Closing and Opening Phase Variability in Dysphonia

Adrian Fourcin(1) and Martin Ptok(2)

(1) University College London
4 Stephenson Way, LONDON NW1 2PE, UK
(2) Medizinische Hochschule Hannover
C.-Neuberg-Str.1 D 30625 Hannover, Germany

(1) a.fourcin@ucl.ac.uk
(2) Ptok.Martin@MH-Hannover.de

Four examples of the use of vocal fold contact phase measurement are discussed for unilateral paresis. In each case this aspect of voice quality is of greater importance than the physical measurement of loudness and pitch related parameters. For three of the cases electro-stimulation has been used as a main part of the treatment. Phonation in both connected speech and, for comparison, in sustained sound production has been used with electro-laryngograph / egg signals providing the basis for measurement. The main new descriptors that have been found to be useful relate to: vocal fold closure and closure duration regularities and distributions; but reference is also made to related measures of peak acoustic amplitude. The new measures described give, in some cases, quite striking results that are of auditory significance and potentially of clinical value.

key words: voice quality; vocal fold closure; regularity; paresis; electro-laryngeal stimulation

Introduction

Moderate unilateral paresis arises from asymmetry of neural control and this can be associated with relatively small differences in vocal fold stiffness, mass and form. In unilateral paralysis these effects will inevitably influence the time course of vocal fold contact during both the opening and the closing phases. For moderate unilateral paresis, however, it is possible that, since the vibratory cycle is reset for each vocal fold closure and closure takes place very rapidly, the closure epochs may occur fairly regularly. The opening phases may, on the other hand, be markedly more irregular since the mucosal wave associated with opening will vary from cycle to cycle as a result of the physical variability of the affected fold and the slower time course of opening can make its movement more susceptible to asymmetry. It could, in consequence, be supposed that although the pitch of the voice with moderate unilateral paresis may be well defined, those aspects of its quality that depend on closed phase definition and regularity will be impaired.

The first of these possibilities has been discussed and explored with clinical voice material using sustained sounds [1,2]. The use of connected speech, however, with much larger data samples for each speaker gives a basis for the use of analyses that are both more reliable and also more relevant to problems of actual voice use in family, work and professional situations.

***Normal Voice** – frequency analysis*
based on the use of connected speech

The rapidly changing resonances and anti-resonances of the vocal tract during normal speech make it difficult to obtain reliable estimates of voice fundamental frequency in ordinary connected speech from the acoustic signal. The difficulty is greater for pathological voice data where averaging techniques can obscure important effects. This problem can be resolved to an important degree by the use of the electro-laryngograph output signal, Lx. Figure 1 shows the result of defining vocal fold contact epochs using the positive going closures indicated by Lx to obtain individual cycle by cycle values of the time separation between successive vf closures, Tx.

The auditory mechanisms that give a sensation of voice pitch depend to an important extent on temporal processing. The representation of Tx intervals at the 1μs level gives a basis for pitch-related calculations that are flexible in display and that can rival the ear's acuity in the relation of frequency, Fx, to pitch.

The percept of pitch depends on an auditory inference of regularity of acoustic stimulation. At the very least, two periods must occur before a pitch percept can be established. Figure 2 shows the results of analysing a three minute sample of voice using Tx samples for every vf closure and for only those cases where three successive closure epochs have defined essentially the same x values. For a good voice there is little difference. This approach is extended in the analysis of Figure 3 where successive vf cycles are plotted against each other. Once more, there is little deviation from the ideal straight line distribution for the good normal voice and the basis for a measure of irregularity is provided when the range of Fx used in contrastive intonation is taken into account by the use of half tone analysis bins.

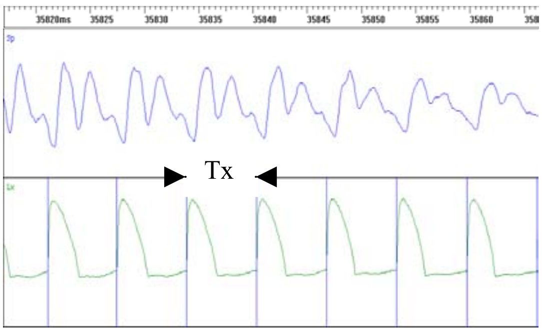

figure 1 The acoustic signal, Sp, above is synchronous with the laryngograph/egg signal, Lx, below; automatically generated markers indicate the closed phase onsets used for measurement.

