

# Are Real-Time Displays of Benefit in the Singing Studio? An Exploratory Study

\*David M. Howard, \*Jude Brereton, †Graham F. Welch, †Evangelos Himonides,  
‡Michael DeCosta, ||Jenevora Williams, and §Andrew W. Howard

\*‡York, †London, ||Guildford, and §Cambridge, United Kingdom

---

**Summary:** This article 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 optimize the technology for this application, but to work alongside teachers and students to study the impact of real-time visual feedback technology use 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, coded for teacher and student behaviors, and all co-researchers were interviewed at the mid- and endpoint 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, scientists and musicians tend to express themselves differently when describing

common experiences, and often their underlying standpoints are different. Although it is not known to what extent these two language codes might be

---

Accepted for publication September 27, 2005.

Presented at the Voice Foundation's 33rd Annual Symposium: Care of the Professional Voice, June 2–6 2004, Philadelphia, Pennsylvania.

From the \*Media Engineering Research Group, Department of Electronics, University of York, Heslington, York, United Kingdom; †School of Arts and Humanities, Institute of Education, University of London, Bedford Way, London, United Kingdom; ‡York, United Kingdom; ||Guildford, United Kingdom; and §SARAND Business IT Solutions, Haverhill, Cambridge, United Kingdom.

Supported by an Innovations grant (B/IA/AN8885/APN15651) from the U.K. Arts and Humanities Research Board (AHRB) in 2003–2004.

Address correspondence and reprint requests to Professor David M. Howard, Head of the Media Engineering Research Group, Department of Electronics, University of York, Heslington, York, YO10 5DD, UK. E-mail: [dh@ohm.york.ac.uk](mailto:dh@ohm.york.ac.uk)

*Journal of Voice*, Vol. 21, No. 1, pp. 20–34  
0892-1997/\$32.00

© 2007 The Voice Foundation  
doi:10.1016/j.jvoice.2005.10.003

reconcilable, there is more widespread use of technology in nonscientific activities, including the arts.

Singing teachers are usually performers or ex-performers, and their pedagogical knowledge is often based on their experience.<sup>1</sup> 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 oral culture, where expertise is handed down from teacher to student generation by generation.<sup>2,3</sup> 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; the teacher may not describe the student's performance appropriately, and/or the student may not understand how to modify his or her singing behavior as intended. 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.

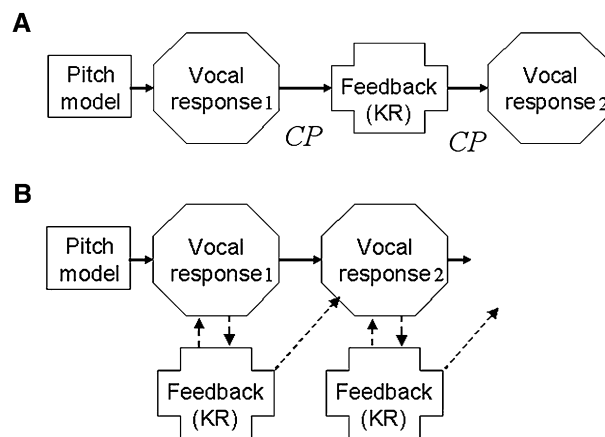
The application and implementation of appropriately designed technology within the singing studio has the potential to provide both the teacher and the student with objective voice quality data for the assessment of progress.<sup>4</sup> Quantifiable parameters have already been identified that vary with training and experience for (1) actors,<sup>5</sup> (2) adult singers,<sup>6</sup> (3) children,<sup>7</sup> as well as (4) girl and boy cathedral choristers.<sup>8</sup>

Real-time visual feedback has been used successfully with primary school children<sup>9,10</sup> and adult singers.<sup>11-13</sup> 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<sup>11,14</sup> into a Windows application (Microsoft Corporation, Redmond, WA) 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<sup>15,16</sup> 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 (eg, verbal or sung), the student makes an attempt to imitate this vocally, and the teacher provides feedback. The gain the student makes by using the teacher's feedback depends on knowledge of what s/he is supposed to be achieving in terms of the



**FIGURE 1.** An illustration of the learning process for pitch in singing based on Welch.<sup>15,16</sup> KEY: (A) the on-going traditional learning process, and (B) the way in which real-time visual feedback can impact the learning process. Abbreviations: KR, knowledge of results from an external source; CP, critical learning period; time is from left to right.

