Music Education Research
Vol. 7, No. 2, July 2005, pp. 225–249

Routledge
Taylor & Francis Group

Real-time feedback in the singing studio: an innovatory action-research project using new voice technology

Graham F. Welch[a]*, David M. Howard[b], Evangelos Himonides[a] and Jude Brereton[b]
[a]*Institute of Education, University of London, UK;* [b]*University of York, UK*

The article reports on a one-year AHRB-funded Innovations project that was designed to evaluate the usefulness, or otherwise, of the application of real-time visual feedback technology in the singing studio. The basis for the research was a multi-disciplinary approach that drew on voice science and acoustics, the psychology of singing and voice education. Participants were based in two different singing studios, one in the north of England and the other in the south. They catered for two different adult singer client groups ranging from skilled amateur to advanced professional. An action-research methodology was adopted in which the two participant singing teachers and their adult students were seen as co-researchers in the research activity. The resultant research data consisted of research diaries, observations and interviews, supplemented by multimedia recordings (audio and video) of actual singing behaviours over time. Data analyses indicate that new technology can impact positively on teacher behaviours and student experiences by providing more meaningful feedback through an enriched pedagogy. This offers the possibility of expanding the professional knowledge and skill base of both groups.

Introduction

Singing is a ubiquitous human behaviour that has a central place in the vast majority of the world's musical cultures, whether traditional, classical (high art) or popular. There are many diverse forms of singing evidenced, but each draws on a basic capacity of the human voice to produce sustained sounds that have social, cultural and personal significance for the performer and the audience (who may be one and the same). Much singing (and musical) behaviour develops within informal settings as part of a nurturing and enculturation process in which the mother's vocal behaviour has a prime influence (Welch, 2005). The infant's earliest vocalisations are essentially melodic and have a bi-potentiality for subsequent speech or singing. Normally, singing behaviours develop and become more advanced, competent and complex across childhood and the majority of young people will enter adolescence

*Corresponding author. School of Arts and Humanities, Institute of Education, University of London, WC1H 0AL UK. Email: g.welch@ioe.ac.uk

ISSN 1461-3808 (print)/ISSN 1469-9893 (online)/05/020225-25

DOI: 10.1080/14613800500169779

with a range of singing competences almost irrespective of the quality of any formal school music education that they have experienced (Mang, 2003; Welch, in press, a, b). Although a small minority of the child and adult populations will have had less positive experiences about their singing and may believe that this means that they are somehow 'less musical' than their peers, there are a range of empirical studies over many years to suggest that singing behaviours can be improved through appropriate experiences (e.g. Welch, 1985, 2001; Obata, 2004; Hall, 2005).

Central to any pedagogical nurturing of singing development is the provision of some form of meaningful feedback to the developing singer. Younger, less experienced singers may have had some initial difficulties in copying exactly the musical features of an example song from the local musical culture. Perceptually, the song text (lyrics) may be the principal focus and the accompanying musical features of the song (pitch, rhythm, melodic structure, phrasing, implied harmony) may be relatively ignored.[1] The resultant singing behaviour appears to be relatively 'untuneful' or lacking in some essential performance features. The teacher's role in such cases is to unpack this complexity and to allow the developing singer to focus on mastery of particular elements, such as melodic contour. With appropriate guidance, including feedback that makes sense in relation to the initial behaviour, it is possible to effect change and improve singing skills, for adults as well as for children (Richards & Durrant, 2003; Obata, 2004; Hall, 2005).

At a more advanced level, there is evidence that teachers of singing customarily use imagery (including kinaesthetic and visual imagery) in teaching vocal technique, often allied to a reliance on sensation and the development of aural awareness (Callaghan, 1998). However, their professional background is usually that of a successful performer, perhaps with some basic knowledge of an underlying voice science that has been gained from casual reading and/or attendance at conferences of professional associations (Callaghan, 1998). Teaching is essentially practice-focused, supported by linguistic imagery and (in some cases) by vocal and postural modelling. These techniques are employed in an iterative process (encouraging and responding to elements of singing behaviour, such as focus melodic fragments, phrases, or complete songs) within the one-to-one context of the lesson in the singing studio. In general, whilst singing pedagogy is relatively poorly documented (in relation to systematic, theoretically founded research, such as demonstrated in recent studies by Mitchell *et al*., 2003; Mitchell & Kenny, 2004), such evidence as does exist indicates that it is often highly idiosyncratic and based on semitransparent methods that are likely to be driven by the teacher's own experience and personal reflections, both as 'learner' and performer. Their pedagogical knowledge is often based on their own experience (Olsson, 1997), but such craft knowledge (cf. Tschannen-Moran *et al*., 1998), although useful for that teacher, may miss certain key features of performance that would be picked up by another teacher. Singing teachers draw on their personal experiences within an essentially hegemonic oral culture (Callaghan, 1998; Potter, 1998). Such experiences dominate and differentiate the language of singing pedagogy literature from that found in texts on the science of singing.

There are two features of the 'traditional' singing lesson that may be potentially confusing or even antithetical to the optimal development of the student, namely the nature of the singing teacher's pedagogic vocabulary and the nature of the teaching process.

The singing teacher's pedagogic vocabulary

Customarily, the teacher of singing uses language from within the art form to promote self-awareness and the possibility of change concerning the student's singing behaviour (e.g. Christiansen, 2005; Keenze & Bell, 2005). Common terms within this professional vocabulary relate to 'support' (for sustained control of the breath), '*appoggio*', 'posture', 'focus' (for a 'good tone'), 'relaxation', 'chest/head' resonance, 'forward placement', 'open throat', '*passaggio*', 'blending registers', '*messa di voce*', 'ring' and 'vowel modification'. These terms have a professional currency, yet as metaphors may also be ambiguous in relation to the underlying coordination of anatomy and physiology, as well as concomitant vocal acoustic output. Detailed investigation of the professional practice of fifty teachers of singing in Australian higher education institutions (Callaghan, 1997) revealed a significant disparity between singing pedagogy and voice science. Although teachers had a strong commitment to experiential learning and individual development, they often demonstrated an incomplete knowledge of underlying vocal physiology and acoustics, leading to the possibilities of misinformation during the singing lesson.

The nature of the teaching process

Certain features of the teaching process may compound any potential weaknesses in the nature of the linguistic discourse within the singing lesson. Typically, the singing teacher is engaged in a psychological interpretation and translation of the student's performance. The teacher's perception of the student's singing behaviour is translated from a musical gesture into linguistic form. The challenge for the student is for teacher's '*post-hoc*' verbal feedback to be interpreted and translated subsequently into an adapted singing performance. This process happens over time as the lesson progresses, normally with the student and teacher taking turns, one initiating and the other responding. A dual possibility thereby exists for the misinterpretation of information (i.e. from student singing behaviour to teacher language and from teacher language to further student singing). This type of pedagogical process can be represented as follows (see Figure 1).

Consequently, anything that can assist the learning process in the provision of more robust, less ambiguous and easily understandable feedback to both teacher and student would seem to be worthwhile. For example, it is possible to modify the feedback process in teaching to enable the learner to have more real-time feedback of the nature of their singing behaviour (see Figure 2). In the late 1980s, the effectiveness of this latter type of feedback was evaluated across an initial seven-week

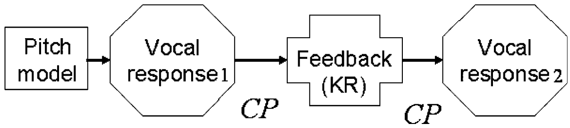

Figure 1. A traditional model of teacher feedback (Knowledge of Results [KR]) during the course of a singing lesson. Feedback from the teacher occurs after the student's vocal response (1) and is designed to have some influence on the subsequent response (2). CP indicates possible critical periods in the processing of information prior to and after the provision of feedback and any subsequent action (Welch, 1983, 1985)

period and then after a further six months with a class of seven-year-old children in a primary school (Welch *et al*., 1989). The class was divided into three matched groups in terms of singing competency, with two experimental groups using a simple computer-based visual feedback system that was designed to foster improvement in their vocal pitch accuracy, whilst a control group followed a more traditional singing curriculum. Results indicated that the provision of a real-time visual metaphor of vocal pitch change on the computer screen enabled children to 'see' and act on the melodic behaviour of their voices. Furthermore, real-time visual feedback promoted increased vocal pitch accuracy compared to a more traditional 'post-hoc' verbal feedback (*pace* Figure 1).