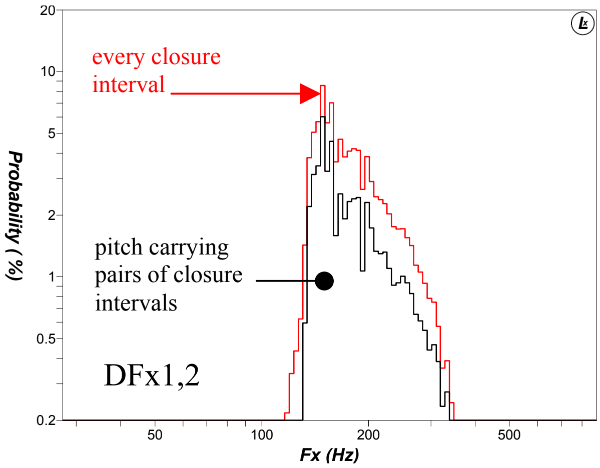

figure 2 DFx1 & 2 – The outer, red, voice frequency probability distribution shows the range of closure intervals for the whole of a 3m read text. The inner, black, distribution shows only the closure interval pairs that fall into the same quarter tone analysis bin.

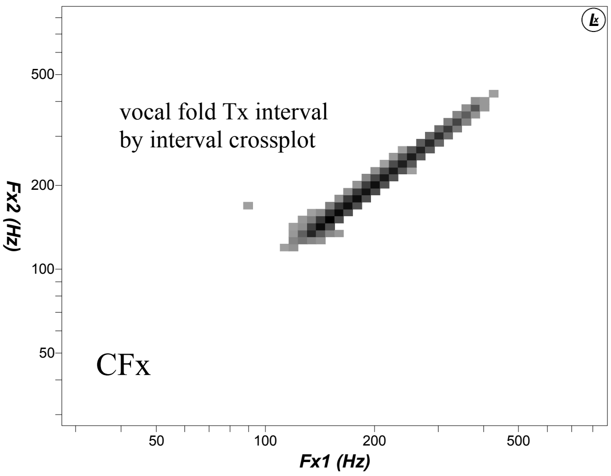

figure 3 CFx – closure interval pairs shown as a crossplot using half tone bins.

***Normal Voice** – closed phase analysis based on the use of connected speech*

Whatever methods of observation and measurement are used, there is inevitably a degree of uncertainty associated with the practical definition of the interval of time between the closure and opening of the vocal folds during phonation. No current technique is perfect and acceptable for routine clinical use. Figure 4 illustrates the simple approach used here in order to obtain relatively robust data in clinical work for the evaluation of the ratio of the interval of time in each vocal fold cycle, between closure and opening of the vf, and the total time for the whole cycle. This particular closed phase ratio is referred to here as Qx, and is based on the same approach to the use of individual vocal fold cycles as was use above for figures 2 and 3.

Figure 5 shows for the normal voice sample of figures 2 & 3, the results of analysing the data as a function of Qx values obtained when every vf cycle is examined individually (the red plot) and also for only those occasions when two successive cycles have fallen into the same pitch analysis bin. Once more, for the normal voice the two distributions are in close remblance and, as before, the ratio of the total number of vf cycles associated with each gives an indication of coherence of control of Qx.

Figure 6 extends the means of showing coherence of control of the closed phase ratio by using the same crossplot technique as was used for figure 3. At present the bin sizes used for the analyses are not defined with the same rigour as was possible for Fx since the control and perception of closed phase effects in running speech is not well established. The parameters chosen are, however, the result of simple clinical exposure and provide a useful basis for cross comparison between different data sets. CQx proves to be valuable in work with paresis.

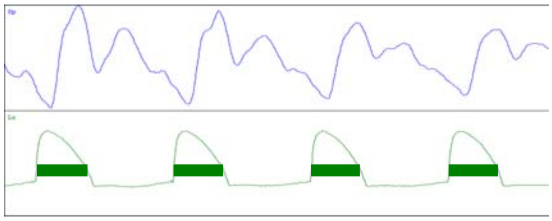

figure 4 Contact phase measurement on the Lx waveform at 70% of the peak to peak level

