Proceedings of the Institute of Acoustics

Real-time visual feedback and the development of singing pitching skills in primary school children

David M Howard, and James AS Angus
Department of Electronics, University of York, Heslington, York YO10 5DD
Tel: 0904 432349; Fax: 0904 432335; e-mail: dmh/jasa@ohm.york.ac.uk

1. INTRODUCTION

It has been estimated that approximately 30% of the adult population in the UK are unable to pitch a note accurately against a reference [1, 2]. We have developed a computer program to enable singing pitching ability to be assessed and developed in primary school children. The system, known as SINGAD (SINGing Assessment and Development), was initially implemented on the BBC microcomputer, but advances in computer technology enabled a number of potential improvements to be implemented and the system was re-written for the Atari range of microcomputers. The Atari version makes use of a MIDI (Musical Instrument Digital Interface) port to play stimuli via a music synthesizer and to record sung responses via a commercial pitch-to-MIDI converter (Roland CP-40).

The Atari version of the SINGAD system has been used to assess the singing pitching skills of a group of primary school children in York on two occasions approximately one year apart. During the intervening year, the classes were divided into those that acted as a control group and those that made use of the real-time visual feedback displays provided by the SINGAD system. Data recorded as a result of the first of these assessments have been presented at previous IOA Speech and Hearing Conferences; overall trends are recorded in Howard et al. [3], and a further analysis of these data provides some clues as to the strategies adopted by children when they attempt to sing three note patterns which are designed to be musical [4]. Further analysis of these data have provided interesting indications of the relative pitching accuracy of the different musical intervals employed in the SINGAD stimulus set [5].

The purpose of this paper is to present the data resulting from the second assessment recording session and to discuss these in relation to results previously presented for the initial assessment and the group (control or SINGAD) that each child belonged to.

2. BACKGROUND

The SINGAD system is designed to enable singing pitching abilities to be assessed quantitatively and to be developed by means of a real-time visual display of voice fundamental frequency (f_0) against time. During the assessment phase a series of assessment trials are presented to the listener who is asked to sing them back. Voice f_0 values are measured and stored for future analysis. During the development phase a real-time trace of f_0 against time is presented. The time and frequency ranges of this plot can be adjusted appropriately for the singing pitching abilities of the user, and a number of pictures can be placed on the screen to provide visual targets for the user to 'hit' with their pitch trace. The usefulness of the SINGAD system in improving pitching skills in the classroom has been established, and it has also been demonstrated that this improvement can take place without the need for teacher intervention [6].

The earlier BBC microcomputer version of SINGAD [7], written in BBC Basic and 6502 assembler, had a number of disadvantages associated with the assessment phase of its operation. Firstly, assessment stimuli were played via the internal loudspeaker and this gave rise to pitching errors, usually of one octave, even with experienced singers. Secondly, data storage was limited by the memory and storage devices available such that only the mean f_0 of 255 cycles could be stored, and the f_0 trace itself could not be viewed for any detailed subsequent analysis. Given the likelihood that an f_0 estimation device will make errors [8, 9], it is essential in practice that any f_0 trace, no matter what device or algorithm produced it, is checked before its f_0 data is analysed and interpreted. Finally, only a fixed set of single note trials were available in the assessment phase which have no particular musical meaning, and therefore it could be argued should not be used in the context of a measurement of singing ability.

The development of the Atari based version [10] enabled these limitations to be addressed as follows. Stimuli are played via the MIDI out port, enabling a standard music synthesizer to be used with user control over the sound to be used. Data storage is not noticeably limited in the context of this application, so all sung voice f_0 data can be stored for subsequent evaluation and analysis. The assessment trials can contain either one, three or five stimuli each, and these can be in the musical context of triads in either a major, minor or pentatonic scale in user-defined key. The three stimuli trials that are available are shown in table 1 in terms of the degree of the scale in relation to the tonic (shown as 'I') and they are designed to be familiar to Western ears. Three stimuli trials were used in a major context in this experiment to maintain the memory task required at a reasonable level. There are five different trials, and these are played in a random order during a SINGAD assessment session either once, twice or three times, making a total of 5, 10 or 15 trials per session. In this work, 15 trials were used with each child.

