Proceedings of the Institute of Acoustics

A LARYNGOGRAPHIC STUDY OF THE SPEAKING AND SINGING VOICES OF YOUNG CHILDREN

PJ White and GF Welch

The Leverhulme Project, Centre for Research and Development in Early Childhood and Primary Music Education, School of Education, Roehampton Institute of Higher Education, London SW15 1RQ

1. INTRODUCTION

The electrolaryngograph has been in use in this and other countries for some time as a method of analysing normal and abnormal voice. It is a non-invasive device which, by the placing of two electrodes bilaterally on the neck at the level of the larynx, delivers a constant voltage across the soft tissues and mucosa of the neck. A fully or partially closed glottis will reduce the impedance, and thus changes in current flow can be observed and measured [1].

Figure 1 shows one cycle of an idealised laryngographic output waveform (Lx). The period of the cycle (Tx) can be measured to produce a fundamental frequency (F_0) estimation ($F_0 = 1/Tx$). The beginning of vocal fold closure is indicated by a sudden sharp rise in the waveform (i), and the peak corresponds to maximum vocal fold contact (ii). From here, as the folds begin to open and the contact area is reduced, the waveform takes a negative direction. At a point on this downward slope the vocal folds are no longer in contact (iii), and the glottis remains open (iv) for the remainder of the cycle. One cycle of vocal fold vibration consists of a closed phase (CP) and an open phase (OP) as shown in the figure (CP + OP = Tx).

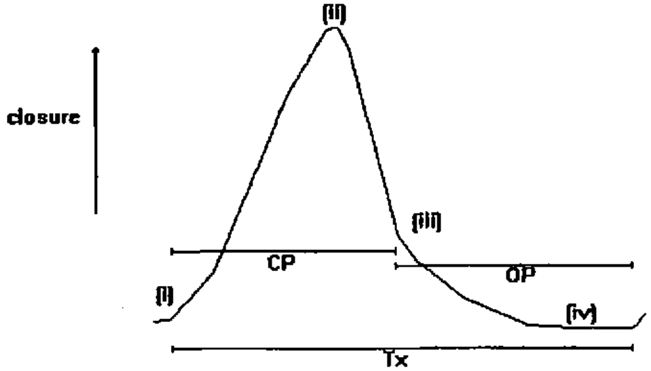

Figure 1 One cycle of an idealised Lx waveform indicating (i) the beginning of vocal fold closure, (ii) point of maximum contact, (iii) no contact - vocal folds open, (iv) folds remain open

Recent studies have suggested that the closed quotient (CQ) (the time for which the vocal folds are in contact during each vibratory cycle expressed as a percentage of the cycle's period, CQ = (CP/Tx) * 100) increases with singing experience and/or training [2, 3, 4 and 5]. Figure 2 shows a few cycles each of (a) an Lx with low CQ, and (b) an Lx with high CQ. The Lx of a trained adult singer would be expected to compare with (b).

(a)

(b)

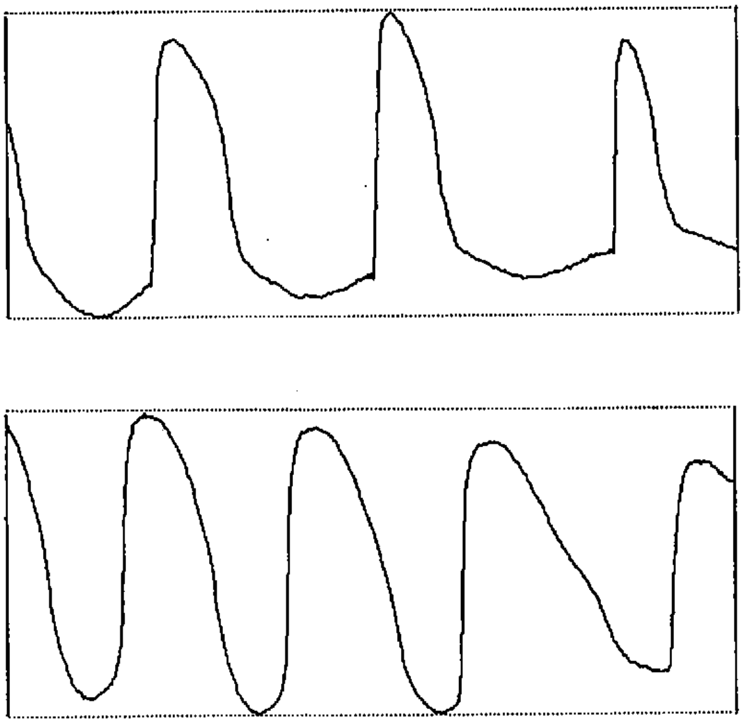

Figure 2 Examples of Lx waveform showing (a) a low CQ, the OP is proportionately greater for each cycle than the CP, and (b) a high CQ, the OP is of a similar length to CP, the folds are remaining closed for a greater proportion of the cycle.

