COMPACT HELIUM SPEECH UNSCRAMBLER USING CHARGE TRANSFER DEVICES

Dr M.A. Jack,* Dr A.D. Milne,† Dr L.E. Virr.Δ

ABSTRACT

This paper describes the principle, design and performance of a compact helium speech unscrambler using charge transfer devices. Helium speech is the term commonly used for the distorted speech uttered by deep-sea divers breathing in an atmosphere which consists of a mixture of helium and oxygen gases. The unscrambler considered here operates in pitch synchronism with the input helium speech waveform and uses time-expansion signal processing to achieve a considerable improvement in the intelligibility of the speech signal.

INTRODUCTION

Deep sea divers breath a mixture of helium and oxygen gases in order to overcome the toxic and narcotic problems produced when breathing high pressure air. However, speech produced in such a high pressure helium/oxygen mixture suffers a unique and severe form of distortion known as 'helium speech'. The main features of the helium speech distortion can be summarised as:

* A linear shift in vowel formant frequencies {1,2}
* Minimal change in fundamental frequency (pitch) {3}
* A non-linear shift in low frequency formants {2}
* Minimal change in formant bandwidths {4}
* Reduction of unvoiced sound intensities {4}

Present helium speech "unscrambler" designs take into account only the first two of these properties {4,5}. In these systems, the distortions present in the helium speech waveform are reduced by signal processing techniques based on spectrum compression - by means of waveform time expansion - in pitch synchronism with each pitch period {5}. Such signal processing requires electronic storage and current systems generally employ digital memory for this function.

The unscrambler design considered here uses similar time expansion techniques but employs analogue charge transfer devices for wave storage with CMOS digital circuitry for control functions and potential realisation as a single integrated circuit. Miniaturisation of the system would permit the unscrambler to be carried by the diver, leading to improvements in through-water, diver-to-diver communication systems.

*Dr M.A. Jack, E.E. Dept., University of Edinburgh, EDINBURGH.
†Dr A.D. Milne, Wolfson Microelectronics Institute, EDINBURGH.
ΔDr L.E. Virr, Admiralty Marine Tech. Est., PORTSMOUTH.

PRINCIPLE OF OPERATION

Figure 1(a) shows the basic form of a voiced helium speech waveform where the amplitude peaks correspond to the start of the pitch intervals. As indicated in the introduction, this pitch interval which is determined by muscular properties of the larynx, shows minimal change from normal air to high pressure helium/oxygen mixture.
Here, the helium speech waveform suffers a shift in (vowel) formant frequencies and the rate of decay of the inter - pitch waveform is correspondingly more rapid than for normal speech. The unscrambler technique considered here consists of storing those sections of the inter - pitch waveform which contain useful information, figure 1(b) and subsequently time-expanding this stored waveform to the general form expected in normal (air) speech, figure 1(c). Figure 1(c) shows that in the long term there is no time-base change (the pitch intervals on input and output remain equal) however, in the short term, between pitch peaks, the time-base is expanded. The discarded sections of input waveform cause little degredation to the speech intelligibility.

The maximum duration of the stored segment is governed by the maximum expected pitch for the input helium speech. The design considered here stores a segment of duration 3ms corresponding to a maximum voiced fundamental frequency of 330 Hz.

The degree of inter- pitch time-base expansion is governed by the specific helium/oxygen mixture being used by the speaker. However, the maximum required time-base expansion is of the order of 3:1. Thus, in order to eliminate loss of any section of input signal which contains meaningful information, four parallel storage channels are required such that one channel is always available to store the signal. More than one channel may be producing an output at any one time, a feature which is acceptable since in normal speech, the sounds produced during succesive pitch intervals tend to superimpose.

SYSTEM DESIGN

Figure 2 shows a schematic block diagram of the compact helium speech unscrambler considered here. The helium speech input is fed to a high-gain pre-amplifier which incorporates high frequency pre-emphasis to compensate for the radiation losses produced at the mouth of a speaker in helium. The pre-amplifier also incorporates an AGC facility (30 db). The pre-amplifier output enters the analogue delay lines (Reticon SAD 512), each of which is capable of storing N =256 samples of the input waveform. As indicated previously, four storage channels are required to permit a maximum time-expansion of 3:1 without loss of information bearing signal.