TONALITY	TRIALS AVAILABLE (Each note shown as degree of scale relative to tonic)				
	1	2	3	4	5
MAJOR	I, III, V	V, III, I	I, V, I	V, I, V	III, III, III
MINOR	I, III, V	V, III, I	I, V, I	V, I, V	III, III, III
PENTA	V, VI, III	VI, V, III	V, III, I	V, III, III	V, V, III

Table 1: Libraries of three stimulus trial patterns available for the three tonality options.

Following an assessment session, cursors are set by hand at the start and end of each stimulus of each trial and summary f_0 statistics (maximum f_0, minimum f_0, mean f_0, f_0 standard deviation, number of f_0 samples, and stimulus duration) are calculated for each trial. In addition, an overall absolute average difference between the sung response and the assessment stimuli is provided for all the trials. All f_0 values can be recorded in either Hz or MIDI note number plus a number of cents (one cent is equal to one hundredth of a semitone). The latter is most perceptually relevant, since it is based on a logarithmic frequency scale and therefore its use is preferred. This information is stored in a binary file used by SINGAD to set up the screens and its internal parameters, as well as in ASCII format which is used for subsequent data analysis.

3. SUBJECTS AND EXPERIMENTAL DETAILS

The pupils who took part in this experiment were from Kingsway Junior School, Clifton, York and they were distributed by year and sex at the time of each recording as shown in table 2. In total, 177 children were recorded in November 1993, and 215 in October 1994.

Date of recording	*11/93*	*10/94*	*11/93*	*10/94*	*11/93*	*10/94*	*11/93*
Class year	*3*	*4*	*4*	*5*	*5*	*6*	*6*
Approx. age (years)	*8*	*9*	*9*	*10*	*10*	*11*	*11*
Girls	24	31	21	36	23	26	26
Boys	24	33	17	38	27	51	15
Total	**48**	**64**	**38**	**74**	**50**	**77**	**41**

Table 2: Subject distribution for both recordings.

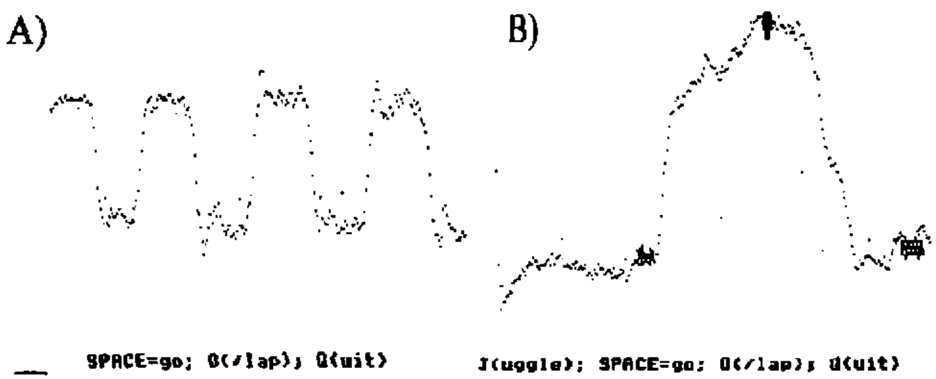

Figure 1: BBC computer SINGAD development screens: (A) imitation of a fire engine, and (B) pitch line and three 'town' targets.

