ON THE TEMPORAL RELATIONSHIP BETWEEN FUNDAMENTAL FREQUENCY AND LARYNX CLOSED QUOTIENT IN INITIAL PLOSIVES

Paul E Garner and David M Howard

Parallel and Signal Processing Applications Research Group,
Electronics Department, University of York, Heslington, York YO1 5DD, UK.

1. ABSTRACT

The use of features such as fundamental frequency (F0) and larynx closed quotient (CQ) to describe human voice source characteristics is well established. In particular, the relationship between CQ and F0, derived from the electrolaryngograph and presented as a scattergram, allows their long-term trends of speech utterances to be investigated. This paper describes a system based on a PC compatible computer controlling an Ariel PC-56D DSP co-processor board, which extracts CQ and F0 and presents a contour rather than a scattergram. Such contours have been used to investigate the temporal course of CQ and F0 for various speech sounds. In particular, voiced and unvoiced plosives in word initial position are considered which exhibit distinct differences in the patterning of CQ and F0. The paper details system operation and discusses results and potential applications.

2. INTRODUCTION

Human voice source characteristics can be identified by various features which can be directly related to the articulation gesture employed in the production of sound. Two such features which are now well established for real-time visual feedback are the vocal fold fundamental frequency (F0) and the larynx closed quotient (CQ) e.g. [Howard, 1994].

In recent years the technique of electrolaryngography for the indirect examination of the vocal fold vibrations has been widely used [Fourcin, 1974]. An analysis of the waveform derived from the electrolaryngograph (Lx) can provide a measure of both F0 and CQ [Abberton et al., 1989].

A common technique used to investigate the relationship between CQ and F0 for long term trends of speech utterances is the presentation of the data in the form of a scattergram (Qx), where the density of points represent the number of occurrences of that particular larynx cycle [Abberton et al., 1989].

This paper describes a system based on a PC compatible computer controlling an Ariel PC-56D DSP co-processor board which, in addition to plotting Qx, also can provide a curve or contour extracted from the Lx waveform. The system has been used to investigate the temporal course of the relationship between CQ and F0 contours for voiced and unvoiced bilabial plosives in word initial position.

The paper describes the operation of the system, pilot study results and details of planned future work.

3. THE MEASUREMENT OF CQ AND F0

The output waveform (Lx) from the electrolaryngograph provides the input to the system. The example Lx waveform shown in figure 1 indicates increasing inter-electrode current flow, or decreasing electrical impedance, as positive-going. Therefore the positive peaks in Lx (increased vocal fold contact[1]) represent the closed phase (CP) of the vocal folds and the troughs the open phase (OP). The boundaries between these two phases are found in the following way. The transition from open phase to closed phase is a well defined event corresponding to the instant when the vocal fold area of contact is increasing most rapidly. This is located as the positive peak in the differentiated Lx waveform (DLx) and the time between these positive peaks is used to define the fundamental period (Tx). There are various methods previously defined for the transition from closed phase to the open phase. Here method a) of Davies et al. is used which can be summarised as follows.

(a) The opening phase corresponds to a position where Lx has reached a value that divides the range from peak to trough of that cycle in a fixed ratio. Previous experimental work [Fourcin 1974] indicated a value of 3/7 for the ratio (see figure 1).

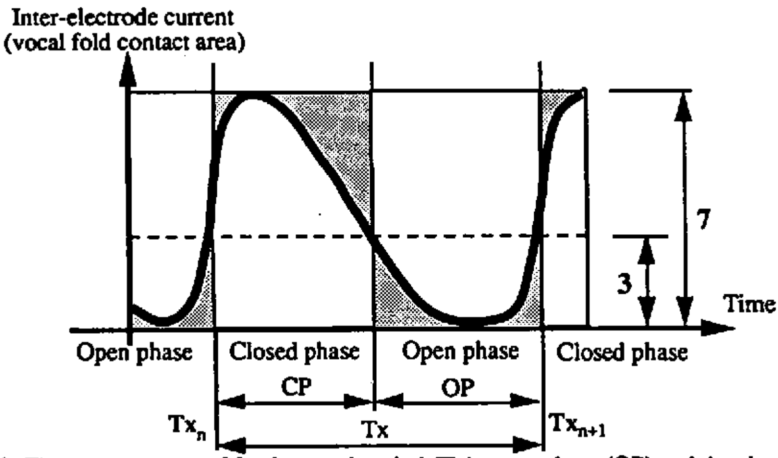

Figure 1: The measurement of fundamental period (Tx), open phase (OP) and the closed phase (CP) from the Lx waveform.

[1] It should be noted that Lx cannot be used to determine whether the vocal fold contact is complete (glottis fully closed) or incomplete (glottal chink).

