DSP-56000 based real-time electrolaryngographically derived closed quotient

David M Howard and Paul Garner
Signal Processing: Voice and Hearing Research Group,
Electronics Department, University of York, Heslington, York YO1 5DD

1. Abstract

Previous work investigating voice source changes with singing training/experience in adults suggests that there are relationships between electrolaryngographically derived vocal fold closed quotient (CQ) measures in speech and singing and the extent of singing training/experience. These data suggest that real-time visual displays of CQ against time could be used beneficially to support students in voice training, which requires a real-time CQ measurement system. This paper reports the development of a CQ measurement system based on the output waveform from the electrolaryngograph (Lx), comprising a PC compatible computer with an Ariel PC-56D DSP co-processor board, which incorporates a Motorola 56001 DSP integrated circuit. The Lx waveform is processed in real-time on a cycle-by-cycle basis to give fundamental frequency and CQ values which are displayed graphically against time in real-time.

2. Introduction

The use of real-time visual feedback of fundamental frequency (F0) against time by speech therapists working on speech production skills of clients is well established [1,2]. We are considering the application of visual displays for professional voice users, particularly singers [3,4]. Previous work with the SINGAD (SINGing Assessment and Development) system [5] has demonstrated that a visual display can enable the development of conscious pitch matching skills in the classroom for five- to seven-year-olds [6].

A measure which has potential application in a visual display for professional voice users is the larynx closed quotient (CQ) derived from the electrolaryngograph [7]. CQ is defined for each cycle of vocal fold vibration as the percentage of that cycle for which the folds are in contact. Howard et al. [8] measured CQ for a group of 18 adult males with varying degrees of singing training/experience, and demonstrated a significant difference in absolute CQ values between the trained and untrained group when singing a two octave ascending and descending G major scale. Those subjects with greater singing training/experience were making use of higher CQ values. Howard et al. [9] report the results of a similar experiment for a group of adult females. There the difference appears not as an absolute CQ difference, but as a difference in the *patterning* of CQ change with F0.

This paper describes the implementation of a real-time CQ and F0 visual display based on a PC compatible computer and a Motorola 56001 co-processor board. Initial results based on displays gained from a singing student are discussed to illustrate the operation of the system.

3. Method

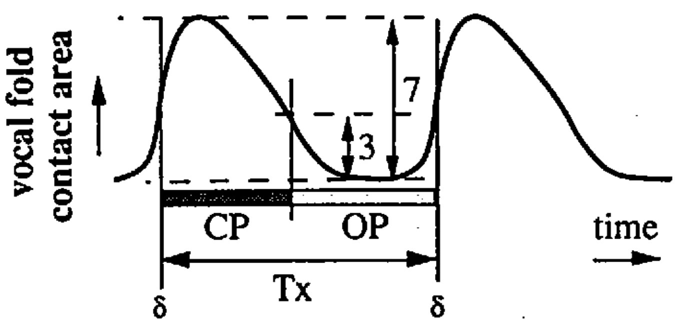

Figure 1: *The measurement of fundamental period (Tx), open phase (OP) and closed phase (CP from the Lx waveform. Tx is the interval between the starts of each CP found as each positive peak (indicated as d) in the time differentiated Lx waveform. The end of CP is found as the point where the negative going Lx crosses 3/7 of that cycle's pk-to-pk amplitude.*

The output from the electrolaryngograph (Lx) provides the input to the system from which larynx closed quotient is to be measured. The derivation of larynx closed quotient is described in Davies et al.[10], and illustrated in figure 1. The start of the closed phase (CP) in each cycle is defined as the instant when vocal fold area of contact is increasing most rapidly. This is readily located as the positive peak in the differentiated Lx waveform (shown as d in figure 1) and the time between each of these positive peaks gives the fundamental period (Tx) and hence the fundamental frequency (F0 = 1/Tx). The end of the CP is defined as the instant when the negative-going portion of Lx crosses a fixed ratio (3/7) of the amplitude of the current cycle. The closed quotient is defined thus:

$$CQ = ((CP) / (Tx) * 100)\ \%$$

The measurement is made using the Ariel PC-56D DSP co-processor board which communicates with, and is controlled by, the host PC compatible computer. An overall software control flow diagram is shown in figure 2. The incoming Lx waveform is sampled at a rate of 20kHz by the analogue to digital converter on the Ariel board. An anti-aliasing filter is present on the Ariel board whose cut off frequency is automatically set to an appropriate value depending on the sampling rate requested. In addition to this an external high-pass filter with a cut off frequency of 20Hz is included to remove any low frequency variations on the LX due to, for example, larynx movement. The main processing is carried out by the 56001 DSP integrated circuit.with the host PC acting in a supervisory role. On command from the PC the 56001 program analyses the incoming sampled Lx waveform to find valid CP and Tx. During this period the PC monitors the keyboard and controls the screen display whilst periodically checking on the state of the 56001, transfering valid CP and Tx data when it is available.

The output from the electrolaryngograph is polarised such that positive going change represents increased area of vocal fold contact The Lx waveform is time differentiated (DLx) and the maximum positive peaks are used to indicate the start of each closed phase and the Tx markers. These peaks are subject to a voiced/voiceless threshold value set to exclude non-valid peaks in

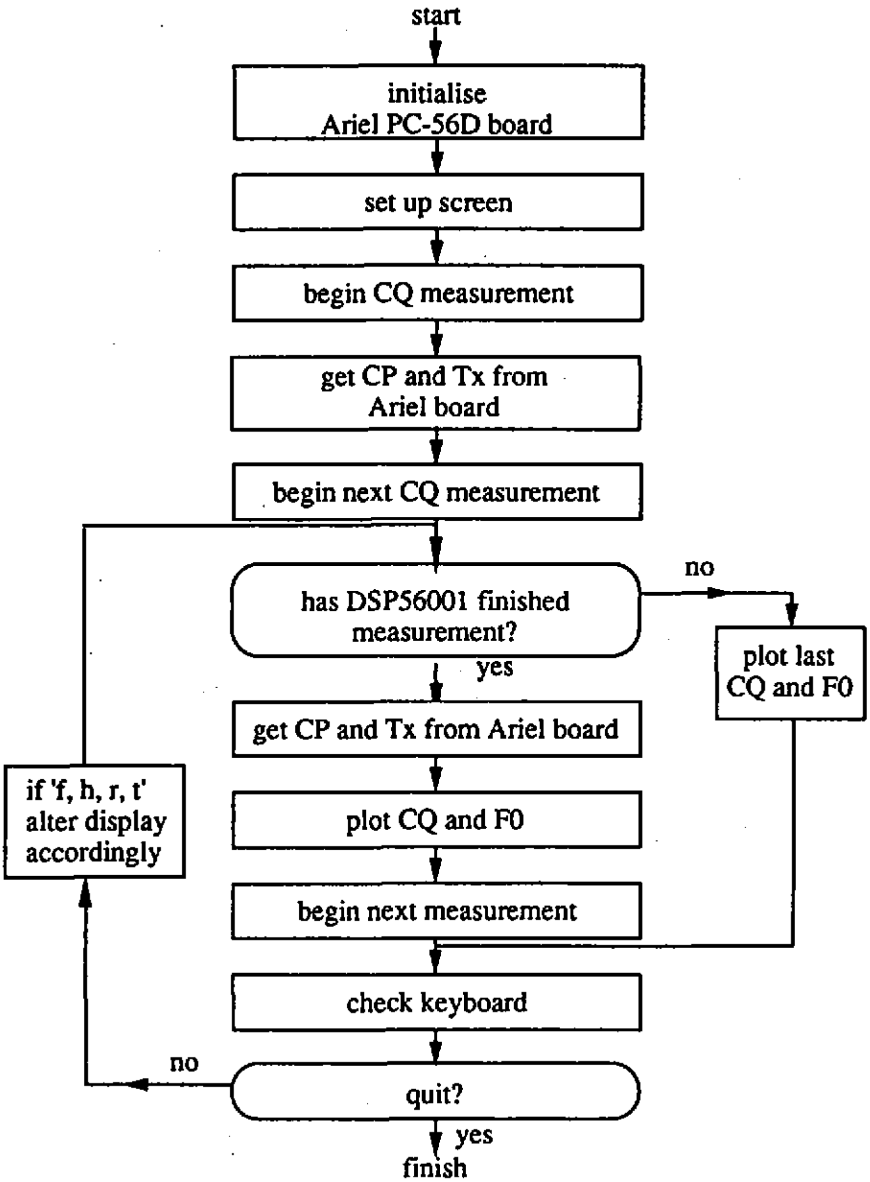

Figure 2: Software flow diagram.

