VOXed: Technology as a meaningful teaching aid in the singing studio

Graham F Welch
School of Arts and Humanities, Institute of Education, University of London, UK
g.welch@ioe.ac.uk http://ioewebserver.ioe.ac.uk/ioe/cms/get.asp?cid=4591&4591_0=4934

Evangelos Himonides
School of Arts and Humanities, Institute of Education, University of London, UK
evangelos@sonustech.com

David M Howard
Media Engineering Research Group, Department of Electronics, University of York, UK
dh@ohm.york.ac.uk http://www-users.york.ac.uk/~dmh8/

Jude Brereton
Media Engineering Research Group, Department of Electronics, University of York, UK
jb64@ohm.york.ac.uk

In: R. Parncutt, A. Kessler & F. Zimmer (Eds.)
Proceedings of the Conference on Interdisciplinary Musicology (CIM04)
Graz/Austria, 15-18 April, 2004 http://gewi.uni-graz.at/~cim04/

Background in vocal pedagogy. The standard pedagogical model employed in the conservatoire studio typically involves weekly/twice weekly lessons with an expert, supported by private practice and performance. The teacher is engaged in a psychological translation of the student's performance, for example by turning musical gestures into language, and the student is engaged in a further translation of the teacher's verbal and visual feedback into adapted singing performance. A dual possibility thereby exists for the misinterpretation of information. Anything that can provide more robust and easily understandable feedback to both teacher and student would seem to be worthwhile, and this forms the basic premise behind this work.

Background in acoustic analysis. The displays implemented include: acoustic pressure waveform, fundamental frequency, spectrogram, spectrum, spectral ratio, vocal tract area, and minimum, mean or maximum vocal tract area against time.

Aims. This paper describes VOXed; an ongoing AHRB-funded research project to evaluate the usefulness or otherwise of real-time visual feedback technology in the singing studio. The primary purpose of the work is not to optimise the technology itself for this application, but to work alongside teachers and students using an action research methodology to study the impact of real-time visual feedback on the students' learning experience. The contributory disciplines are voice sciences, psychology and education.

This paper reports the multidisciplinary framework that underpins the design of a new research project into the use of technology to support singing development.

The VOXed project is supported by the UK's Arts and Humanities Research Board (AHRB) under an Innovation Award numbered B/IA/AN8885/APN15651.

Additional information about the VOXed project can be found on the world-wide-web at <http://www.voxed.org>.

The nature of singing

Probably because of the ubiquity and bi-potentiality of the human voice for speech and singing (both in reception as well as production - see Welch, in press), the outputs of the vocal instrument are central components in many of the worlds' diverse performing arts. Examples include virtually all the musics of Africa in which singing is often the core group activity, the hugely popular Bollywood genre of *filmi* music from the

Indian subcontinent, indigenous musics, such as the traditional 'throat musics' of Southern Siberia, Mongolia and Tibet in which two musical lines are sung simultaneously by a single voice, as well as the musical narrative forms of Japan, such as *Nohgaku* and *Shinnai*, which challenge a bi-polar Western conceptualisation of vocal behaviour as either singing or speech.

The vocal instrument

Underpinning this worldwide use of the voice for musical performance and communication is a common anatomy and physiology shaped by biological maturation that is interfaced with experience, cultural imperative and tradition. The vocal instrument consists of three basic elements (see Figure 1): (i) the respiratory system to provide the energy source for the voice, (ii) the vocal folds within the laryngeal assembly which vibrate in the airstream to generate the basic sound and (iii) the vocal tract (essentially the spaces above the larynx – the pharyngeal space within the neck and the oral cavity, sometimes complemented by the nasal cavity) which shapes the sound (*cf* Welch & Sundberg, 2002). To make voiced sound, the respiratory system compresses the lungs to generate an upward flowing airstream which 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.

Pitch communication. Changes in vocal pitch are a product of variations in the mass and length of the vibrating vocal folds that arise from the relative interactive contraction of two sets of internal laryngeal muscles. The contractual dominance of one set of muscles (cricothyroids) has the effect of stretching and lengthening the vocal folds to create a longer, thinner, more taut muscular system. The lengthened folds tend to vibrate more quickly in the airstream and produce a perceptibly higher pitch. Conversely, when the other set of muscles are dominant, located within the vocal folds (thyroarytenoids), their contraction reduces the folds' length and increases their vibrating mass, resulting in a slower vibratory pattern with a perceptibly lower pitch (Welch & Sundberg, *op cit*).