In November 1993 and October 1994 the recordings were carried out over two consecutive days using two Atari SINGAD systems in opposite corners of the school library, each under the supervision of one of the authors. Each system made use of a commercially available pitch-to-MIDI converter (Roland CP-40) to provide its f_0 input for the sung responses, and note stimuli were played via piano patches on a Roland music synthesizer. Audio was provided via Boss micro monitors. Whenever possible, each subject observed the previous subject carrying out the test as a way of becoming familiar with the experimental procedure. Subjects were asked to keep an eye on the green/red signal level indicator on the pitch-to-MIDI converter when singing their responses in order to ensure that: (a) their voice was being recorded, and (b) they were not overloading the device. This had a further benefit in that it provided a focus of attention away from local distractions and disturbances. The level indicator on the pitch-to-MIDI converter was also monitored by the experimenter.

Major tonality three stimulus trial patterns with each trial being played three times in the key of Bb major were used. This key was selected as it is a comfortable range for children to sing. Stimuli were played at a rate of 1 per second in each trial, with each stimulus lasting for 0.8s with an inter-stimulus time of 0.2s. Each subject therefore responded to a total of 45 stimuli. In the first assessment experiment, a total of 7,965 notes (45*177) were analysed, and in the second a total of 9,675 notes (45*215) were analysed.

During the period between the recordings, the classes involved were either part of the 'control' or the 'SINGAD' group. The control group did not make use of the SINGAD system (except for the two assessment recordings). Pupils in the SINGAD group were allowed to work with the BBC version of SINGAD development in pairs as time permitted. The BBC version was employed due to the availability of BBC computers in the school. The development screen consists of a real-time pitch line which tracks from left to right with time. Figure 1 shows two example BBC SINGAD development screens; one is a 'plain' screen on which a fire engine siren has been imitated (the use of such imitations is very useful to encourage pitch movement), and the other has three picture targets which have been hit. Pupils directed their own use of the development software which has been designed with the familiar BBC-style user-interface.

4. RESULTS

The sung responses were analysed by placing cursors by hand at the start and end of each of the three notes of each stimulus. Figure 2 shows example extracts from SINGAD assessment screens for five trials by a particular subject. The panel in the top left hand part of the figure shows the each of the 15 trials, with the stimuli indicated by the small black squares. A square at the bottom indicates Bb3 (Bb below middle C), one in the middle indicates D4 (D above middle C), and those at the top indicate F4 (F above middle C). Vertical dashed lines are the cursors, horizontal dotted lines represent the positions of the stimuli played, the trial number is shown at the mouse arrow cursor in each case, and the f_0 sung response is plotted between each pair of cursors. The first experiment required 15,930 (177*45*2) hand positioned cursor placements, and the second required 19,350 (215*45*2)! The same researcher set all cursors in both experiments to ensure consistency.

The number of individuals who attended *both* recordings, and whose data could therefore be analysed for the purposes of this experiment, were as follows:

31 pupils (16 girls; 15 boys) were present in year 3 (1993) and year 4 (1994)

33 pupils (17 girls; 16 boys) were present in year 4 (1993) and year 5 (1994)

37 pupils (15 girls; 22 boys) were present in year 5 (1993) and year 6 (1994).

Summary f_0 statistics for the data between each pair of cursors is stored as an ASCII file, and these were transferred to a Macintosh computer to be analysed using Excel via a custom macro. Each pupil who was present for both recordings sang 45 notes in each session, each of which was compared with the appropriate stimulus. Results are plotted in figure 3 as the mean of the 45 absolute f_0 differences between each sung note and its stimulus for the November 1993 recording (X-axis) against the equivalent measurement for the October 1994 recording (Y-axis), by sex, for each school year and by group (control or SINGAD). A diagonal line (X=Y) has been added to indicate where 'no change' in pitching ability between the recordings would be indicated. Points falling below this line imply a pitching ability improvement, and those above it imply a worsening in pitching ability.

