Journal of Voice
Vol. 9, No. 2, pp. 163–172

Variation of Electrolaryngographically Derived Closed Quotient for Trained and Untrained Adult Female Singers

David M. Howard

Parallel and Signal Processing Applications Research Group, Department of Electronics, University of York, Heslington, York, England

F0 - fundamental frequency for singing, the range typically varies by gender with men having a lower F0

Summary: The derivation of larynx closed quotient (CQ) measures from the electrolarynogograph output is discussed and data are presented for a group of trained and untrained adult female singers (N = 26) for a sung two-octave G major scale. Statistically significant trends are observed between the trained and untrained groups that suggest for the trained group: (a) CQ tends to be lower for pitches below D4 and higher for pitches above B4, and (b) the gradient [CQ/log(F0)] tends to correlate positively with the number of years singing training/experience. These data are compared with those reported previously for an adult male group, and it is suggested that CQ could be a useful parameter to include in a real-time visual display for singing training. **Key Words:** Electrolaryngography—Singing—Vocal efficiency—Closed quotient.

The formal training of the singing voice is traditionally associated with qualitative evaluation of the vocal output by a teacher who makes much use of imagery in various forms to achieve a particular sung sound. Whereas such techniques will continue to be basic to achieving the sung sound appropriate for a particular voice, increased awareness amongst those presenting themselves for vocal training of the physical nature of the vocal mechanism results in demands for further information to supplement the use of imagery. In practice, the responses to such demands are so often ill-informed since the craft of the singing teacher is usually based on experience that has been handed down over many teacher–pupil generations.

The introduction of microcomputers and noninvasive techniques for monitoring aspects of vocal output have meant that it is now possible to analyse large quantities of data without interfering with vocal production. More recent technological advances are enabling such analyses to be carried out in real time, and visual displays of a number of quantifiable aspects of vocal output are becoming viable. As it becomes clearer which parameters vary as the professional voice develops, the possibility exists for the provision of real-time visual displays for use during the training process. Such displays will never *replace* the voice teacher since there are aspects of their craft, such as musical expression, working with an accompaniment, stagecraft, confidence building, dramatic presence, making best use of the performance space acoustics, etc., that are by their nature, qualitative processes.

This paper explores the nature of data derived from measures that could be gathered in real-time and used in visual displays. The data relate to electrolaryngographically derived closed quotient measures for adult female trained and untrained singers. In particular, patterns of variation of closed quotient with fundamental frequency are discussed that have a high correlation with the number of years singing training/experience, and these are compared to previously reported data for an adult male group.

BACKGROUND

The electrolaryngograph (1) enables vocal fold contact area to be monitored noninvasively, giving

Accepted August 12, 1994.
Address correspondence and reprint requests to Dr. D. M. Howard at Parallel and Signal Processing Applications Research Group, Department of Electronics, University of York, Heslington, York Y01 5DD, England.

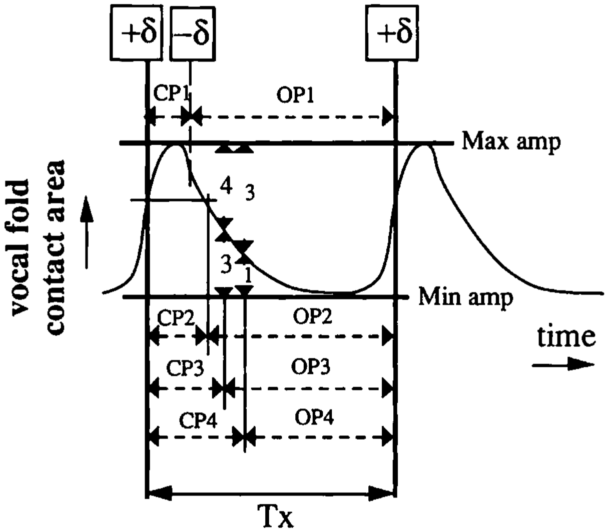


FIG. 1. Example electrolaryngograph output waveform (Lx) showing the basis for the measurement of Tx and four methods for the measurement of closed phase (CP) and open phase (OP) from Lx. +δ, instant of maximum positive peak in differential of Lx; −δ, instant of maximum negative peak in differential of Lx; Tx, fundamental period of Lx (F0 = 1/Tx); CP1/OP1, method (b) of Davies et al. (21)—end of CP defined at instant of maximum rate of vocal fold opening; CP2/OP2, method (c) of Davies et al. (21)—end of CP defined at position where negative going Lx crosses the amplitude level at which the start of the CP has been defined for that cycle; CP3/OP3, method (a) of Davies et al. (21)—end of CP defined at position where negative going Lx crosses amplitude threshold set at 3:7 of that cycle's peak-to-peak amplitude; CP4/OP4, method of Orlikoff (22)—end of CP defined at position where negative going Lx crosses amplitude threshold set at 1:4 of that cycle's peak-to-peak amplitude.

a basis for the quantification of aspects of vocal fold vibration during voiced speech sounds and sung notes. Two electrodes are placed superficially on either side of the neck of the subject at the level of the larynx, and a constant amplitude high-frequency voltage is applied. The electrolaryngograph monitors electrical impedance changes between the electrodes as the vocal folds vibrate by detecting the current flowing between the electrodes.

The electrolaryngograph output waveform (Lx) represents current flow between the electrodes, which will be greater as the vocal fold contact area increases and less as they part. Figure 1 shows a few cycles of Lx. This interpretation has been confirmed by synchronous observation of the Lx waveform alongside other techniques, for example: high speed larynx photography (2–4), an adapted high-voltage X-flash imaging system (5,6), and computer-simulated Lx waveshapes based on models of vocal fold vibration during phonation (7–9). It should be noted that the electroglottograph (EGG) output waveform, which is equivalent to Lx in terms of its experimental derivation, is usually plotted as the inverse. Comprehensive reviews of electrolaryngograph/EGG operation can be found in Baken (10) and Childers and Krishnamurthy (11).