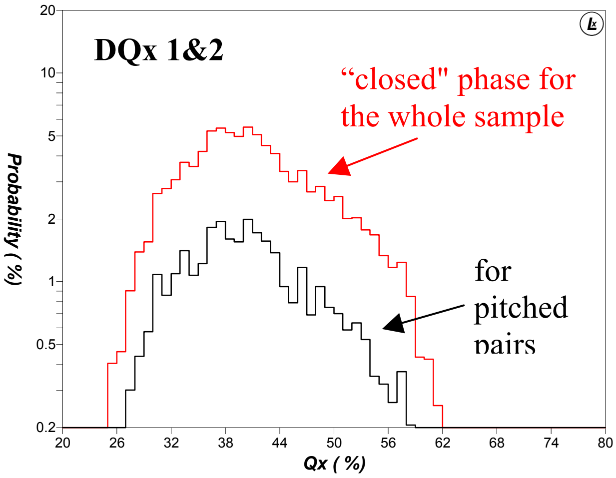

figure 5 Distributions of closed phase ratios for the three minute sample of figure2 & 3 and using the same pitch seeking basis for the indication of the pitch related components of the voice sample

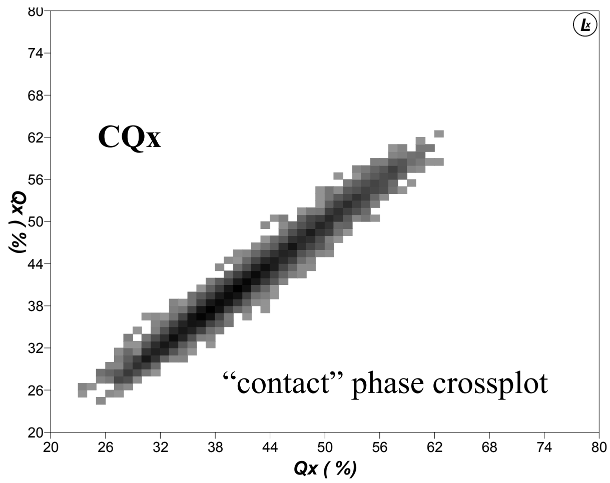

figure 6 Contact phase crossplot – successive vocal fold contact intervals, as shown in figure4 , are plotted against each other to give a digram representation of the degree of coherence that exist between adjacent vocal fold vibratory cycles not now in respect of their timing but their correspondence in closed phase durations.

CASE 1

Spontaneous remission of vocal fold paralysis

Following the removal of a thyroid gland carcinoma, this 54 year old male speaker had essentially complete paralysis of the left vocal fold and reduced mobility in the right. The associated very rough voice quality was not completely explicable in terms of pitch irregularity. One month following the initial operation, however, he had spontaneously recovered complete right fold mobility and a degree of effective left fold function.

Voice Frequency range

Figure 6 shows the results of individual vocal fold period analyses for a standard read text using the German version of the North Wind and the Sun. DFx1distributions are in close correspondence for the two conditions in regard to their shape and in each case cover essentially the same ranges (0·39 and 0·33 octaves). There is a significant lowering of the mean vf frequency from 188Hz to 145Hz. The DFx2 distributions correspond closely to the shapes of their parent DF1 plots and this indicates that the intrinsic regularity of vf vibration is good. These results do not correspond to the auditory impression of a marked perceived roughness that would have a high rating for G in the GRBS scale.

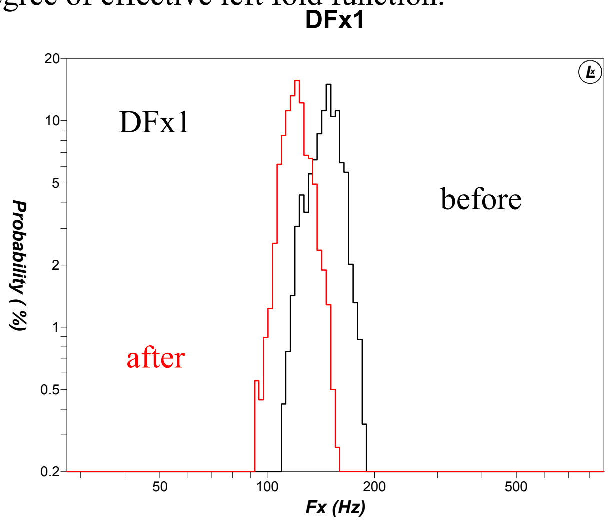

figure 7 Vocal fold frequency ranges, DFx1, before and after admission, show a reduction in mean frequency but no difference in range

This is also the conclusion to be drawn from the examination of the CFx crossplot for the initial condition. The irregularity index of 6·5% is not at a value that ordinarily would be associated with a high level of poor voice quality; and there is relatively little change from the initial to the final CFx irregularity value of 7·4% – which although slightly greater is associated with a very good voice quality.

