Are real-time displays of benefit in the singing studio?
An exploratory study

David M Howard(1), Jude Brereton(1), Graham F Welch(2), Evangelos Himonides(2), Michael DeCosta(3), Jenevora Williams (4) and Andrew W Howard(5)

(1) Media Engineering Research Group, Department of Electronics, University of York, Heslington, York, UK
(2) School of Arts and Humanities, Institute of Education, University of London, Bedford Way, London, UK
(3) Professional singing teacher, York, UK
(4) Professional singing teacher, Guildford, UK
(5) SARAND Business IT Solutions, Haverhill, Cambridge, UK.

All correspondence to:
Professor David M Howard
Head of the Media Engineering Research Group
Department of Electronics
University of York,
Heslington
York
YO10 5DD
UK

TEL: +44 1904 432405
FAX: +44 1904 432335
EMAIL: dh@ohm.york.ac.uk
URL: http://www-users.york.ac.uk/~dmh8

This paper was presented at the Voice Foundation’s 33rd Annual Symposium: Care of the Professional Voice, Philadelphia, 2004.

---

1 The research was undertaken under an Innovations grant (number B/IA/AN8885/APN15651) from the UK Arts and Humanities Research Board (AHRB) in 2003-2004.
ABSTRACT

This paper reports on an exploratory research project to evaluate the usefulness or otherwise of real-time visual feedback in the singing studio. The primary purpose of the work was not to optimise the technology itself for this application, but to work alongside teachers and students to study the impact of real-time visual feedback on the students’ learning experiences. An action research methodology was used to explore the benefit of real-time displays over an extended period. The experimental phase of the work was guided by a Liaison Panel of teachers and academics in the areas of singing, pedagogy, voice science, speech therapy and linguistic science. Qualitative data were collected from eight students working with two professional singing teachers. The teachers and students acted as co-researchers under the action research paradigm. Teachers and students alike kept journals of their teaching and learning experiences. Singing lessons were observed regularly by the research team and all co-researchers were interviewed at the mid- and end-point of the project. The use of technology had a positive impact on the learning process and this is evidenced through case study data.

KEY WORDS

technology in singing lessons, singing pedagogy, voice analysis
INTRODUCTION

In general, science and musical performance tend to use different language codes and coding for knowledge and often their ontological standpoints are different. Whilst it is not known to what extent these two language codes are reconcilable, there is no longer a widespread culture of technology phobia in non-scientific fields of human endeavour and the benefits from the application of technology have been demonstrated in many fields, including the arts.

Singing teachers are usually performers or ex-performers themselves and their pedagogical knowledge is often based on their own experience [1]. Craft knowledge such as this, although useful for the individual teacher, may mean that certain key features of performance are missed, which might be picked up by another teacher. Singing teachers draw on their personal experiences within an essentially hegemonic oral culture [2, 3]. Such experiences dominate and differentiate the language of singing pedagogy literature from that found in singing science texts.

The standard pedagogical model employed in the conservatoire studio typically involves weekly/twice weekly lessons with an expert, supported by private practice and performance. The teacher is engaged in a psychological translation of the student’s performance, for example by turning musical gestures into language, and the student is engaged in a further translation of the teacher’s verbal and visual feedback into adapted singing performance. A dual possibility thereby exists for the misinterpretation of information. Anything that can provide more robust and easily understandable feedback to both teacher and student would seem to be worthwhile, and this forms the basic premise behind this work.
The application and implementation of appropriately designed technology within the singing studio has the potential to provide both the teacher and student with relatively objective voice quality data for the assessment of progress [4]. Quantifiable parameters have already been identified that vary with training and experience for: (a) actors [5], (b) adult singers [6], (c) children [7], as well as (d) girl and boy cathedral choristers [8].

Real-time visual feedback has been previously used successfully with primary school children [9, 10] and adult singers [11-13]. This previous work has suggested that simple displays of a small number of analysis parameters are usually the most effective when learning performance skills. Our experience also suggests that technological applications are only of potential benefit if they are easy to use by non-specialists and provide information that is meaningful, valid and useful. Such robust information, associated with appropriate knowledge of results can then underpin feedback to provide assessments that are more accurate in formative and summative evaluation contexts.