Appropriate interpretation of the Lx (or EGG) waveform enables aspects of the nature of vocal fold vibration to be investigated on a cycle-by-cycle basis, such as the fundamental period, and hence fundamental frequency (F0), the open and closed phases of each vibratory cycle, and the rate of glottal closing and opening. Examples of current applications include the monitoring of vocal fold vibratory patterns in normal speech (12) and disordered speech (13), real-time F0 visual feedback systems (14,15). The provision of a real-time F0 input for assessment and rehabilitation of clients using advanced speech pattern extraction hearing aids (16), reference F0 analysis (17), for example, for use in the development of microphone-based F0 estimation systems (18), and larynx closed-phase linear predictive coding (LPC) speech analysis techniques (19,20).

This work makes use of Lx to provide a measure of electrolaryngographically derived larynx closed quotient (CQ), which is defined as the percentage of each cycle for which the folds are in contact.[1] Its experimental derivation can be described with reference to Fig. 1. The polarity of the Lx is first checked to ensure that positive changes reflect increased inter-electrode current flow.[2] The detection of the start of the closed phase is based on the assumption that the vocal fold contact area changes more rapidly when it is increasing than when it is decreasing (i.e., the folds snap together more rapidly than they part). These points on the Lx waveform can be readily located after time differentiating the Lx waveform, and finding the positive peaks. The time between these peaks is used to provide a measure of Tx (the fundamental period of Lx).

[1] The Lx (EGG) waveform represents changes in current flow between the electrodes that most agree is due to changes in vocal fold contact area. However, the waveform neither indicates whether the glottal area is zero nor whether the vocal folds are maximally apart. For the purpose of this work, CQ measure is defined in terms of its derivation from Lx and it should not be interpreted, for example, as a measure of the time in each cycle for which the glottal area is zero.

[2] The polarity of waveforms played back from different tape recorders is not always predictable depending on whether an odd or an even number of inverting amplifier circuits exist in the signal path. This can vary if, for example, different tape recorders are used to play back data tapes on different occasions.

These points are also used to define the start of the closed phase (CP) in each cycle.

Davies et al. (21) suggest three methods for signalling the end of the CP in each Lx cycle: (a) the instant when the negative-going Lx waveform crosses a fixed ratio (3:7) of the current cycle's amplitude, (b) the instant of the maximum negative peak in the time differential of the Lx waveform, and (c) the instant where the negative-going Lx waveform crosses the amplitude level at which the start of the CP has already been defined for that cycle. In this work, method (a) is used to define CP on a cycle-by-cycle basis. CQ is then found as

$$CQ = [(CP / Tx) \times 100]\%$$

Orlikoff (22) makes use of a measure of CP, termed *contact phase,* which is derived as follows. The maximum in the time-differentiated EGG waveform is used to indicate the instant of glottal closure, and the instant where the EGG wave crosses 25% of the current cycle's amplitude defines the instant of glottal opening. Since Orlikoff is making use of a Fourcin electrolaryngograph, polarising the output as is conventional for Lx, this corresponds to method (a) of Davies et al. but with a different amplitude threshold.

Krishnamurthy and Childers (19) used the minimum in the differentiated EGG waveform as the instant of glottal closure, and the maximum in the differentiated EGG waveform as the instant of glottal opening in order to define the CP for larynx closed-phase LPC of speech. This corresponds directly with method (b) of Davies et al. They used a Fourcin electrolaryngograph and plotted their output with the polarity convention of EGG. Reliable closed-phase LPC analysis requires that the analysis window be defined only during the CP. The use of method (b) is more appropriate here, since it will generally find CP values that are slightly smaller than those from method (a). (This can be inferred by observation from Figs. 1 and 2.)

Method (a) of Davies et al. generally provides the smoothest CP variation with time. This effect is demonstrated in Fig. 2 for the word "nine" spoken by an adult woman, which shows Tx and CP measured by methods (a), (b), and (c) respectively. Although all three methods give similar patterns of CP change with time, methods (a) and (b) also give similar CP values, whereas method (c) gives CP values that are considerably lower. The use of an amplitude ratio in method (a) has the additional advantage that it can be adjusted to optimise the CQ measurement against other techniques. Increasing the amplitude ratio will tend to decrease the measured CP values. Orlikoff chose 25% as his baseline because "this represented the lowest level that would ensure freedom from waveform artefacts often associated with the EGG 'minimal contact phase.' " Davies et al., (21) chose 3:7 as their ratio based on "experimental work by one of the authors (A. J. Fourcin)," but quote 3:10 or 0.3 in their Fig. 1. Reference to their source code shows that the ratio they are actually using in practice is 3:7, as illustrated in Fig. 1.