More recently, the significance of real-time feedback in musical learning through the use of technology has been explored in studies of physiological self-regulation of conservatoire students, embracing biofeedback and neurofeedback (Gruzelier & Egner, 2004) and also in studies of expressivity in performance (Juslin *et al*., 2004). Collectively, these studies indicate that feedback can be beneficial in the reduction of muscular tension in string and wind players, as well as increased attention and enhanced emotional expressivity in performance.

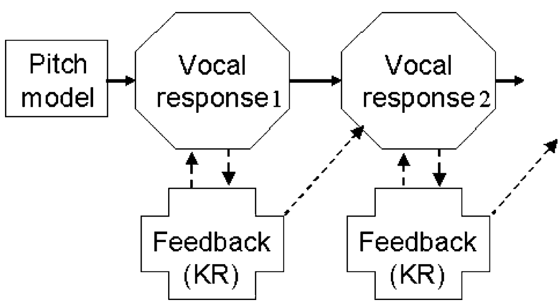

Figure 2. A real-time model of feedback (Knowledge of Results [KR]) during the course of a singing lesson. Feedback is provided in real-time during the student's vocal response (1) and is designed to be immediately accessible in order to be able to influence the subsequent response (2). Possible critical periods in the processing of information prior to the provision of feedback and any subsequent action are eradicated in this model (Welch, 1983, 1985; Welch *et al*., 1989)

Intrapersonal and interpersonal feedback in singing

Feedback in singing is both intrapersonal and interpersonal. In addition to the interpersonal feedback from teacher to student and vice versa, or extrinsic feedback from some other source (such as technology), the developing singer communicates intrapersonally in a variety of ways related to the nature and design of the internal feedback system. Feedback can be auditory, visual, tactile, kinaesthetic or vestibular (Welch, 1985; Gabrielsson, 2003). As well as being used as a basis for ongoing perceptual interaction with the surrounding world, feedback is also used in the construction of individual musical identity (Hargreaves *et al*., 2002). At one level, there is an internal psychological feedback system that is essentially outside conscious awareness and which relates to a moment-by-moment self-monitoring of singing behaviour (cf. the enactment of 'vocal plan formation'—Peretz & Coltheart, 2003). A schema theory of singing development (Welch, 1985) proposed that any initiation of a specific singing behaviour (termed 'voice programme'), such as copying or reproducing an external song model, would generate internal system expectations of proprioceptive and exteroceptive feedback that are compared to the actual feedback received from the sense receptors and auditory environment respectively (including both bone- and air-conducted sound).

This internal feedback system also provides the basis for self-reflective psychological judgments as to the 'appropriateness' of any given example of singing behaviour, such as its correspondence to an external song model or to an internal mental representation of a target melody's key, tonal relationships, loudness and/or timbre. In the absence of evaluative feedback from an external source (termed 'Knowledge of Results' [KR]), the singer has to make their own judgment of the 'appropriateness' of their sung response compared to their internal model. This comparison is likely to depend on the relative developments within and between their 'musical' and 'phonological' lexicons (cf. Peretz & Coltheart, 2003) as conscious judgments, in the sense of being able to articulate the self-judgment, require the use of language. Furthermore, there is also evidence that accurate reproduction of songs from the dominant culture requires a combination of a range of musical and linguistic skills (Davidson, 1994; Welch *et al*., 1996, 1997, 1998).

In some cases, there will be a realisation of a mismatch between the intended and actual singing behaviour and a subsequent correction can take place. Awareness, however, is not a necessary guarantee of improved vocal accuracy or singing development. 'Out-of-tune' singing can persist, for example, because singers do not know how to change their behaviour, even though they may realise that something is 'incorrect' or 'inappropriate'. It can also persist because there is no awareness that their singing behaviour needs to change. At a conscious, reflective level, the singer's intrapersonal communication (verbal and non-verbal) is a form of self-monitoring that is essential for the development of skilled performance behaviour of diverse pieces in a wide variety of acoustic contexts. Adjustments, both mental and in physical coordination, may need to be made as the performer moves from the individual context of the singing studio to the more public rehearsal environment and on to the demands of the actual performance when stress levels may be higher (Gabrielsson, 1999) due to the efferent stimulation of the adrenal gland (Rossi, 1993; Thurman, 2000; Sapolsky, 2003).

In addition, there are other context effects. Performance behaviours are subject to social and cultural imperatives, as shown in classical singing styles by a shift in emphasis from vocal agility in the eighteenth century to vocal resonance in the late nineteenth century (Mason, 2000) and by different cultural stylistics in operatic performance (Rosselli, 2000).

Practice, particularly deliberate practice, may be regarded as an essential feature of intrapersonal monitoring of singing competency and the development of performance expertise (cf. Lehmann, 1997). Nevertheless, whilst it is possible for an individual to become more and more consistent in their singing behaviour, the absence of meaningful extrinsic qualitative feedback from an outside source can lead to the rehearsal of minor (or even significant) idiosyncratic behaviours (termed 'subjective reinforcement'—Welch, 1985) that interfere with optimal sound quality.

The vocal instrument

If the provision of real-time visual feedback in the singing studio is to be effective, it needs to relate the resultant acoustic output as closely as possibly to the behaviour of the underlying anatomy and physiology of the voice mechanism. Whilst there is an extensive research literature on the acoustics of singing (e.g. Sundberg, 1987; Howard, 1999), the nature of the scientific discourse often uses a different language code to that of the performing artist or singing teacher. Furthermore, scientific research continues to explore the boundaries of established knowledge, often in a laboratory setting, and there can be a considerable time lag between research, publication and its application in real-world contexts. Additionally, the definition of 'quality' in singing continues to be somewhat elusive, although there is evidence that particular acoustic features are likely to find collective favour perceptually by singing pedagogues, even if they may disagree as to the nature of the underlying vocal behaviour (Mitchell & Kenny, 2004).

The vocal instrument consists of three basic elements (see Figure 3): (i) the respiratory system provides the energy source for the voice, (ii) the voice source itself, which are the vocal folds within the laryngeal assembly that vibrate in the airstream from the lungs to generate the basic sound and (iii) the vocal tract, which is essentially defined by the spaces above the larynx—the pharyngeal space within the neck and the oral cavity, sometimes complemented by the nasal cavity—which shape the sound (cf. Welch & Sundberg, 2002). To make voiced sound, the respiratory system compresses the lungs to generate an upward flowing airstream that sets the edges of the vocal folds in vibratory motion, resulting in pulsed sound waves that travel (mainly) through the vocal tract to be radiated outwards from the lips. Changes in *vocal pitch* are a product of variations in the mass

