NEW LARYNGOGRAMS OF THE SINGING VOICE

DAVID M HOWARD GEOFFREY A LINDSEY

Department of Phonetics and Linguistics
University College London
Great Britain

ABSTRACT

Laryngographic techniques evolved in speech analysis are extended in the present work to the analysis of the singing voice. Attention is focused on laryngographically derived measures of vocal fold open-phase times. Specifically, the measure of open quotient (open phase time over whole period time) provides a quantitative parameter for the characterisation of voice production differences between speakers, trained singers and untrained singers.

INTRODUCTION

The laryngograph [1] has been used for many years as a tool for the analysis of normal and pathological speech as well as of the singing voice. More recently, detailed studies have investigated changes in the laryngograph output waveform (Lx) on a cycle-by-cycle basis, with a view to correlating these with the acoustic output from the vocal tract [2]. The present work is designed to develop a new series of laryngographically based plots which can be utilised on a routine basis for speaking and singing voice research and, potentially, in visual displays to be developed to give feedback in singing/speaking voice production training.

DATA, SUBJECTS, AND RESULTS

Four adult male singers took part in the experiment. Two are 'trained', having had formal voice training and extensive solo performance experience; one of these is a baritone (GW) and the other a tenor (SB). The other two have choral singing experience and vocal ranges in the middle (baritonal) range for men; one of these (DH) is an experienced amateur musician who has received vocal training for a short period, and the other (GL) is untrained. GL made recordings in both a 'natural,' informal style and in a quasi-trained style which he adopts for choral performance.

The subjects were digitally (PCM) recorded onto videotape in the anechoic room at UCL, with the output from a high quality condenser microphone (Sp) on one channel and the laryngograph output waveform (Lx) on the other.

The data consisted of:
1) a reading of a phonetically balanced passage lasting approximately two minutes;
2) five monophthongal vowels, in the environments /b_d/ and /m_n/, spoken with falling intonation and then sung on C (256Hz), C (128Hz), E (330Hz) and E (165Hz);
3) major scales exhibiting each singer's range, sung on the vowel /a/; and
4) a performance of "God Save the Queen" starting on G (192Hz).

The analyses consist of the following (see Figs. 1-6):
a) the speech pressure waveform (Sp);
b) the laryngograph output waveform (Lx), derived by measuring the current passing through the throat between two voltage driven electrodes placed on the wings of the thyroid cartilage; the peaks thus correspond to the maxima of vocal fold closure and the valleys to maxima of glottal opening, in each cycle [1];
c) vertical period markers derived from Lx (Tx) [3];
d) a plot of the logarithm of fundamental frequency (derived on a period by period basis from Tx) against time (Fx); and
e) two plots of the open quotient (OQ), which is defined as the duration of the open phase of each cycle divided by the duration of the whole of that cycle, calculated from Lx by different methods -- the first method (OQ1) takes the point of vocal fold closure in a glottal cycle to be the peak of the differentiated Lx waveform, and the point of opening to be the minimum in the differentiated Lx waveform; whilst the second (OQ2) takes the upper 70% of the peak-to-peak amplitude of each cycle to represent closure of the glottis and the lower 30% to represent an open glottis (see [3] for a full description)

Figs. 2 to 6 show these analyses plotted for the first note of the final occurrence of the word "God" (E 330Hz) in the sung performances of "God Save the Queen" by each of the subjects. Fig. 1 shows equivalent plots for the vowel [a] spoken by a female subject with high fall intonation. In each figure pertaining to singing, the OQ plots derived by both methods remain relatively steady; mean values for the portion shown are tabulated in Fig. 7. This table also gives summary statistics relating to the speaking voice of each subject based on the reading of the passage (section 1 of the recording) which have been derived from a second order fundamental frequency distribution plot [4]. This plot consists of a histogram of the number of consecutive pairs of Tx period values which fall within the same histogram 'bin'. The summarised statistics show the number of such pairs in the plot for that subject (under the heading "samples"), the Fx range at the 0.1% probability level, and the modal Fx value. Mean values of OQ calcutated by the two methods are also given in the table.