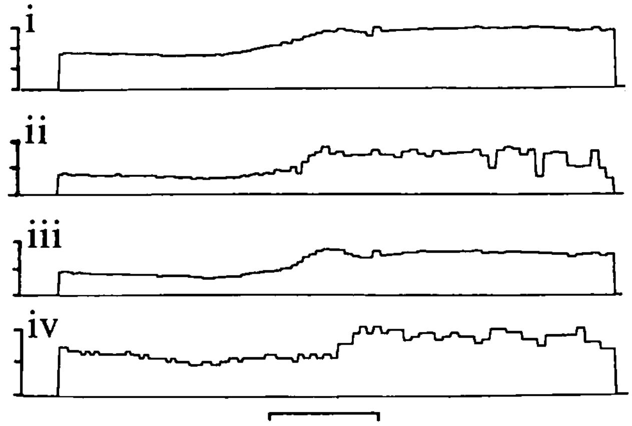


FIG. 2. Plots for "nine" spoken on a falling intonation (rising fundamental period) by a woman: (i) fundamental period against time; (ii) closed phase (CP) measured by method (b) of Davies et al. (21) against time—end of CP defined at instant of maximum rate of vocal fold opening [CP1/open phase (OP)1 in Fig. 1]; (iii) CP measured by method (a) of Davies et al. (21) against time—end of CP defined at position where negative going Lx crosses amplitude threshold set at 3:7 of that cycle's peak-to-peak amplitude (CP3/OP3 in Fig. 1); and (iv) CP measured by method (c) of Davies et al (21) against time—end of CP defined at position where negative going Lx crosses the amplitude level at which the start of the CP has been defined for that cycle (CP2/OP2 in Fig. 1). (Tx ticks: 0.0, 2.5, 5.0, 7.5 ms; CP ticks: 0.0, 2.5, 5.0 ms.; Time axis calibration marker: 50 ms)

In order to gain some validation of the CQ measurement technique, CQ values gained using method (a) of Davies et al. have been shown to correlate well with CQ data measured by hand annotation of the output glottal flow waveform from an automatic speech pressure waveform inverse-filtering technique (23). In practice, if the only requirement is to look for trends in CQ, then the actual ratio used is not important provided that it is kept constant. In this respect, Kitzing (24) reports that his closed-time EGG measurements could differ by up to 15% from photoglottographic measurements. Childers et al. (25) report estimates of open quotient (OQ) values (OQ = 100 − CQ) measured from the time differential of EGG (DEGG) and from ultra-high-speed laryngeal film glottal area in which the average difference between the two is 11% for their group of four normal speakers and 15.4% for their group of eight pathological speakers. OQ measures from DEGG were generally lower than those from the film. Their results also suggest that these differences increase with loud or high-pitched voices.

A pilot study (26) demonstrated that for four male singers there was a tendency for electrolaryngographically derived OQ values to be lower (or CQ higher) with increased singing training/experience. A later study (23) considered data gathered from 18 male singers with varying degrees of singing training/experience and found a highly statistically significant increase in the measured CQ with increasing numbers of years singing training/experience. These measures were based on a two-octave ascending and descending G major scale.

Howard et al. (23) suggest that an increase in CQ with singing training/experience could be interpreted in terms of a more physically efficient voice usage as follows: (a) the time for which an acoustic path to the lungs via an open glottis exists is reduced, resulting in a reduction in the total acoustic energy transmitted to the essentially anechoic environment of the lungs (known as subglottal damping) where it would be lost to the listener, (b) less stored lung air is vented in each cycle due to the decreased open phase, thereby improving the efficiency of power source usage and enabling notes to be held for a longer time, and (c) the perceived voice quality is less breathy.

The purpose of this paper is to report electrolaryngographically derived CQ measures for adult female singers with varying degrees of singing training/experience. Reference is made to similarities and differences between the male and female groups.

SUBJECTS AND DATA ANALYSIS

Twenty-six adult female singers took part in the experiment (F1–F26). Each subject indicated the extent (in years) of her formal singing training (regular formal voice teaching, which for these subjects was in the Bel Canto tradition) and experience (professional or amateur performance experience as a soloist and/or in choirs). These are shown in Table 1.

The data recorded were as follows: (a) name of subject and date of recording; (b) a read passage lasting ~2 min; (c) the following spoken with a falling intonation: **BOOED, BEAD, BAD, BUD, BED, BIRD, BARD, BOARD, MOON, MEAN, MAN, MUN, MEN, MERN, MARN,** and **MORN**; (d) the utterances used in (c) sung on C (256 Hz), E (330 Hz), C (512 Hz), and E (660 Hz); and (e) a two-octave ascending and descending G major scale from G (196 Hz) on the vowel of *spa* with each note lasting ~⅓ s in which subjects could breath as desired.[3]

The output waveform from an electrolaryngograph (Lx) and the speech pressure waveform (Sp) from a Bruel and Kjaer reference ½-inch condenser microphone were recorded on the two channels of a stereo digital audio tape (DAT) recorder (Sony TCD-D10) in a sound-isolated recording room at University College London. The use of a digital recorder and a reference microphone ensured that the undesired effects of low-frequency phase distortion in the recording process were avoided (27).

Appropriate electrolaryngograph electrode positioning was ensured by means of viewing the Lx waveform on an oscilloscope during the recording. The electrodes were set on the subject's neck for maximum Lx signal amplitude. The oscilloscope display was monitored throughout the recording to ensure that the Lx amplitude remained as appropriate as possible given the wide pitch range employed in this recording. This was of particular importance since some of the untrained subjects made use of large vertical larynx movements in order to sing the higher pitches demanded, and the Lx waveform amplitude was greatly diminished. In these cases, the electrodes were repositioned for the upper octave of the scale to maintain an appropriate Lx waveform amplitude, and the subject was asked to support the electrodes over their larynx by moving the electrodes with the larynx with the thumb and index finger of one hand. (The use of one hand ensures that any relative movement between that hand and the neck causes both electrodes to move together that is preferred to independent electrode movement.) For some of the untrained subjects, it was

EGG less useful for untrained singers at higher pitches due to really bad form that causes unusable Lx outputs

[3] Some of the less experienced singers were unable to sing accurately a two-octave G major scale because of either the extreme of pitch range involved, or difficulties with intonation control or a lack of pitch-matching skills. These subjects were encouraged to produce notes across as wide a range as possible, which in most cases covered the pitch range of a two-octave G major scale. The lack of intonation accuracy in the sung scale was not a drawback for this experiment, since the requirement is for CP measures across a wide F0 range.