Vocal loudness is mainly a 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). It seems likely that subtle changes in loudness during a sung phrase are the product of rib cage movement, whereas the movement of the abdominal muscles provides a more general 'platform' for the diaphragm to act on (Hixon & Hoit, 1999).

Vocal colouring (such as in the vowels of speech and the different timbres of the singing voice) is a result of 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).

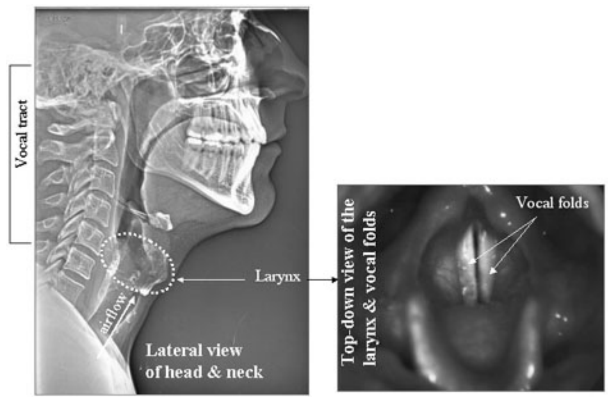

Figure 1. Anatomical structure of the singing voice.

Voiced sounds are acoustically rich, having many harmonics above the fundamental frequency. Acoustically, the vocal tract can be conceived as having several interconnected chambers, each of which individually and collectively filters and modifies the sound generated by the two sets of laryngeal muscles to create particular voice qualities. In addition, the tongue modifies the spaces in the oral cavity and upper pharynx (oropharynx) to create a complex variety of different sounds. The overall effect is to shape a unique vocal output for each individual that is capable of being measured as a 'voice print', much in the same way that each finger has its own unique finger print (Howard et al, 1993).

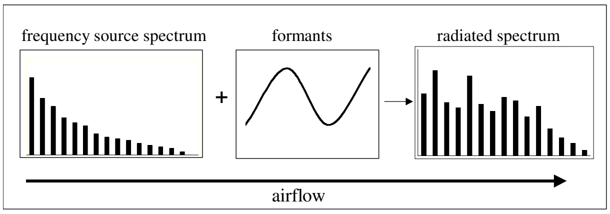

Figure 2. : The source-filter model of voiced sound.

Formants. Nevertheless, a general feature of voiced sound is that there are peaks in the spectrum of the sound that is radiated from the lips. These peaks are known as formants, created by vocal tract resonances that appear at certain frequencies and which enhance particular harmonics of the complex waveform emanating from the vibration of the vocal folds – the source-filter model (see Figure 2). 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 and tongue shape respectively. The relationships between the other three formants (F3, F4 & F5) relate primarily to vocal colour and also 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 (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.

Formants and singing pedagogy. Much conventional singing teaching relies on metaphor derived from formant manipulation to shape vocal behaviour (even though teacher and student may not be aware of the basic acoustic explanation), presumably in the hope that the student will be able to use the guidance to self-monitor singing quality in rehearsal and performance. For example, vocal timbre that is perceived as characteristically 'dark' in tone quality has formants that are relatively lower in the spectrum compared to voices whose quality is described as 'light' (Sundberg, 1970). 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 raising spectrally the first two formants (F1 and F2) and increasing the relative power of the 'singer's formant'.

Diverse nature of singing. Although the vocal timbre associated with such formant clustering is a characteristic of singing in Western high art music, it would likely be perceived as 'inappropriate' in the performance of other vocal genres. For example, a clustering of the first, second and third formants is a characteristic of the timbre of indigenous 'throat music' (Levin & Edgerton, 1999), whereas 'country singing' is more similar acoustically to speech, with the greatest energy focused on the lowest two formants. Likewise, sung performance with a 'belt' or 'show' singing style (so-called because of its high intensity in stage performance) is comparable acoustically to loud speech. Compared to classical Western-style singing, each of these diverse singing genres relies on different co-ordinated manipulations of the vocal system (respiratory system, vocal folds and vocal tract) in order that their characteristic timbres may be produced. In general, an 'untrained' singer tends to use or rely on habitual speech co-ordinations, often resulting in upper sung pitches that can only be produced with higher lung pressures and relatively extreme and effortful muscular tension.