DISCUSSION AND CONCLUSIONS

In spoken data, evidence has been found [2] for a 'preferred' value of Fx towards the lower end of a speaker's overall speaking range, at which the vocal folds vibrate with optimal efficiency and vigour. At this value of Fx, open and closed phases are approximately equal and the peak-to-peak amplitude of the Lx waveform tends to be at its maximum. Independent evidence that this is indeed a preferred frequency of vocal fold vibration is found in Fx distributions for stretches of continuous speech which have their modes near this Fx value. It is plausible to suggest that this Fx value is utilised by the speaker as a departure/arrival point for prosodic excursions and as a locus for neutral, non-pitch prominent syllables. Fig. 1 illustrates this for the speech of a woman (EA), whose Fx range is measured as 118Hz to 371Hz and whose Fx mode is 147Hz (Fig. 7); this speaker's OQ1 and OQ2 values pass through the 50% level around the point at which her Fx contour passes through 147Hz.

It has also been found [3] that for many speakers closed phase duration varies considerably more with Fx than does open phase duration. That is, the open phase remains comparatively constant while the closed phase is shortened as Fx rises and lengthened as Fx falls. Thus the closed phase is roughly the inverse of Fx, while the open phase is rather steadier. This can be seen in the plots for our speaker (Fig. 1) where OQ values, calculated by both methods, tend to fall with the falling Fx.

The majority of a singer's pitch range makes use of Fx values which are high relative to the normal speaking pitch range. In view of the preceding observations on speech, one might expect raised OQ values at sung pitches at the higher end of the range. This does not appear to be the case with our trained singers. Figs. 2-6 show OQ values for the note E (330Hz), which is well above the upper limit of the measured Fx speaking range for each subject: Fig. 7 gives a summary of their speaking ranges and modes. The singers are ordered by experience, and our speaker (EA) is included as the final table entry for comparison. This ordering of subjects corresponds closely to the trends in OQ1 and OQ2 values for the singers.

Subjects GW and SB are professionally trained with many years solo performance experience. GW is also a singing teacher. In both cases the mean OQ values, calculated by both methods, are markedly lower (30% to 38%) than the values found for the other subjects (49% to 75%). Subject DH, who has had some singing training, and extensive choral experience at an amateur level has mean OQ values close to 52%. Subject GL, who has had no formal singing training, recorded productions in two maners, an informal style (GL(U)) and a quasi-trained style (GL(T)), and it is clear that he attains more appropriate OQ values when he adopts his choral style. The mean OQ values given for our speaker (EA) must be interpreted wih reference to Fig. 1, since it is clear that the OQ values vary over a wide range with Fx change: her OQ values are above 50% during the first portion of the utterance and they descend below 50% towards the end as her Fx is lowered.

These data suggest that there is a clear trend towards lowered OQ values with increased singing training and performance experience. This is of note because speech data suggest that raised

OQ values correlate with raised Fx, and therefore lowered OQ in singing must presumably be a direct effect of a trained style of singing. This implies that an important mechanism involved in note productions by a trained singer is the considerable lengthening of the closed phase in each larynx cycle with respect to the open phase. This has two main consequences: firstly the voice quality becomes less breathy; and secondly the longer closed phase ensures more prolonged substantial acoustic output from the vocal tract as the coupling-in of the subglottal cavities (and the associated increase in acoustic damping) occupies less of the cycle. Thus the singer makes use of a natural acoustic consequence of an action which presumably requires no additional pulmonic energy, to achieve this increase in output.

Future work in this area will include a wider range of subjects with various levels of singing training and experience in order to evaluate the robustness of these measures. The possibility also exists of a new form of visual display to aid the singer, which could complement the singing assessment and development (SINGAD) system currently aimed at note pitching by children from five years upwards [5]. This system makes use of a microcomputer and a specially developed hardware interface, based on [6], which estimates fundamental frequency from an acoustic input; it also allows work on vibrato, note onset and offset, and pitch systems of different musical traditions. A new OQ component along the lines discussed above, would make a significant contribution towards establishing a comprehensive and coherent tool for students of singing.