The calculation of the closed quotient is judged here to be more appropriate than the open quotient (the percentage of the period during which the vocal folds are open, and therefore precisely complementary to CQ) 'to the extent that real-time displays of these values can potentially be used to give visual feedback in voice training, since it appears that positive CQ changes are generally desirable' [3].

All the published studies to-date have been based on adult subjects who were asked to perform a number of speaking and singing tasks designed to contrast differences in style between the two types of voice. The findings implied that formal voice training produced a lengthening of the time the vocal folds were closed during each cycle with respect to the OP, leading to three main effects which can be seen to relate to the desired effect of singing training:

* an increase in the proportion of the cycle in which the glottis remains closed ensures a greater acoustic output from the vocal tract due to a proportionate decrease in the open phase and the acoustic damping effects of the subglottal cavities associated with it;
* the reduced open phase enables valuable stored lung air to be used at a slower rate, the effect being more energy-efficient as well as allowing notes to be held for a longer time;
* the voice has a less breathy quality.

Few studies exist that have attempted to quantify children's voice qualities, or gender differences. In order to shed some light on the possible effects of training, the authors of this paper undertook a study of young children's voices to see if there was a similar CQ variation in a sample child population. The present study is based on the methodology previously used in the adult studies, contrasting the quantification of larynx closed quotient changes in the spoken and singing voices of five-year-old children. The aim was to gain some objective measures of the child speaking and singing voice at source, with a comparison of these styles of voice, and of gender.

2. METHOD

2.1 Subjects

A class of children was selected from each of two schools in the greater London area. These schools were considered to include a cross-section of children representing such criteria as sex, ethnic background and social class[1]. Recording took place in the Autumn term of 1991 when most children had not yet reached their sixth birthday.

2.2 Procedures

A room in each school had been set aside for the interviews. The children were brought into the room in small groups, where the equipment was demonstrated and explained. The testing was carried out individually on children taken voluntarily in turn from the group. Each child was encouraged to take part, but was given the option not to.

---

[1] These schools are currently participating in a large three-year study of musical development in young children between the ages of three and eight years. Most of the children in the present study were part of the larger research project cohort and had therefore become familiar with the researcher through previous visits.

The subject was positioned on a chair in view of an oscilloscope giving real-time output from a portable electrolaryngograph. A wide frequency response cardioid tie-clip microphone was attached to the child's clothes at a distance of approximately 15 cms below the mouth, and was connected to the left channel of a Casio DA-7 portable digital audio tape (DAT) recorder. Digital recordings ensured that the effects of low-frequency phase distortion were avoided [6]. The output from the laryngograph was recorded onto the right channel of the DAT recorder. Care was taken to ensure correct positioning of the electrodes at either side of the neck at the level of the larynx (see *Discussion*).

A small number of the children had to be eliminated from the study where the Lx trace on the oscilloscope was seen to be poor and unreliable. These children were allowed to continue with the tasks but were not recorded. Adult-sized electrodes were used in all cases as they were found to produce the maximum amplitude response during the recording compared to the smaller child-sized electrodes. Each subject was asked to (a) sing a song s/he knew well covering a large part of the vocal range, and (b) to produce natural, spontaneous speech to include the whole range of frequencies of his/her modal speaking voice. As it was considered necessary to include a certain amount of flexibility to obtain the best results, the subjects were asked to sing *Twinkle, twinkle little star* or, if unknown or unwilling, a song of her/his choice, and to relate a recent outing or event, discuss a class activity, or describe a scene in a picture. Each child was recorded for a total period of no longer than five minutes.

2.3 Analysis

The resulting recordings were checked for distortion by listening separately to both the speech and the Lx recordings on the tape, and by using a Masscomp 5600 computer. The Speech Filing System (SFS) set of programs [7] enables the recorded data to be displayed on screen in small samples to check for such distortions as, for example, noise and excessive movement of the electrodes. Only those recordings considered to be unaffected by any of the above were used in the data analysis. The number of subjects included in the analysis totalled 48: female = 18, male = 30.

A sample of approximately 30 seconds of speech and 20 seconds of singing was analysed for each child and a cycle-by-cycle CQ calculation made from the output laryngograph waveform. The fundamental frequency and, from the CP calculations, the closed quotient values for each cycle of input data were stored and later processed to produce the mean CQ values (and standard deviations) for each sample.

3. RESULTS

A total of 96 data files were created (48 subjects * 2 voice types), and each was analysed separately using the Masscomp 5560 as described above. Table 1 shows the CQ values for (a) female, (b) male, and (c) male-female groups giving the mean and standard deviation (SD) for each group.