DSP-56000 BASED CLOSED QUOTIENT

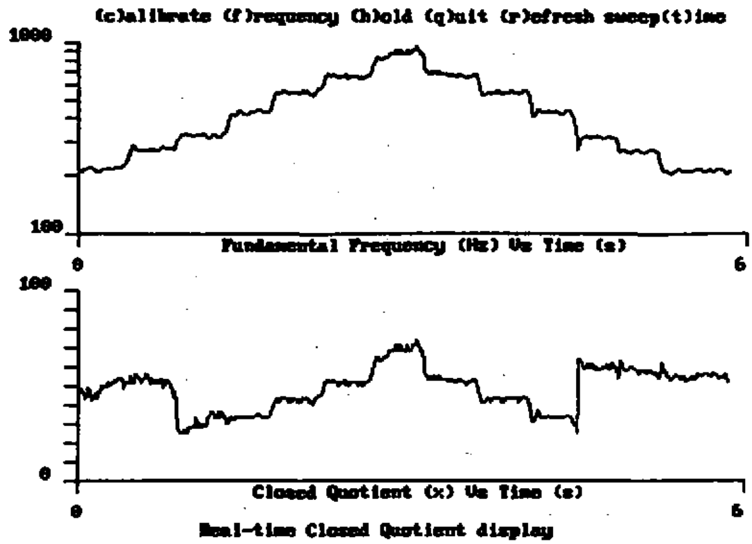

Figure 3: Real-time visual display screen output for an adult female singing a two octave ascending and descending A major arpegio from A(220Hz). (UPPER TRACE: F0 in log Hz; LOWER TRACE: CQ in %)

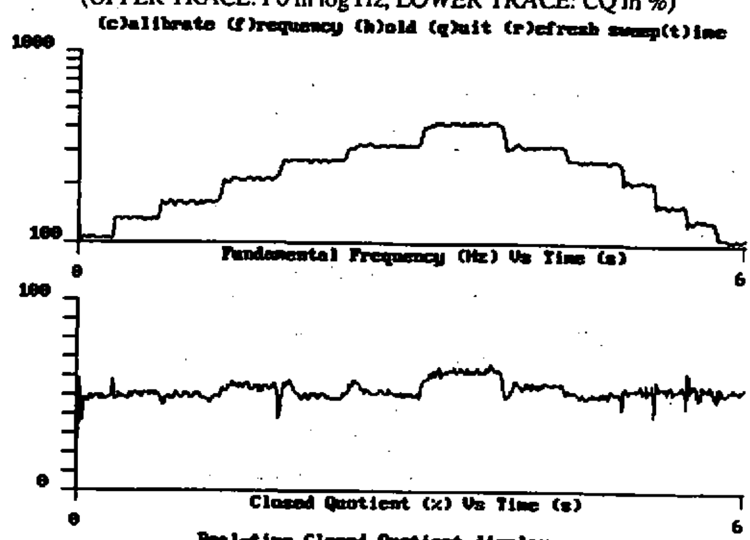

Figure 4: Real-time visual display screen output for an adult male singing a two octave ascending and descending A major arpegio from A(110Hz). (UPPER TRACE: F0 in log Hz; LOWER TRACE: CQ in %)

DLx due to noise. This threshold is calculated from the incoming sampled Lx waveform and then applied to the succeeding data samples in order to find the first valid Tx marker. Once a valid positive peak in DLx has been found, the system is inhibited from finding another peak for a fixed time (known as the refractory time) to prevent retriggering on peaks due to noise and local overshoots of the Lx waveform. The refractory time can be altered and it has a default setting of 1ms, the period of 1kHz which we consider to be a reasonable upper F0 range for our singing analyses[1].