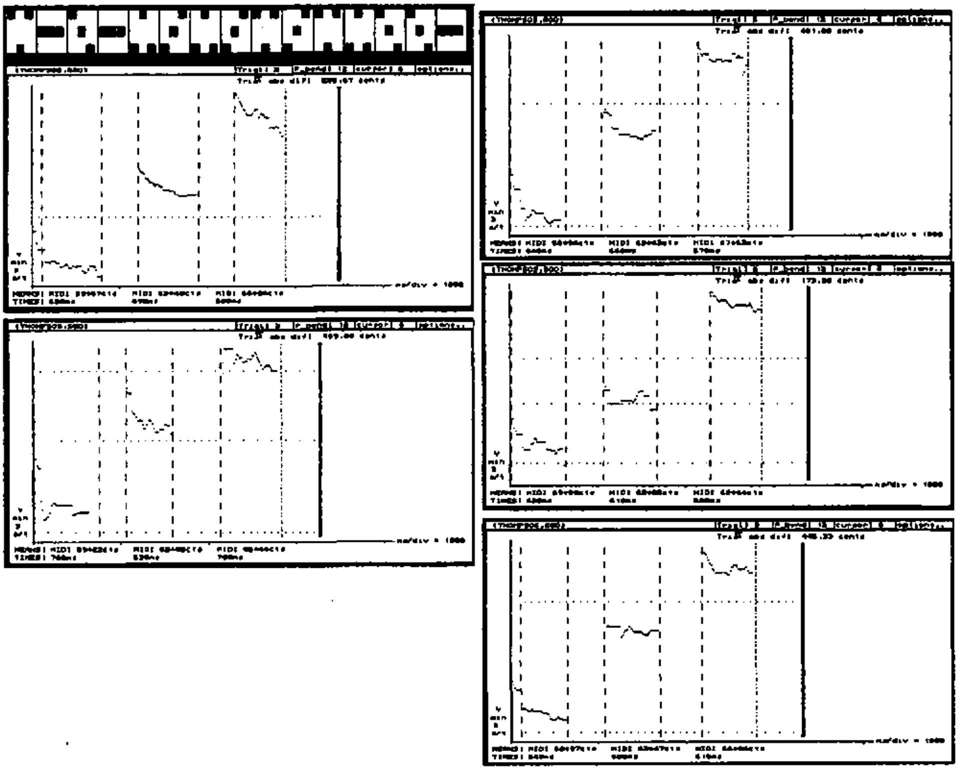

Figure 2: SINGAD assessment screen for a child whose sung responses were similar in relative pitch to each other, no matter what the stimulus.

Each row of figure 3 relates to one group of pupils who were respectively in November 1993 and October 1994 in years 3 then 4, 4 then 5 and 5 then 6. The left hand plot in each row is for the control group and the right hand plot is for the SINGAD group. In the years 3 and 4 control group, a number of pupils have improved in pitching ability without the use of SINGAD, but this is also true of a number of pupils who made use of SINGAD. In both cases, considerably more pupils improved their pitching skills than became worse. Whether the use of SINGAD had an effect is not easy to judge, but it would appear that there was some unevenness in the pitching abilities of the control and SINGAD groups, the latter being generally better at the start. (It should be noted that the groupings were decided upon prior to the first recording, so no quantitative pitching accuracy data were available.) The improvement demonstrated by pupils in the SINGAD group were towards mean pitching accuracies of less than one semitone (100 cents), potentially much more musically useful, which is not true for the control group.

In the years 4 and 5 the greatest improvement is seen in the control group, where all but 2 pupils have improved. As a group, those that made use of the SINGAD development software were considerably poorer in pitching accuracy tan the control group, with errors mostly greater than 200 cents. The plots do suggest however, that more of the girls have improved compared with the boys, but these errors are very large and there may be considerable guesswork in the sung responses or sung responses that remain static irrespective of the stimuli. The assessment screens shown in figure 2 are for one occurrence of each of the 5 trials (the trials are plotted along the top of the figure and are 2, 3, 5, 6, 9 in this case), where the subject has sung the same ascending pattern of three notes to each trial. The accuracy of these responses can be seen in the means quoted under each plot for each of the three stimuli of each trial.