TABLE 1. *Singing training/experience of all subjects and third octave analysis CQ data*

CQ mean
(CQ standard deviation)

Subject no.	No. of years training	No. of years experience	Band 1	Band 2	Band 3	Band 4	Band 5	Band 6
F1	30	30	20.29	22.53	25.96	29.99	29.01	31.65
			(2.23)	(1.98)	(2.14)	(3.31)	(4.85)	(3.01)
F2	25	25	26.18	26.24	29.93	30.48	34.97	40.81
			(2.50)	(1.83)	(1.90)	(2.50)	(3.14)	(3.61)
F3	20	20	24.02	25.58	30.08	33.23	35.92	33.00
			(1.22)	(2.14)	(1.66)	(1.58)	(2.52)	(3.27)
F4	15	20	25.20	29.67	31.20	32.40	36.42	34.69
			(2.38)	(2.19)	(3.76)	(2.94)	(2.94)	(3.93)
F5	12	12	27.75	33.36	36.99	45.18	53.55	60.94
			(3.16)	(3.64)	(5.61)	(4.75)	(3.03)	(4.58)
F6	8	10	17.82	21.71	23.82	25.70	29.87	29.26
			(1.67)	(1.95)	(1.70)	(3.63)	(2.66)	(3.63)
F7	8	8	29.67	19.90	20.70	23.37	29.72	38.00
			(4.48)	(4.14)	(2.22)	(3.69)	(2.09)	(3.77)
F8	7	20	30.32	34.97	41.51	49.79	54.28	49.16
			(2.76)	(3.06)	(3.03)	(5.52)	(4.84)	(5.01)
F9	7	17	16.53	21.31	24.67	25.95	28.91	29.23
			(1.99)	(2.89)	(2.18)	(1.95)	(2.50)	(2.65)
F10	5	7	28.90	33.01	39.56	42.34	34.03	35.18
			(4.33)	(5.63)	(3.27)	(6.49)	(3.24)	(5.19)
F11	5	5	28.15	30.08	34.99	41.40	50.12	56.79
			(1.46)	(1.65)	(2.11)	(3.80)	(3.59)	(3.17)
F12	2	3	25.89	27.17	27.04	27.64	31.01	28.36
			(2.84)	(3.16)	(3.03)	(3.17)	(4.32)	(3.82)
F13	0	13	23.91	28.00	31.18	29.28	25.75	25.44
			(2.53)	(1.52)	(1.35)	(3.16)	(2.92)	(4.93)
F14	0	10	29.81	30.07	29.79	30.78	26.21	22.35
			(3.73)	(3.74)	(3.14)	(3.69)	(3.36)	(1.47)
F15	0	8	27.92	27.93	29.71	24.24	25.01	17.00
			(4.26)	(3.23)	(3.10)	(3.09)	(3.85)	(0.0)
F16	0	6	21.95	23.55	24.71	24.35	21.94	27.69
			(3.52)	(3.89)	(4.82)	(5.47)	(5.99)	(6.39)
F17	0	4	27.46	27.93	27.37	25.25	19.69	23.78
			(3.31)	(3.38)	(3.01)	(3.50)	(5.33)	(7.27)
F18	0	3	34.70	37.57	39.25	39.11	32.45	31.27
			(3.24)	(2.24)	(23.21)	(2.64)	(5.19)	(4.64)
F19	0	3	32.36	34.47	30.80	28.72	23.65	17.83
			(3.32)	(3.60)	(5.04)	(3.43)	(4.91)	(4.58)
F20	0	2	29.15	30.63	28.00	28.08	28.96	31.06
			(3.31)	(2.48)	(2.54)	(2.67)	(1.55)	(3.75)
F21	0	2	24.98	27.55	30.33	30.98	29.32	24.67
			(2.95)	(3.51)	(5.01)	(5.20)	(6.28)	(4.34)
F22	0	0	29.33	30.62	31.85	27.61	37.49	54.41
			(7.82)	(9.14)	(10.69)	(7.43)	(9.98)	(7.95)
F23	0	0	26.20	28.94	35.05	42.10	5.00	37.01
			(4.43)	(2.93)	(7.00)	(4.82)	(6.13)	(5.03)
F24	0	0	31.02	36.20	37.27	35.64	31.06	34.50
			(3.29)	(4.32)	(4.68)	(4.56)	(4.10)	(9.19)
F25	0	0	24.43	40.32	41.50	43.51	40.86	38.99
			(13.88)	(4.09)	(3.09)	(2.88)	(4.20)	(3.43)
F26	0	0	27.00	29.22	30.65	27.88	22.73	30.40
			(3.81)	(4.22)	(5.68)	(6.61)	(5.61)	(8.24)
	Wilcoxin	T1	110	119	99	69	39[a]	39[a]
	Wilcoxin	T2	58	49[a]	69	99	129	129

F1 to F12: trained subjects; F13 to F26: untrained subjects.
Data are in hertz [mean(SD)].
[a] Significant *T* values.

not possible to obtain a usable Lx recording towards the upper end of the pitch range. They are identified below.

Trained subjects, on the other hand, tend to keep their larynxes lowered throughout the sung range and for them it was particularly important to observe the oscilloscope and set an electrode position appropriate for singing rather than speaking. Some trained subjects felt too constricted using the elastic neck band to support the electrodes, and they were instructed to support the electrodes with the thumb and forefinger of one hand. (This was found to be a particular problem during the sung scale, which made use of the widest pitch range.)

There was no direct monitoring of sound pressure level during the recordings. Subjects were instructed to sing at a level that they found to be comfortable for the task in hand.