The current exploratory project, known as “VoxEd”, adapted voice analysis and display techniques developed previously [11, 14] into a Windows application in a user-friendly environment, enabling real-time feedback of moment-by-moment activities of the vocal system. Such real-time feedback has already been shown to enhance the teaching and learning experience of singers by contributing positively to the reiterative cycle of performance and feedback in singing pedagogy.

Welch [15, 16] proposed a model based on a psychological analysis of vocal pitching to characterise the role of feedback in the learning process as illustrated in Figure 1. During the traditional teacher-student interaction, a target is provided by the teacher (e.g. verbal or sung), the student makes an attempt to imitate this vocally, and the teacher provides feedback. The gain the
student makes by utilising the teacher’s feedback depends on knowledge of what s/he is supposed
to be achieving in terms of the intended outcome, and the degree of correspondence between the
student’s subjective assessment and the external assessment from the teacher, the latter referred to
as “knowledge of results” or “KR”. Following the feedback, the student makes another attempt
(see Fig. 1-A), and this process of attempt followed by feedback forms the backbone of the
traditional singing pedagogical procedures.

By complete contrast, real-time visual feedback offers the possibility of providing feedback
during the student’s vocal response. The effect of modifications can be observed immediately and
concurrently with the vocal response (see Fig. 1-B). Apart from the more obvious advantage of
removing the time lag between a vocal response and the feedback, real-time provision enables the
student to make another attempt straight away as scrutiny of the feedback provided during the
previous attempt gives an immediate indication as to what needs to be altered or remain.

** FIGURE 1 ABOUT HERE **

PURPOSE

The VOXed project sought to evaluate the usefulness of technology in the singing studio and
whether this was of benefit to the teaching and learning of singing. The key research question
concerned how to provide real-time visual feedback embracing qualitative support in a singing
studio context which would be accessible, valid, robust, meaningful and useful to both teacher
and student alike in the teaching and learning of singing.

The main purpose of the project was not to design the technology itself for this application (given
the time constraints within the twelve-month externally-funded AHRB project), but rather to
work together with a Liaison Panel, singing teachers and students to adapt real-time visual feedback displays for voice which were already available. The adapted software was then provided on laptop computers with attached web-cams to two singing teachers for use over an extended period, and its impact on the singing pedagogical process was monitored with the teachers and students acting in the role of co-researchers (see below).

METHOD

Given the desire to root this work within current pedagogical practice in the singing studio, the first activity on this project was a one day workshop with a Liaison Panel made up of the authors, singing teachers, and invited colleagues whose research involves speech, singing, psychology, linguistics, vocal health, engineering, and education. The purpose of this workshop was to explore and share current singing pedagogical practice amongst the Liaison Panel, to introduce Liaison Panel members to examples of existing voice technology that could be used to provide visual feedback when vocalising, to review existing displays that might be useful in the context of the singing studio and to produce a specification for the software to be employed in the project. This meeting was organised with an agenda consisting of a set of specific research foci as shown in table 1.

**TABLE 1 ABOUT HERE**

The Liaison Panel also considered how real-time visual feedback might be able to assist the teaching of the many and varied aspects of vocal technique which have to be addressed by singing teachers. Aspects that were suggested by the Liaison Panel included the following:
- Voice quality (e.g. bright, mellow, breathy);
- Consonants;
- Vowel quality and length;
- Pharyngeal widening;
- Larynx position;
- Tongue positioning for vowels;
- Jaw position for pitch;
- Legato/staccato singing;
- Registers;
- Resonance;
- Head/neck alignment;
- General posture;
- Breathing.

This list was not intended to be an exhaustive inventory of all aspects of singing technique. Rather, it provided a summary of essential aspects which the Liaison Panel felt occurred frequently as pedagogical issues, and which were perceived as having the potential to benefit most through the use of a real-time visual display in the studio. It was acknowledged that there are other very important aspects of singing pedagogy for which real-time visual displays may have relatively little to offer to the singer, such as building a repertoire, sight-singing, use of language, working with a conductor, stagecraft and performance practice. Such areas are likely to remain the sole preserve of the singing teacher, albeit with increased support from researchers [17, 18].
Two experienced professional singing teachers were involved in the project as co-researchers, one based close to London in Guildford, and the other in York. Each teacher nominated four students to act as fellow co-researchers; two of these to use the VoxEd software and two to act as ‘control’ subjects to provide an indication of the efficacy and practice of the teacher’s current teaching techniques in the absence of technology.