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." After the feedback, the student makes another attempt (see Figure 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 Figure 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.

### PURPOSE

The VOXed project sought to evaluate the use and usefulness of a specially designed technology package 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.

Although the main purpose of the project was not to design the technology for this application (given the time constraints within the 12-month externally funded AHRB project), this was a subsidiary research question as no existing package was available that was suitable for the prime purpose. The lead researchers worked together with a liaison panel, singing teachers, and students to adapt real-time visual feedback displays for voice that were already available. The adapted software was then provided on laptop computers with attached Web cameras 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

The project was innovative in design through its application of social science and arts performance research methodologies in an applied scientific context. Action research<sup>17,18</sup> was selected as a key element to ensure full participant engagement throughout the project. Action research empowers participants to feel that they have some ownership of the research; their views are taken into account in a formal manner as part of the data collection process.

Three research questions were addressed, as follows:

1. What are key features of the teaching of singing in a studio setting that might be addressed by the use of technology?
2. Given the possibility of defining such features, is it possible to design a technology package that could be used?
3. What are the effects of technology use in the singing studio on teacher and student behaviors?

The first question was addressed through an initial consultation with a group of experts (the liaison panel) and subsequently in a short-term longitudinal study in two singing studios. The second was addressed through an iterative process between the core research team and two teachers over a period of 4 months. The third question was addressed by systematic observation, interviews, and analysis of the journals of the participants who were acting as co-researchers.

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 technology 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.

#### Research question 1: key features of teaching singing in a studio

Given the desire to root this work within current pedagogical practice in the singing studio, the first

**TABLE 1.** *Project Research Foci (RF1-RF7)*

RF1	The extent to which technology will be accepted in the studio.
RF2	The ease-of-use of the technology in the studio and for private practice.
RF3	The nature of the data offered by the technology.
RF4	The integration of the data into singing teaching and learning.
RF5	The ease with which the data can be interpreted and used.
RF6	How the technology impacts on the learning and teaching experience.
RF7	Potential perceived threats posed through the use of technology.

activity on this project was a 1-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 among the liaison panel, to introduce liaison panel members to examples of existing voice technology that could be used to provide visual feedback when vocalizing, 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 organized with an agenda consisting of a set of specific research foci as shown in Table 1.

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 that have to be addressed by singing teachers. Aspects that were suggested by the liaison panel included the following:

- Voice quality (eg, bright, mellow, breathy)
- Consonants
- Vowel quality and length
- Pharyngeal widening
- Laryngeal position
- Tongue positioning for vowels
- Jaw position for pitch
- Legato/staccato singing
- Registers
- Resonance
- Head/neck alignment
- General posture
- Breathing

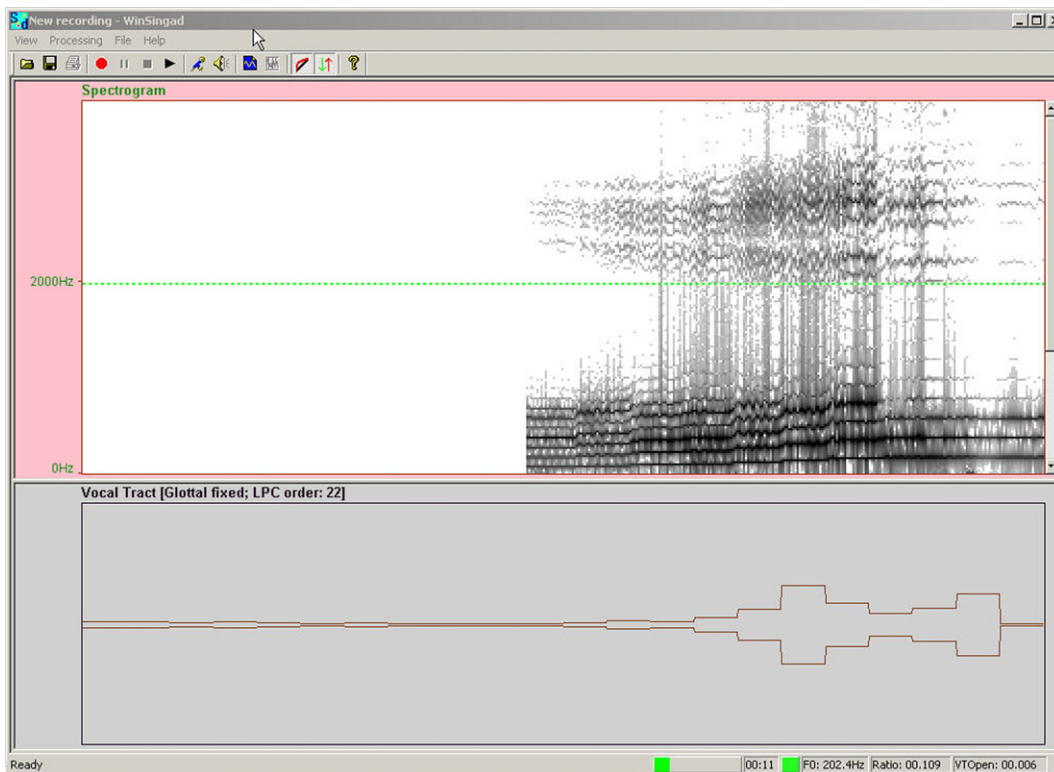
This list was not intended to be an exhaustive inventory of all aspects of singing technique. Rather,