In essence, the overall vibrating dimensions of the vocal folds at any given age, coupled to the degree to which they can be stretched/ lengthened or contracted/shortened, underlie the voice's basic pitch range (*tessitura*) and form the physical basis for the conventional 'labels' that are applied to singing voices, such as soprano, alto, tenor or bass.

Feedback in singing

The developing singer communicates intra-personally in a variety of ways related to the nature of the feedback system. Feedback can be auditory, visual, tactile, kinaesthetic or vestibular (Welch, 1985; Gabrielson, 2003) and it is used in the construction of individual musical identity, both in the sense of 'identity in music' – as a musician – as well as in the sense of 'music in identity' – as a feature of an individual's overall personal 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 the singing behaviour (*cf* 'vocal plan formation' - Peretz & Coltheart, 2003). In the first months of infancy, this system is being developed in the vocal behaviours that are the precursors of spontaneous singing and early speech, prior to their use in the emergence of a 'coalescence between spontaneous and cultural songs' (Hargreaves, 1996:156) from the age of two onwards.

A schema theory of singing development (Welch, 1985) proposed that any initiation of a specific singing behaviour (termed 'voice programme' in the original model), such as copying an external song model, would generate expectations of proprioceptive and exteroceptive feedback that are compared to the actual feedback received from the sense receptors and auditory environment (as both bone and air conducted sound) respectively. This internal motor behaviour feedback system also provides the basis for self-reflective psychological judgements 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'), the singer has to make their own judgement 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 lexicon' and 'phonological lexicon' (*cf* Peretz & Coltheart, *op cit*), in the sense that accurate reproduction of songs from the dominant culture requires the 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 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 intra-personal communication 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 individuality of the singing studio to the more public rehearsal environment, as well as in relation 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 intra-personal communication and the development of performance expertise. Lehmann (1997) suggests that there are three necessary mental representations involved, namely concerning the desired performance goal, the current performance and the production of the music.

Singing pedagogy

The conservatoire tradition embodies an individualized approach, meaning that in the majority of situations, it essentially involves one teacher working with one particular student at a time, basing the lesson plan on 'perceived' needs (i.e. what the singing teacher believes that the needs are). The vast majority of teachers would have been (or are) performers themselves and this ‘craft knowledge’ experience is likely to colour and individualise their singing pedagogy. Consequently, there are many different discourses evidenced about the teaching of singing, both in professional journals, in private conversations and in the wide range of terminology that teachers employ.. In general, singing pedagogy is relatively poorly documented (in relation to systematic, theoretically founded research), often highly idiosyncraticand quite often based on semi-transparent methods that are likely to bedriven by socio-cultural fashion interwoven with inherited empirical methods. Typically, the teacher is engaged in a psychological translation of the student's performance, for example by turning musical gestures into language, and the student is engaged in a further translation of the teacher's verbal and visual feedback into adapted singing performance.

A dual possibility thereby exists for the misinterpretation of information. Anything that can provide more robust and easily understandable feedback to both teacher and student would seem to be worthwhile, and this forms the basic premise to investigate the use of technology in the signing studio.

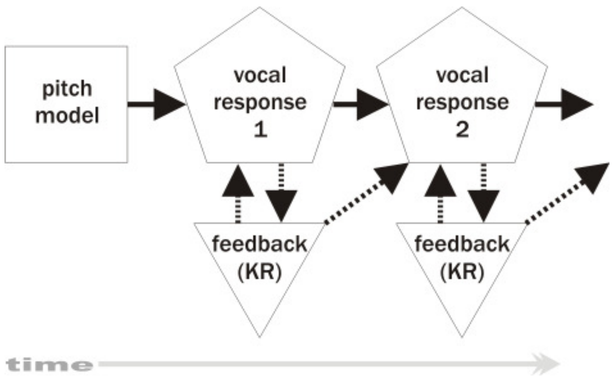

Figure 3. The way in which real-time visual feedback can impact the learning process. KR = knowledge of results from an external source (Welch, 1985).