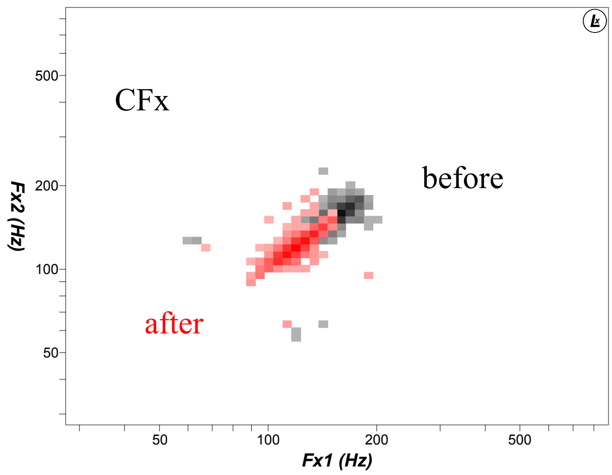

figure 8 CFx is a crossplot of the relation between adjacent vocal fold periods for the whole of the recorded sample and this figure shows both the small change in frequency range of figure 6 and also a very small difference in vocal fold vibrational irregularity between the “before” and “after” conditions – 6·4% before and 7·4% after.

This marked difference makes it necessary to look for the physical factors that might be correlated with these very striking voice quality contrasts and an obvious possibility, in the light of the initial discussion, is to examine the structure of control of the closed phase of vocal fold fold vibration during sample data sets obtained for the two conditions. The Lx data lends itself to this type of multiple use.

Cycle to cycle closed phase control

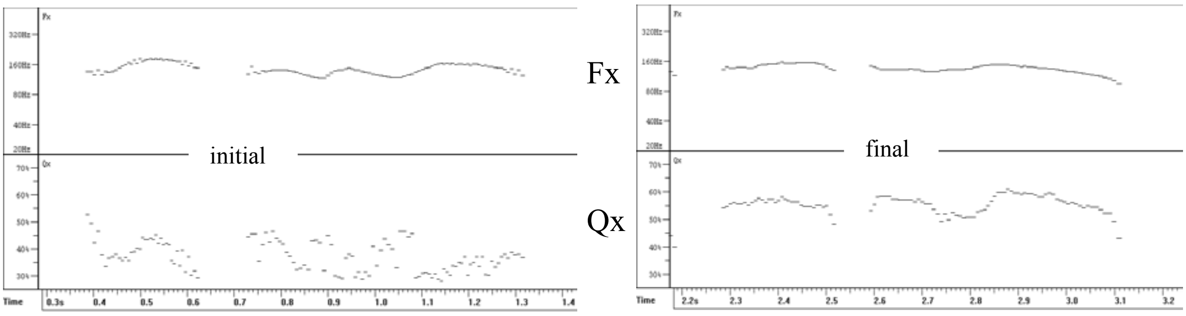

figure 9 the two sets of graphs represent Fx and Qx for the excerpt "wind and sun" taken from the original recordings used for the analyses of figures 6 & 7

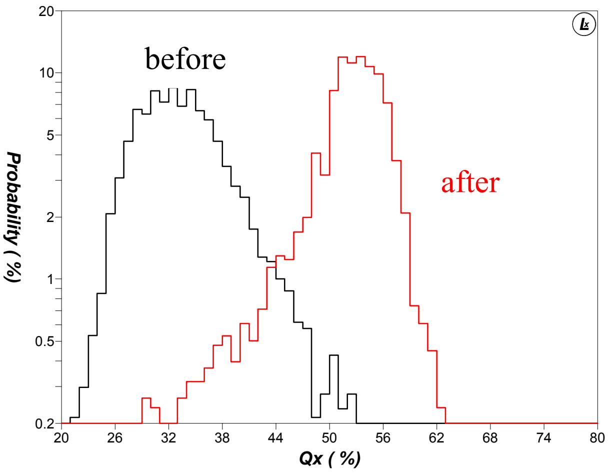

figure 10 A very large improvement in closed phase range and control has occurred with the mean closed phase ratio increased from 33% to 51%

The closed phase analyses of DQx1 and 2 in figure 5 for a normal voice were derived from Qx data derived from the Lx waveform as shown in figure 4. Qx is simply a function of time and it can be used interactively for therapy or for the purposes of comparison. In figure 9 the pairing of the initial and final versions of the same utterance shows quite strikingly how it could be that a voice with good vocal fold frequency control could nevertheless have a perceptually rough character. The two Qx traces are evidently different in texture. They are also evidently different in the average values of Qx with which they are associated. The initial condition has a much lower average value of Qx. This corresponds to the average values of closed phase ratio of 33% and 51% which are the means of the two distributions of Qx in Figure 10.