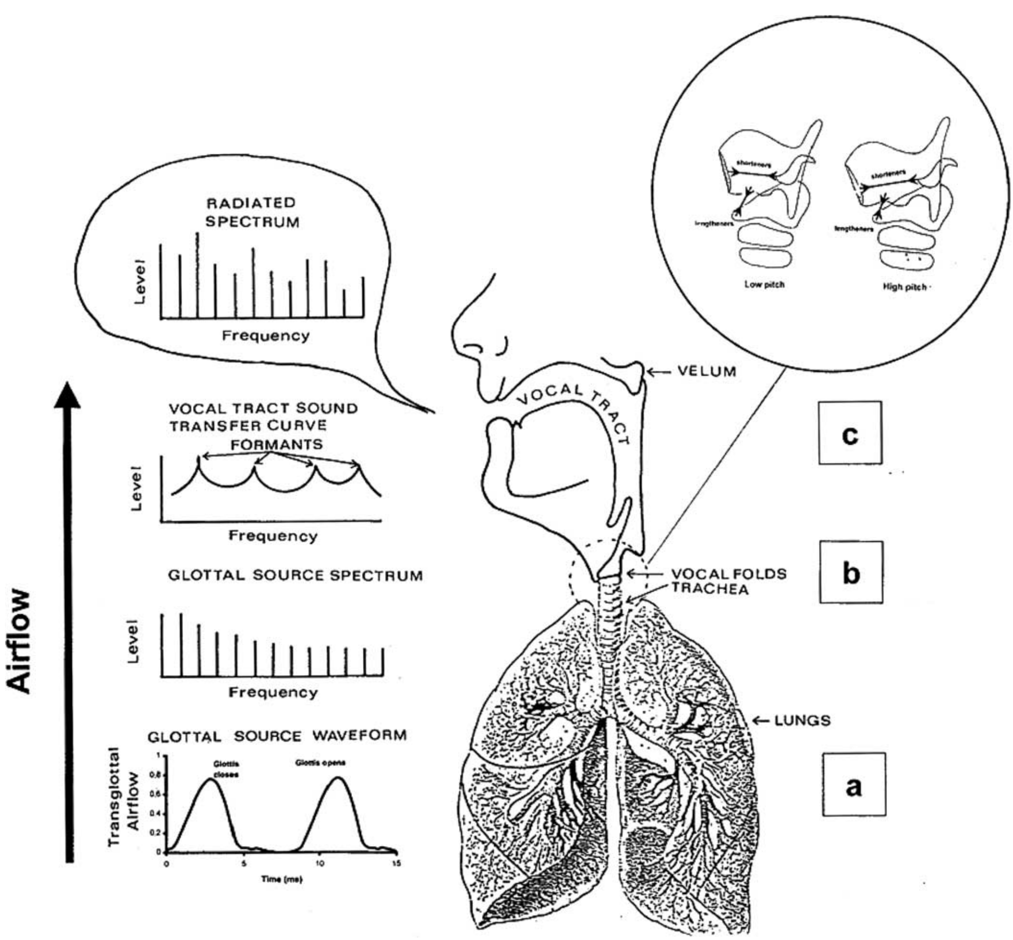

Figure 3. The basic anatomy and physiology of singing and its acoustic effects derived from (a) the respiratory system, (b) the vocal folds and (c) the vocal tract (Welch & Sundberg, 2002, adapted from Sundberg, 1987)

and length of the vibrating vocal folds that arise from the relative interactive contraction of two sets of internal laryngeal muscles (see top right insert in Figure 3).

Changes in *vocal loudness* are mainly the result of changes in air pressure from the lungs: the higher the pressure, the louder the voice. Professional singers are very consistent in their use of the respiratory system, but there is no standard single type of breathing behaviour across singers (Thomasson, 2003). Vocal *timbre* is a product of the interaction between the vibratory motion of the vocal folds, allied to a particular vocal tract configuration that amplifies or dampens certain components of the resultant complex sound wave, i.e. enhancing or suppressing some of the simultaneously sounding pure tones (cf. Sundberg, 1996).

A general feature of voiced sound is that there are peaks in the spectrum of the sound that is radiated from the lips (see Figure 3). These peaks are known as *formants*, created by vocal tract resonances that appear at certain frequencies which enhance particular harmonics of the complex waveform emanating from the vibration of the vocal folds. There are five formants that are crucial to vocal communication and our perception of voiced sound. The relationship between the lowest two formants (F1 & F2) gives rise to our labelling of sounds as 'vowels' and are generally dependent on jaw opening, lip protrusion, larynx height and tongue shape respectively. The relationships between the other three formants (F3, F4 & F5) relate primarily to vocal colour and also to the carrying power of the voice. When the vocal tract is configured to cluster these three upper formants together (usually by opening the pharynx and lowering the larynx), there is a particular energy peak created, known as the singer's formant cluster (Sundberg, 1974, 1987). This is a form of natural amplification that allows the singer's voice to be heard with relatively little effort above the sound of a full orchestra. It is also the most sensitive region perceptually in the auditory spectrum (Hunter & Titze, 2005). The relative spectral placement and strength of the formants is also implicated in perceptions of the 'placement' of the singing voice, such as being either 'forward' ('in the mask') or 'backward' (Vurma & Ross, 2003). 'Forward placement' is usually regarded as a more ideal vocal quality for classical singing performance (Emerich *et al*., 1997) and can be achieved by increased jaw opening and moving the tongue forward, thus increasing the energy of the first two formants (F1 and F2) in the frequency spectrum and the relative power of the 'singer's formant cluster' (F3, F4 and F5).

It is one thing to be able to read and understand the above explanation, but it is a different challenge to be able to recognise these features and to manipulate and sustain optimal singing behaviour systematically in the rehearsal studio and performance venue. The discourse of the singing lesson is unlikely to embrace detailed features of voice acoustics. It is important, therefore, that singing teachers and their students have the opportunity to know that they are sharing common perceptions and conceptualisations of the student's singing behaviours.

***VOXed*: the application of technology for real-time feedback in the singing studio**

Research evidence indicates that it is possible to quantify and display aspects of voice production and to identify specific features that vary with age, sex and experience in diverse populations, such as for actors (Rossiter *et al*., 1995a), adult singers (Rossiter *et al*., 1995b, Rossiter & Howard, 1998; Howard, 1995; Sundberg, 1996) and trained and untrained child singers (Welch & White, 1993/4; Welch & Howard, 2002). Furthermore, as mentioned above, real-time visual feedback has been previously used successfully in an exploratory study with primary schoolchildren (Howard & Welch, 1993; Welch *et al*., 1989) and adult singers (Rossiter *et al*., 1996; Thorpe *et al*., 1999).

However, with the partial exception of two primary school studies (Welch, 1985; Welch, Howard & Rush, 1989), the provision and use of feedback technology had tended to be driven by the technology designers within carefully controlled parameters. In the past, participants have tended to be placed in the role of relatively passive recipients in researcher-controlled, laboratory-driven evaluations rather than being put in the position where they are able to be active agents with significant control of the technology and its use.

Accordingly, in order to research alternative pedagogical approaches and to explore the extent to which advanced voice science technology could be incorporated in a meaningful way in the singing studio, research funding was made available by the UK's Arts and Humanities Research Board (AHRB) under its 'Innovations' programme. The project (*VOXed*) was funded for twelve months across the academic year 2003–2004. This included one academic term of singing studio-based development of the selected technology, followed by one term of consistent system use and evaluation.

The project had several novel features: (i) an action-research methodology was adopted in which the teachers and the students had complete control over the use, or otherwise, of the technology during their lessons; teachers and students were in the formal position of co-researchers and were regularly consulted by the main project team; (ii) although there was not time to design new technological applications, the researchers worked closely with an expert panel of singing teachers, engineers and voice scientists to identify potentially useful forms of feedback from a range of possible systems and then employed an expert software designer to create an integrated software interface (termed *WinSingad*). The system was carefully designed to run on any modern desktop or laptop PC (windows)-based platform. The software interfaced with the PC's native soundcard and was capable of analysing sound captured by a microphone. An external plug-and-play web camera was also used for visual feedback regarding the singer's posture and general physical gestures whilst performing. The software was designed with a standard and advanced set of menu options that allow the user to decide the particular focus for the real-time visual feedback by selecting one or more of a number of display panels (see example in Figure 4).[2] The panel displays provide information on the input waveform, fundamental frequency against time (for information on vocal pitch and melody), short-term spectrum, narrow band spectrum and spectral ratio against time (for information about the quality and consistency of vocal timbre, as well as the carrying power of the voice), and vocal tract areas and mean/min vocal tract area against time (for information about the relative size of inner resonating cavities). In practice, subsequent use indicated that singing teacher participants tended not to employ the full range of displays on screen at once, but were more selective (see below).

The research project embraced seven main research questions: (i) the extent to which teachers and students accepted and made use of technology in the studio; (ii) the ease-of-use of the technology, both in the studio and elsewhere for private

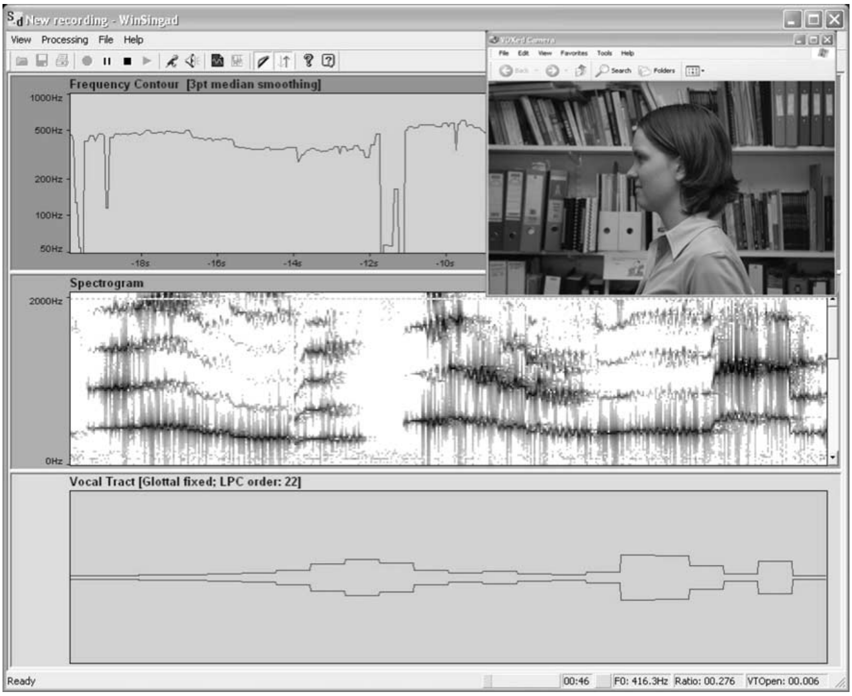

Figure 4. An example of the WinSingad screen designed for the VOXed project showing three panels from the available menu, plus a web-cam picture. In this example, the upper panel provides a display of a frequency (pitch) contour, the middle panel displays spectral energy (vocal timbre) over time in a spectrogram, whilst the lowest display models the vocal tract. The web-cam is usually located to provide a side-view of the singer's posture. All input (digital audio & video) can be saved and replayed

practice; (iii) the nature of the data offered by the technology; (iv) the extent to which the data could be integrated into singing teaching and learning; (v) the readiness with which the data could be interpreted and utilised; (vi) whether the technology overly intruded into the learning and teaching experience; and (vii) whether there was any potential/actual perceived threat posed to the teacher and/or the student by the use of technology.