The use of real-time visual feedback enables feedback to be provided during the student’s vocal response, enabling modifications to be made immediately and their concurrent effect to be observed (see Fig. 3). Apart from the more obvious advantage of removing the time lag between a vocal response and the feedback that is inevitable without real-time provision, the student is able to make another attempt immediately, based on appropriate observations of the feedback provided during the previous attempt.

Quantifiable parameters have been identified that vary with training and experience for: (a) actors (Rossiter *et al*, 1995), (b) adult singers (Rossiter *et al*, 1998; Howard, 1995), as well as (c) girl and boy cathedral choristers (Welch & Howard, 2002). Real-time visual feedback has been previously used successfully with primary school children (Howard & Welch, 1993; Welch *et al*, 1989) and adult singers (Rossiter *et al*, 1996; Thorpe *et al*, 1999). Our experience 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 can then underpin feedback to provide more accurate formative and summative assessments.

VOXed

The VOXed project focuses on the evaluation of the usefulness (or not) of real-time visual feedback technology in the singing studio. The question that the research team is attempting to answer is whether technology can be incorporated in a 'meaningful' way in the singing lesson studio. Thus the primary purpose of this project is not to design and/or optimise the technology itself for this application, but to work together with the instructors and students and monitor the impact of real-time visual feedback on the whole teaching-learning *modus operandi*.

The action research methodology that has been employed for the VOXed project can be divided in two primary milestones, first, a design phase that included a one day workshop with an advisory group, and second, the analysis of a substantial amount of (ongoing) observational data.

Workshop. A one-day workshop was held with an advisory group of singing teachers, the authors, and interested colleagues whose research involves speech, singing, psychology, linguistics, vocal health, engineering, and education. The purpose of this event was to introduce the panel to examples of existing technology in singing feedback and to seek opinion and information regarding the potential usefulness of that technology within the conservatoire context.

The panel reviewed a wide variety of existing technology and software applications. There was general agreement that teachers and students would be helped if they had better access to information concerning posture (the singer's external framework), gesture (the activity of the voice source) and vocal quality (the shaping of the vocal output). Specific research questions included:

- the extent to which teachers and students will accept and make use of technology in the studio;
- the ease-of-use of the technology, both in the studio and elsewhere for private practice;
- the nature of the data offered by the technology;
- how the data can be integrated into singing teaching and learning;
- the readiness with which the data can be interpreted and utilised;
- whether the technology overly intrudes into the learning and teaching experience;
- any potential perceived threat posed to the teacher and/or the student by the use of technology.

Observational data. The research team is gathering observational data and feedback from two separate singing teachers (one in York and one in Guildford). Each teacher is working with four students acting as co-researchers. Of these, two students are using the VOXed software in lessons and two are acting as 'control' students, working without the software, respectively. The VOXed team is gathering statistical data from the software itself for each lesson, as well as qualitative data from observation of the lessons, and interviews and journal records of the teachers and the students.

The technology

Fundamental frequency against time. The measurement of fundamental frequency (f0) has been the subject of considerable research (e.g. Thorpe *et al*, 1999). No one technique exists that is accurate for all subjects, covering the complete human pitch range uttered in any acoustic. The choice of a technique should be matched to the situation where it is to be used. A real-time display must not exhibit any delay to the user, it should be accurate operating over a wide f0 range for singers, of the order of C2(65Hz) to C6(1047Hz). Consequently, a peak-picking system has been employed that was originally developed in analogue form for use in cochlear implants (Howard, 1989), and subsequently applied in the SINGAD system (Howard *et al*, 1993; Welch *et al*, 1989). Each of the elements of its circuit has been implemented in C++, and an example plot of f0 against time is shown in Figure 4.

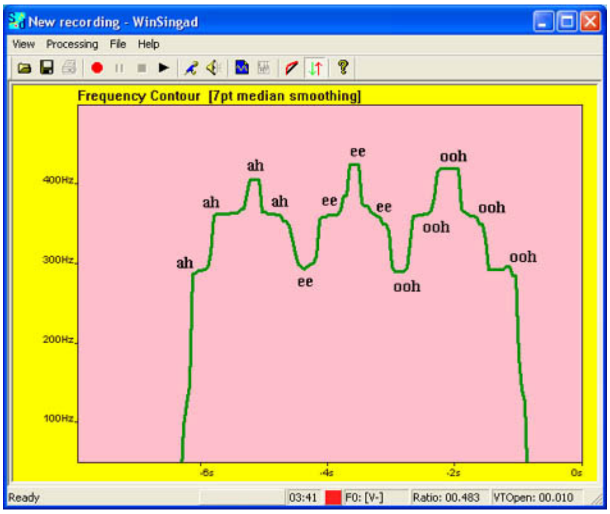

Figure 4. Screen capture of fundamental frequency over time during a soprano's warm-up session (consecutive arpeggio triplets in different vowels).