(a)

Female	CQ (%) Mean	(SD)
Speech	42	(13.14)
Singing	38	(10.99)

(b)

Male	CQ (%) Mean	(SD)
Speech	37	(13.35)
Singing	37	(11.38)

(c)

Male and female	CQ (%) Mean	(SD)
Speech	39	(13.27)
Singing	37	(11.52)

Table 1 *Showing mean CQ values, and standard deviation, for (a) females (n = 18), (b) males (n = 30) and, (c) males and females combined (n = 48).*

From these data female subjects in the sample have a slightly lower mean CQ for singing (38%) than for speech (42%). As a group females have a similar width of range (speech = 30-60%, singing = 27-54%) for both vocal activities, but the range is shifted down slightly during singing to embrace more lower and less higher CQ values. Males have the same mean CQ for singing as for speech (37%) and there is very little difference in the overall ranges. In this sample, females have a higher mean CQ (42%) for speech than males (37%), but for singing, both groups have similar mean CQ values (females = 38%, males = 37%). The range of CQ values is altogether higher for females than for males in both speech and singing. Sample males also exhibit a narrower range than females for both voice types.

As a combined group (males/females), the mean CQ is slightly lower for singing than for speech. There is relatively little difference in values obtained for the two voice types for the males. The females, on the other hand, show a greater *change* in use of CQ between speech and singing.

The mean scores for the two voice conditions were subjected to a statistical analysis (Mann-Whitney U test) for both conditions within each group. The results were not significant. The same tests were then applied to the mean scores for each condition between the female and male groups. The female groups showed a significantly higher mean CQ for speech than the male group ($p < .01$).

4. DISCUSSION AND CONCLUSIONS

Although the application of a non-invasive measure of vocal fold physiology is ideal for research purposes there are difficulties associated with such a study of children. Young children are known to sometimes exhibit gross vertical movements of the larynx, effectively displacing the electrodes for some utterances, particularly those of a high fundamental frequency. An attempt was made to position the electrodes to gain the best results for what was considered to be the modal speaking F_0 range. It is possible, therefore, that some voiced sounds are recorded via the microphone which do not register on the laryngograph. The electrodes were repositioned one or more times during the interview if necessary. Another problem associated with laryngographic recordings is the high degree of fatty tissue in the necks of some children (and adults) which can make a reasonable Lx output response difficult or impossible to obtain.

The adult studies have included tasks such as reading a prose passage in order to obtain uninterrupted natural speech, singing two octaves of a major scale, and reproducing specified vowel sounds at given frequencies. None of these tasks are feasible for such young children. Simpler methods were therefore devised for this study to elicit the required responses of any variation in speaking and singing voices. In a study of 18 adult males [4], it was noted that 'some of the untrained singers ... apparently tend to use singing CQ values that are lower than their speech CQ values'. The same appears to be true of the voices of the female group included in this study. These subjects used a more breathy phonation overall for singing than for speech. The male group showed similar use of CQ for both voice conditions.

Future studies will aim to compare age groups, and the characteristics of the Lx waveform in pre- and postpubescent children. Also, the voices of children who have received formal singing training will be compared to the present data in an attempt to establish any links with training and/or experience with an increase in CQ values, as previously seen in the studies of adult voices.

5. ACKNOWLEDGEMENTS

This study forms part of a research project funded by the Leverhulme Trust. The authors would also like to thank the Phonetics and Linguistics Department of University College London for the use of their computer facilities, and Dr David Howard for his valuable advice.

6. REFERENCES

[1] FOURCIN AJ and ABBERTON ERM, 'First applications of a new laryngograph', *Medical and Biological Review*, 21 pp172-82 (1971)

[2] HOWARD DM and LINDSEY GA, 'New laryngograms of the singing voice', *Proceedings of the XIth International Congress of Phonetic Sciences*, USSR, Tallinn, 5 pp166-9 (1987)

[3] LINDSEY GA and HOWARD DM, 'Larynx excitation in the singing voice', In *Speech, Hearing and Language: work in progress*, 3, University College London pp169-77 (1989)
[4] HOWARD DM, LINDSEY GA and ALLEN B, 'Toward the quantification of vocal efficiency', *Journal of Voice*, 4 (3) pp205-12 (1990)
[5] HOWARD DM, LINDSEY GA and PALMER S, 'Larynx closed quotient measures for the female singing voice', *Proceedings of the XIIth International Congress of Phonetic Sciences*, Aix-en-Provence, 5 (5) pp10-13 (1991)
[6] HOLMES JN, 'Low frequency phase distortion of speech recordings', *Journal of the Acoustical Society of America*, 50 pp747-9 (1974)
[7] HUCKVALE MA, BROOKES DM, DWORKIN LT, JOHNSON ME, PEARCE DH and WHITAKER L, 'The SPAR Speech Filing System', *Proceedings of the European Conference on Speech Technology*, 1 pp305-17 (1987)