it provided a summary of essential aspects that the liaison panel felt occurred frequently as pedagogical issues, and that 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.<sup>19,20</sup>

### **Research question 2: design of a software package**

A decision was taken early on to make the software available for any PC compatible machine running Windows XP, NT, or 98 that has any installed audio card for audio data input and output. The software developed specially for the VOXed project was written as a Windows application using Microsoft *Visual Studio C++*, and it is known as *WinSingad*. This name betrays its origins in earlier software written for the Atari<sup>9</sup> and BBC<sup>21</sup> microcomputers, known as *SINGAD* (SINGing Assessment and Development), which was designed 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, to aid rather than in any way to compromise the knowledge of results for the user (Figure 1). *WinSingad* makes available several analyses of individual parameters plotted against time that can be set for use singly or in any combination, depending on the specific



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

vocal activity being supported; a full description of the system is given in Howard et al.<sup>22</sup> These analyses were chosen with reference to the comments made by members of the liaison panel at their first meeting with respect to displays that were likely to be useful for supporting the key features of the teaching of singing considered in question 1.

The graphical display of each analyzed 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 colors 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.<sup>23</sup>

The eight display panels currently available in *WinSingad* are as follows:

1. Input acoustic pressure waveform against time
2. Fundamental frequency against time
3. Short-term spectrum
4. Narrow band spectrogram
5. Spectral ratio against time
6. Vocal tract area
7. Mean/min vocal tract area against time
8. Real-time Web camera window (this could be saved to disk for future playback)

The teachers, therefore, could 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 (Figure 2), and the Guildford-based teacher used the spectrogram display and the Web camera, either together on one screen or separately on different occasions (Figure 3).

The core research team was particularly interested in the thoughts and feelings of the teachers and learners as they engaged with the technology over a period of 4 months. Teachers and students provided the core team with continuous critical evaluation with regard to the implementation of the software in terms of its operation, appropriateness, context for use, and user friendliness. This resulted in 10 software updates being undertaken during the lifetime of the project, enabling modifications to be made in rapid response to user comments on the software operation as well as bug-fixing. These included:

- User-controlled variation of the colors 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)

### DATA ANALYSIS

The evidence base collected in this project that provided the raw material for data analysis was available in several forms as indicated in Table 2. The collection and analysis of these data provided the answer to research question 3.

#### Research question 3: effects of technology on students and teachers

Singing lessons were regularly observed and video recorded by the research team, and students and