There were two contrasting sites for the research, one in the north of England (York) and one in the south (Guildford). The project was designed so that each location had one volunteer singing teacher and a group of four participant students. Of these, two students volunteered to use the VOXed software in lessons and two acted as 'controls', in the sense that the latter were to have their lessons without the software. As well as data on singing saved by the software itself, the VOXed team gathered qualitative observational data during the lessons, supplemented by semi-structured interview data outside lessons and teacher and student journal records.

Research outcomes

Teacher and student behaviours in sample lessons

In addition to the journals of the perceptions of teachers and students, detailed real-time observations of sample studio-based lessons for each teacher were supplemented by subsequent analyses of digital video recordings of the same lessons. A second-by-second log was created of teacher and student behaviours across the lesson. The lessons ranged from 2614 seconds (43.56 min) to 3490 seconds

Table 1. Analyses of teacher time in seconds and percentages (%) in example lessons, with and without technology. Lesson 1 was an early fieldwork observation prior to technology being introduced (and assumed to be a baseline teacher behaviour for each focus student). Lesson 2 was a later observation when the use of VOXed technology was established. Lesson 2 is subdivided to show the percentages of time that the teacher had the technology available, differentiated by whether it was used (+tech) or not (−tech) within that lesson

	Teacher 1 (north): analysis of lesson time					
	Lesson 1		**Lesson 2**			
	No technology		*With* technology			
Activity	minutes	*% of time*	minutes (−tech)	*% of time (−tech)*	minutes (+tech)	*% of time (+tech)*
Accompanying	**26.1**	*59.9*	**21.3**	*36.5*	**0.0**	*0*
Conversation	**10.3**	*23.7*	**4.8**	*8.1*	**12.4**	*21.3*
Demonstration	**0.0**	*0*	**0.3**	*0.5*	**2.9**	*4.9*
Instruction	**1.8**	*4.1*	**0.2**	*0.3*	**0.8**	*1.3*
Listening	**0.6**	*1.5*	**0.0**	*0*	**1.4**	*2.4*
Marking a feature	**3.5**	*8*	**3.9**	*6.7*	**1.4**	*2.4*
Not related to lesson	**1.2**	*2.8*	**6**	*10.3*	**3.1**	*5.3*
Grand Total	**43.6**	*100*	**36.4**	*62.4*	**22.0**	*37.6*
	Teacher 2 (south): analysis of lesson time					
	Lesson 1		**Lesson 2**			
	No technology		*With* technology			
Activity	minutes	*% of time*	minutes (−tech)	*% of time (−tech)*	minutes (+tech)	*% of time (+tech)*
Accompanying	**13.9**	*26.2*	**2.6**	*4.8*	**0.3**	*0.5*
Conversation	**14.4**	*27.2*	**3.5**	*6.5*	**5.9**	*10.8*
Demonstration	**0.0**	*0*	**2.8**	*5.1*	**3.7**	*6.8*
Instruction	**8.6**	*16.3*	**7.1**	*13*	**18.2**	*33.6*
Listening	**4.8**	*9*	**0.0**	*0*	**0.6**	*1.1*
Marking a feature	**6.4**	*12*	**4.4**	*8.1*	**0.7**	*1.3*
Not related to lesson	**4.9**	*9.3*	**1**	*1.9*	**3.5**	*6.5*
Grand Total	**52.9**	*100*	**21.4**	*39.4*	**32.8**	*60.6*

Table 2. Proportions of two students' time in seconds and percentages (%) in two sample lessons, with and without technology. Lesson 1 was an early fieldwork observation prior to technology being introduced (and assumed to be a baseline for student behaviour), Lesson 2 was a later observation when the use of VOXed technology was established. Lesson 2 is subdivided to show the percentages of time that the student made use of the technology (+tech) or not (−tech) during that lesson

	Student 1 (north): analysis of lesson time					
	Lesson 1		**Lesson 2**			
	No technology		*With* technology			
Activity	minutes	*% of time*	minutes (−tech)	*% of time (−tech)*	minutes (+tech)	*% of time (+tech)*
Conversation	**10.3**	*23.7*	**9.0**	*15.5*	**2.2**	*3.8*
Listens	**5.3**	*12.1*	**0.7**	*1.2*	**1**	*1.7*
Performs	**26.7**	*61.3*	**40.4**	*69.5*	**3.6**	*6.2*
Not related to lesson	**1.2**	*2.9*	**0.3**	*0.5*	**0.9**	*1.6*
Grand Total	**43.6**	*100*	**50.5**	*86.7*	**7.7**	*13.3*

	Student 2 (south): analysis of lesson time					
	Lesson 1		**Lesson 2**			
	No technology		*With* technology			
Activity	minutes	*% of time*	minutes (−tech)	*% of time (−tech)*	minutes (+tech)	*% of time (+tech)*
Conversation	**14.4**	*27.2*	**6.9**	*12.7*	**1.7**	*3.1*
Listens	**15**	*28.3*	**6.3**	*11.6*	**5.9**	*10.9*
Performs	**18.6**	*35.2*	**11.7**	*21.6*	**17.3**	*32*
Not related to lesson	**4.9**	*9.3*	**2.5**	*4.6*	**1.9**	*3.5*
Grand Total	**52.9**	*100*	**27.3**	*50.5*	**26.8**	*49.5*

(58.16 min). These microanalyses were then synthesised into larger, discrete behavioural categories and recorded as minutes and percentages of the total lesson time (see Tables 1 and 2). The categorisation drew on previous research that had employed specially devised music pedagogy observation protocols (Kennell, 1997; Lanipekun, 2004 (unpublished manuscript); Ward, 2004).

Teacher behaviour was classified into seven main types:

- accompanying the student,
- conversation,
- demonstration,
- instruction,
- listening,
- marking a feature and
- activity not related to the lesson (such as non-singing focused discussion).

When the technology was fully available and each teacher had been inducted into its use, towards the end of Term 1 and throughout Term 2 of the fieldwork phase, a further distinction was made between behaviour that made use of technology and that which did not (see Lesson 2 columns in Table 1). The table provides examples of how the two participant teachers used their time in sample lessons. 'Lesson 1' presents an analysis of a lesson that was sampled (observed and videotaped) early in the fieldwork, prior to the VOXed technology being established. This serves as a 'baseline' for comparison with the same teacher's behaviour with the same student later in the fieldwork ('Lesson 2') when the technology was fully available and being used at their discretion.

The two tables (Table 1 and Table 2) contain time analyses from different moments in the fieldwork for the two teachers and one of each of their students (both tenors and with common elements in their lesson repertoire). Two lessons were selected as examples for each of these particular students, one from the opening phase of the fieldwork (Lesson 1), one from the latter phase when the technology was fully established (Lesson 2).