The closed phase (CP) can be expressed as a ratio to the overall Lx period as shown below and is usually referred to as the closed quotient.

$$\text{Closed Quotient} : CQ = \left(\frac{CP}{Tx} * 100 \right)\%$$

Also the fundamental frequency is given by:

$$\text{Fundamental Frequency} : F0 = \frac{1}{Tx}$$

These two parameters, CQ and F0, form the basis for the contour displays described below.

4. THE CQ AND F0 CONTOUR

The measurement of CQ and F0 for the contour display is performed by the combination of a PC compatible computer controlling an Ariel PC-56D DSP board which incorporates a Motorola 56001 DSP integrated circuit [Howard & Garner, 1992]. The Ariel board contains a switchable input pre-amplifier, an analogue to digital converter whose sampling rate is set to 19.6kHz and additional program and data memory. An overall system diagram is shown in figure 2.

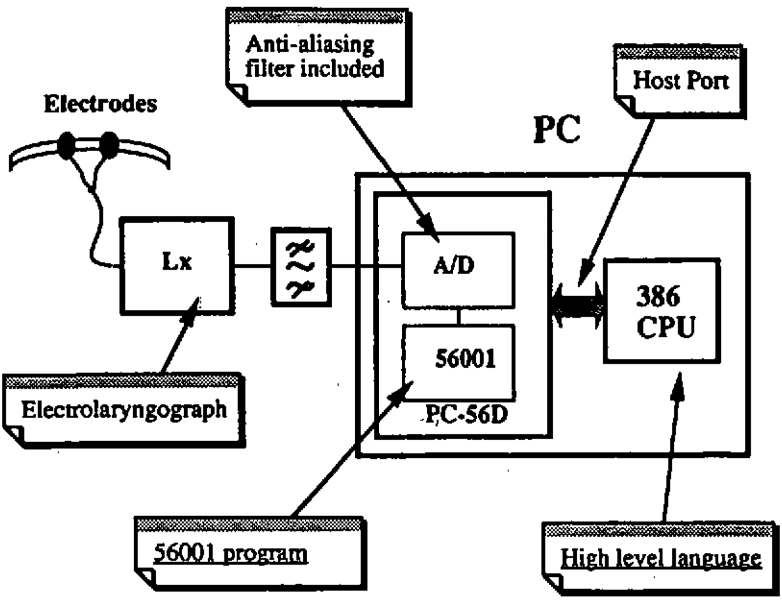

Figure 2: The overall system block diagram.

There are always two programs required for the CQ/F0 contour display: a high level language program for the PC and a signal processing program for the PC-56D board.

The signal processing program has the task of extracting CP and Tx from the electrolaryngographically derived Lx waveform. If a valid Tx marker is not detected within 33ms of the previous one (i.e. F0 < 30Hz) the system assumes that a voiceless section of the speech has been encountered. The CP and Tx data along with markers indicating voiceless periods are transferred to the PC via the host port.

The high level language program has various tasks to perform. First, it initialises the PC-56D, down-loads the signal processing program and also provides the user interface. In addition it takes the transferred CP and Tx data and converts this to CQ and F0. The CQ:F0 data pairs are collected in sequence, in groups separated by the voiceless periods detected by the signal processing program. This effectively splits the utterance into periods of voiced speech. These periods of CQ:F0 data are subjected to a cubic spline interpolation algorithm which provides the necessary interpolated co-ordinate data points required to plot the data

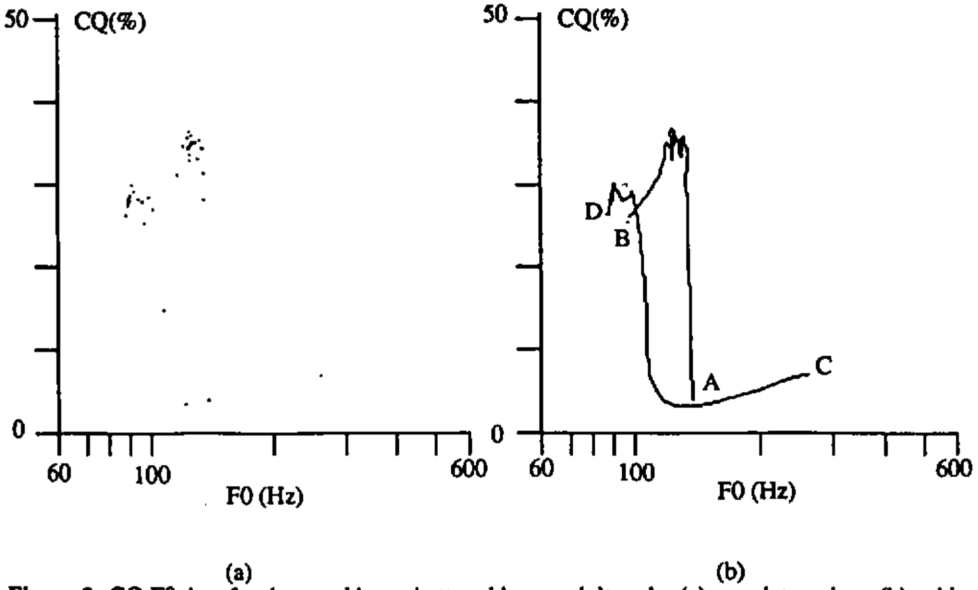

Figure 3: CQ:F0 data for the word 'paper' uttered by an adult male: (a) raw data points, (b) cubic spline interpolated contours.