Data analysis was carried out with use of a real-time CQ analysis system implemented by Howard and Garner (28), which makes use of an Ariel PC-56D card in a PC-compatible computer. Data are captured to 14-bit accuracy at a sampling rate of 20 kHz. CQ analysis is carried out by the Motorola 56001 digital signal processor with use of method (a) of Davies et al. (21) and linear quantisation between samples either side of the 3:7 threshold to define the end of the CP. Results are sent to the host PC for storage and plotting. An example screen plot of log(F0) against time and CQ(%) against time is shown in Fig. 3 for an ascending two-octave scale for a trained subject (F11). Plots such as these, but with expanded time axes, were used to check the integrity of Lx data in the upper octave.

RESULTS

The data for all subjects were plotted as scattergrams (Qx) of CQ(%) against F0 (log Hz). Figures 4 and 5 show example Qx plots for a trained (F5) and an untrained (F19) subject, respectively. All the Qx plots exhibited an overall pattern of variation with an essentially linear variation of CQ with log(F0)

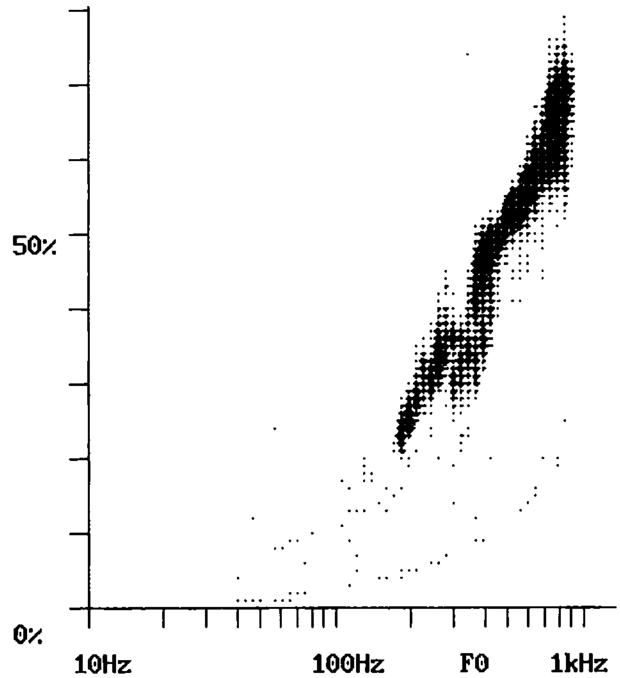


FIG. 4. Scattergram of closed quotient (%) against log(F0) for a two-octave ascending G major scale sung by a trained subject (F5).

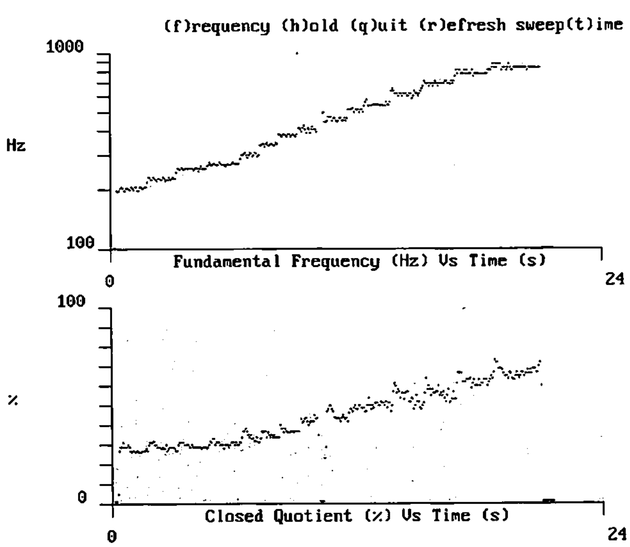


FIG. 3. A plot of a two-octave ascending G major scale sung by a trained subject (F11): log(F0) against time (upper), and closed quotient measured by method (a) of Davies et al. (21) against time (lower).

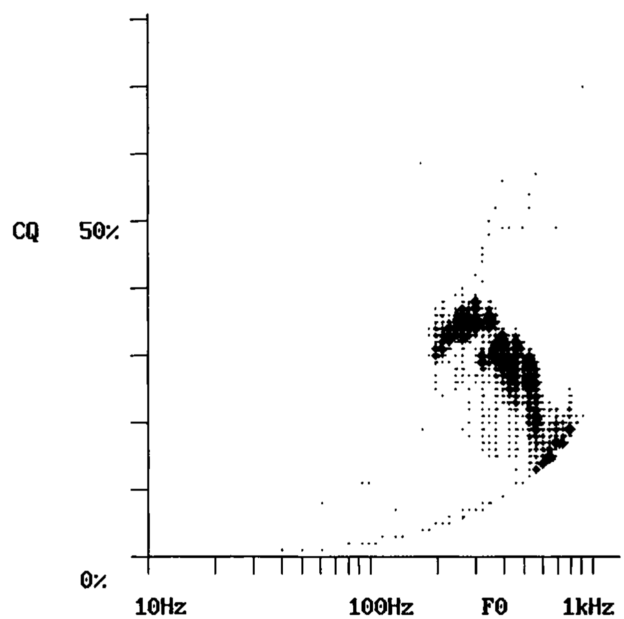


FIG. 5. Scattergram of closed quotient (CQ, %) against log(F0) for a two-octave ascending G major scale sung by an untrained subject (F19).

with two F0 sub ranges approximately either side of G4. In order to enable a statistical investigation of differences between the trained and the untrained groups, two strategies were adopted: (a) a third octave analysis of CQ mean values, and (b) a piece-wise linear CQ:F0 trend analysis within a high and low F0 sub range. These are described later.

Third octave band analysis

The CQ data for each subject was divided into six third octave bands. Third octave bands were chosen given their perceptual salience. The six bands were delimited by the following frequency bounds: 196, 247, 311, 392, 494, 622, and 784 Hz. A program was written in C to calculate the mean and standard deviation CQ values in each band, which are recorded in Table 1.