Several key features emerge from the time analyses of the teacher behaviours:

- There are significant differences between the basic pedagogical approaches of the two teachers, based on a comparison of the proportions of their lesson time allocated to categories from sample lessons in the opening phase of the fieldwork (i.e. comparing time data in seconds in Lesson 1 for each teacher, $\chi^2 = 14.41$, *df* 5, $p < .025$).[3] Teacher 1 makes use of an accompanist, particularly at the start of the

Table 3. The number of occurrences (as a %) of each type of pedagogical behaviour for the two teachers in Lesson 2 (where the technology was available for use)

Activity	**Teacher 1 (north) Occurrence of different types of pedagogical behaviour (%)**		**Teacher 2 (south) Occurrence of different types of pedagogical behaviour (%)**	
	Lesson 2 (% of Events) *With* technology		Lesson 2 (% of Events) *With* technology	
	⊗	+	⊗	+
Accompanying	2.1	0.0	10.1	1.4
Conversation	12.4	22.9	7.2	14.5
Demonstration	4.2	12.5	7.4	14.5
Instruction	2.1	2.1	10.2	14.5
Listening	0.0	10.4	0.0	1.4
Marking a feature	6.3	8.3	10.2	2.9
Not related to lesson	4.2	12.5	1.4	4.3
(sub-total%)	31.3	68.7	46.5	53.5
Grand Total (%)	100		100	

Key: ⊗ *= technology available, but not used;* + *= technology available and used* lesson to 'warm-up' the singer. Teacher 2, however, spends a greater proportion of the available time using formal instruction and in marking particular features during the lesson.

- Taken as a whole, the effects of having the technology available in the singing lesson varied with the individual teacher (i.e. as evidenced in a comparison of time use in Lesson 1 'without technology' with Lesson 2 'with technology' overall for each teacher). Although Teacher 1 took opportunities to comment and demonstrate using the feedback technology (and see Table 3), there was no significant difference in the categorization of this teacher's time across the two sample lessons ($\chi^2 = 10.41$, *df* 6, n.s.). In contrast, Teacher 2 had a large increase in the amount of time spent in demonstration and instruction using the technology ($\chi^2 = 26.45$, *df* 6, $p < .001$).
- Furthermore, although both teachers made various changes to their lesson time activities with the availability of the new technology, they continued to be significantly different from each other in their pedagogical approach during Lesson 2 ($\chi^2 = 41.63$, *df* 6, $p < .001$). Teacher 2 tended to pause the student's singing to have a shared focus on particular features of their performance by referring to the VOXed screen display (coded as 'instruction'), supported by demonstrations of particular vocal technique. Teacher 1, however, tended to spend more time watching the VOXed screen across the lesson, but to make much shorter reference to it through instruction or demonstration.
- Overall, although the technology was available throughout Lesson 2, Teacher 1 (who became a keen advocate for its use) only made reference to it for approximately one-third of the available time (37.6%) in this lesson. Somewhat paradoxically, although Teacher 2 reported that she was quite selective in her employment of the feedback technology, she spent nearly two-thirds of the time (60.6%) in Lesson 2 in activities that embraced its use (see Table 1 and also see Table 3 and its discussion below).

The digital video recordings of the lessons were also analysed for the categorisation of student time (see Table 2). The second-by-second microanalyses of each student's behaviour were converted into minutes and classified under four main headings:

- conversation (interacting with the teacher),
- listening,
- performing and
- other activities that were not directly related to the lesson.

As might be expected, there are several correspondences between the students' experiences and the behaviours of their teachers.

- A comparison of the two lessons (one for each student) from the opening phase of the fieldwork (Lesson 1) suggests that there were significant differences between the two students in their non-technologically supported lesson activities ($\chi^2 = 8.17$, *df* 3, $p < .05$). Student 1 spent nearly two-thirds of his time performing

(61.3%) and relatively little time listening (12.1%). This has a close correspondence to the time analysis for the same lesson for Teacher 1 (for example, see time proportions for 'instruction' and 'marking a feature' in Table 1). In contrast, Student 2 spent approximately equal amounts of time engaged in performing (35.2%), listening (28.3%) and in conversation (27.2%), with similar correspondences to his teacher's behaviour (see Teacher 2, Table 1).

- Notwithstanding the significant differences between the two students in their overall use of lesson time and the correspondingly significant differences between the teachers' pedagogical behaviours, there were no significant differences between the proportions of time of Lessons 1 and 2 for each student. Student 1 spent the majority of the time in each of the two lessons in performance (Lesson 1 = 61.3%, Lesson 2 = 75.7% (combined)) and much less time in conversation and listening ($\chi^2 = 4.09$, *df* 3, n.s.). In contrast, Student 2 spent relatively equal time in each of the two lessons across the three categories of activity. Although Student 2 spent more time in performing in Lesson 2, the overall difference was non-significant ($\chi^2 = 4.03$, *df* 3, n.s.).
- Although both students experienced a certain consistency in the pattern of time use during their two lessons, there was evidence of the use and impact of technology in Lesson 2. Student 1 spent 13.3% of the available time using the feedback technology, including 6.2% of the time whilst performing. The impact of the technology for Student 2, however, was much more in evidence, with half of the available time (49.5%) spent in technology-supported learning. Student 2 spent proportionately more of the performing time with technology (32%) than without (21.6%). Overall, there were significantly different experiences between the two students during the lesson with technology ($\chi^2 = 20.86$, *df* 9, $p < .025$).

Overall, both teachers made use of the VOXed technology during their teaching during the second term, the main pedagogical data collection phase in the project. The data for the sample lessons with technology (Lesson 2) indicate that the ratio of lesson time spent with the VOXed system was 1:2 for Teacher 1, compared with a ratio of 2:1 for Teacher 2 (extrapolated from the percentages for Lesson 2 in Table 1). Some of this time was spent monitoring their student's sung performance behaviours during the lesson, as evidenced by the differences between the teacher time ratios compared with the lower ratios of student time with the technology (1:7 for Student 1, but 1:1 for Student 2—see Table 2).

However, notwithstanding the evident differences between the two teachers in their pedagogical behaviour and its impact on the use of student time, the analyses suggest that each teacher had a particular pedagogical strategy that embraced the use of technology (such as in marking features, demonstrating a model behaviour, or using a display feature as a catalyst for discussion). This is illustrated by a reanalysis of the observational data (notes and video) in relation to the number of occurrences of each type of teacher behaviour. Using this form of analysis, although Teacher 1 spent approximately one-third of Lesson 2 using the technology, this actually accounted for two-thirds of the teacher's pedagogical behaviours (see Table 3 and compare with Table 1), with the largest number of events (22.9%) focused on teacher-student conversation. In comparison, whilst Teacher 2 spent nearly two-thirds (60.6%) of the Lesson 2 time using technology, the numbers of pedagogical events were more evenly divided (53.5% occurrences used technology). However, in comparing the relationship event incidence to time use, Teacher 2 focused equally on conversation, demonstration and instruction (14.5% each, Table 3), whilst spending the greatest amount of time on instruction (33.6%, Table 1).

Other recent research (e.g., Ward, 2004) has confirmed that talking is normally the dominant behaviour of instrumental teachers during a lesson. Similarly, there is no significant difference between these two teachers in their bias towards using talk in their lessons ($\chi^2 = 0.06$, *df* 1, n.s.). In the present study, whilst the use of an accompanist at the beginning of each lesson by Teacher 1 to 'warm up' the student accounted for the largest proportion of the opening lesson (59.9%, Table 1), teacher 'talk' (conversation, instruction and marking a feature) was the second largest (35.8%, Table 4). Furthermore, the 'talk' proportion increased overall for Teacher 1 in Lesson 2 (40.1%, Table 4), but most of this was technology focused. Similarly, talk was the main feature of the teaching style of Teacher 2, but also increased with technology use (55.5% to 73.3% overall, Table 4).

General outcomes

Other general findings in relation to the initial aims of the project are as follows:

- The two participant teachers and their students were positive about the benefits of having feedback technology available in the singing studio. This was evidenced in several ways, two of which stand out. In the original proposal, the action-research methodology was designed for two teachers to use the VOXed system with two of their students whilst continuing to teach two others using their customary (non-technological) pedagogical approach. This was in order to provide comparison

Table 4. Teacher talk, with and without the use of technology

	Teacher talk					
	Lesson 1		Lesson 2			
	minutes (−tech)	*% of time*	minutes (−tech)	*% of time*	minutes (+tech)	*% of time*
Teacher 1 talk	15.6	35.8	8.8	15.1	14.6	25
Teacher 2 talk	29.4	55.5	15	27.6	24.8	45.7
Combined:	*Conversation Instruction Marking a feature*	*Conversation Instruction Marking a feature*	*Conversation Instruction Marking a feature*	*Conversation Instruction Marking a feature*	*Conversation Instruction Marking a feature*	*Conversation Instruction Marking a feature*

data on the possible effects of using technology on their teaching. However, Teacher 1 was so convinced of the usefulness of the feedback technology that he began to use it with all his students in the main fieldwork phase (Term 2). He stated that it would have been 'unethical' to do otherwise because it was making an important contribution to their learning. Additionally, although Teacher 2 also began with the agreed research pattern of participant technology use, in the second phase of the fieldwork she was also observed to use the technology with larger numbers of students as part of group workshops.