The plots for years 5 and 6 suggest that the use of SINGAD might have had some benefit. In this case, the control and SINGAD groups are more evenly matched in terms of the spread of pitching accuracy in the earlier recording, and more pupils are below the diagonal line and towards that lower error end in the SINGAD group compared to the control group.

5. DISCUSSION AND CONCLUSIONS

All these results need to be considered in the light of the more detailed analysis of the first recording [3, 5] which demonstrated that pitching abilities were better in the higher years anyway, as these skills developed with age. One would therefore expect there to be an improvement over one year for some children as a result of normal classroom teaching and exposure to and experience of singing. There is some evidence to suggest that there is some benefit from use of the SINGAD development software: in years 3 and 4 improvements in the SINGAD group have been towards the more musical useful end at under one semitone, and in years 5 and 6, greater degrees of pitching improvement are observed. In years 4 and 5 the two groups are not well matched at the start, and the improvements in the control group are more marked here than for the years 5 to 6 control group.

The results of this experiment in which the singing pitching abilities of children at a York primary school were assessed by means of a computer-based system on two occasions a year apart suggest that there may be some benefit from using real-time visual feedback of voice f_0. The trend may, however be somewhat obscured because there will be some improvement over a year for some pupils in any case, as demonstrated by previous more detailed analyses of the first set of data. Further work is required to ascertain whether the absolute mean is a reasonable measure to use as a summary statistic, the extent to which the sung responses from some subjects are not affected by the stimuli by being either essentially random or static, and whether statistical significance can be tested given the non-normal nature of the data distributions.