A one-tailed Wilcoxon's rank sum test was carried out for each band between the trained and untrained groups. In Bands 5 and 6 the trained group have statistically significantly higher mean CQ values than the untrained group ($p < 0.025$ in both cases). In Band 2, the untrained group have statistically significantly higher mean CQ values than the trained group ($p < 0.05$), a trend that is further reflected in Band 1 but without statistical significance. This relationship is summarised graphically in Fig. 6 where the mean of CQ means is plotted in each band for both groups.

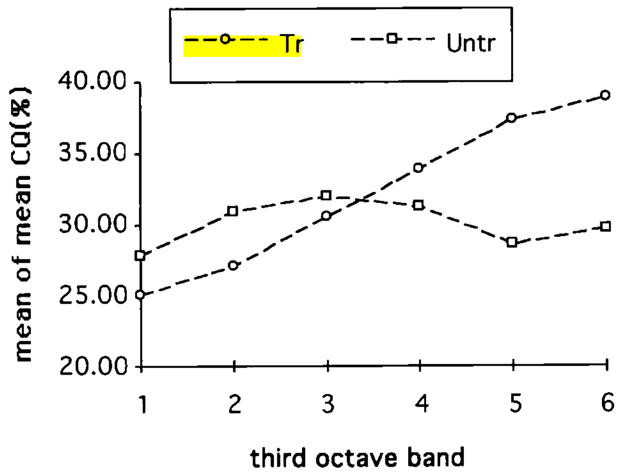
CQ:F0 trend analysis

The Qx scattergrams for all subjects were studied, and a piece-wise linear trend analysis was carried out in the low octave (LO) and high octave (HI) as follows:

(a) The F0 values were noted at which turning points occurred, and the mean and log standard deviation values were calculated. The mean was 404 Hz (approximately G4), the lower standard deviation bound was 354 Hz (approximately F4), and the upper was 492 Hz (approximately B4).

(b) The CQ:F0 data were split for each subject into a LO and HI piece-wise linear F0 range, LO: 174–354 Hz; and HI, 492–880 Hz. Where necessary to maintain a piece-wise linear relationship, these bounds were fine-tuned on an individual subject basis.

(c) Linear regression lines in the HI and LO



CQ means for third octave - useful for groundtruth

FIG. 6. Mean of closed quotient (CQ) means in each third octave band for the trained (Tr) and the untrained (Untr) groups. (Statistically significant differences exist between the groups for Bands 2, 5, and 6.)

ranges were calculated for each subject using SPSS. The gradient {b = CQ/log(F0)}, and test statistic {$t = b$/(standard error of b), $p < 0.00005$ in all cases} were recorded.

(d) The HI data for three subjects whose Lx data were of too low an amplitude for reliable CQ:F0 analysis were rejected (see Table 2).

The gradient values are recorded in Table 2. That the gradients associated with the trained group are higher than those of the untrained group with statistical significance is shown by a one-tailed Wilcoxon's rank sum test: LO range ($p = 0.025$), and HI range ($p < 0.025$). Spearman's rank correlation coefficient was also calculated for each range for the subject group taken as a whole in order to assess whether greater years of singing training/experience correlates with CQ/F0 gradient. Subjects were first ranked by number of years training and further subdivided by the number of years experience. A statistically significant trend resulted: LO F0 range ($\rho = 0.40551$, $n = 26$, $p < 0.05$), and HI F0 ($\rho = 0.33515$, $n = 23$, $p < 0.05$).

DISCUSSION AND CONCLUSIONS

These results for women can be summarised within the two-octave pitch range considered as follows: (a) CQ tends to be reduced for pitches below D4 and increased for pitches higher than B4 with training; and (b) the CQ/F0 gradient within the pitch

TABLE 2. *Singing training/experience of all subjects and high-octave (HI) and low-octave (LO) band CQ/log (F0) gradient analyses*

could be useful for groundtruth, but I don't really understand the gradient

Subject no.	No. of years training	No. of years experience	LO gradient CQ/log(F0)	HI gradient CQ/log(F0)
F1	30	30	25.08	91.44
F2	25	25	13.84	71.32
F3	20	20	38.51	13.08
F4	15	20	30.76	7.36
F5	12	12	39.72	69.04
F6	8	10	29.10	−29.27
F7	8	8	16.53	73.50
F8	7	20	55.21	−35.66
F9	7	17	45.36	7.27
F10	5	7	51.56	−14.96
F11	5	5	26.33	76.57
F12	2	3	10.15	24.17
F13	0	13	37.15	−30.15
F14	0	10	5.58	−27.27
F15	0	8	12.16	−51.33
F16	0	6	17.48	Lx poor
F17	0	4	7.435	Lx poor
F18	0	3	22.39	−49.59
F19	0	3	5.58	−47.68
F20	0	2	−0.23	13.58
F21	0	2	29.8	−31.89
F22	0	0	23.30	Lx poor
F23	0	0	30.15	−9.44
F24	0	0	34.26	−28.83
F25	0	0	29.16	−8.17
F26	0	0	24.38	−6.45
Wilcoxin		T1	45[a]	24[a]
Wilcoxin		T2	123	108

CQ, closed quotient; Lx, electrolaryngograph output waveform.
[a] Statistically significant *T* values.

ranges: G3 to G4, and B4 to G5 tends to correlate positively with the number of years singing training/experience.