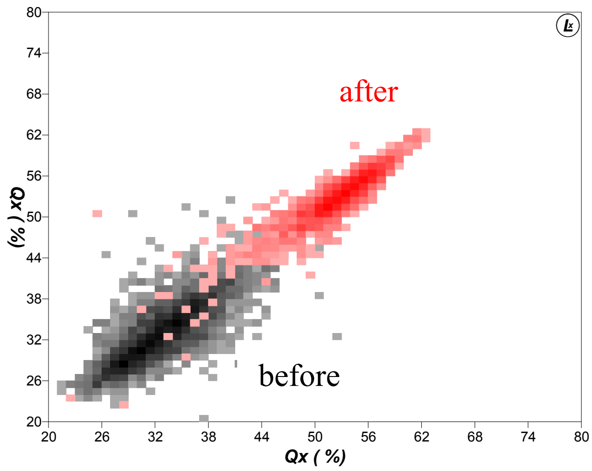

figure 11 CQx is a crossplot derived from the measurement of Lx based closed phase intervals. these two superposed distributions correspond to the histograms of figure7 and indicate both the improvement in range control and also the substantial improvement in the control of closed phase regularity – from 51% to 40%.

The perceived roughness of the initial voice condition is also associated with a greater spread of the CQx distribution. For the normal voice of figure 6 the spread of the crossplot corresponds to an index of 29% whilst for the "before" condition shown here in figure 11 the index is 51% and this is linked to the very low average Qx also shown in figure 8.

Here the very substantial improvement in voice quality has occurred spontaneously. It is of special interest to examine other rough voice cases of this type – with therapy.

CASE 2

Progressive improvement of contact phase control

Following partial thyroidectomy, this 24 year old woman had almost complete left vocal fold paralysis and a poor speaking voice quality that was partly explained by a 16% irregularity index. She received electro-stimulation therapy for her left vocal fold and her irregularity index in running speech decreased to 6·8%, this change did not, however, totally account for the improvement and since her voice became substantially less breathy it was important to measure any changes in her overall control of closed phase ratio during the course of treatment.

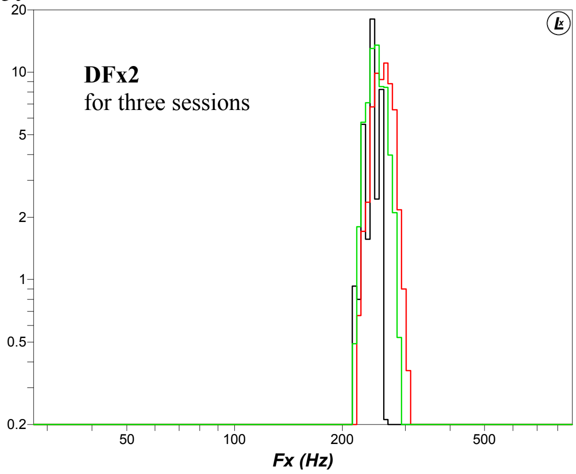

figure 12 Second order, DFx2, analyses for three clinical sessions show little change in mean frequency and not a substantial alteration in the range.

The voice quality improvement noted with this patient was monitored using the data obtained from the routine Sp and Lx recordings and their associated DFx and CFx analyses. Figure 12 is a particular example. It was not clear, however, that the changes were entirely due only to the reduction in temporal irregularity and this appears to be confirmed by the DQx1 analyses shown in figure 13. Overall from stage 1 to stage 3 there is a change from a mean Qx of 31% to 48% and this reduction in breathy voice quality is accompanied by a striking increase in Qx regularity.

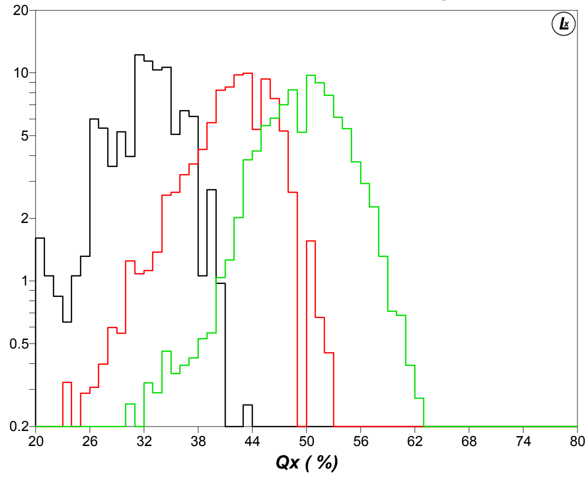

figure 13 Three superimposed DQx1 distributions show a progressive improvement in the control ot the closed phase ratio from clinical session to session with electro-laryngeal stimulation therapy.