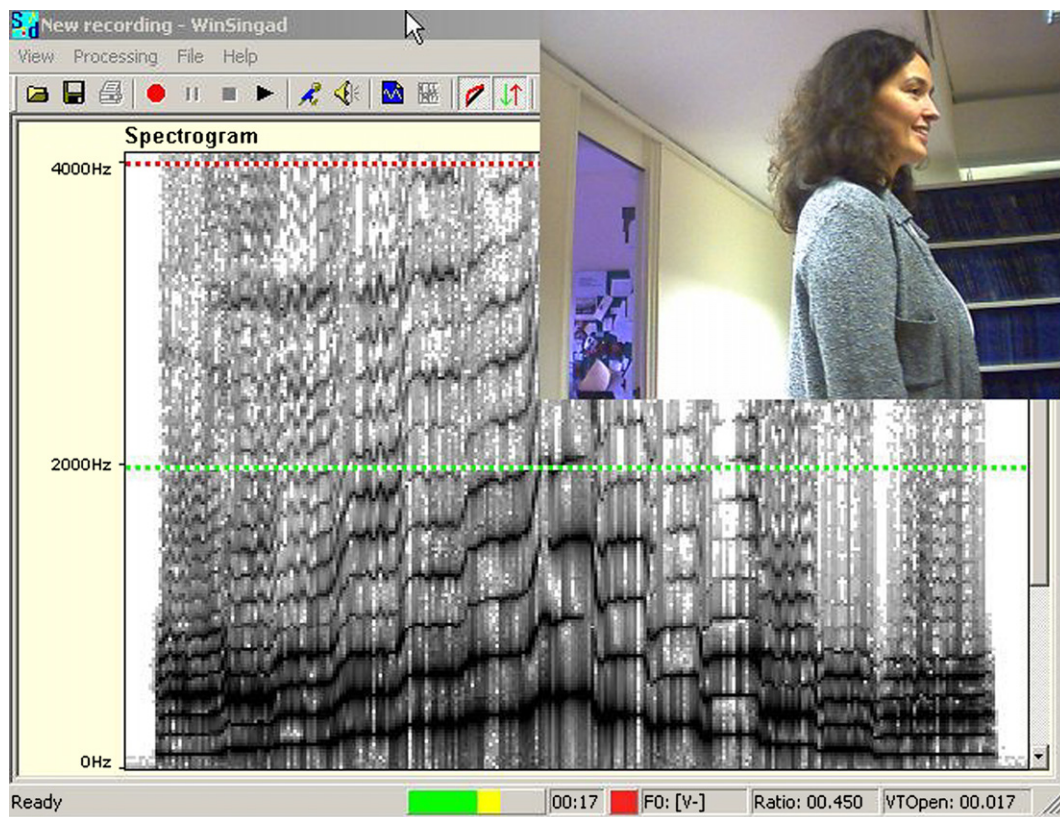


FIGURE 3. WinSingad spectrogram display for an arpeggio on [a:] with a Web camera image superimposed.

TABLE 2. Project Evidence Base

Who	Observational Data	Qualitative Data	Recorded Data	Quantitative Data
Liaison panel		Panel members' input at initial and final meetings	Video recording of initial and final liaison panel meetings	
Teachers	Observation of sample lesson by teacher	Teacher journals	Video recording of sample lesson by teacher	Second-by-second analysis of activities during lessons
		Teacher interview	Video recordings of student lessons with/without software	
Students	Observation of student lessons with/without software	Student journals	Sound files recorded through software	
		Student interviews	Video recordings of student lessons with/without software	Second-by-second analysis of activities during lessons

teachers 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 dataset.

The evidence base collected in this project was varied as indicated in Table 2 and reflected the multi-method research design. Analysis of the data was undertaken in various ways.

Teacher and student lesson diaries and feedback interviews were transcribed, and again salient features were noted. Several variables in the data were observed as follows:

1. Two teachers:
  - a. different histories and professional performer and teacher backgrounds
  - b. one male and one female
2. Two different pedagogical approaches:
  - a. Using *WinSingad* to monitor the majority of singing behaviors in lessons
  - b. Using *WinSingad* to illustrate some specific point
  - c. Use of particular *WinSingad* display panels
3. Different *WinSingad* usage behavior:
  - a. Regularly usage
  - b. Irregular usage
  - c. Usage dependent on categories of student, eg, professional, amateur

4. Displays within *WinSingad*:
  - a. Teachers have particular pedagogical foci
  - b. Tendency to have preferred *WinSingad* display panels
5. Musical repertoire:
  - a. Tendency to favor Western high art/classical music
6. Students' backgrounds:
  - a. Age
  - b. Gender
  - c. Previous singing experience
  - d. Current vocal needs
  - e. Upcoming performance schedule