Spectral ratio against time. A key element in singing training is that of voice projection, and one acoustic consequence of this is the appearance of a peak in the output spectrum in the region 2.5kHz to 4kHz, known as the singer's formant (e.g. Sundberg, 1987). The ratio of the energy in this band to the energy in the total signal is calculated. This measurement is constrained between 0 and 1 providing the full band extremes encompass the singer's formant band. In this implementation, these are set to (100Hz to 4000Hz) and (2500Hz to 4000Hz) respectively. These values can be changed by the user.

Fig. 5 shows an example plot of this ratio against time for an arpeggio in the sound 'ii' sung in a projected and non-projected style.

Figure 5. Dual-panel screen capture. In the bottom panel a ratio against time display plots the difference between a projected (trained) /i/ sound arpeggio and a non-projected one.

Vocal tract area. A display of the vocal tract area can be obtained via a lattice filter model derived from a linear predictive analysis of the vocal output (Rossiter *et al*, 1995). This models the vocal tract in terms of the areas (or diameters/radii) of a set of equal length tubes between the glottis (space between the vocal folds) and the lips. Figure 6 shows an example vocal tract area display for a sung /a:/ vowel, where the glottis and lips are at the left and right edges of the display respectively. There are, however, limitations associated with this representation. Firstly, it strictly only models non-nasal voiced sounds, due to the assumptions employed in linear prediction.

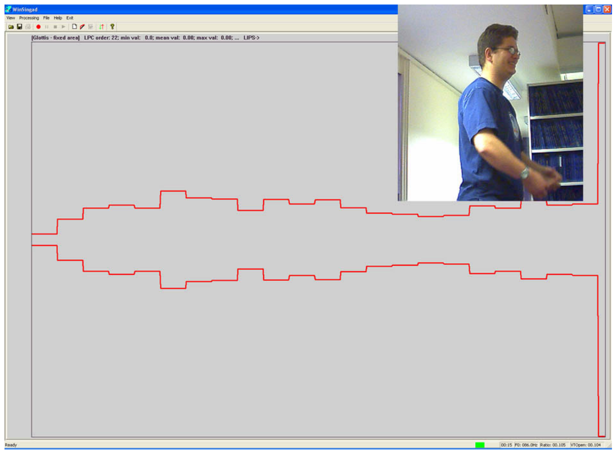

Figure 6. Dual-panel screenshot. Example vocal tract area display for a sung /u:/sound. The glottis and the lips are at the left and right hand side of the plot respectively. Both the singer and the tutor can monitor posture and possible neck tension at the same time (camera shot).

Secondly, the output area values have no absolute area reference, and therefore they are arbitrary. They are usually therefore normalized either to a fixed glottis width (this is adopted in Figure 6), or to a fixed maximum value. Finally, there are situations where more than one set of tube areas provides a solution, and results can be presented that could not be articulated by a human vocal tract. Due to the integrated nature of the solution process, it is not obvious how it might be constrained, for example, to vocal tract configurations that are physically possible. It is for this reason that summary plots of the average, minimum or maximum vocal tract area against time are being incorporated in the latest version of the software.

Summary vocal tract area against time. The mean, minimum and maximum vocal tract area is calculated for each frame of input data, and these can be plotted against time. An important aspect of singing training relates to the degree of perceived openness of the vocal tract, or the degree of constriction, and it is suggested that some indication of this may be given through reference to minimum vocal tract area against time.

Side view camera. Singers often make use of a mirror during training for feedback on their posture.

With a computer display, it is possible to make use of a camera with the result displayed on screen. We are employing a camera to enable singers to view their posture, for example, using a side view to enable the straightness of their spine to be observed.

The screen is placed at head height to encourage a vertical head position.