At the beginning of the clinical sessions the CQx measure of closed phase regularity was 60% at the next session this had diminished to 42% and at the final session reported here it had stabilised at but with a corresponding change to a markedly less breathy character. This set of different steps towards a less pathological voice is illustrated in figure 14 by the series of CQx crossplots.

For this case three complementary changes have taken place: temporal regularity has been increased; the breathy quality has been reduced and the irregularity of closed phase cycle to cycle variation has also been reduced. Connected speech measurement indices have proved more useful than vowel descriptors.

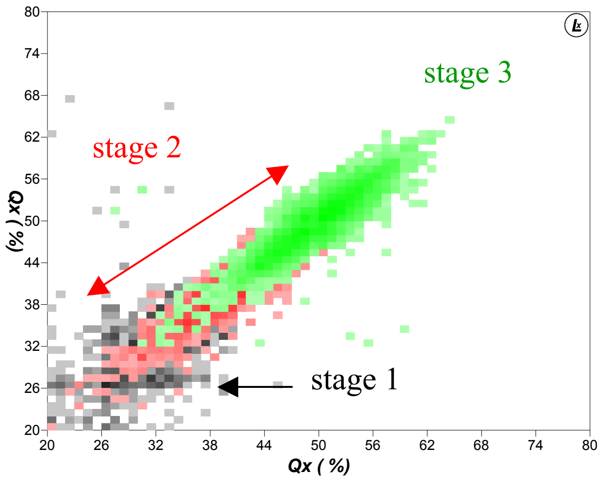

figure 14 CQx for each of the successive stages of therapy show that both the range also the regularity of closed phase control improve consistently

CASE 3

Improvement in contact phase range and regularity

Following a viral infection, this 54 year old woman experienced gradually increasing difficulty in speaking. She was found to have left vocal fold paresis and, in the absence, after careful clinical examination, of other clinical symptoms, she was given a course of electro-laryngeal stimulation over period of two months. A take-home system was used with regular hospital visits to monitor progress and to fine-tune the programme of therapy. There was consistent voice improvement with her irregularity index decreasing from 15% to 9%. This change was not sufficient, however, to account for the total improvement and closed phase measurements were also made.

Figure 15 shows the two second order vf frequency distributions that are linked to the measurements of temporal closure irregularity. They indicate a relatively good control of voice pitch but they are, by their nature, not capable of showing other aspects of voice quality that might account for the marked voice improvement that occurred.

DQx distributions showed a gradual improvement in overall closed phase control but they did not reveal the other major change in progress that is shown by the CQx crossplots of figures 16 and 17. Figure 16 shows not only that there is a bimodal distribution in closed phase control initially but also that there is a marked degree of Qx irregularity for the upper end of the Qx distribution. As therapy (and time) progressed, voice quality improved and the Qx distribution became much less bimodal and upper values of Qx were far more regular in their cycle to cycle variation. Overall Qx irregularity descended from 66% to 38%. Simple mean values of critical parameters are not capable of conveying the importance of many aspects of voice and the structure types of Figure 16 tell far more than their mean Qx values of 44% and 55%.