The action research methodology employed for the VoxEd project had two main phases:

1. Introduction and design;
2. The gathering and analysis of a substantial amount of observational and participant data.

Phase one included the initial Liaison Panel meeting and an induction session for each teacher and their student co-researchers to introduce them to the Action Research Methodology [19, 20] and the software. Action Research was chosen as we wanted to empower participants to feel that they had direct control of the technology and its use. We were also very interested in their thoughts and feelings as teachers and learners as they engaged with the technology over a period of several months. Teachers and students were encouraged to provide the authors with continuous critical evaluation with regard to the implementation of the software in terms of its operation, appropriateness, context for use and user friendliness. As a result, ten software updates were undertaken during the lifetime of the project for bug-fixing and as a direct response to user comments, including:

- user-controlled variation of the colours associated with each display panel;
- inclusion of lines on spectrograms and spectra to show ratio frequency limits;
- the use of the spacebar to start and stop recordings;
- user setting of plot line thicknesses;
• automatic filename generation when saving audio files  
  (to reduce keyboard strokes and assure data integrity).

Singing lessons were regularly observed and video recorded by the research team and students and teachers alike were asked to keep journals of their teaching and learning experiences. At the end of the project, teachers and students were interviewed about their experiences. The software design allowed sound files to be recorded, saved and stored for playback or analysis at a later date and so these provided an additional data set.

A large amount of case study data was collected during the life-span of the exploratory project. This provided a rich evidence base of qualitative data as summarised in Table 2.

** TABLE 2 ABOUT HERE **

** WinSingad Software and Hardware **

The software used in the VOXed project was specially written as a Windows application using Microsoft Visual Studio C++ and it is known as WinSingad. It runs on any PC compatible machine under Windows XP, NT, or 98, and makes use of any installed audio card for audio data input and output. It has its origins in earlier software written for the BBC and Atari microcomputers [20, 9 respectively], known as SINGAD (SINGing Assessment and Development), to enable pitch matching skills to be assessed and developed in primary school children. The basic design philosophy behind WinSingad is to provide information in a manner that is easy to understand, clear and uncluttered, in order to aid rather than in any way to compromise the knowledge of results for the user (see Fig. 1).
WinSingad makes available a number of analyses of individual parameters plotted against time that can be set for use singly or in any combination, depending on the specific vocal activity being supported; a full description of the system is given in [22]. The graphical display of each analysed parameter is organised as a separate display panel within the main WinSingad window, and each display panel can be visible or hidden, moved up or down the screen relative to other display panels, enlarged with respect to any other selected display panels, changed in terms of the colours used for the plot, background, axes and text, changed in terms of the plot line thickness; in each case, any user controllable parameters relating to the processing algorithm can be accessed. It should be born in mind, though, that parameters that can be altered can have an effect on the output display that is inappropriate, and can lead to erroneous data interpretation [23].

The eight display panels currently available in WinSingad are:

- Input acoustic pressure waveform against time;
- Fundamental frequency against time;
- Short-term spectrum;
- Narrow band spectrogram;
- Spectral ratio against time;
- Vocal tract area;
- Mean/min vocal tract area against time;
- Real-time web camera window;

The teachers were, therefore, able to choose to use just a single display panel, or any combination of display panels up to having all eight display panels open together. In practice, teachers generally made use of just one or two of the display panels at any given time, enlarging the one of
specific interest as desired. The two participant teachers typically made use of different configurations of display panels; the York-based teacher generally used the spectrogram display panel and the vocal tract display (see Fig. 2), and the Guildford-based teacher used the spectrogram display and the web-camera, either together on one screen or separately on different occasions (see Fig. 3).

**FIGURES 2 AND 3 ABOUT HERE**

**DATA ANALYSIS**

The data collected in this project was quite varied as indicated in table 2. Analysis of the qualitative data was undertaken in various ways. Video recordings of singing lessons with and without the software were viewed and the salient features of each lesson mapped to a time-line and coded. All specific use of the WinSingad software in the teaching and learning sequence was noted.

Teacher and student lesson diaries and feedback interviews were transcribed and again salient features were noted. A number of variables in the data were observed as follows.