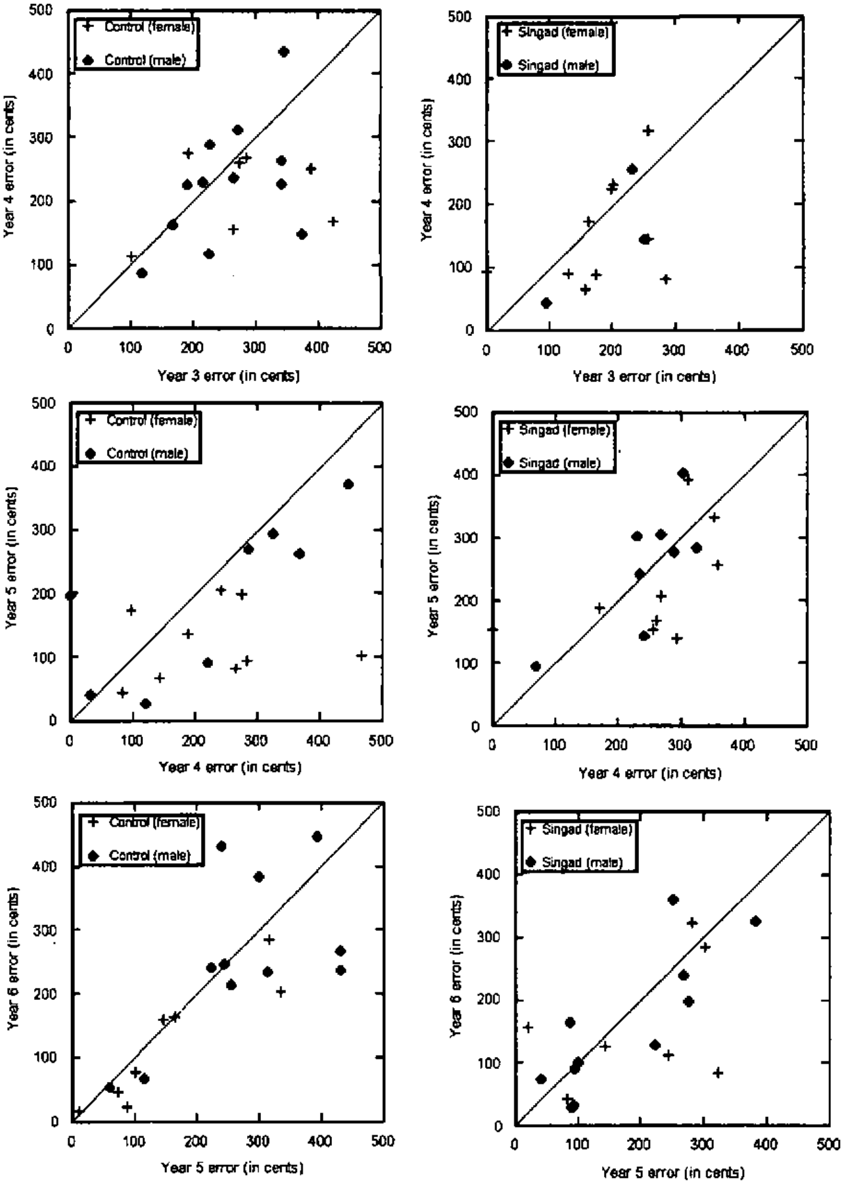

Figure 3: Absolute average mean f_0 differences for each pupil present at both recordings.

The development of singing pitching skills in primary school children

Informal feedback from the children themselves suggested that they found SINGAD fun to use and that it did encourage them to explore their vocal abilities somewhat. The SINGAD system is being developed for PC compatible machines with a new f_0 estimation circuit [11] to increase its availability. The potential benefit is considerable, even if it only serves to instil some measure of increased vocal confidence.

6. ACKNOWLEDGEMENTS

The authors would like to thank staff and pupils at Kingsway Junior School, Clifton, York who made it possible to carry out this experiment.

7. REFERENCES

[1] Welch, G.F. (1979). Poor pitch singing: a review of the literature, *Psychology of Music*, **7**, (1), 50-58

[2] Welch, G.F. (1997). The developing voice, In: Thurman, L. and Welch, G.F. (Eds.), *Bodymind and voice: foundations of voice eductaion*, Iowa: National Centre for Speech and Voice, 481-494.

[3] Howard, D.M., Angus, J.A.S,. and Welch, G.F. (1994). Singing Pitching Accuracy from Years 3 to 6 in a Primary School, *Proceedings of the Institute of Acoustics*, **16**, (5), 223-230.

[4] Angus, J.A.S., and Howard, D.M. (1996). Singing pitching accuracy in children aged 7 to 11, *Proceedings of the Institute of Acoustics*, **18**, (9), 357-364.

[5] Howard, D.M., and Angus, J.A.S. (1998). A comparison between singing pitching strategies of 8 to 11 year olds and trained adult singers, *Logopedics Phoniatrics Vocology*, **22**, (4), 169-176.

[6] Welch, G.F., Howard, D.M., and Rush, C. (1989). Real-time visual feedback in the development of vocal pitch accuracy in singing, *Psychology of Music*, **17**, 146-157.

[7] Howard, D.M., and Welch, G.F. (1989). Microcomputer-based singing ability assessment and development, *Applied Acoustics*, **27**, (2), 89-102.

[8] Hess, W. (1983). *Pitch determination of speech signals*, Berlin: Springer-Verlag.

[9] Howard, D.M. (1998). Practical voice measurement, In: *The voice clinic handbook*, Harris, T., Harris, S., Rubin, J.S., and Howard, D.M. (Eds.), London: Whurr Publishing Company.

[10] Howard, D.M., and Welch, G.F. (1993). Visual displays for the assessment of vocal pitch matching development, *Applied Acoustics*, **39**, (3), 235-252.

[11] Angus, J.A.S., Garner, P.E., and Howard, D.M. (1998). Fundamental frequency estimation for use with a singing pitching development system for primary school children, *Proceedings of the Institute of Acoustics*, this volume.