The pre-amplifier output further appears at the pitch detector. In view of the reduced envelope decay times of the speech waveform in in helium, a simple peak detector circuit can be successfully employed for pitch synchronisation. For unvoiced sounds, characterised by a noise-like waveform, the pitch detector operates continuously.

The cycle of operation of the process is as follows. The output from the pitch detector indicates the start of a pitch period. The input helium speech - of bandwidth up to 16 kHz - is read into one of the four channels at a clock rate of 85 KHz for an interval of T = 3ms. At the end of this 3 ms period, the clock frequency for this channel is reduced by a factor which is dependent on the helium/oxygen mixture being used. The stored signal is read out at a lower rate, with an attendent bandwidth compression to the normal 3-4 KHz speech bandwidth. On detection of the start of the subsequent pitch interval, the clock and signal multiplexers change over and the next channel reads in the helium speech.

PERFORMANCE

Figure 3(a) shows a section of the input helium speech waveform (85 m depth) with figure 3(b) indicating the pitch detector output defining a 3 ms storage interval from the pitch peak. The loss of signal in the discarded part of each pitch period causes minimal degradation in the output speech. The pitch period in figure 3 is seen to be of the order of 7 ms.

Figure 4(a) shows another section of helium speech waveform (85 m depth) with the corresponding time expanded output from one of the four channels. Here the inter - pitch time-base expansion is seen to be of the order of 2:1.

Figure 5 compares the input helium speech waveform (85 m depth), Figure 5(a) with the corresponding unscrambled output, Figure 5(b). Note that the pitch interval (approx 6 ms) remains unchanged although the inter-pitch time-base expansion is of the order of 2:1.

Subjective listening tests of this equipment were performed by a panel of non-expert (in both a phonetic and a diving sense) listeners. The test material consists of a section of helium speech recorded at 85 m depth which comprised 15 individual phrases spoken once only. The intelligibility to the "raw" helium speech was zero. Figure 6 indicates the measured intelligibility of the unscrambler output - each listener writing down each phrase as he understood it. Clearly the mean performance of the listeners improved as they became accustomed to the unscrambled output and intelligibility figures of 90% are indicated. Table 1 compares the system parameters of this compact analogue unscrambler with current digital realisations and emphasises the engineering advantages offered by the charge transfer device design.

CONCLUSIONS

The helium speech unscrambler design reported here offers distinct engineering and operational advantages in comparison to currently available (digital) realisations and the use of analogue charge transfer device technology with associated CMOS logic offers a design which is suited to realisation as a single integrated circuit.

ACKNOWLEDGEMENTS

The authors wish to acknowledge helpful contact with Prof H. Hollien, University of Florida, U.S.A. This work has been carried out under MOD contract NCW381/1256.

REFERENCES

1. Beil, R.G. 'Frequency analysis of vowel sounds produced in a helium rich atmosphere' J.A.S.A. Vol. 34. 1962 pp 347-349.

2. Holywell, K. & Harvey, G. 'Helium speech' J.A.S.A. Vol 36. 1964 pp 210-211.

3. Hollien, H. et al. 'Voice fundamental frequency levels of divers in helium oxygen speaking environments' Biomedical Research Col. 4. No 2, 1977 pp 199-207.

4. Giordano, T.A. et al 'Helium speech unscramblers - a critical review of the state-of-the-art' Trans. IEEE Vol. AU21. No 5. 1973. pp 436 - 444.

5. Gill, J.S. et al, British patent no. 1321313, June 1970.

FEATURE	ANALOGUE	DIGITAL
Input bandwidth	16 kHz	12 kHz
Power consumption	150 mW	1200 mW
Physical size	10 cb. in.	1160 cb. in.
Weight	0.5 lb	17 lb
Standby power	30 hrs	10 hrs
Data frame (T)	3 ms	2.5 ms

Table 1 Comparison of system parameters.