- two teachers, each with:
  - different histories and professional performer and teacher backgrounds
  - one male and one female
- two different pedagogical approaches:
  - using WinSingad to monitor all singing behaviour in lesson
  - using WinSingad to illustrate some specific point
  - use of particular WinSingad display panels
• different WinSingad usage behaviour:
  – regularly usage
  – irregular usage
  – usage dependent on categories of student e.g. professional, amateur

• displays within WinSingad
  – teachers have particular pedagogical foci
  – tendency to have preferred WinSingad display panels

• musical repertoire
  – tendency to favour Western high art/classical music

• students’ backgrounds
  – age
  – sex
  – previous singing experience
  – current vocal needs
  – upcoming performance schedule.

DISCUSSION

Careful analysis of the wealth of qualitative data gathered during the life-span of the project made it possible to begin to answer the specific research foci identified at the beginning of this project. Initial analysis demonstrated differences in the way that the two teacher co-researchers worked in their respective studios and made use of the technology as part of their singing pedagogy. For example, one teacher habitually used the display at all points throughout a specific lesson, whereas the other teacher would use the technology at specific points in a lesson, usually to illustrate or demonstrate certain aspects of vocal technique.
Results from the varied data sources are set out below to illustrate the discussion of research foci, such as diaries (coded D, with teacher/student number and lesson number indicated), video recordings of lessons (coded V, with teacher/student number and lesson number indicated), feedback interviews (coded I, with teacher/student number indicated) and Liaison Panel meetings (coded P, with meeting number indicated). The York- and Guildford-based teachers are coded T1 and T2 respectively.

This discussion is organised against the research foci that underpinned the work.

RF1. The extent to which technology will be accepted in the studio

Almost universally, the teachers and students involved in the project accepted and made regular use of the software in the singing studio. There were a few instances where reservation was expressed by individual students at the outset, but this turned out to be a consequence of initial uncertainty and apprehension over what was required of them and how the technology might be used in this situation. Over the period of the project, there was virtual unanimity on the usefulness of the technology in the lesson context and its user friendliness.

The two teachers used different elements of the technology for differing purposes and this related directly to their own preferred teaching strategy. Teacher T1 made use the software with students of all skill levels, and soon after the commencement of the project, he was using the software not only with the students involved as co-researchers, but also with the majority of his other students. He did this on ethical grounds because he believed that all the students would benefit from access to the technology outputs. Teacher T2 also thought that all students could gain much from using the software, but felt that it would not be accepted so readily by some less confident amateur
singers whose main reason for attending singing lessons was enjoyment, rather than as part of ongoing professional training.

Both teachers asked if it would be possible to continue using the technology after the end of the project, and it is hoped that follow-on can be organised. The majority of students exposed to the software were keen to use it in subsequent lessons, and many were keen to find out more about how it worked and/or what it was showing. Individual students also asked whether it would be possible to acquire copies of WinSingad for use in their private practice, and this has promoted the release of a freeware version².

T1: “But you understand what you’re doing now. So when you’re looking at that if you wish to, if you want to increase the resonance or see what you’re doing yeah?”

STUDENT: “mmm.”

T1: “Then you know when you think there [teacher points to front of head] then that will brighten the resonance the more intensely you think into there the brighter we get that yeah? [teacher points to singers formant region on spectrogram display] and this of course is the cords itself; the more fundamental sound which you’ve been doing beautifully right from the beginning. Do you follow me?”

STUDENT: “Yeah.”

T1: “So if you were singing and you knew the song and you weren’t looking at that [teacher points to sheet music] you could actually.”

STUDENT: “You could follow that.” [Student points to screen.]

T1: “Or you could actually make it do what you wanted it to do.”

[V:T1S4L1]

Student comments after using the software include:

---

2 ² A freeware version of the WinSingad software, offering a restricted set of four panels only, is available for download (www-users.york.ac.uk/~dmh8).
“I think it would be beneficial to singers, because it’s putting in literally black and white what hasn’t been there in black and white before” [I:S6].

“What I’ve found most fascinating about it was that is was able to distinguish difference in harmonics and resonance with basically any note.” [I:S5]

RF2. The ease-of-use of the technology in the studio and for private practice

The teacher co-researchers considered themselves not to be particularly computer literate; indeed, neither had ever made use of an external microphone or web camera previously. Apprehension was expressed by both teachers about having to contend with a computer in the context of a singing lesson, where their attention is fully focussed on the progress of the session itself. However, with minimal training, both found that they were able to use the laptop-based software successfully in the studio, and the systems never crashed during a singing lesson. The action research protocol used in this project meant that the teachers and student co-researchers were encouraged to give early and on-going feedback about the software and its use and to have some control over the process.