After finding a valid Tx marker, the 56001 stores the incoming Lx waveform up to the next Tx marker. As Lx is stored, the maximum amplitude is sought followed by the minimum amplitude for that cycle. Then the amplitude threshold, established as 3/7 of the peak-to-peak amplitude of each cycle, can be found and used to find the end of the closed phase for that cycle. The CP length is the difference between its end and start times, and the values of CP and Tx are passed to the host PC computer, where CQ is found and a real-time display of F0 and CQ is presented.

The software provides the following options for the user:
(i) the F0 axis, which is logarithmic, can be altered between 100Hz-1kHz, 10Hz-1kHz, and 100Hz-10kHz[2] (by pressing 'f');
(ii) the trace can be held on the screen (by pressing 'h');
(iii) the screen can be cleared (refreshed) before the next trace (by pressing 'r');
(iv) the sweep time can be varied between 6s, 12s and 18s (by pressing 't');
(v) the program can be quit (by pressing 'q').

Figures 3 and 4 show screen displays of F0 and CQ for an adult female and an adult male respectively singing two octave A major arpeggii, ascending and descending. The female start note is A(220Hz), and that for the male is A(110Hz). The F0 trace is plotted logarithmically so that the visual representation closely approximates the F0 percept, between the frequency bounds selected by the user (see (i) above), and CQ is plotted on a linear 0%-100% axis.

4. Discussion

In each plot, the pattern of notes for each arpeggio is clearly to be observed in the F0 plot (upper trace in each figure). The CQ plot for the female (see figure 3) has a value around 50% for the first two notes, then it jumps down to around 30% and rises in discreet steps to the top note, from where it falls in steps until the last three notes. The CQ plot for the male (see figure 4) remains essentially constant, between approximately 47% and 60% throughout. The top note, A (440Hz), has the higher CQ around 60%, with the remaining notes having values which are very close.

The wide F0 range demanded of the singing voice as compared to the speaking voice results in the vocal folds adopting different modes of vibration across different parts of the frequency range, often referred to as vocal *registers*. However "the terminology used for registers is very

---

[1] Top C for a soprano (and few sopranos can sing top C comfortably!) has a nominal equal-tempered frequency of 1024Hz. At present in our work, we use the G below this as our upper pitch bound, which has a nominal frequency of 784Hz.
[2] To enable F0 to be analysed above 1kHz, the refractory time would require alteration accordingly. This range has been made available to enable high soprano voices (those "in alt" and "in altissimo") to be analysed in the future.

confusing" [11, p50]. A register break is usually clearly heard, especially in the untrained voice, as an abrupt change in phonatory quality if a series of notes progressing upwards (or downwards) is sung over a wide pitch range.

As our female singer was singing the arpeggio, register breaks could clearly be heard at the two instants where there are abrupt changes in CQ, with the notes between the breaks being heard as having a similar voice quality. For our male singer, the most obvious change in vocal quality was for the top note which is towards the extreme of his range.

It should be noted that these observations can only be treated as preliminary, since the display has only very recently been fully operational. However, it is of interest to note that these initial results support findings relating to the effects of singing training/experience on adult males and females. A pilot study of 4 adult males [12] and a later study of a group of 18 adult males [8] have demonstrated that closed quotient for trained and untrained singers remains essentially constant with F0, and the greater the singing training/experience, the higher the CQ used. Our male singer (one of the authors) has received some vocal training and has extensive choral singing experience. For the arpeggio plot in figure 4, his CQ mean is approximately 53%, and the data of Howard et al. [8] found mean CQ values between 35% (untrained) and 63% (trained).