- The teacher co-researchers did not consider themselves to be particularly computer-literate and had no experience of feedback technology as a teaching tool prior to participation in the research project. Nevertheless, despite some initial misgivings about whether it was possible to focus on a student's singing whilst using the technology, each teacher reported subsequently that they found the system to be user-friendly and not a distraction or hindrance. Both commented that they were very confident in being able to use the technology as needed to support learning. As part of the action-research methodology, decisions of how and when to use the technology were entirely at the teachers' discretion. Lesson observations indicated that teachers interpreted the selected displays in differing, but meaningful, ways and perceived themselves to be successful in integrating the additional feedback into their teaching.
- Each teacher developed their own strategy for technology use and both teachers tended to favour particular feedback display options. Teacher 1 made use of the vocal tract display, whilst Teacher 2 did not. Teacher 1 preferred to use the spectrogram option in full colour, whereas Teacher 2 used the same display option in black and white. For Teacher 1, the full colour version matched his verbal descriptions of vocal timbre, such as 'warm' or 'bright'. Teacher 1 used the technology with all the students, whereas Teacher 2 had a strategy in which it was used with more experienced students because 'beginning students' were perceived to need to have mastered 'certain basic skills'. Notwithstanding our concern to provide relatively simple displays that allowed the user to have a specific aspect of singing as the focus (such as pitch, timbre and resonance), both teachers enjoyed and made full use of the most complex display, the spectrogram (see middle panel example in Figure 4). Teachers used this screen option in a wide variety of ways, including illustrations of vowel quality, length of vowels and consonants, vocal register transitions and interactions between vocal loudness (intensity) and tone quality. Teacher 2 also used the spectrogram innovatively in order to provide visual feedback on a student's breathing behaviours, such as for breath placement during long sung phrases. Both teachers rarely used the fundamental frequency contour display option, probably because their students were perceived as already competent in this singing behaviour.
- As mentioned earlier, the limited time frame for the project (12 months) meant that there was a need to concentrate on an adaptation of the best of existing voice technology software, rather than in seeking to develop completely new feedback displays. The prime focus was in the provision of a portfolio of feedback options

that did not require a detailed technical background in voice science for use by voice pedagogues in the singing studio. These aims were successful in that the team, supported by a workshop with the expert panel, were able to create the initial *WinSingad* feedback technology package (Howard *et al*., 2004) in the space of a few months prior to the commencement of the fieldwork. Subsequently, the two teachers required only short induction periods to familiarise themselves with the technology. Furthermore, comments from both teachers and their students indicated that the screen displays were perceived as accessible and useful, not least because the visual feedback system provided an opportunity to share some aspect of the student's singing behaviour, either as it happened in real-time, or as a 'frozen' display for subsequent commentary, analysis and discussion.

- Although the system was designed to provide visual feedback of the singer's vocal acoustic behaviours, it is the nature of the on-screen information to be visual metaphors. However, unlike the verbal metaphors that have been reported as characterising the teaching of singing in the studio (Callaghan, 1997), these visual metaphors can be either observed in real-time or held in time whilst under discussion. This does not remove the possibility of ambiguities in perception, but the system does allow the teacher and student to have a more common evidence base on which to share their individual understandings of the singing behaviour under focus. For example, it is possible for the teacher to model and the student to copy and the two sung behaviours to be frozen on the screen and compared. Shared subjectivity was also evidenced in the teachers' use of spectrogram data. Both teachers tended to treat the on-screen information as visual patterns associated with particular vocal behaviours in order to focus their students attention on how these patterns could be manipulated systematically (Howard *et al*., 2005).

Discussion

As outlined in the introduction, there are at least two major challenges that face the teacher of singing when attempting to foster the development of their students. Firstly, the linguistic metaphors for vocal behaviours that are employed (by either teacher or student or both) may be misunderstood and, secondly, the ongoing temporal nature of the teaching process relies on multiple successive translations of student singing behaviour and spoken commentary. The student's current knowledge, understanding and skill levels are embedded in their singing behaviours and reflected in the nature of the performance strategy, effort and engagement that they bring to the lesson. A key task is for the student's 'personal understanding' to be matched against the 'target understanding' presented by the teacher (Entwistle & Smith, 2002), who, in turn, draws on their individual expert subject knowledge, as well as their attitudes and beliefs about the teaching and learning process. In such a pedagogical context, the application of multimedia technology allows any particular moment of the student's singing behaviour to be captured and 'frozen' in time. This facilitates the possibility of a better relative match between a teacher's target understanding (their specific teaching aim at that moment) and the personal understanding of an individual student. In the present study, this relationship is evidenced in the participants' comments that the feedback technology provided a stable external and shared focus on which to base their interactions.

Furthermore, there is an extensive research literature on effective instrumental (including vocal) practice (Hallam, 1997; Barry & Hallam, 2002) that embraces the earlier concepts of 'deliberate practice' (Ericsson *et al*., 1993) and 'formal practice' (Sloboda *et al*., 1996). The literature indicates that optimal practice conditions include opportunities to: (i) develop appropriate auditory schemata that focus on elements of the music (phrases, sections) as well as the whole, and (ii) experience models of desired behaviour. In the case of the VOXed technology, for example, the 'modelled' behaviour can equally be from the student (as displayed on the laptop computer screen) as from the teacher. It is also likely that practice will be more systematic and motivated in the presence of Knowledge of Results (KR) (see introduction).

At least two other aspects of the research literature on practice relate to the use of the VOXed multimedia technology to the singing studio, namely students' likely practice histories and teachers' instructional style. With regard to the former, a study at the Oslo conservatorium of music (Jørgensen, 2000) revealed that 40% of new students over three successive years reported that their former teachers had put 'very little' or 'no' emphasis on practice behaviour. This is in contrast with the findings of a USA-based study (Barry & McArthur, 1994) in which the majority of teachers reported that they 'always' or 'almost always' discussed the importance of practice and specific practice techniques with their students. In such a context of possible contradictory perceptions, the introduction of a robust multi-sensory feedback system of singer behaviour should ensure that imagery, commentary and singing are more likely to be integrated and for the student to understand how best to achieve the requisite 'target' behaviour. With regard to the second, notwithstanding the evidence that it is normal for there to be a lot of teacher talk in instrumental teaching, there is also evidence that the teacher's instructional style, i.e., what they actually do during the lesson (such as demonstrating a particular technique or having the student try a particular approach) can be more important than what is said (Barry, 2000, as cited in Barry & Hallam, 2002). If this is the case, then having real-time auditory and visual feedback of the student singer's behaviour (and feedback that can be replayed back at any point) should also ensure a greater cohesiveness between teacher talk, instructional style and student singing behaviour. This was evidenced in transcripts of teacher-student dialogue from the VOXed recordings (Howard *et al*., 2005).

Conclusion

Overall, although this was an exploratory 12-month project with no guarantee at its outset that the proposed application of multimedia technology would actually be useful in the individualised context of a professional singing studio,[4] the evidence suggests that the outcomes were positive.

- Within the time available, it was possible to design a new integrated software package (*WinSingad*) that was based on a diverse range of existing voice displays.
- This multimedia package was developed with support from a specially convened expert panel and iterative feedback from the teacher and student co-researchers during the opening (pilot) phase.
- Subsequently, the participant teachers and their students reported clear benefits in using the feedback technology within a lesson and also over a series of lessons and were very supportive of the principle of technology use in a studio setting. Confirmatory evidence was recorded in the observations of other members of the research team.
- The wealth of data that emerged (including observational field notes, digital audio-visual recordings, semi-structured interviews, expert panel responses and journal entries) required the (ongoing) development of a combination of qualitative and quantitative data analysis procedures by the research team. This has included a categorisation of teacher and student time that was based on second-by-second analysis of their respective behaviours from the digital video recordings and which is reported for the first time in this paper.
- Analyses of individual sample lessons near to the beginning and end of the fieldwork revealed that each teacher had their own, individual pedagogical strategy that was relatively consistent over time (Table 1) and which was also reflected in analyses of the students' experiences (Table 2).
- Teacher 1 initially tended to prioritise accompanied student singing, supported by teacher-student conversation. Although this remained the case over time, the advent of the technology brought about relatively more teacher-student interaction, including the increased used of teacher demonstration.
- In contrast, Teacher 2 engaged in more talk-based activities throughout, but with increased emphases on teacher instruction and student performance when using the technology (Tables 1 and 2).
- An analysis of the number of pedagogical events during the sample lessons indicated that technology use accounted for two thirds of Teacher 1's teaching behaviours (Table 3), but only one-third of the lesson time (Table 1). This suggests that his interventions, although numerous, were relatively short in duration. In contrast, Teacher 2 had a general bias towards an increase in both teaching events and time use when the technology was employed (Tables 1 and 3).
- Categorical analyses of the teachers' pedagogical behaviours revealed a contrast between perception and practice. Although Teacher 2 reported more selective use of the feedback technology than Teacher 1, in reality she used the system for a greater proportion of the teaching time in the lesson that was sampled than Teacher 1.