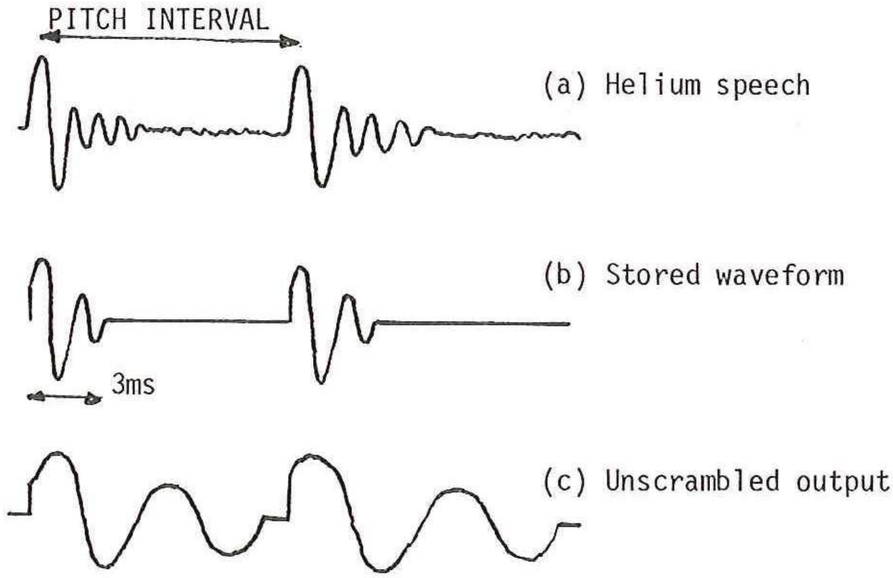

Figure 1 Principle of helium speech unscrambler operation.

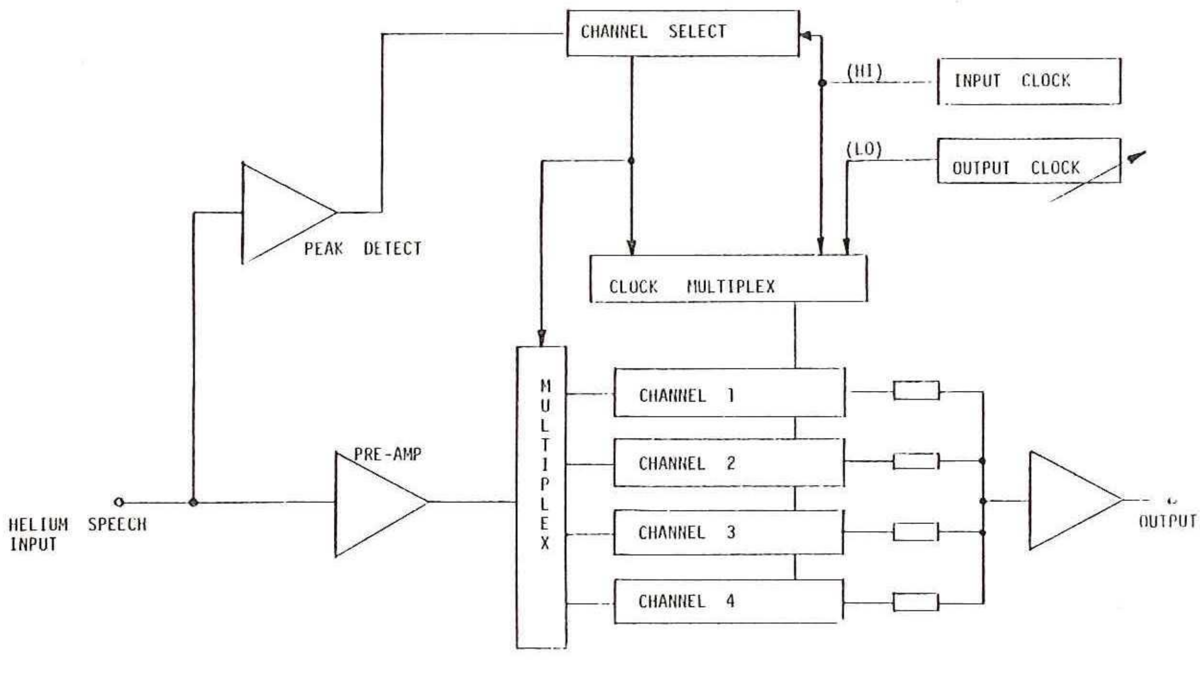

Figure 2 System block diagram.