graphically. This is illustrated in figure 3. The display shown in figure 3(a) shows a plot of the total received CQ and F0 data for the utterance 'paper' plotted as single data points. Figure 3(b) shows the cubic spline interpolated contours for the same utterance. The two curves AB and CD, which make up the display, represent the voiced vowel sounds /eI/ and /ə/ of 'paper' respectively. The gap in the curves between B and C represents the second voiceless bilabial plosive /p/ in 'paper'.

5. RESULTS

The system is presently being used to examine the relationship between CQ and F0 for various speech sounds. Initially word initial voiced and unvoiced plosives were investigated and the pilot study results are presented here.

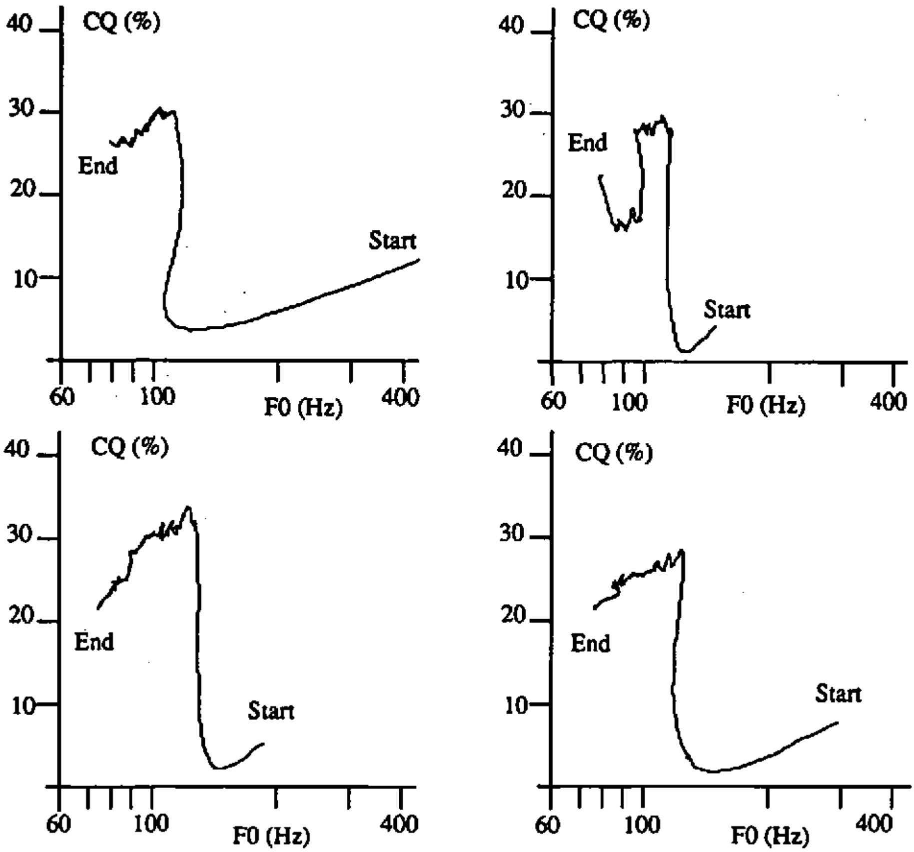

Figure 4: Four utterances of the word 'power' (same speaker as in figures 3 and 5)

Figure 4 shows four utterances of the word 'power' and figure 5 shows four utterances of the word 'bower', all uttered by one of the authors. The direction in which the contours are plotted is shown by indicating the start and end points.

