Proceedings of The Institute of Acoustics

SPEECH FORMANT SHIFTS IN HYPERBARIC HELIUM

S.W. BEET
ROYAL SIGNALS AND RADAR ESTABLISHMENT, MALVERN.
C.C. GOODYEAR
DEPT. OF ELEC. ENG., UNIVERSITY OF LIVERPOOL.

1. Introduction

To avoid the adverse physiological effects of breathing air at high pressures, deep-sea divers breathe a mixture of helium and oxygen. One result of this is a marked shift of the frequencies of the vocal tract resonances so that the speech becomes unintelligible. The main factor which determines the new formant frequencies is the higher phase velocity of sound in the gas and, taken by itself, this would result in a scaling of the formant frequencies by a constant factor, independent of pressure for a given mixture. It has been observed, however, that this factor is a function of frequency and pressure [1-3]. In this paper we describe a new method for estimating formant frequency shifts and a signal processing technique for restoring the speech to a more natural form.

2. Theoretical estimate of the formant shifts

A vocal-tract resonance will occur whenever the wavelength of a component of the excitation signal corresponds to some physical dimension of the tract. Consequently, the centre frequency, f_H, of a formant under hyperbaric conditions can be related to that in atmospheric air by the equation

$$\frac{f_H}{f_A} = \frac{c_H(f_H)}{c_A(f_A)} \tag{1}$$

where the subscripts H and A denote properties under hyperbaric conditions and atmospheric air respectively, f is the formant frequency and c(f) is the phase velocity of sound in the vocal tract.

Lin and Morgan [4] have analysed acoustic wave propagation along a fluid-filled thin-walled elastic tube. Their model is restricted to sinusoidal wave propagation and to pressure distributions with axial symmetry along the tube. These are not unreasonable assumptions for the lower formants ofthe vocal tract. In their model the principal equation satisfied by the possible phase velocities, c, is as follows :

$$1 - \frac{1}{\omega^{*2}} + \frac{\nu_t^2}{\omega^{*2}(1-c^{*2})} - \frac{\xi\,\omega^{*2}\,(1-c^{*2})}{c^{*2}[c^{*2} + \xi\eta\,\omega^{*2}(1-c^{*2})]} = \frac{\alpha^2\, J_0(ka)}{ka\, J_1(ka)}$$

where

$$\omega^* = \frac{a\,\omega}{c_t} \qquad c^* = \frac{c}{c_t}$$

$$\xi = \frac{h^2}{12a^2} \qquad \eta = \frac{2}{1-\nu_t}$$

and

$$\alpha = \sqrt{\frac{\rho_f \; a}{\rho_t \; h}} \qquad k = \omega \sqrt{\frac{1}{c_f^2} - \frac{1}{c^2}}$$

In which the subscript t refers to a property of the tube and f to a property of the fluid; a is the radius to the centre of the wall, h is the wall thickness, ρ is a density, ν is Poisson's ratio and ω the angular frequency of the propagating wave.

All the parameters required to solve this equation are available from the direct vocal tract measurements of Ishizaka, French and Flanagan [5]. Using these and an assumed tissue density of 1000 kg m^{-3}, we have solved equation (2) for the phase velocity of the dominant propagation mode as a function of frequency for various gas mixtures and pressures. The results are shown in figure 1 and the corresponding relationship between formant frequencies in atmospheric air and in hyperbaric heliox are shown in figure 2.

The calculated shifts agree surprisingly well with the measured values of Belcher and Hatlestad [3], especially in view of our simplified model of the vocal tract as a cylindrical cavity with uniform wall properties along its length and the assumption that the wall is thin. An earlier model, due to Fant [1], regards the vocal tract wall as an inertial acoustic load but ignores its elasticity. Fant's model also requires a semi-empirical constant, the closed-lip resonance, to be chosen to fit the measured data.