Teacher T1 reported at the end of the project that the software was easy to use and that he felt confident in setting it up and interpreting the displays with his students. He mainly used just two of the eight available display screens [I:T1]. Teacher T2 reported that the technology was easy to use and that navigation around the software was ‘surprisingly simple’ [I:T2]. The software overall was thought to be robust by both teachers and they both especially liked the in-built feature that the software would open up in exactly the same configuration as it was when last shut-down.
Neither teacher thought that students should use the software unsupervised or at home for private practice [P2]. However, they did indicate that if improvements were made to some of the displays and more pedagogical information were made available readily within the software package, then this might make it applicable in the future for supporting private practice.

RF3. The nature of the data offered by the technology

The data and feedback offered by the software was deliberately not tailored to a specific pedagogical model and was kept general enough to fit idiosyncratic teaching methodologies (as exampled by the teacher commentary above). Lesson observations indicated that each teacher interpreted the selected displays in differing but meaningful ways and successfully integrated the additional feedback into their teaching. In essence, the nature of the data may be considered as a variety of visual metaphors related to underlying physiological behaviours and acoustic outputs.

The vocal tract display was judged to be somewhat ambiguous by teacher T2, although teacher T1 was happy working with it and he used it regularly [P2].

Whilst teacher T2 used the spectrogram display in black and white, teacher T1 preferred to use this display in colour as this matched his thinking about vocal sound in terms of “dark warm sound” or “bright sound” [P2]. This clashed with previous thinking about interpreting colour versus black and white displays from a technical point of view [23], but it was clear that for this teacher in this situation, it was an appropriate use of the display.

Liaison Panel members felt that some singing teachers might need help in understanding and using some of the more complex display screens [P2]. In this research project, the software was given minimal background technical information relating to the analysis and nature of the data.
offered by the software, but it was agreed that further help and non-mathematical descriptions of
the science behind the displays and its relationship with vocal physiology and musicianship
would be appropriate for future releases of the software.

**RF4. The integration of the data into singing teaching and learning**

The two teachers used the software in different ways. Teacher T1 used it with all students, either
during the complete lesson or at specific moments to confirm the teacher’s or student’s
perceptions, to illustrate a particular point of vocal technique or to highlight a specific issue
arising from a particular note or phrase. Teacher T2 used the software at most points in the
singing lesson to address a number of different aspects of singing technique. However, she did
not use it with all students and was particularly keen not to expose it to beginners until they had
mastered sufficient basic skills. The spectrogram display was used to illustrate a number of facets
including: vowel quality, vowel length, consonant length, breath timing and breath length during
long phrases and vocal register transitions. The spectrogram was also found to be useful by this
teacher to illustrate constituent tone quality over a wide range of dynamics.

The web-camera proved useful for observing not only neck and head alignment, but also pharynx
widening, larynx position and jaw position during vowels [D:T2]. Indeed, just placing the web-
camera in position on its stand seemed to have a positive effect on the student’s posture, even if
the camera display screen was not in use.

*He said that he wasn’t looking but he changed his posture anyway.* [D:T2S5L1]

Teacher T2 made little use of the vocal-tract display panel as she felt that it was not consistent
enough in its output (i.e. the teacher felt that changes she had made to her own vocal tract shape
during sung vowels were not reflected accurately enough by the vocal tract area display on the screen). Teacher T1 on the other hand, made regular use of the vocal-tract display panel alongside the spectrogram display with certain students, especially when illustrating placement and vowels shaping. In general, teacher T1 used the spectrogram and ratio displays to monitor the student’s vocal performance throughout the lesson, freezing the screen at various points to illustrate moments where comment was felt to be needed to encourage improvement of a specific issue or to reinforce good performance practice. This is illustrated in Fig. 4 when the focus on vocal ‘brightness’ was perceived to have been achieved during the singing of a rising scale.

**FIGURE 4 ABOUT HERE**

Both teachers rarely used the fundamental frequency contour display. This was a consequence of a competent singing standard already having been achieved by the students involved in the project. Both teachers felt that this display would be potentially very useful with beginners or with students who had difficulty with note pitching, or exhibited sliding or swooping between notes [I:T1 and T2].