In addition, digital video recordings of singing lessons with and without the software were made. These were viewed, and the salient features of each lesson mapped to a time-line and coded. Four video recordings of lessons, two for each teacher, each working with the same student, were subjected to a detailed second-by-second analysis of teacher and student behavior. One lesson for each teacher did not involve the use of the *WinSingad* system because these lessons occurred early in the work before the software had been made available, which is referred to as "Lesson 1." The other lesson for each teacher was with the same student in each case, and it occurred later in the project when the *WinSingad* system was used, which is referred to as "Lesson 2." The results for Lesson 1 provide

a control (no technology) observation against which the results for Lesson 2 (when technology was available) can be directly compared. The detailed analysis of these lessons is reported in Welch et al,<sup>24</sup> and a summary is provided in Table 3. For notational convenience, the York teacher and Guildford teacher are coded T1 and T2, respectively.

Welch et al<sup>24</sup> carried out a statistical analysis of the data presented in Table 3, and they note the following:

- The basic pedagogical approach of each teacher is different as shown by a comparison of the time spent during Lesson 1 on the various activities (chi-square analysis based on an analysis of time spent on each activity in minutes:  $\chi^2 = 14.41, df 5, P < 0.025$ ). T1 uses an accompanist to “warm-up” the singer. T2 makes more use of formal instruction and of marking particular features during the lesson.
- The teachers responded differently to having the technology available (compare the percentages in Lesson 1 (where there was no technology) with Lesson 2 when technology was used). There was no significant difference in the distribution of time by T1 across the two sample lessons (chi-square analysis based on an analysis of time spent on each activity in

minutes:  $\chi^2 = 10.41, df 6, n.s.$ ), whereas T2 exhibits a large increase in the percent of time demonstrating and instructing with the technology (chi-square analysis based on an analysis of time spent on each activity in minutes:  $\chi^2 = 26.45, df 6, P < 0.001$ ).

- The teachers continued to be significantly different in terms of their pedagogical approach during Lesson 2 when the technology was available (chi-square analysis based on an analysis of time spent on each activity in minutes:  $\chi^2 = 41.63, df 6, P < 0.001$ ).
- It is of note that when the technology was available throughout Lesson 2, T1 (who became a keen advocate for its use) only made reference to it for approximately one third of the available time (37.6%), whereas T2, who reported that she was selective in her employment of the feedback technology, spent nearly two thirds of the time (60.6%) in Lesson 2 using it.

### 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

**TABLE 3.** Percentages of Time Spent by Each Teacher on Different Activities During Two Lessons Are Presented

Activity	T1 (York)		T2 (Guildford)			
	Lesson 1	Lesson 2	Lesson 1	Lesson 2		
	No Technology Available % of Lesson	Technology Not Being Used % of Lesson	Technology Being Used % of Lesson	No Technology Available % of Lesson	Technology Not Being Used % of Lesson	Technology Being Used % of Lesson
Accompanying	59.9	36.5	0.0	26.2	4.8	0.5
Conversation	23.7	8.1	21.3	27.2	6.5	10.8
Demonstration	0.0	0.5	4.9	0.0	5.1	6.8
Instruction	4.1	0.3	1.3	16.3	13.0	33.6
Listening	1.5	0.0	2.4	9.0	0.0	1.1
Feature marking	8.0	6.7	2.4	12.0	8.1	1.3
Not related to lesson	2.8	10.3	5.3	9.3	1.9	6.5
<b>Total</b>	<b>100.0</b>	<b>62.4</b>	<b>37.6</b>	<b>100.0</b>	<b>39.4</b>	<b>60.6</b>

Notes: Lesson 1 occurred early in the project before *WinSingad* had been introduced and provides a control that is indicative of typical teaching behavior without technology. Lesson 2 occurred later when the use of *WinSingad* was well established, and it shows the proportion of time spent using and not using the technology. Data are summarized from Ref. 23.



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.

This discussion is organized against the research foci that underpinned the work, and the examples from the varied data sources that are set out below to illustrate the discussion of research foci include:

- 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.

### **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 of 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 most of his other students. He did this on ethical grounds because he believed that all students would benefit from access to the technology outputs.

Teacher T2 also thought that all students could gain much from using the software, but he 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 on-going professional training.

Both teachers asked whether it would be possible to continue using the technology after the end of the project, and it is hoped that follow-on can be organized. Most 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 version. [A freeware version of the *WinSingad* software, offering a restricted set of four panels only, is available for download (<http://www-users.york.ac.uk/~dmh8>).]