2. Signal processing to restore formants

Many techniques have been developed for improving the intelligibility of helium speech. Some of these have recently been reviewed by Jack and Duncan [6]. Processors of modest complexity operate using a waveform stretching technique which provides a linear frequency scaling of the spectral envelope. From the discussion above, however, it is clear that a non-linear frequency compression is required at larger depths. This can be achieved using frequency domain processors but these generally incur a large computational cost. The authors have recently described [4] a frequency domain technique using linear prediction which is relatively simple and which should be capable of implementation using two or three VLSI signal processing devices.

The overall structure of the processor is shown in figure 3. The adaptive all-zero filter consists of a cascade of second-order sections with pre-quantised coefficients. The form of second-order section used allows the radius and angle of the z-domain zeroes to be updated independently during adaption. At the end of each data block the values acquired are passed to the slave filter and held fixed while this filter processes the same block of data to produce a spectrally whitened residual signal. The filter coefficients are also modified using a small set of look-up tables to give new values corresponding to a normal atmospheric environment. The new values are used in the inverse filter whose input is the residual. This produces a signal with properly scaled formant frequencies and bandwidths, close to those which would have been produced in normal atmospheric air and with the pitch unchanged. The overall spectral trend, however, becomes distorted, especially at frequencies close to half the sampling rate and a fixed filter was included to compensate for this. This filter was designed to correct the trend for a neutral vowel, but gives satisfactory shaping for all speech sounds.

SPEECH FORMANT SHIFTS IN HYPERBARIC HELIUM

This technique has given good results in computer simulations using recorded examples of diver's helium speech but has not yet been tested in real time. The maximum depth for the speech examples available to us was around 200 m and it was not clearly apparent that non-linear scaling improved intelligibility in this range. The method does, however, provide a straightforward means of controlling formant bandwidth. In hyperbaric helium, the lower formant in particular become broadened due to higher radiation from the vocal tract. Modifying the look-up tables to restore appropriate bandwidths was found to improve the speech quality.

References

1. G. FANT, J. LINDQUIST, B. SONESSON and H. HOLLIEN 1971 Underwater Physiology, Proc. 4th Symp. on Underwater Physiology, Pennsylvania, 293-299, "Speech distortion at high pressures".

2. C.T. MORROW 1971 Journ. Ac. Soc. Amer., 50(3), 715-728, "Speech in deep submergence atmospheres".

3. E.O. BELCHER and S. HATLESTAD 1983 Journ. Ac. Soc. Amer., 74(2), 428-432, "Formant frequencies, bandwidths and Q's in helium speech".

4. T.C. LIN and G.W. MORGAN 1956 Journ. Ac. Soc. Amer., 28(6), 1165-1176, "Wave propagation through fluid contained in a cylindrical elastic shell".

5. K. ISHIZAKA, J.C. FRENCHH and J.L. FLANAGAN 1975 IEEE Trans. ASSP-23, 370-373, "Direct determination of vocal tract wall impedance".

6. M.A. JACK and G. DUNCAN 1982 Radio and Electronic Engineer, 52(5), 211-223, "The helium speech effect and electronic techniques for enhancing intelligibility in a helium-oxygen environment".

7. S.W. BEET and C.C. GOODYEAR 1983 Electronic Letters, 19(11), 408-410, "Helium speech processor using linear prediction".