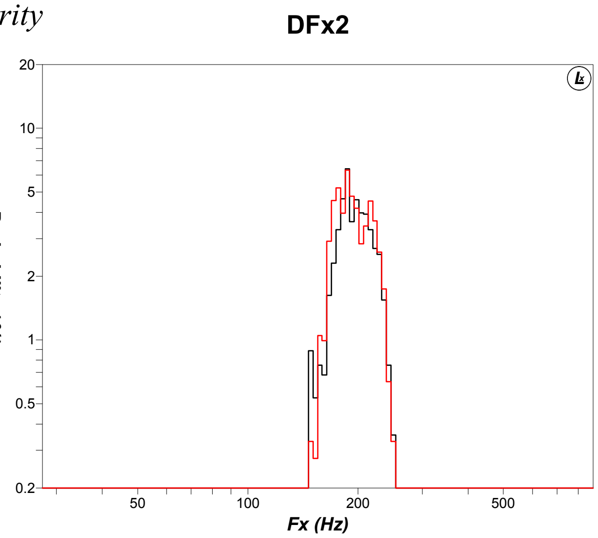

figure 15 Second order vf frequency plots

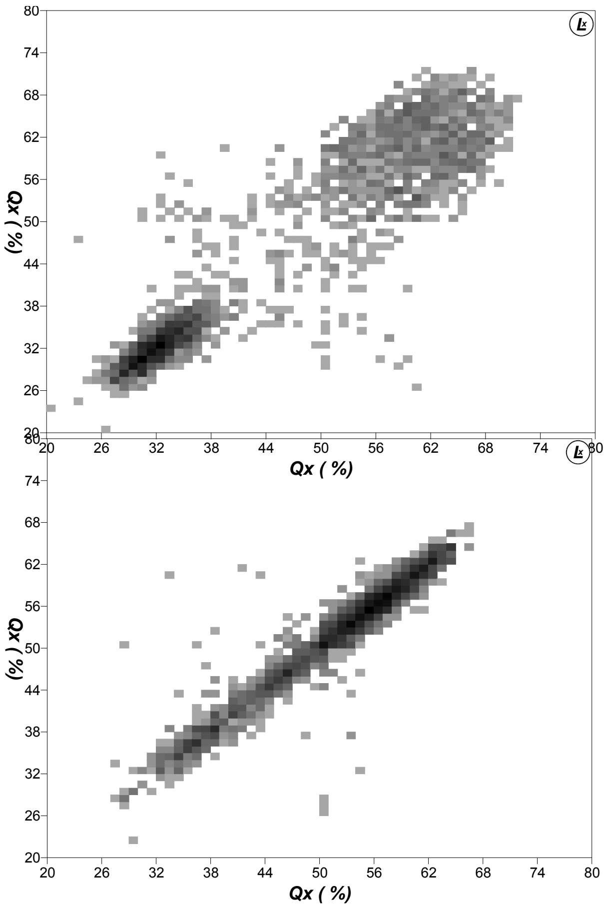

figure 16 Closed phase ratio crossplots, CQx, showing the difference in structure and regularity that accompanied the marked voice improvement of this speaker.

CASE 4

Specific voice quality improvement in a singer's voice – for closure control and regularity but not for pitch.

This 45 year old woman singing teacher presented with a feeling of voice failing after use. Initial measurements showed that her voice quality was excellent both in respect of temporal irregularity (5·7%) and closed phase control (31% irregularity) – although she had the beginning of the increase in closed phase at the top of her speaking frequency range that may be associated with the voice of the older woman.

The previous three cases have illustrated different but closely related pathologies that all involved an important contribution from the relative lack of closed phase control. In the present instance the speaker has a very good voice to begin with and standard measures were not capable of showing any deficiency. The control of pitch range, regularity and amplitude were all quite normal although there was a sustained complaint of inadequacy.

Closer examination revealed a slight laryngeal rotation as the result of a cervical problem and a possible slight loss of right vocal fold tension.

Clinical sessions and home based therapy did result in some improvement. Figures 18 and 19 give an indication of the nature of one aspect of the change. Initially closed phase ratio control was linked to normal cycle to cycle variability of about 31%. Finally this variability was reduced to 19%, a high degree of consistency of control.

In effect this speaker sought a performance level that corresponded to her professional requirements and was at a significantly greater level of competence than is ordinarily achievable. The present analysis methods provide a useful approach in these cases.