- Notwithstanding the large amount of teacher talk in the sampled lessons (Table 4), both students spent an increased proportion of the lesson time performing when the technology was being used (Table 2).

The prime purpose of the project was not to have the outcomes led by the perceptions and needs of the voice science community, nor to optimise the voice analysis display software, but rather to enable the expert singing teacher to have maximum oversight in establishing whether or not such technology could be useful in the professional studio. Given that the research outcomes to-date are favourable, further work now needs to be done to develop the software system further in the light of these initial field trials, to continue to ensure that 'users' are fully involved in the design and evaluation process and to expand the user base to encompass a wider population of teachers, students and musical genres. The action research methodology continues to offer great potential in legitimising the teachers' and performers' 'voice' in such development.

Accordingly, in addition to the ongoing data analyses, a new network (*Netvotech*) of voice scientists, teachers and performers is being established with funding under the Engineering and Physical Sciences Research Council's (EPSRC) 'Culture and Creativity' programme (2005–2007) to stimulate ongoing discussion and research between voice professionals in the arts and science fields. *Netvotech*[5] is a network that is focused on a consideration of how technology might be used to enhance and foster the healthy human voice in performance. At the moment, there is evidence from this innovatory study with VOXed that real-time feedback displays can be of benefit in the singing studio. This is a positive basis for future exploration.

Acknowledgements

The *VOXed* project was supported by the UK's Arts and Humanities Research Board (AHRB) under their 'Innovation Award scheme' (reference B/IA/AN8885/APN15651). Additional information about the VOXed project can be found on the world-wide-web on the project website: http://www.VOXed.org

Notes

1. Recent neuropsychobiological research (Peretz & Coltheart, 2003; Peretz *et al*., 2004) supports the findings of previous educational and psychological studies of the possible divergence in attention between song text and music. Current neurological evidence indicates that sung text is governed by the language processing system that mediates normal speech and that any disruption to the speech system can leave melodic expression intact (Peretz *et al*., 2004, p. 386).
2. A freeware version of the *WinSingad* software, offering a fixed number (4) of panels, is available for download (http://www.sarand.com/prd/winsingad/winsingad.html).
3. All chi-square analyses are based on data in minutes.
4. The VOXed project was funded under the AHRB's Innovations Innovation Awards scheme. This was designed to support research in all areas of the arts and humanities that was

'notably innovative and of very high quality, but where the concepts may be speculative or the outcomes uncertain'. On 1 April 2005, the Arts and Humanities Research Board (AHRB) became the Arts and Humanities Research Council (AHRC).

5. Contact either Professor David Howard (email: dh@ohm.york.ac.uk) or Professor Graham Welch (email: g.welch@ioe.ac.uk), the co-directors of the network, for information. The *Netvotech* website is currently under development and will be available in the summer of 2005.

Notes on contributors

Professor Graham Welch holds the Institute of Education, University of London Established Chair of Music Education and is Head of the Institute's School of Arts and Humanities. He is Chair of the Society for Education, Music and Psychology Research (SEMPRE), recent past Co-Chair of the Research Commission of the International Society for Music Education (ISME). He also holds Visiting Professorships at the Universities of Sydney, Limerick and Roehampton, the Open University and the Sibelius Academy in Finland. Research and publications embrace a variety of aspects of musical development and music education, teacher education, psychology of music, singing and voice science, special education and disability.

Professor David M. Howard gained a First Class Honours Degree in Electrical and Electronic Engineering from University College London in 1978 and a PhD from the University of London in 1985 on cochlear implants. He became a lecturer in speech and hearing sciences at UCL in 1979 and he moved to York in 1990. He gained a Personal Chair in Music Technology in 1996. His research interests include the analysis and natural synthesis of singing, music and speech. Outside work, David finds time to conduct a local 12-strong choir from the tenor line, he sings as a Deputy Songman in York Minster, and he plays the pipe organ.

Evangelos T. Himonides is a Lecturer in Music Technology Education and Research Officer in the School of Arts and Humanities, Institute of Education. His doctoral research is focused on the voice and psychoacoustics. His multi-disciplinary research background is reflected in his two undergraduate degrees (music; and information technology (multimedia)—the latter at first class honours) and his MA in Choral Education. Current research projects and related publications embrace the teaching of singing, music technology, voice science, human-computer interaction and virtual learning environments with multimedia. When time is available, Evangelos works as a freelance web-developer, recording engineer, session electric guitarist and vocalist.

Jude Brereton completed her first degree in German and Linguistic Science at the University of York. After graduating in 1995 she taught German and Music before continuing her studies at Trinity College, Dublin, where she gained an MPhil in Speech and Language processing in 1999. Returning to York in 2000 to co-ordinate the Millennium Mystery Plays she ran her own event management company and took her present post as Research Assistant in 2003.

Jude combines academic work with many other interests, which include playing the viola, recorder and bass viol, choral conducting, teaching music and singing.

References

Barry, N. H. (2000) Behind closed doors: what really goes on in the practice room, paper presented at meeting of the Southern Chapter of the College Music Society, Lafayette, LA.

Barry, N. H. & Hallam, S. (2002) Practice, in: R. Parncutt & G. McPherson (Eds) *The science and psychology of music performance* (New York, Oxford University Press), 151–165.

Barry, N. H. & McArthur, V. (1994) Teaching practice strategies in the music studio: a survey of applied music teachers, *Psychology of Music,* 22(1), 44–55.

Callaghan, J. (1997) The relationship between scientific understandings of voice and current practice in the teaching of singing in Australia. Unpublished PhD thesis, University of Western Sydney, Nepean.

Callaghan, J. (1998) Singing teachers and voice science—an evaluation of voice teaching in Australian tertiary institutions, *Research Studies in Music Education,* 10, 25–41.

Christiansen, L. (2005) Messa di voce and dynamic control, *Journal of Singing,* 61(3), 269–270.

Davidson, L. (1994) Songsinging by young and old: a developmental approach to music, in: R. Aiello & J. Sloboda (Eds) *Musical perceptions* (New York, Oxford University Press), 99–130.

Emerich, K. E., Baroody, M. M., Caroll, L. M. & Sataloff, R. T. (1997) The singing voice specialist, in: R. T. Sataloff (Ed.) *Professional voice: the science and art of clinical care* (New York, Raven Press), 735–753.

Entwistle, N. & Smith, C. (2002) Personal understanding and target understanding: Mapping influences on the outcomes of learning, *British Journal of Educational Psychology,* 72, 321–342.

Ericsson, K. A., Krampe, R. T. & Tesch-Römer, C. (1993) The role of deliberate practice in the acquisition of expert performance, *Psychological Review,* 100, 363–406.

Gabrielsson, A. (1999) The performance of music, in: D. Deutsch (Ed.) *The psychology of music* (2nd edn) (London, Academic Press), 501–602.

Gabrielsson, A. (2003) Music performance research at the millennium, *Psychology of music,* 31(3), 221–272.

Gruzelier, J. H. & Egner, T. (2004) Physiological self-regulation: biofeedback and neurofeedback, in: A. Williamon (Ed.) *Musical excellence* (New York, Oxford University Press), 197–219.

Hall, C. (2005) Gender and boys' singing in early childhood, *British Journal of Music Education,* 22(1), 5–20.

Hallam, S. (1997) What do we know about practicing? Towards a model synthesizing the research literature, in: H. Jørgensen & A. C. Lehmann (Eds) *Does practice make perfect? Current theory and research on instrumental music practice* (Oslo, Norges Musikkhøgskole), 179–231.

Hargreaves, D. J., Miell, D. & MacDonald, R. A. R. (2002) What are musical identities and why are they important?, in: R. A. R. MacDonald, D. Hargreaves & D. Miell (Eds) *Musical identities* (Oxford, Oxford University Press), 1–20.

Howard, D. M. (1995) Variation of electrolaryngographically derived closed quotient for trained and untrained adult singers, *Journal of Voice,* 9(2), 163–172.