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:

“*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 it 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. However, with minimal training, both found that they could 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 built-in 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 were deliberately not tailored to a specific pedagogical model; the software is provided as a tool to be

used when the teacher feels it is appropriate. The output from the software was kept general enough to fit idiosyncratic teaching methodologies (as exemplified 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 the teaching process.

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].

Although teacher T2 used the spectrogram display in black and white, teacher T1 preferred to use this display in color 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 color 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, users were given minimal background technical information relating to the analysis and nature of the data offered by the software, and the liaison panel suggested 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 several 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 several 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.

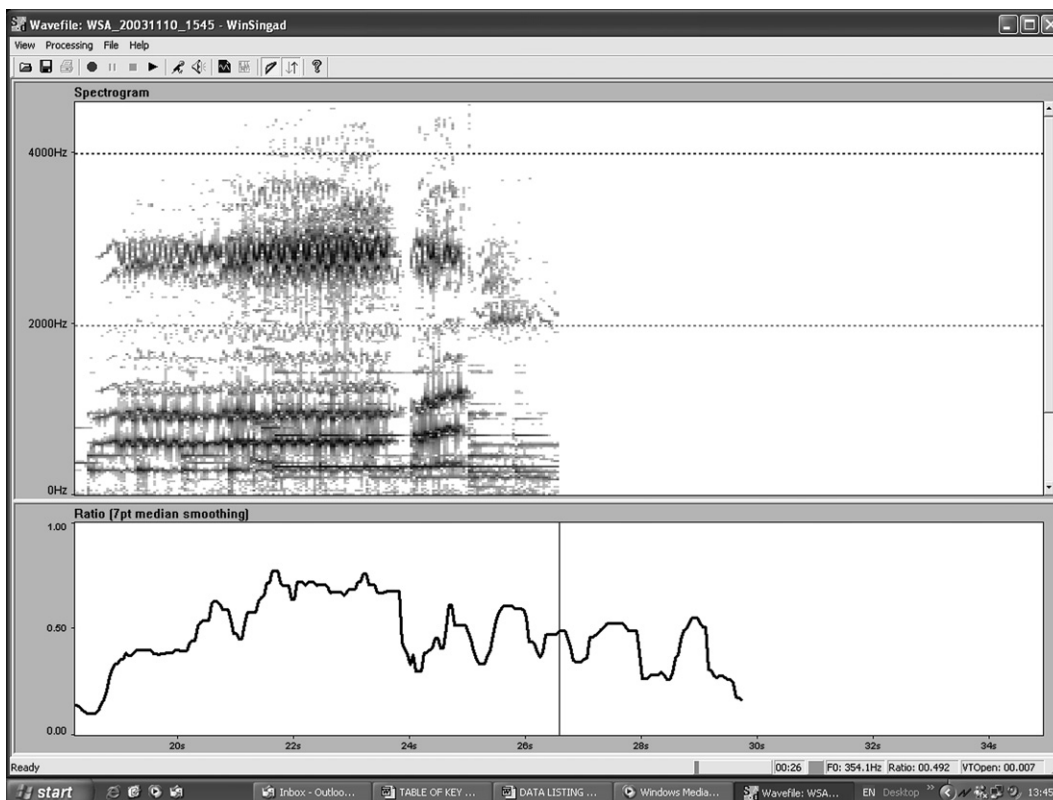
The Web camera proved useful for observing not only neck and head alignment, but also pharynx widening, larynx position, and jaw position during vowels [D:T2]. Indeed, just placing the Web camera in position on its stand seemed to have a positive effect on the student's posture, even if the camera display 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 (ie, 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 Figure 4 when the focus on vocal "brightness" was perceived to have been achieved during the singing of a rising scale.

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



**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]}.

who 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 utilized**

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<sup>23</sup> [P1], and most 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, and "bright" or "dark" tone colors.

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 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 recognize that it backed-up the singing teaching that they had already received. For example, by forming a particular thought, the student could 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) while using the software and working with teacher T1 and a piano accompanist (P1), who was 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 (ie, "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 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. As 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 use 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, after 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**

At the outset, several issues concerning the application of technology in the singing studio were identified as research foci, after an initial liaison panel meeting. The purpose of this work was not to optimize the displays provided by the technology; rather it was to establish whether technology could be usefully 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.