Fig.1

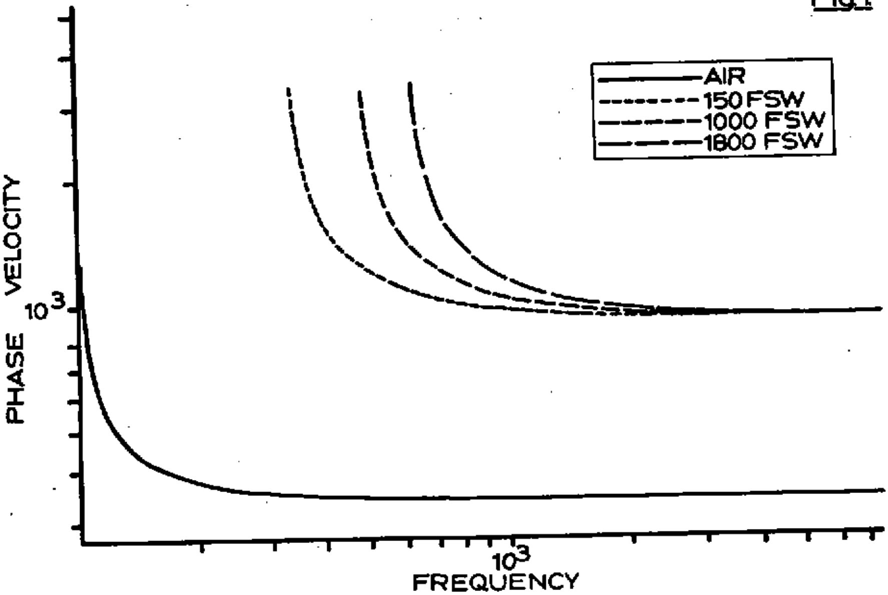
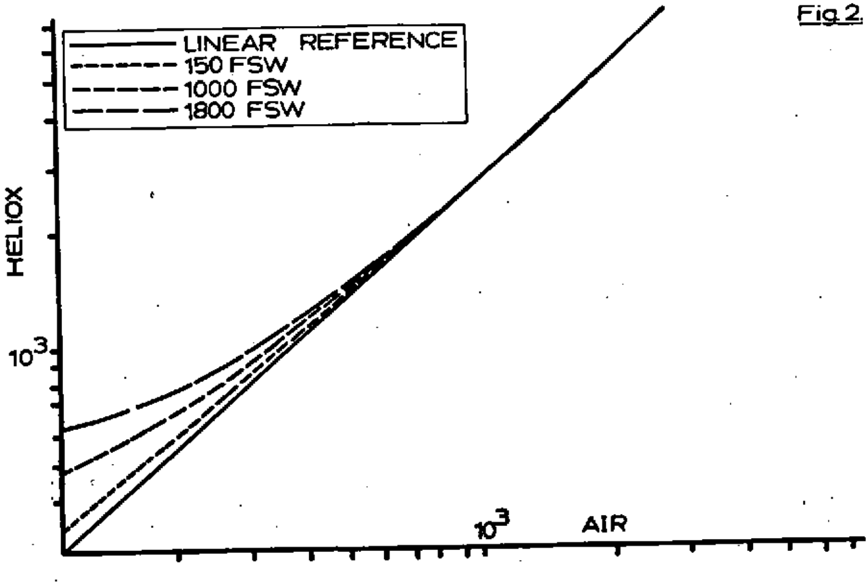

AIR
150 FSW
1000 FSW
1800 FSW
PHASE VELOCITY
10^3
10^3
FREQUENCY

Fig 2.


LINEAR REFERENCE
150 FSW
1000 FSW
1800 FSW
HELIOX
10^3
10^3
AIR

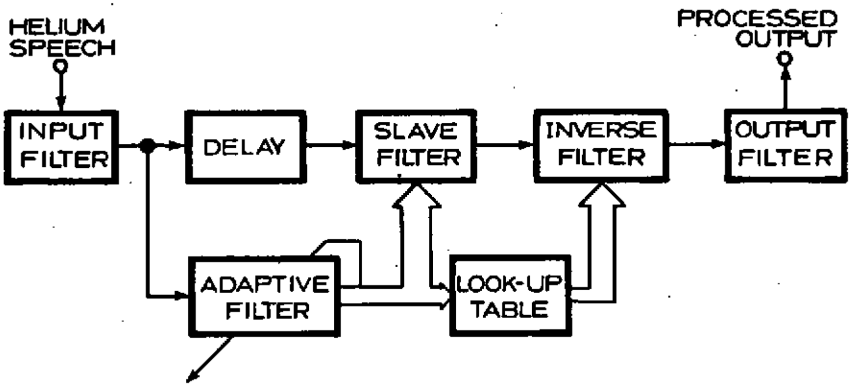

HELIUM SPEECH
PROCESSED OUTPUT
INPUT FILTER
DELAY
SLAVE FILTER
INVERSE FILTER
OUTPUT FILTER
ADAPTIVE FILTER
LOOK-UP TABLE

FIG.3