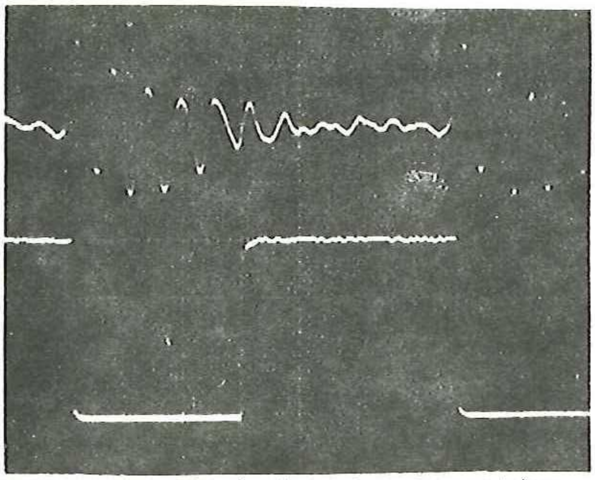

Figure 3 Pitch detector operation.
(a) Input helium speech 500mV/div.

(b) Peak detector output 5V/div

Horizontal scale 1μs/div.

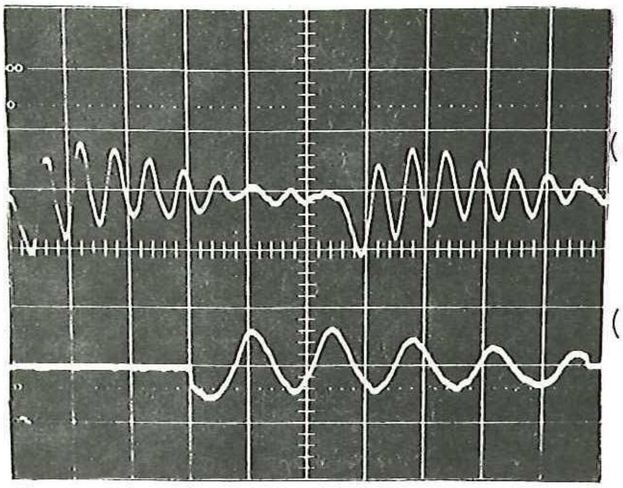

Figure 4 Channel output.
(a) Input helium speech 500mV/div.

(b) Channel output 500mV/div.

Horizontal scale 1μs/div.

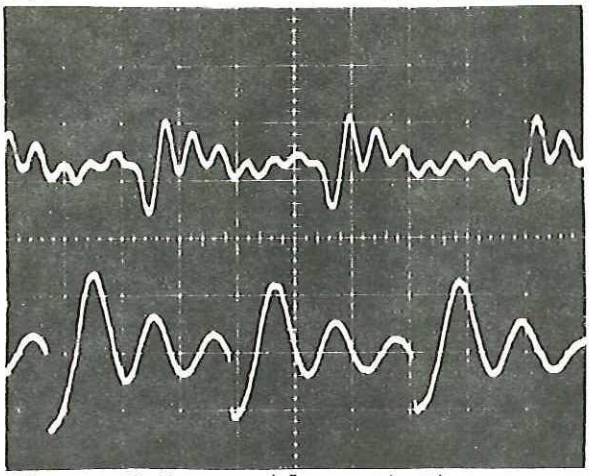

Figure 5 Unscrambler output.
(a) Input helium speech 500mV/div.

(b) Final output 500mV/div.

Horizontal scale 2μs/div.

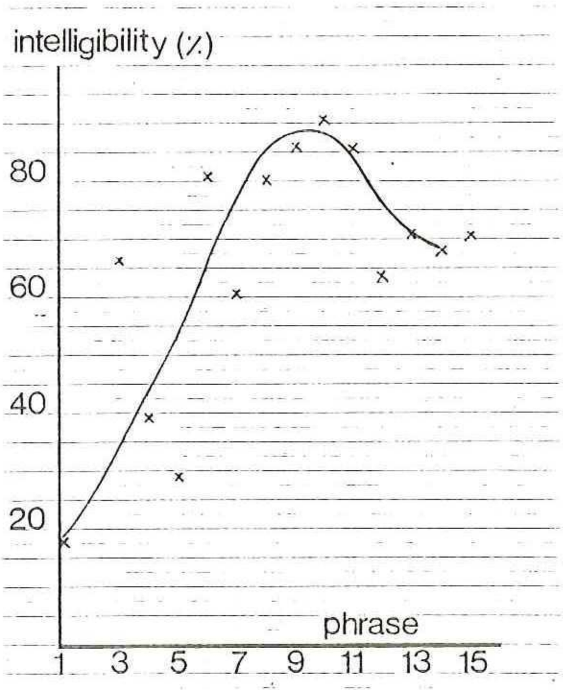

Figure 6 Intelligibility test results.