Later versions of the software included the capability to record, save data and play back audio data. This served as a useful way of logging a student’s vocal performance over time for progress monitoring or to enable the pinpointing of specific points in a piece or exercise which the teacher wanted to discuss in more detail, either earlier in the current lesson or from a previous lesson.

Teacher T1 felt that it was important that a student was interrupted as soon as possible after a mistake had been made in order that feedback could be given. He found it very useful to be able to freeze the screen easily in order that an extended discussion could follow the teacher on any particular point of technique in the student’s performance [I:T1].
RF5. The ease with which the data can be interpreted and utilised

Data interpretation was discussed at length during the Final Liaison Panel meeting. One Panel member surmised that “teachers are interpreters of information and the way they translate knowledge relates to their own experiences” [P2].

Initially, it was thought that a simple visual display would be the most useful [P1 & 19] and the majority of the display panels available in WinSingad were initiated with this specification. In practice, the real-time spectrogram, which is one of the most complex visual displays offered, was the display most commonly used by both teachers. However, although there is a large amount of detail in a spectrogram, much of the interpretation by the teachers was based on a holistic approach to the observation of the patterns making up the spectrogram rather than an in-depth analysis of micro-level detail. In particular, spectrograms were used to provide feedback on breathing patterns, phrasing, vocal quality, projection, dynamics and vibrato. Teachers employed their musical performance knowledge to interpret patterns on the spectrogram to facilitate feedback and discussion of issues relating, for example, to legato singing, consonant sloppiness, “bright” or “dark” tone colours.

The potential and actual “ambiguity” of the displays to the teacher and student were perceived as both a strength as well as a weakness, in that the interpretation of a specific display by the teacher, although not always strictly correct in its scientific justification, may nevertheless prove useful to the student and help the student to improve. We suggest that that the strengths of any such ambiguity in terms of what was available to the learning process probably outweigh the weaknesses. In the words of teacher T2: “it doesn’t matter whether the student understands why
thick or thin bands appear on the spectrogram, but I can direct what I want him to achieve, e.g. I want more wiggles” [I:T2].

Teacher T1 agreed that all pupils who used the software were enthusiastic about its use as they could recognise that it backed-up the singing teaching that they had already received. For example, by forming a particular thought, the student was able to alter the image on the screen; this seems to correspond with the teacher’s pedagogical approach which involves thinking the “right things” to make the “right sound” [I:T1].

A specific illustration emerged from the second lesson of student four (S4) whilst using the software and working with teacher T1 and piano accompanist (P1), also a singing teacher who contributed to the teaching process. Here the teacher was concerned by the lack of perceived resonance in the first sung syllable of the word “Jerusalem” and suggested that the student (a tenor) uses a more open vowel for the first syllable (i.e. “jar” instead of “jer”). The student attempts this phrase again and the teacher monitors the output on the spectrogram display.

T1: “That wasn’t what I expected.”
P1: “I think [student 4] is not really saying ‘ah’; it's an ‘ay’ shape.”
S4: “I’m saying what?”
P1: “You’re saying an ‘ay’ shape on Jerusalem rather than ‘ah’.”
S4: “Instead of Jar?”
P1: “So it’s not gonna come out as inadequate as ‘er’ would sound”

[V:T1S4L2]

The teacher was surprised by what was displayed on the spectrogram as it is not what he expected. The student attempts the same phrase again, but this time the piano accompanist points
out that the student is not pronouncing the first syllable vowel exactly as directed by the teacher. This example illustrates that the visual feedback software cannot be a replacement for the listening skills of the teacher, although it can help the teacher to monitor what is happening during the student’s vocal performance.

RF6. How the technology impacts upon the learning and teaching experience

There is no evidence to suggest that the technology overly intrudes into the teaching and learning experience. Indeed, it was found so helpful by teacher T1 that he used it with many more of his students than just the four designated as co-researchers. Since the technology had demonstrated its potential efficacy to students, the teacher felt it was his ethical duty to use it with other students. Indeed, he also reported that occasionally students would ask to use the software during lessons when he was not intending to make use of it.

The action research methodology and the portfolio of screen displays (with their options) enabled all participants to decide where and when they made use of the technology in the teaching and learning process. No prescription as to when or how the software was to be used was either offered by the research team, or designed into the software. However, the software was designed such that, on exit, it automatically saved the current configuration of display panels that make up the screen display, along with all user settings, so that it started up in the same condition for the next teaching and learning session.