Several 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 recognize 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 analyzed.
- The development of the underlying science of vocal technique needs further expansion.
- The development of a system to identify “healthy” voice use 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 modeling 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 armory to be used as needed to keep the instrument in proper healthy working order.

**Acknowledgments:** The authors are indebted to the students who took part in this experiment, whom they thank most sincerely. They also thank the reviewers for their most helpful and insightful suggestions on earlier drafts.

## REFERENCES

1. Olsson B. The social psychology of music education. In: Hargreaves DF, North AC, eds. *The Social Psychology of Music*. Oxford, UK: Oxford University Press; 1997:290–305.
2. Callaghan J. Singing teachers and voice science—an evaluation of voice training in Australian tertiary institutions. *Research Studies Music Ed*. 1998;10:25–41.
3. Potter J. *Vocal Authority*. Cambridge, UK: Cambridge University Press; 1998.
4. Welch GF. The assessment of singing. *Psychol Music*. 1994;22:3–19.
5. Rossiter DP, Howard DM. Observed change in mean speaking voice fundamental frequency of two subjects

- undergoing voice training. *Logoped Phoniatr Vocol.* 1998; 22:187–189.
6. Howard DM. Variation of electrolaryngographically derived closed quotient for trained and untrained adult female singers. *J Voice.* 1995;9:163–172.
  7. Howard DM, Angus JAS. A comparison between singing pitching strategies of 8 to 11 year olds and trained adult singers. *Logoped Phoniatr Vocol.* 1997;22:169–176.
  8. Welch GF, Howard DM. Gendered voice in the cathedral choir. *Psychol Music.* 2002;30:102–120.
  9. Howard DM, Welch GF. Visual displays for the assessment of vocal pitch matching development. *Appl Acoust.* 1993; 39:235–252.
  10. Welch GF, Howard DM, Rush C. Real-time visual feedback in the development of vocal pitch accuracy in singing. *Psychol Music.* 1989;17:146–157.
  11. Rossiter D, Howard DM, DeCosta M. Voice development under training with and without the influence of real-time visually presented feedback. *J Acoust Soc Am.* 1996;99: 3253–3256.
  12. Thorpe CW, Callaghan J, Doorn JV. Visual feedback of acoustic voice features for the teaching of singing. *Australian Voice.* 1999;5:32–39.
  13. Nair G. *Voice Tradition and Technology: A State-of-the-Art Studio.* San Diego, CA: Singular; 1999.
  14. Rossiter D, Howard DM. ALBERT: a real-time visual feedback computer tool for professional vocal development. *J Voice.* 1996;10:321–336.
  15. Welch GF. Variability of practice and knowledge of results as factors in learning to sing in-tune. *Bull Council Res Music Ed.* 1985;85:238–247.
  16. Welch GF. A schema theory of how children learn to sing in tune. *Psychol Music.* 1985;13:3–18.
  17. Bray JN, Lee J, Smith LL, Yorks L. *Collaborative Inquiry in Practice: Action, Reflection, and Making Meaning.* Thousand Oaks, CA: Sage; 2000.
  18. Burnaford GE, Fischer J, Hobson D, eds. *Teachers Doing Research: The Power of Action Through Inquiry.* 2nd ed. Mahwah, NJ: Lawrence Erlbaum Associates; 2001.
  19. Williamon A, ed. *Musical Excellence, Strategies and Techniques to Enhance Performance.* Oxford: Oxford University Press; 2004.
  20. Juslin PN, Sloboda JA, eds. *Music and Emotion.* Oxford: Oxford University Press; 2001.
  21. Howard DM, Welch GF. Microcomputer-based singing ability assessment and development. *Appl Acoust.* 1989; 27:89–102.
  22. Howard DM, Welch GF, Brereton J, Himonides E, DeCosta M, Williams J, Howard AW. WinSingad: a real-time display for the singing studio. *Logoped Phoniatr Vocol.* 2004;29:135–144.
  23. Howard DM. The real and non-real in speech measurements. *Med Eng Phys.* 2002;24:493–500.
  24. Welch GF, Howard DM, Himonides E, Brereton J. Real-time feedback in the singing studio: an innovatory action-research project using new voice technology. *Music Ed Res.* 2005;7:225–249.