A study of adult females [9] suggests that the more trained female voice exhibits a change in the *patterning* of CQ with F0, with CQ tending to rise with F0, across all or part of their pitch range. The mean CQ exhibits no obvious trend with singing training/experience, and it varies widely between different subjects. Our female subject has had some singing training, and it can be seen from figure 3 that her CQ rises with F0, a suggested feature of a more trained female voice. However, one of the main objectives of singing training (for males and females) is to *cover* any register breaks so that they become essentially inaudible to the listener. In this case, her register breaks are very audible and clearly shown on the CQ display.

The purpose of producing this real-time display is to enable CQ/F0 changes to be monitored while listening to the vocal output with a view to incorporating use of the display as a formal voice training procedure. Changes in CQ have also been noted when singing in different styles [13]. Listening to the sung sound is the main channel by which a singing teacher operates, and responses are made based on qualitative judgements of that sound, drawing on the teacher's own background and experience. The object of this work is to provide data on the quantitative elements of the voice training process and to implement appropriate real-time displays which could be used by the pupil to support the work of the teacher both in lesson time and during practice time. Thus the teacher will be given more time to explore other key aspects of voice training which have to be based on qualitative judgements, such as musicality and performance skills.

5. Conclusions

The implementation of a real-time display of electrolaryngographically derived closed quotient and fundamental frequency has been described, which makes use of an Ariel PC-56D DSP co-processor board and a PC compatible computer. The display has been briefly tried with a male and a female singer, each having received some voice training. The results for an ascending and descending two octave arpeggio illustrate CQ/F0 changes which support recent experimental findings in a study of the effects of vocal training. The future intention is that this display will be used as a real-time visual feedback system during vocal training for professional voice users, and a pilot study is planned for 1993.

6. Acknowledgements

The authors would like to thank Michelle Evans for singing for us and agreeing to have her data plotted and to David Rossiter for useful comments on an earlier draft. This work is supported by a York University Innovation and Research Pump Priming grant, and SERC research grant number GR/H14878.

7. References

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

[2] A J FOURCIN. 'Laryngographic assessment of vocal fold vibration', In: B WYKE, (Ed.) 'Ventilatory and phonatory control systems', Oxford: Oxford University Press (1974).

[3] D M HOWARD, & J COLLINGSWORTH. 'Voice source and acoustic measures in singing', *Acoustics Bulletin*, 17 (4) pp5-12 (1992).

[4] D M HOWARD, & D ROSSITER. 'Results from a pilot longitudinal study of electrolaryngographically derived closed quotient for adult male singers in training' *these proceedings* (1992).

[5] D M HOWARD, & G F WELCH. 'Microcomputer-based singing ability assessment and development', *Applied Acoustics*, 27 pp89-102 (1989).

[6] G F WELCH, D M HOWARD, & C RUSH. 'Real-time visual feedback in the development of vocal pitch accuracy in singing', *Psychology of. Music*, 17 pp146-157 (1989).

[7] A J FOURCIN, & E R M ABBERTON, 'First applications of a new laryngograph',*Medical and Biological Review*, 21 pp172-182 (1971).

[8] D M HOWARD, G A LINDSEY, & B ALLEN. 'Towards the quantification of vocal efficiency ', *Journal of Voice*, 4 (3) pp205-212 (1990). (See also Errata, *Journal of Voice*, 5 pp93-95 (1991)).

[9] D M HOWARD, G A LINDSEY, & S K PALMER. 'Larynx closed quotient measure for the female singing voice', *Proceedings of the 12th International Congress of Phonetic Sciences, France: Aix-en-Provence*, 5 pp10-13 (1991).

[10] P DAVIES, G A LINDSEY, H FULLER, & A J FOURCIN. 'Variation in glottal open and closed phase for speakers of English', *Proceedings of the Institute of Acoustics*, 8 pp539-546 (1986).

[11] J SUNDBERG. 'The science of the singing voice', Dekalb: Northern Illinois University Press (1987).

[12] D M HOWARD, & G A LINDSEY. 'New laryngograms of the singing voice', *Proceedings of the 11th International Congress of Phonetic Sciences, USSR: Tallinn*, 5 pp166-169 (1987).

[13] D M HOWARD. 'Quantifiable aspects of different singing tyles - A case study', *Voice - The Journal of the British Voice Association*, 1 (1) pp47-62 (1992).