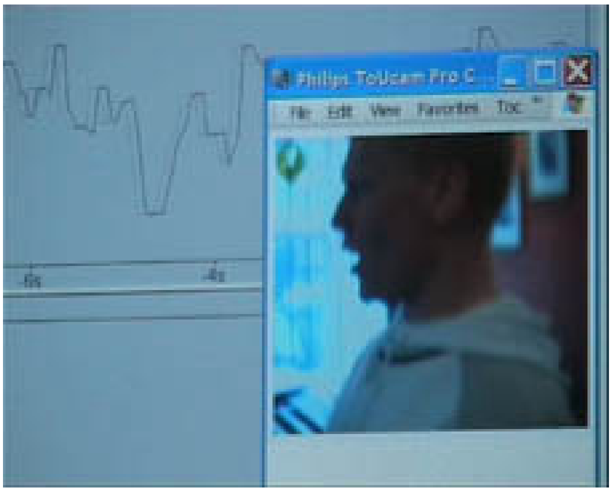

Figure 7. Screen captured from a current case study tutorial. The singer monitors the intonation of a high pitched note whilst checking his posture.

Although the research is ongoing and not due to finish until the summer of 2004, there is already evidence emerging from the action research that such technology can be useful in singing development. In particular, participants (students and teachers) report benefits in having both real-time feedback and also the possibility of data capture for subsequent playback and discussion (Howard *et al*, 2004).

Acknowledgments. The authors thank the singing teachers and other professional colleagues who contributed to the initial workshop.

References

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.

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* (pp. 735-753). New York: Raven Press.

Gabrielsson, A. (1999). The Performance of Music. In D. Deutsch (Ed.), *The Psychology of Music* (2nd ed., pp. 501-602). London: Academic Press.

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

Hargreaves, D. J. (1996). The development of artistic and musical competence. In I. Deliege & J. Sloboda (Eds.), *Musical Beginnings* (pp. 145-170). Oxford: Oxford University Press.

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* (pp. 1-20). Oxford: Oxford University Press.

Hixon, T. J., & Hoit, J. D. (1999). Physical examination of the abdominal wall by the speech-language pathologist. *American Journal of Speech-Language Pathology, 8*, 335-346.

Howard, D. M. (1989). Peak-picking fundamental period estimation for hearing prostheses. *Journal of the Acoustical Society of America, 86*(3), 902-910.

Howard, D. M., Hirson, A., French, J. P., & Szymanski, J. E. (1993). A survey of fundamental frequency estimation techniques used in forensic phonetics. *Proceedings of the Institute of Acoustics, 15*(7), 207-215.

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

Howard, D. M., Welch, G. F., Brereton, J., Himonides, E., & Howard, A. W. (2004). WinSingad: A real-time display for the singing studio. *ms submitted for publication*.

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, UK: Psychology Press.

Levin, T. C., & Edgerton, M. E. (1999). The throat singers of Tuva. *Scientific American, 281*(3), 70-77.

Mason, D. (2000). The teaching (and learning) of singing. In J. Potter (Ed.), *The Cambridge Companion to Singing* (pp. 204-220). Cambridge: Cambridge University Press.

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

Rosselli, J. (2000). Song into theatre: the beginnings of opera. In J. Potter (Ed.), *The Cambridge Companion to Singing* (pp. 83-95). Cambridge: Cambridge University Press.

Rossi, E. L. (1993). *The Psychobiology of Mind-Body Healing* (Revised ed.). New York: W.W. Norton.

Rossiter, D. P., Howard, D. M., & Comins, R. (1995). 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. (1995). 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.

Sundberg, J. (1970). Formant structure and articulation of spoken and sung vowels. *Folia Phoniatrica, 22*, 28-48.

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, pp. 1095-1104). Berlin: Springer.

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* (pp. 61-67). Iowa: National Center for Voice and Speech.

Vurma, A., & Ross, J. (2003). The perception of 'forward' and 'backward placement' of the singing voice. *Logopedics Phoniatrics Vocology, 28*(1), 19-28.

Welch, G. F. (1985). Schema Theory of How Children Learn to Sing In-tune. *Psychology of Music, 13*(1), 3-18.

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* (pp. 253-268). 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. (in press). Singing as communication. In D. Miell, R. A. R. MacDonald & D. J. Hargreaves (Eds.), *Musical Communication*. Oxford: OUP.