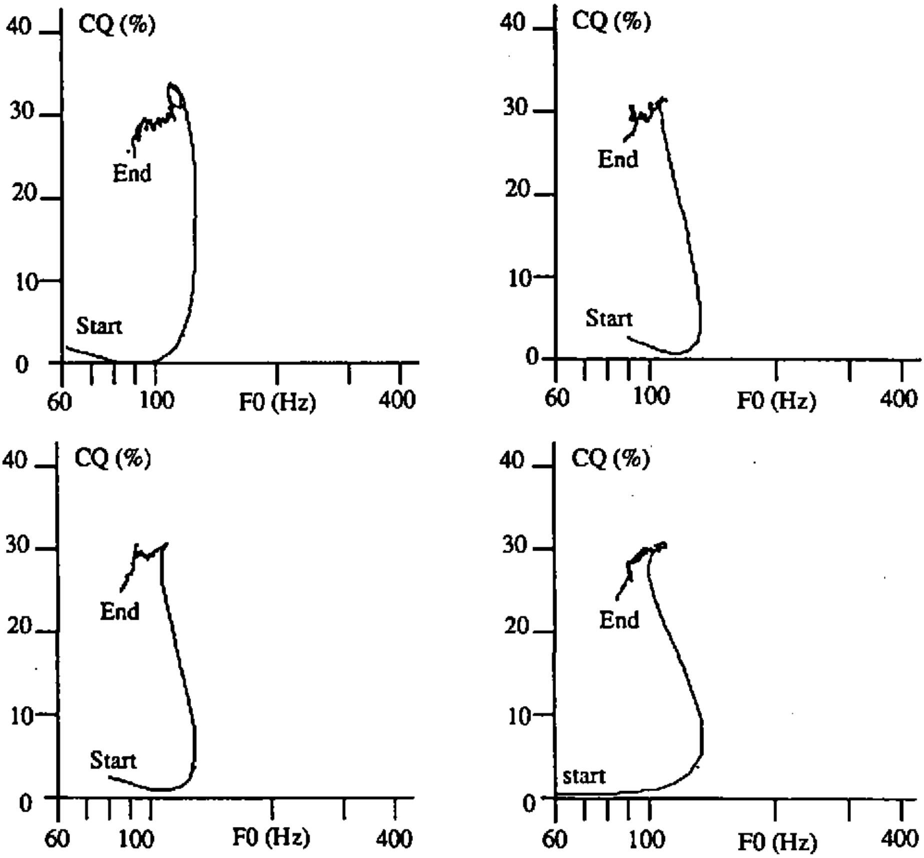

Figure 5: Four utterances of the word 'bower' (same speaker as in figures 3 and 4).

Although the differences between the two contours for 'power' and 'bower' are quite distinct care must be taken in the interpretation of this type of display.

The initial results from the contour extraction system imply there is a distinct variation in contours between word initial voiced and voiceless combinations. A more detailed study is in progress to determine the relationship between CQ and F0 for a selection of speech sounds for a number of male and female speakers.

One application immediately suggests itself. If a unique relationship between CQ and F0 for particular speech sounds can be established it could provide new rules for speech synthesis systems which could result in a more natural sound. At present, most synthesis-by-rule systems incorporate few or no rules for dynamic voice source control.

6. CONCLUSIONS

A system which extracts contours indicating the relationship between CQ and F0 has been developed. This system has initially been used to investigate the differences between word initial voiced and unvoiced plosives and distinct differences have emerged in the patterning of CQ and F0. This pilot study is to be broadened to include other voiced/voiceless pairs for a number of male and female speakers. The CQ and F0 relationship obtained from this study will be used to provide dynamically changing excitation in a speech synthesiser and any resulting changes in naturalness will be investigated.

7. REFERENCES

ABBERTON, E.R.M., HOWARD, D.M. & FOURCIN, A.J. (1989). Laryngographic assessment of normal voice: A tutorial, *Clinical Linguistics and Phonetics*. 3, 281-296.

DAVIES, P., LINDSEY, G.A., FULLER, H. & FOURCIN, A.J. (1986). Variation in glottal open and closed phases for speakers of English, *Proceedings of the Institute of Acoustics*, 8, (7), 538-546.

FOURCIN, A.J. (1974). *Laryngographic assessment of vocal fold vibration*, In B Wyke, (Ed), Ventilatory and Phonatory control systems, Oxford: Oxford University Press.

FOURCIN, A.J. & ABBERTON, E.R.M. (1971). First applications of a new laryngograph, *Medical and Biological Review*, **21**, 172-182.

HIRSON, A. & FAWCUS, R. (1991). Visual feedback in the management of dysphonia, In: M. Fawcus, (Ed) Voice disorders and their management, 2nd edn, London: Chapman Hall.

HOWARD, D. M. (1994). Real-time Visual Displays in Speech and Singing, *Defence Science Journal*, 43, 3, pp 211-221.

HOWARD, D. M. & GARNER, P.E. (1992). DSP-56000 based real-time electrolaryngographically derived closed quotient, *Proceedings of the Institute of Acoustics*, 14, 375-381.

Motorola *DSP56000/DSP56001, Digital signal processor user's manual*, [Rev.1]. Motorola.

Ariel corp. (1991). *User's manual for the PC-56D, DSP co-processor board*. Second edition. Ariel Corporation.