In an equivalent investigation for adult men (n = 18), Howard et al. (23) noted a significant difference between the mean CQ values for their trained and untrained groups (two-tailed Mann-Whitney U test, p = 0.001), and that CQ appeared to vary directly with log(F0) as a function of singing training/experience. Spearman's rank correlation coefficient for their male data ($\rho = 0.7234$, $n = 18$, $p < 0.005$) suggests that greater years of singing training/experience correlates with higher CQ values. For men, the following conclusions therefore suggest themselves: (a) CQ remains essentially constant with F0, and (b) CQ means tend to correlate positively with the number of years singing training/experience.

The wide F0 range demanded of the singing voice as compared with the speaking voice results in the vocal folds adopting different modes of vibration to enable them to operate over a wide phonation frequency range. These modes of vibration are related to the notion of voice *registers,* for which "the terminology . . . is very confusing" (29). The notion of the voice registers is supported, for example, by what is heard when an untrained singer sings a series of notes progressing upward or downward in pitch over a wide F0 range. There will be groups of adjacent notes that are heard as having very similar phonatory qualities, and abrupt changes may be heard between them at what are termed the "voice transition" or register "break" points. There is often a range of pitches for which registers overlap where singers can sing the same note using different registers.

A number of female subjects exhibit turning points in their Qx plots (mean = 404 Hz, or approximately G4). This could be indicative of a chest-middle register break point. Sundberg (29, p. 51) gives the following values "in the neighbourhood of which" the ranges of register overlap occur for women: chest-middle, G (400 Hz); and middle-head, E (660 Hz). He further states that "These ranges of register overlap, and the register boundaries vary substantially among individuals." One goal of singing teaching is to make the breaks between vocal registers perceptually inaudible to the listener, and more experienced singers learn to "cover" or "darken" the tone at or around their break points and also to modify their vowel sounds. To darken the tone, the volume of the pharynx is maintained by keeping the larynx in a lowered position (29,30).

By way of a summary, Fig. 7 shows idealised Qx plot shape tendencies based on a two-octave adult female and male data as a function of the number of years singing training/experience. The CQ data for the men increase with movement along the singing ability continuum, remaining essentially flat with F0. The Qx shapes for the women increase in gradient with singing training/experience. All the gradients tend to be positive in the lower octave and those for the untrained tend to be negative in the higher octave, and the turning point occurs around G4.

Although these data suggest that there are differences in electrolaryngographically derived CQ with F0 in singing between subjects with varying degrees of singing training/experience, they cannot be a basis for establishing whether these differences occur as a direct result of singing training/experience. Such confirmation will be gained only with longitu-

groundtruth concern - a followup paper that more confidently defines this relationship and studies CQ of singers *being* trained over time will be important for justifying our groundtruth data good inspiration for visualizing CQ improvement over time

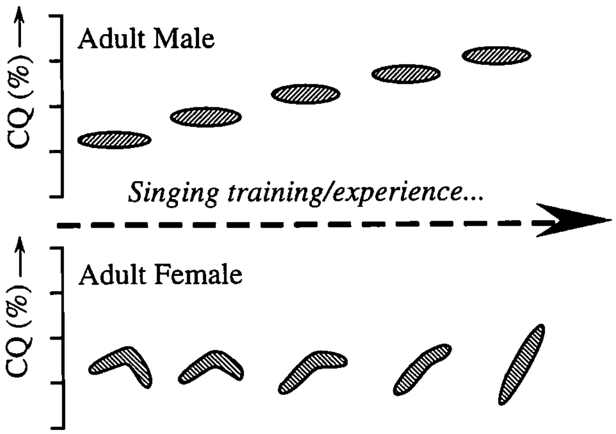


FIG. 7. Suggested idealised changes in closed quotient (CQ): log(F0) scattergram shapes with singing training/experience for: men (upper), and women (lower).

dinal observations of subjects whose voices are being trained (31). The suggestion that particular trends in these data may relate to voice "breaks" requires confirmation based on perceptual evaluation of the data by trained singers. An informal listening by eight singing teachers to the scales sung by five of these subjects produced little consensus. The listening panel suggested that two-octave scales were not the most appropriate exercises to use. Further work in this area is required.

One potential application of knowledge of CQ/F0 variation with singing training is as a real-time visual display for use in voice teaching. In addition, such a display could be used to develop stylistic techniques since the variation of CQ with F0 has been shown to change with singing style (32). It should be noted that there are other factors which cause changes in CQ, for example, changes in loudness, and the use of pressed phonation which is generally associated with a long larynx closed phase in each cycle (33), which must be taken into account. Such a display would, for example, complement other displays such as the SINGAD (SINGing Assessment and Development) system (34,35), which has demonstrated its usefulness as a real-time visual display in the classroom for the development of conscious pitching ability for 5- to 7-year-old children.

Changes in CQ with F0 characterise just one aspect of the professionally trained singing voice, but there are a number of other features that are equally, if not more important. Real-time visual displays that quantify parameters of the trained professional singing voice and display their characteristics in a meaningful manner could be used by students to improve the efficiency of their practice. Their teachers will then be able to devote more time to the more qualitative aspects of the professional singer's training, such as stagecraft, communication with the audience, the development of musicality, and working with conductors, orchestras, and keyboard accompanists, where experience and musical taste are the final arbiters.

Acknowledgment: The author would like to thank all the singers for their time and willingness to be recorded for these experiments, the reviewers for their time and most helpful comments, and Paul Garner for software help. This work was supported by Science and Engineering Research Council research grant number GR/H14878.