Both teachers used the technology during all aspects of singing lessons, including: vocal warm up, vocal exercises and the rehearsal of specific pieces. In no aspect of the singing lesson was the software reported to be an intrusion. On the contrary, its impact was essentially perceived as positive by participants. In addition, the authors believe that continued use of the technology was
also beneficial in expanding the understanding of the teachers concerning the acoustic realities being displayed.

Teacher T2 felt that using the software should be perceived as a natural progression in the use of other teaching aids. So, although the main tools used by a singing teacher are their ears, eyes and voice, the teaching process can often be enhanced by the use of mirrors, tape recorders and video cameras. The used of a computer to provide real-time visual feedback is a logical next step in this series. Additionally, the use of such technology in singing lessons can also “fire the imagination and keep the student interested” [P2], simply because it involves the use of a computer.

Teacher T1 felt that using the technology had not overtly altered his pedagogical approach or made any difference to his teaching process. On the other hand, teacher T2 reported that her teaching had become more “logical and objective” and made her think in more physical terms about what she was doing [I:T2].

RF7. Potential perceived threats posed through the use of technology

Both teachers expressed initial reservations about being in control of a computer during their lessons, as neither felt particularly confident with such technology. However, following an initial introduction and demonstration session with a researcher, and the presence of a researcher in a few lessons, both began to make use of the system on a regular basis and their confidence rapidly grew.

One student initially saw the technology as a threat to achieving his specific goal of preparing a piece for performance in the short time-scale available.
“He asked not to use it in a particular lesson, as he felt it would take more time and he had lots of music to get through” [I:T2].

However, after he had had a demonstration and seen other students working with the technology, he then saw it as a benefit and asked to use it also.

**CONCLUSIONS**

This paper has considered the use of technology in singing lessons from a variety of standpoints, related to its overall acceptance, its usefulness, and the relationship of the teacher and students to this additional feedback sources. The project deliberately employed an action research methodology in which the teachers and students acted as co-researchers and were seen to be in control of this additional feedback source. Previous research on the application of real-time visual feedback informed the design of the software that was specially developed for this project and known as WinSingad. WinSingad allows the user to select individual displays, which are presented in the program window as individual display panels. Each display panel can be configured in terms of the colour of the plot line, text, axes and background, the vertical position of each display panel relative to other display panels, and any processing parameters relating to the analysis shown within the display panel. One display panel can be made larger than the others for a more detailed view. All these settings are saved when the user exits from the program and this is the condition in which the program is launched next time.

A number of issues concerning the application of technology in the singing studio were identified as research foci, following an Initial Liaison Panel Meeting. The purpose of the work was not to optimise the displays provided by the technology; rather it was to establish whether or not technology could be usefully be applied in the context of singing lessons. These research foci
were considered during the application of WinSingad with two professional singing teachers using an action research methodology where the teachers, their students and the researchers contributed equally to the data gathering and evaluation process. Overall, fears about using a computer in this context were fully allayed, and all teachers and students involved indicated keenness to continue to use the software during the trial and after it run its course.

A number of issues have been identified with respect to any future more widespread application of technology in the singing studio as follows.

- Both teachers involved in this study are familiar with and have craft-based knowledge of vocal physiology and voice acoustics. We recognise that other singing teachers may not have such background and experience, so future versions of any software applied in singing lessons should be sufficiently flexible to address the likely variable knowledge base of the user groups.
- Help “balloons” or equivalent would provide the user with information that they could access as desired about functions, menus, displays and the science behind the visual feedback.
- The vocal tract area display needs further development to make it more robust and more representative of the true shape of the vocal tract.
- The accommodation of metaphors in singing pedagogy needs further research. Teachers’ choices of linguistic terms may be idiosyncratic, even when observing the same acoustic phenomena being displayed. More research is needed to understand this pedagogical variety and its potential impact on the understanding of acoustic displays.
- The nature and process of translation of psychologically-focused instructions given by a teacher to the student needs to be better understood and analysed.
• The development of the underlying science of vocal technique needs further expansion.
• The development of a system to identify “healthy” voice use in order to reduce threats to vocal performance would be a significant enhancement.
• A follow-on research project that extends the group of co-researchers would provide additional experience on which to build. The use of an expert advisory Liaison Panel was considered to be a real strength of the research design.
• A robust method should be developed for performance profiling of students’ singing abilities.
• Investigation of the applicability of new signal processing algorithms for this application, such as human hearing modelling and wavelet analysis, might enhance the information being provided.