Howard, D. M. (1999) The human singing voice, *Proceedings of the Royal Institution of Great Britain,* 70, 113–134.

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

Howard, D. M., Welch, G. F., Brereton, J., Himonides, E. & Howard, A. W. (2004) WinSingad: a real-time display for the singing studio, *Logopedics Phoniatrics Vocology,* 29, 135–144.

Howard, D. M., Brereton, J., Welch, G. F., Himonides, E., DeCosta, M., Williams, J. & Howard, A. W. (2005) Are real-time displays of benefit in the singing studio? An exploratory study. Manuscript submitted for publication.

Hunter, E. J. & Titze, I. R. (2005) Overlap of hearing and voicing ranges in singing, *Journal of Singing,* 61(4), 387–392.

Jørgensen, H. (2000) Student learning in higher instrumental education: who is responsible?, *British Journal of Music Education,* 17(1), 67–77.

Juslin, P. N., Friberg, A., Schoonderwalt, E. & Karlsson, J. (2004) Feedback learning of musical expressivity, in: A. Williamon (Ed.) *Musical excellence* (New York, Oxford University Press), 247–270.

Keenze, M. & Bell, D. (2005) Teaching breathing, *Journal of Singing,* 61(4), 371–377.

Kennell, R. (1997) Towards a methodology of vocal pedagogy research, in: B. Roberts (Ed.) *The phenomenon of singing* (St. Johns, Newfoundland, Memorial University Press), 129–137.

Lehmann, A. C. (1997) The acquisition of expertise in music: efficiency of deliberate practice as a moderating variable in accounting for sub-expert performance, in: I. Deliege & J. Sloboda (Eds) *Perception and cognition of music* (Hove, Psychology Press), 161–187.

Mang, E. (2003) Singing competency of monolingual and bilingual children in Hong Kong, in: L. C. R. Yip, C. C. Leung & W. T. Lau (Eds) *Curriculum innovation in music* (Hong Kong, Hong Kong Institute of Education), 237–242.

Mason, D. (2000) The teaching (and learning) of singing, in: J. Potter (Ed.) *The Cambridge companion to singing* (Cambridge, Cambridge University Press), 204–220.

Mitchell, H. F. & Kenny, D. T. (2004) The impact of 'open throat' technique on vibrato rate, extent and the onset of classical singing, *Logopedics Phoniatrics Vocology,* 29(4), 171–182.

Mitchell, H. F., Kenny, D. T., Ryan, M. & Davis, P. J. (2003) Defining 'open throat' through content analysis of experts' pedagogical practices, *Logopedics Phoniatrics Vocology,* 28(4), 167–180.

Obata, C. (2004) A practical study of the singing training and mental treatment to overcome "Onchi" (translation from the original Japanese) Unpublished PhD thesis, Tokyo Gakugei University, Japan.

Olsson, B. (1997) The social psychology of music education, in: D. J. Hargreaves & A. C. North (Eds) *The social psychology of music* (Oxford, Oxford University Press), 290–305.

Peretz, I. & Coltheart, M. (2003) Modularity and music processing, *Nature Neuroscience,* 6(7), 688–691.

Peretz, I., Gagnon, L., Hébert, S. & Macoir, J. (2004) Singing in the brain: insights from cognitive neuropsychology, *Music Perception,* 21(3), 373–390.

Potter, J. (1998) *Vocal authority* (Cambridge, Cambridge University Press).

Richards, H. & Durrant, C. (2003) To sing or not to sing: a study on the development of 'non-singers' in choral activity, *Research Studies in Music Education,* 20, 78–89.

Rosselli, J. (2000) Song into theatre: the beginnings of opera, in: J. Potter (Ed.) *The Cambridge companion to singing* (Cambridge, Cambridge University Press), 83–95.

Rossi, E. L. (1993) *The psychobiology of mind-body healing* (Revised edn) (New York, W. W. Norton).

Rossiter, D. P., Howard, D. M. & Comins, R. (1995a) Objective measurement of voice source and acoustic output change with a short period of vocal tuition, *Voice,* 4(1), 16–31.

Rossiter, D. P., Howard, D. M. & Downes, M. (1995b) A real-time LPC-based vocal tract area display for voice development, *Journal of Voice,* 8(4), 314–319.

Rossiter, D. P., Howard, D. M. & DeCosta, M. (1996) Voice development under training with and without the influence of real-time visually presented biofeedback, *Journal of the Acoustical Society of America,* 99(5), 3253–3256.

Rossiter, D. P. & Howard, D. M. (1998) Observed change in mean speaking voice fundamental frequency of two subjects undergoing voice training, *Logopedics Phoniatrics Vocology,* 22(4), 187–189.

Sapolsky, R. (2003) Taming stress, *Scientific American,* 289(3), 67–75.
Sloboda, J. A., Davidson, J. W., Howes, M. J. A. & Moore, D. G. (1996) The role of practice in the development of performing musicians, *British Journal of Psychology,* 87, 287–309.
Sundberg, J. (1974) Articulatory interpretation of the 'singing formant', *Journal of the Acoustical Society of America,* 55, 838–844.
Sundberg, J. (1987) *The science of the singing voice* (DeKalb, IL, Northern Illinois Press).
Sundberg, J. (1996) The human voice, in: R. Greger & U. Windhorst (Eds) *Comprehensive human physiology* (vol. 1) (Berlin, Springer), 1095–1110.
Thomasson, M. (2003) *From air to aria. Relevance of respiratory behaviour to voice function in classical western vocal art* (Stockholm, Royal Institute of Technology).
Thorpe, C. W., Callaghan, J. & van Doorn, J. (1999) Visual feedback of acoustic voice features for the teaching of singing, *Australian Voice,* 5, 32–39.
Thurman, L. (2000) The human endocrine system, in: L. Thurman & G. F. Welch (Eds) *Bodymind and voice: foundations of voice education* (Iowa, National Center for Voice and Speech), 61–67.
Tschannen-Moran, M., Woolfolk Hoy, A. & Hoy, W. K. (1998) Teacher efficacy: its meaning and measure, *Review of Educational Research,* 68(2), 202–248.
Vurma, A. & Ross, J. (2003) The perception of 'forward' and 'backward placement' of the singing voice, *Logopedics Phoniatrics Vocology,* 28(1), 19–28.
Ward, V. (2004) The performance teacher as music analyst: a case study, *International Journal of Music Education,* 22(3), 248–265.
Welch, G. F. (1983) Improvability of poor pitch singing: experiments in feedback. Unpublished PhD thesis, University of London.
Welch, G. F. (1985) A schema theory of how children learn to sing in-tune, *Psychology of Music,* 13(1), 3–18.
Welch, G. F. (2001) *The misunderstanding of music* (London, Institute of Education).
Welch, G. F. (2005) Singing as communication, in: D. Miell, R. MacDonald & D. Hargreaves (Eds) *Musical communication* (New York, Oxford University Press), 239–260.
Welch, G. F. (in press, a) The musical development and education of young children, in: B. Spodek & O. Saracho (Eds) *Handbook of research on the education of young children* (Mahwah, NJ, Lawrence Erlbaum Associates Inc).
Welch, G.F. (in press, b). Singing and vocal development, in: G. McPherson (Ed.) *The child as musician: a handbook of musical development* (New York, Oxford University Press).
Welch, G. F. & Howard, D. M. (2002) Gendered voice in the Cathedral choir, *Psychology of Music,* 30(1), 102–120.
Welch, G. F., Howard, D. M. & Rush, C. (1989) Real-time visual feedback in the development of vocal pitch accuracy in singing, *Psychology of Music,* 17, 146–157.
Welch, G. F., Sergeant, D. C. & White, P. J. (1996) The singing competences of five-year-old developing singers, *Bulletin of the Council for Research in Music Education,* 127, 155–162.
Welch, G. F., Sergeant, D. C. & White, P. J. (1997) Age, sex and vocal task as factors in singing 'in-tune' during the first years of schooling, *Bulletin of the Council for Research in Music Education,* 133, 153–160.
Welch, G. F., Sergeant, D. C. & White, P. J. (1998) The role of linguistic dominance in the acquisition of song, *Research Studies in Music Education,* 10, 67–74.
Welch, G. F. & Sundberg, J. (2002) Solo voice, in: R. Parncutt & G. E. McPherson (Eds) *The science and psychology of music performance: creative strategies for teaching and learning* (New York, Oxford University Press), 253–268.
Welch, G. F. & White, P. (1993) The developing voice: education and vocal efficiency—a physical perspective, *Bulletin of the Council for Research in Music Education,* 119, 146–156.