ACKNOWLEDGEMENTS

The authors would like to thank Graham Welch and Simon Bainbridge for their freely given time in making recordings, Peter Davies for his OQ analysis package and his help with software modifications, and David Smith for the Tx processing system.

REFERENCES

[1] Fourcin, A.J., and Abberton, E.R.M. (1971). "First applications of a new laryngograph", Medical and Biological Review, 21, 172-182.

[2] Lindsey, G., Davies, P., and Fourcin, A. (1986). "Laryngeal coarticulation effects in English VCV sequences", IEE Conference Proceedings 258, 99-103.

[3] Davies, P., Lindsey, G., Fuller, H., and Fourcin, A.J. (1986). "Variation in glottal open and closed phase for speakers of English", Proceedings of the Institute of Acoustics, 8, 539-546.

[4] Fourcin, A.J. (1981). "Laryngographic assessment of phonatory function", The American Speech Language Hearing Association (ASHA) Reports, 11, 116-127.

[5] Howard, D.M., Welch, G.F., Gibbon, R.R., and Bootle, C.M. (1987). "The assessment and development of singing ability - initial results with a new system", Proceedings of the Institute of Acoustics, 9, in press.

[6] Howard, D.M. and Fourcin, A.J. (1983). "Instantaneous voice period estimation for cochlear stimulation", Electronics Letters, 19, 76-78.

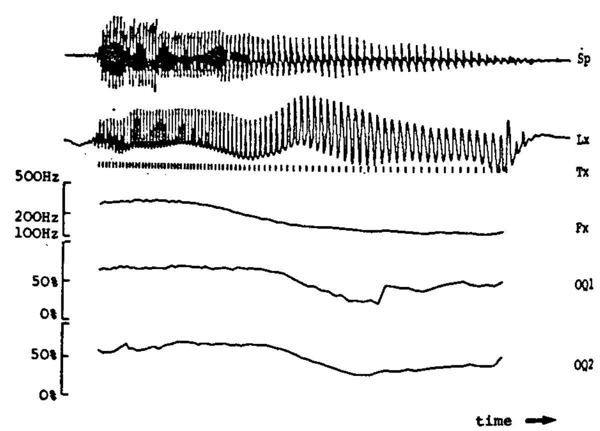

FIGURE 1: Subject EA spoken utterance [a] on a low fall

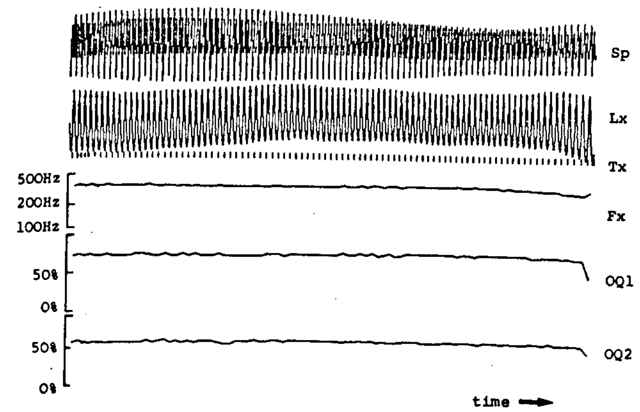

FIGURE 2: Subject GL(U) sung note: (GL in 'natural' style)

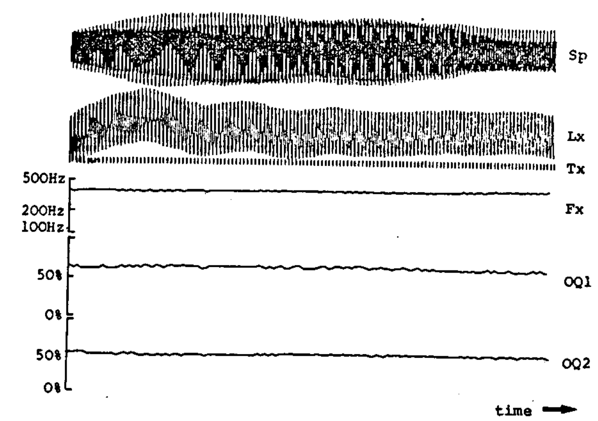

FIGURE 3: Subject GL(T) sung note: (GL in quasi-trained style)