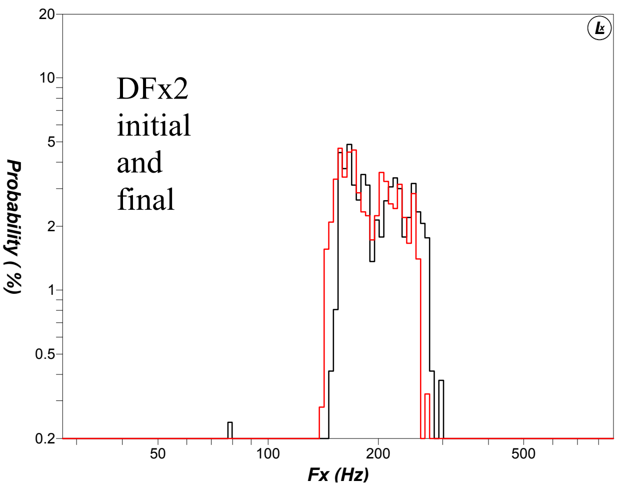

figure 17 Second order vf frequency plots

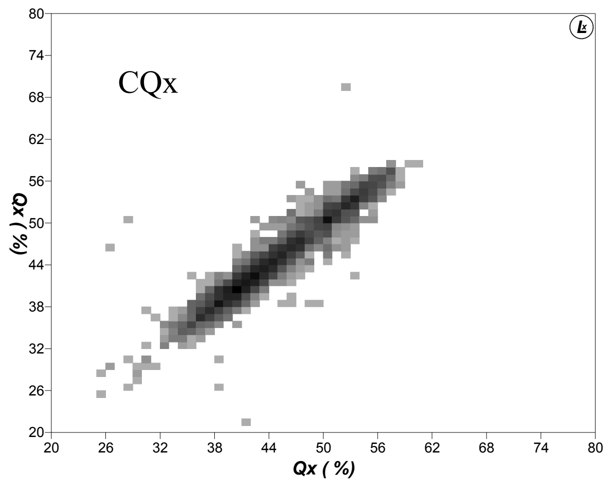

figure 18 Initial closed phase ratio crossplot

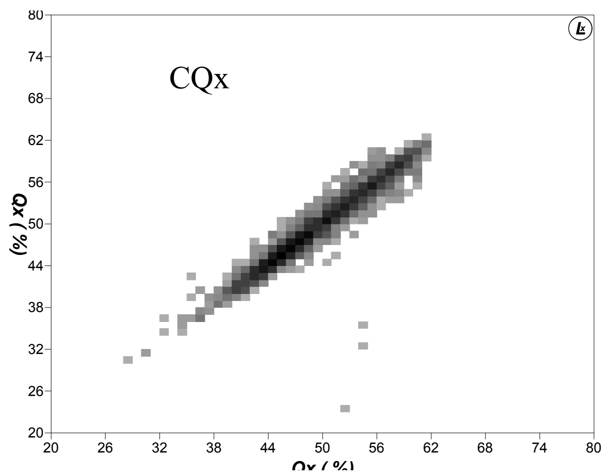

figure 19 Final closed phase ratio crossplot

Summary Table

speaker and condition	sustained vowel			connected speech – read text						
	jitter	shimmer	HNR	irregularity indexes			mean		range	
	%	%	dB	Fx%	Ax%	Qx%	Fx Hz	Qx %	Fx oct.	Qx %
normal 60 yr F	0·36	2·8	28·4	3·2	2·9	29	188	41	0·96	32–53
1 initial 54 yr M	2·0	10·2	24·6	6·5	3·6	50·7	144·7	33·5	0·39	28–41
1 final	0·3	3·3	32·5	7·4	8	29·6	118·2	51·5	0·33	46–57
2 initial 24 yr F	5·1	6·7	33	15·9	4·9	59·7	236·5	31·5	0·31	27–37
2 final	0·8	3	37·8	4·3	3·8	43·3	243	48·5	0·24	43–56
3 initial 54 yr F	1·1	3·3	35·3	8·3	3·6	66·1	187·7	44·5	0·5	30–64
3 final	0·5 2·9	3·8 4·6	38·5 37·3	9·1	2·9	38·2	182·3	50·5	0·49	37–61
4 initial 45 yr F	0·2	2·8	36·7	5·7	6·2	31·5	187·7	44·5	0·69	37–54
4 final	0·4	2·9	38·6	4·9	5·3	19·5	182·3	48·5	0·66	43–57

Note Fx is the frequency value derived from the instantaneous period Tx, where Tx is the time measured between successive vocal fold closures;

Ax is the pressure-positive first peak of the acoustic speech signal immediately following vocal fold closure;

Qx is the ratio of the closed phase interval of vocal fold vibration to Tx, derived from the width of the closure peak of the laryngograph signal between the estimated instant of closure and a point on the opening phase that is 70% down from the peak to peak value of that peak;

Irregularity indexes are derived from the crossplots of the respective parameters

References

[1] Lin-Chuang-Er-Bi-Yan-Hou-Ke-Za-Zhi: Computer based voice analysis for the normal adult. Chines Journal of clinical otorhinolaryngoglogy { Lin-chuang-er-bi-yan-hou-ke-za-zhi}; 12(10): 448-50; 1998

[2] Chernobel'skii,-S-I: Differential diagnosis of unilateral paresis and paralysis of the larynx; { K differentsial'noi diagnostike odnostoronnikh parezov i paralichei gortani}; Vestn-Otorinolaringol; (2): 24-6; 1996

Analyses were made using Laryngograph ® Speech Studio – www.laryngograph.com