplitudes are about $\|Z_r\|^2(c/\sigma\ell_t)$, where ℓ_t is the tube length (5).

ANALYSIS RESULTS

The results obtained by exercising the model in normal and heliox atmospheres as described above are compared in Figures 3-5. Each figure plots a characteristic of a heliox formant as a function of the frequency of the corresponding formant in air. In each, the result is shown for helium fractions p_h of 0.95, 0.987, and 0.992, corresponding to depths of 75, 300, and 500 meters if PO_2 = 0.4 ATA.

Figure 3 relates the formant frequency in air and heliox. The solid lines are the prediction of our model, while the dashed lines represent Fant and Lindquist's results. Close inspection shows that our curve is of the same quadratic form as Equation (1). However, the predicted shift at low frequencies is significantly less than the dashed curve at extreme depths, indicating a lower effective F_0. The difference between the solid and dashed curves is due to inclusion of the glottal and radiation impedances Z_g and Z_r and the effect of wall vibration on the characteristic impedance Z_0.

Figure 4 gives the ratio of a heliox formant bandwidth to that of the corresponding formant in air. The limiting value at high frequency is just α, but obviously the bandwidth in the first formant region increases by much more than α in helium speech. This behavior has recently been confirmed by measurements on real helium speech (6). Failure to correct this phenomenon is one of the major deficiencies of current unscramblers.

Figure 5 shows that the amplitude of the upper formants relative to the lower formants is greater in heliox than in air by 4 to 10 dB. This result contradicts the common belief that successive formant amplitudes fall off more rapidly. The explanation of this discrepancy is uncertain. It may be due to masking of the high frequency emphasis by poor response of the diver microphones in hyperbaric conditions at high frequencies.

It is apparent that the formant bandwidths and amplitudes of the acoustic tube in heliox correspond to a simple frequency "stretch" $F_h = \alpha F_a$ at the higher frequencies, but that the low frequency behavior deviates drastically from this simple behavior at great depths. The deviation is due mostly to a large increase in vocal tract wall vibration. This vibration is much more pronounced in heliox than in air because of the density increase. Increased glottal losses are a secondary contributor to the deviation.

CONCLUSION

Our analysis shows that Fant and Lindquist's formula for formant shifts in heliox somewhat overstates the shift at low frequencies. It also shows that first formant bandwidths can be expected to increase dramatically, while the spectral balance should tip towards more high-frequency emphasis. Although the uniform tube model is extremely simplified, there is ample reason to believe that these results apply well to real helium speech. Specific numerical results should, however, be used with caution. They depend not only on the precise atmospheric conditions, but also on the particular estimates of acoustic properties of body tissues and structures given by Flanagan (3).

References

1. U. S. Navy Diving Manual, volumes 1 (1978) and 2 (1981). Dept. of the Navy, Washington, D. C. (available through U.S. Govt. Printing Office.)

2. M. A. Richards, "Helium Speech Enhancement Using the Short-Time Fourier Transform," IEEE Trans. Acoustics, Speech, and Signal Processing, vol. ASSP-30(6), pp. 841-853, December 1982.

3. J. L. Flanagan, Speech Analysis, Synthesis, and Perception, second edition. Springer-Verlag, New York, 1972.

4. G. Fant and J. Lindquist, "Pressure and Gas Mixtures on Divers' Speech," Quarterly Progress and Status Report STL-QPSR-1/1968, Speech Transmission Laboratory, Royal Institute of Technology, Stockholm, Sweden, pp. 7-17, 1968.

5. M. A. Richards and R. W. Schafer, "Acoustic Tube Analysis of Helium Speech," submitted to Journal of the Acoustical Society of America.

6. E. O. Belcher and S. Hatlestad, "Formant Frequencies, Bandwidths, and Qs in Helium Speech," Journal of the Acoustical Society of America, vol. 74(2), pp. 428-432, August 1983.

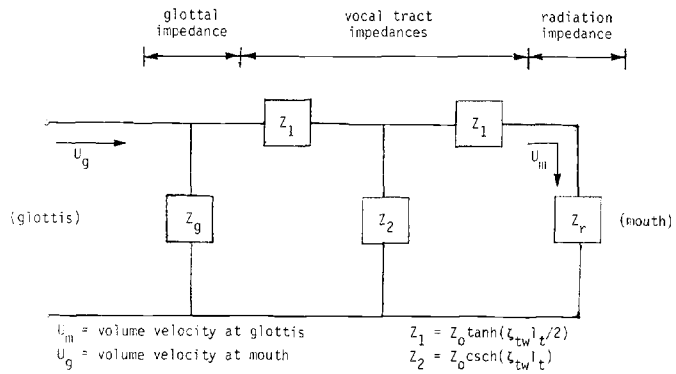

Figure 1. Transmission line analog of the vocal tract (after (3))