The application of real-time visual displays in the professional singing studio has been shown to be of potential and actual benefit to the teaching and learning process, and not to intrude deleteriously into the quality of teacher-student interaction. Ultimately, the intention is that the application of technology in this way will serve to support the teaching and learning process for those issues for which it is intended. Technology will never replace the professional singing teacher working on the key issues that make one singer stand out above another, such as stage presence, musical interpretation, repertoire planning, advice on which roles to perform, working with conductors and orchestras and rehearsal technique. Technology will be but another tool in the armoury to be used as needed to keep the instrument in proper healthy working order.

ACKNOWLEDGEMENTS
The VOXed project is supported by the UK’s Arts and Humanities Research Board (AHRB) under an Innovation Award numbered B/IA/AN8885/APN15651. The authors are indebted to the students who took part in this experiment, whom they thank most sincerely.

REFERENCES


FIGURE AND TABLE CAPTIONS

Figure 1: An illustration of the learning process for pitch in singing based on Welch [15], [16]

KEY: (A) the on-going traditional learning process, and (B) the way in which real-time visual feedback can impact the learning process.

KR = knowledge of results from an external source; CP = critical learning period; time is from left to right.

Figure 2: WinSingad displays: spectrogram (upper) and vocal tract area (lower) for a scale on /u:/.

Figure 3: WinSingad spectrogram display for an arpeggio on [a:] with a webcam image superimposed.

Figure 4: Spectrogram (upper screen) and ratio (lower screen). {Teacher T1: “Can you see, you got the brightness in your rising scale as well” [V:T1S4L1]}

Table 1: Project research foci.

Table 2: Project evidence base.
Figure 2: An illustration of the learning process for pitch in singing based on Welch [15], [16]

KEY: (A) the on-going traditional learning process, and (B) the way in which real-time visual feedback can impact the learning process.

KR = knowledge of results from an external source; CP = critical learning period; time is from left to right.
Figure 2: WinSingad displays: spectrogram (upper) and vocal tract area (lower) for a scale on /aː/. 
Figure 3: WinSingad spectrogram display for an arpeggio on [a:] with a webcam image superimposed.
Figure 4: Spectrogram (upper screen) and ratio (lower screen). [Teacher T1: “Can you see, you got the brightness in your rising scale as well” [V:T1S4L1]}
<table>
<thead>
<tr>
<th>RF1</th>
<th>The extent to which technology will be accepted in the studio.</th>
</tr>
</thead>
<tbody>
<tr>
<td>RF2</td>
<td>The ease-of-use of the technology in the studio and for private practice.</td>
</tr>
<tr>
<td>RF3</td>
<td>The nature of the data offered by the technology.</td>
</tr>
<tr>
<td>RF4</td>
<td>The integration of the data into singing teaching and learning.</td>
</tr>
<tr>
<td>RF5</td>
<td>The ease with which the data can be interpreted and utilised.</td>
</tr>
<tr>
<td>RF6</td>
<td>How the technology impacts upon the learning and teaching experience.</td>
</tr>
<tr>
<td>RF7</td>
<td>Potential perceived threats posed through the use of technology.</td>
</tr>
</tbody>
</table>

*Table 1: Project research foci.*
<table>
<thead>
<tr>
<th>Who</th>
<th>Observational Data</th>
<th>Qualitative Data</th>
<th>Recorded Data</th>
</tr>
</thead>
<tbody>
<tr>
<td>LIAISON PANEL</td>
<td>Panel members’ input at initial and final meetings</td>
<td>Video recording of Initial and Final Liaison Panel meetings</td>
<td></td>
</tr>
<tr>
<td>TEACHERS</td>
<td>Observation of sample lesson by teacher</td>
<td>Teacher journals</td>
<td>Video recording of sample lesson by teacher</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Teacher interview</td>
<td>Video recordings of student lessons with/without software</td>
</tr>
<tr>
<td>STUDENTS</td>
<td>Observation of student lessons with/without software</td>
<td>Student journals</td>
<td>Sound files recorded through software</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Student interviews</td>
<td>Video recordings of student lessons with/without software</td>
</tr>
</tbody>
</table>

*Table 2: Project evidence base.*