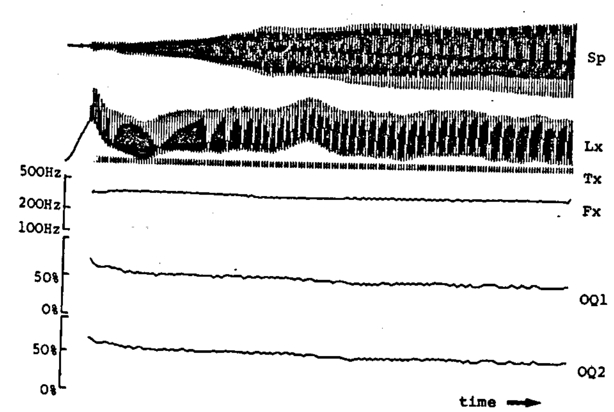

FIGURE 4: Subject DH sung note: (slightly trained baritone)

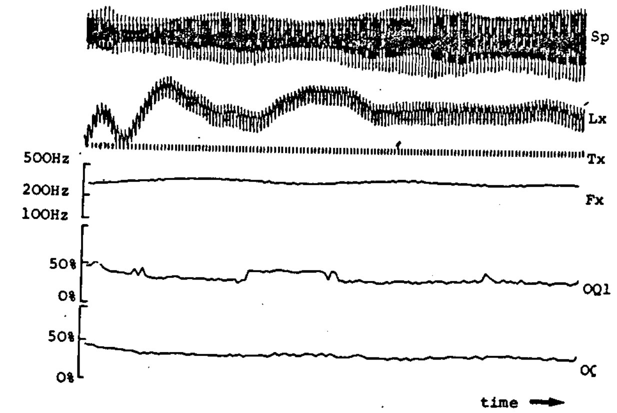

FIGURE 5: Subject SB sung note: (trained tenor)

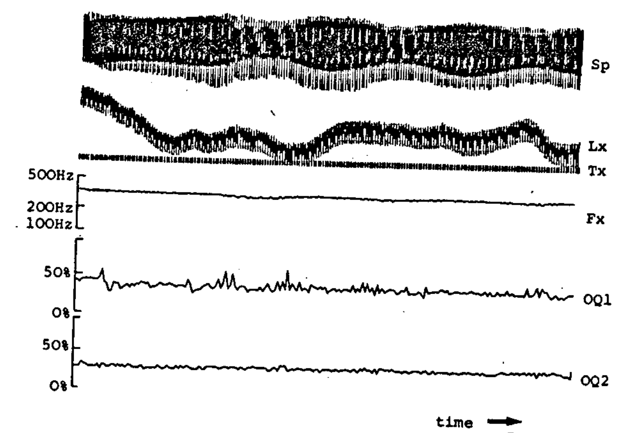

FIGURE 6: Subject GW sung note: (trained baritone)

	2nd order Fx distribution statistics - read passage			Open Quotient mean values	
Subject	samples	0.1% range	Mode	OQ1	OQ2
GW	2529	89Hz-312Hz	149Hz	37.5%	38.1%
SB	2539	104Hz-280Hz	123Hz	36.9%	35.5%
DH	2694	104Hz-280Hz	162Hz	52.1%	52.5%
GL(T)	3181	82Hz-312Hz	116Hz	64.6%	49.8%
GL(U)	3181	82Hz-312Hz	116Hz	74.6%	60.3%
EA	2758	118Hz-371Hz	147Hz	60.8%	59.4%

NOTES: OQ1 measured by differentiated Lx method
OQ2 measured by 70% to 30% Lx amplitude method
OQ measures for utterances in figures 2-6

FIGURE 7: Table of summary speech Fx statistics and mean open quotient values for all subjects