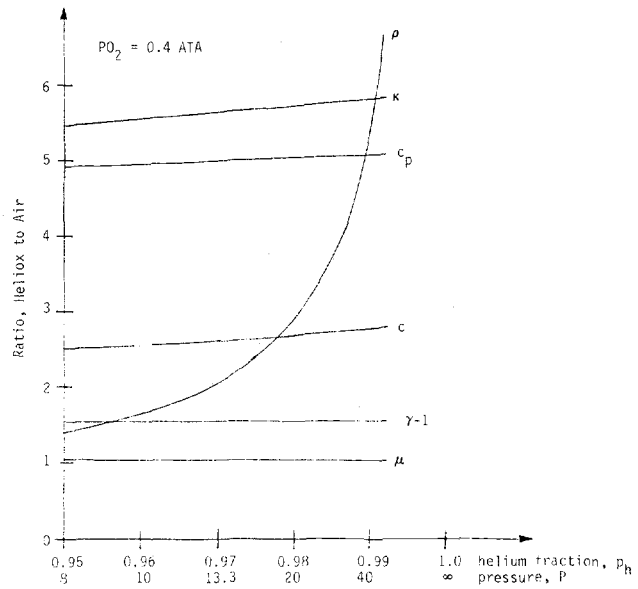

Figure 2. Ratios of gas properties of heliox and air

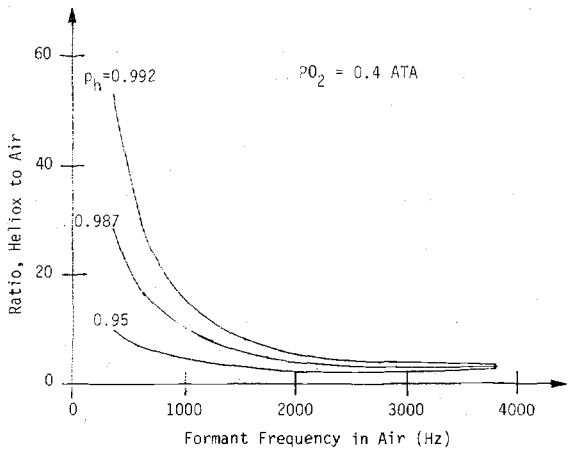

Figure 4. Ratio of formant bandwidths in hyperbaric heliox and air

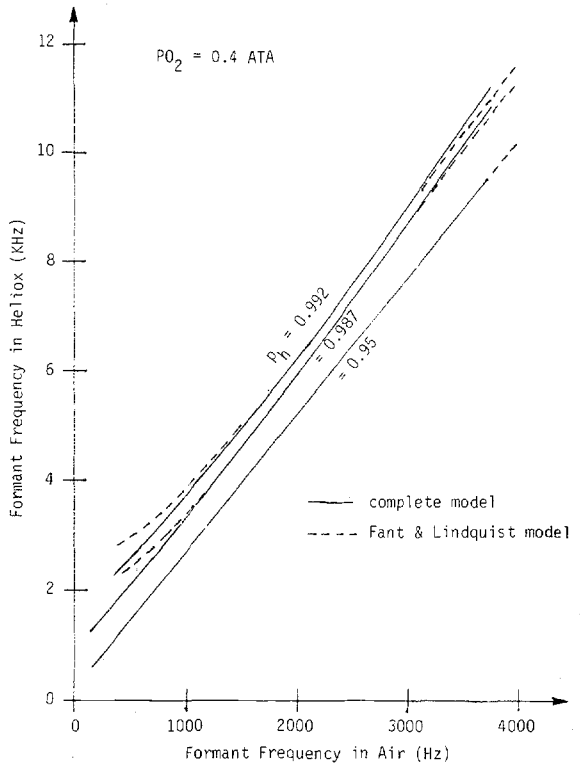

Figure 3. Effect of hyperbaric heliox on formant frequencies

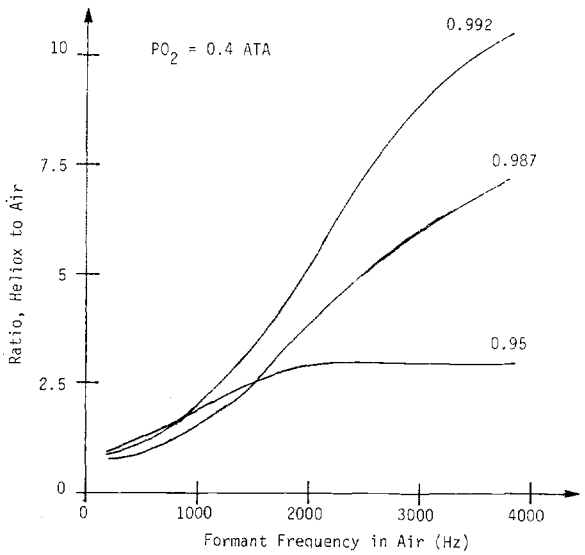

Figure 5. Ratio of formant amplitudes in hyperbaric heliox and air