REFERENCES

1. Fourcin AJ, Abberton ERM. First applications of a new laryngograph. *Medical and Biological Illustrated* 1971;21:172–82.
2. Fourcin AJ. Laryngographic assessment of vocal fold vibration. In: Wyke B, ed. *Ventilatory and phonatory control systems*. Oxford: Oxford University Press, 1974.
3. Baer T, Löfquist A, McGarr NS. Laryngeal vibrations: a comparison between high-speed filming and glottographic techniques. *J Acoust Soc Am* 1983;73:1304–8.
4. Gilbert HR, Potter CR, Hoodin R. The laryngograph as a measure of vocal fold contact area. *J Speech Hear Res* 1984;27:178–82.
5. Noscoe NJ, Fourcin AJ, Brown MA, Berry RJ. Examination of vocal fold movement by ultra-short pulse X radiography. *Br J Radiol* 1983;56:641–5.
6. Fourcin AJ. Electrolaryngographic assessment of phonatory function. *Journal of Phonetics* 1987;14:435–42.
7. Titze IR. Parameterization of the glottal area, glottal flow, and vocal fold contact area. *J Acoust Soc Am* 1984;75:570–80.
8. Titze IR, Baer T, Cooper D, Scherer R. Automatic extraction of glottographic waveform parameters and regression to acoustic and physiologic variables. In: Bless DM, Abbs JH, eds. *Vocal fold physiology: contemporary research clinical issues*. San Diego: College Hall, 1984:146–54.
9. Childers DG, Hicks DM, Moore GP, Alsaka YA. A model for vocal fold vibratory motion, contact area, and the electrolaryngogram. *J Acoust Soc Am* 1986;80:1309–20.
10. Baken RJ. *Clinical measurements of speech and voice*. London: Taylor and Francis, 1987.
11. Childers DG, Krishnamurthy AK. A critical review of electroglottography. *Crit Rev Biomed Eng* 1985;12:131–61.
12. Abberton ERM, Howard DM, Fourcin AJ. Laryngographic assessment of normal voice: a tutorial. *Clinical Linguistics and Phonetics*. 1989;3:281–96.
13. Kitzing P. Clinical applications of electroglottography. *Journal of Voice* 1990;4:238–49.
14. Hirson A, Fawcus R. Visual feedback in the management of dysphonia. In: Fawcus M, ed. *Voice disorders and their management*. 2nd ed. London: Chapman Hall, 1991.
15. Fourcin AJ, Abberton ERM. The laryngograph and the Voiscope in speech therapy. In: Loebell E, ed. *Proceedings of the 16th International Congress of Logopaedics and Phoniatrics*. Basel: Karger, 1976:116–22.
16. Abberton ERM, Fourcin AJ, Rosen S, et al. Speech perceptual and productive rehabilitation in electro-cochlear stimu-

lation. In: Schindler RA, Merzenich MM, eds. *Cochlear implants*. New York: Raven Press, 1985:527–37.
17. Hess W, Indefrey H. Accurate pitch determination of speech by means of a laryngograph. *Proceedings of the International Conference on Acoustics Speech and Signal Processing, ICASSP-84*. Piscataway: 1984;1:1–4.
18. Howard DM. Peak-picking fundamental period estimation for hearing prostheses. *J Acoust Soc Am* 1989;86:902–10.
19. Krishnamurthy AK, Childers DG. Two-channel speech analysis. *IEEE Transactions on Acoustics Speech and Signal Processing*. 1986;ASSP-34:730–43.
20. Howard DM, Brookes DMB, Chan DSF. Dynamic excitation control in parallel formant speech synthesis. *Proceedings of the 7th FASE Symposium, Edinburgh*. 1988;3:1123–31.
21. Davies P, Lindsey GA, Fuller H, Fourcin AJ. Variation in glottal open and closed phase for speakers of English. *Proceedings of the Institute of Acoustics*. 1986;8:539–46.
22. Orlikoff RF. Assessment of the dynamics of vocal fold contact from the electroglottogram: data from normal male subjects. *J Speech Hear Res* 1991;34:1066–72.
23. Howard DM, Lindsey GA, Allen B. Towards the quantification of vocal efficiency. *J Voice* 1990;4:205–12. (See also Errata. *J Voice* 1991;5:93–5.)
24. Kitzing P. Simultaneous photo- and electroglottographic measurements of voice strain. In: Titze IR, Scherer RC, eds. *Vocal fold physiology*. Denver: The Denver Center for the Performing Arts 1983;221–9.
25. Childers DG, Hicks DM, Moore GP, Eskenazi L, Lalwani A. Electroglottography and vocal fold physiology. *J Speech Hear Res* 1990;33:245–54.
26. Howard DM, Lindsey GA. New laryngograms of the singing voice. *Proceedings of the 11th International Congress of Phonetic Sciences, USSR: Tallinn*. 1987;5:166–9.
27. Holmes JN. Low frequency phase distortion of speech recordings. *J Acoust Soc Am* 1975;50:747–9.
28. Howard DM, Garner P. DSP-56000 based real-time electrolaryngographically derived closed quotient. *Proceedings of the Institute of Acoustics* 1992;14:375–82.
29. Sundberg J. *The science of the singing voice*. Dekalb: Northern Illinois University Press, 1987.
30. Bunch MA, Sonninen A. Some further observations on covered and open voice qualities. *National Association of Teachers of Singing Bulletin* 1977;34:26–30.
31. Howard DM, Rossiter D. Results from a pilot longitudinal study of electrolaryngographically derived closed quotient for adult male singers in training. *Proceedings of the Institute of Acoustics* 1992;14:529–36.
32. Howard DM. Quantifiable aspects of different singing styles—a case study. *J Voice* 1992;1:47–62.
33. Sundberg J, Gauffin 00. Waveform and spectrum of the glottal voice source. In: Lindblom B, Öhman S, eds. *Frontiers of speech communication research*. London: Academic Press, 1979:301–20.
34. Howard DM, Welch GF. Microcomputer-based singing ability assessment and development. *Applied Acoustics* 1989;27:89–102.
35. Welch GF, Howard DM, Rush C. Real-time visual feedback in the development of vocal pitch accuracy in singing. *Psychology of Music* 1989;17:146–57.