Learning Musical Instrument Skills Through Interactive Sonification

Sam Ferguson
Acoustics Research Laboratory
Faculty of Architecture
The University of Sydney
sferguson@arch.usyd.edu.au

ABSTRACT

Interactive visualisation methods have been used to understand real-time acoustic analysis for the purpose of learning musical skills. However, interactive sonification has not often been considered, perhaps because it is assumed the musician cannot concentrate simultaneously on two sounds – their instrument's sound, and the sonified information. However, whilst some finesse is required in designing sonification algorithms so that they interact with the musician's sound in a controlled manner, there possibly are particular advantages to adopting the sonification approach. This research reports on a suite of interactive sonification algorithms for communicating real-time acoustic analysis results to singers and instrumentalists.

Keywords

Interactive Sonification, Sonification, Sound Visualization, Music Education

1. INTRODUCTION

Research interest in using real-time acoustic analysis to provide feedback to training singers and musical instrumentalists has increased in recent times. However, with any use of acoustic analysis in real-time the communication method employed is crucial if humans are to interpret the information, as the results of acoustic analysis are both cryptic to analyse and often consist of huge numbers of results updating hundreds of times per second. It is very difficult to use this data without some form of visualisation, and many researchers are interested in developing efficient and effective methods for this purpose. Most of these efforts have been visually based and have avoided other methods of information communication.

Welch was one of the first to implement a system for providing visual feedback regarding pitch to primary school aged students of singing. He showed that there were measurable improvements possible by comparing three school classes that were taught to sing using different methods of feedback (including his own visual pitch tracking system) against each other [25]. Both he and Howard have developed this research systematically, recently culminating in the *VOXed* project [23, 24], and the visual feedback system *WinSingAD* [7].

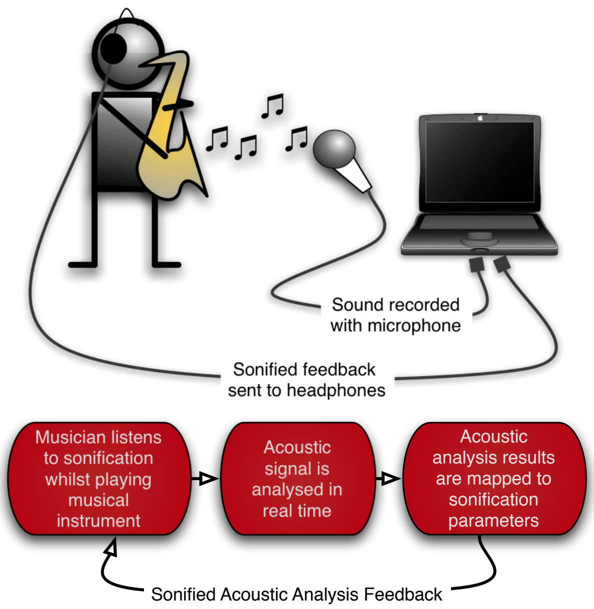

Figure 1: Acoustic analysis results can be sonified to assist with musical instrument practice.

Thorpe, van Doorn, Callaghan and Wilson have developed similar commercial software, which they call *Sing and See*.They have investigated the dynamics of computer based feedback systems extensively with encouraging results [20, 3]. Their system provides feedback concerning the acoustic features of the voice, incorporating a variety of pitch displays and both a 2-dimensional and 3-dimensional (i.e. spectrogram) representation of the magnitude spectrum of the acoustic input.

Ferguson *et al* have built a prototype musical sound visualisation that does not use typical acoustic displays [4]. It incorporates principles of information visualisation, and attempts to present the most relevant aspects of the musical sound with the greatest immediacy by using familiar metaphors. It is aimed at instrumentalists as well as singers.

These feedback systems are designed for a very specific situa-

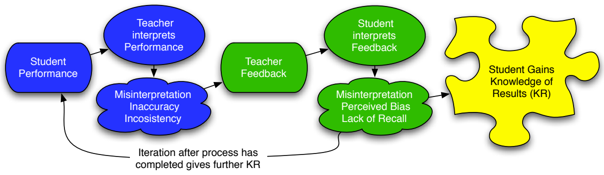

Figure 2: The dynamics of tuition can be investigated in terms of an iterative feedback loop.

tion, and for a particular style of student. The features that are highlighted are mostly 'technical' instrumental skills, and may be more important for specific styles of music (e.g. classical, orchestral, jazz). A successful modern student requires extensive technical control of their instrument in order for them to approach the nuances of musical interpretation exceptionally.

In this research I have drawn upon my knowledge of both acoustic and psychoacoustic research, as well as my experience teaching and being taught musical instruments (mostly woodwind). By attempting to solve some of my own problems maybe a useful approach to musical tuition can be created. It is my hope that the frustrations felt by Rostvall and West [17] can be to some degree alleviated by employing alternative and complementary methods in musical teaching and practise.

2. ITERATIVE FEEDBACK LOOPS

A common way for beginning students to learn a musical instrument is to engage a teacher to explain the specifics of how to physically play the instrument, whilst also often attaining basic musical understanding. A combination of demonstrations and verbal explanation provide a very flexible method of communication allowing the student gain a level of proficiency with their instrument, as well as with general musical skills. However, when the student is attempting to gain higher levels of skill the teacher usually takes a slightly different role, providing verbal feedback to the student about aural and musical impressions of their sound, such as the intonation, sound quality and rhythmic elements. Whilst this is the primary method of musical tuition in use today, there are several difficulties associated with this model of tuition, some of which may be ameliorated to some extent by the complementary use of acoustic analysis. One of the major difficulties is the small amount of time in which they may receive feedback from their teacher. Other difficulties can be seen by analysing the iterative feedback loop that is often present in tuition more closely. This iterative feedback loop incorporates four main parts: firstly the student performs while the master listens; secondly the master interprets the performance and looks for opportunities to improve the student's performance; thirdly the master provides their feedback verbally (predominantly) and the student listens to the feedback; and lastly the student attempts to interpret the master's feedback. The process iterates when the student attempts an improved performance based on the feedback received. This iterative process can incorporate errors, due to partial misinterpretation, incorrect or inconsistent judgement, inconsistency of physical and acoustic conditions, or perceived bias, at each of these stages (Figure 2).

A distinction needs to be made at this stage, because most of the skills we are interested in are made up of various sub-skills. For instance, in attaining fine control of pitch there are two skills necessary: an aural ability to discern pitches, and the psychomotor skills to produce the pitch discerned as correct [3]. The two skills are interdependent, and as such it can be difficult to attain adequate physical ability without adequate aural ability. This aural ability relies on feedback, as the student can not initially discern how the pitch they have produced compares with the pitch they were aiming for [9]. The term 'Knowledge of Results' (KR) has been used to describe the understanding the student receives from the teacher about the sound result they are actually achieving [7]. Feedback is essential to produce KR and the time delay between performance and feedback is crucial to the effectiveness of feedback. In a traditional tuition situation a student's performance for their teacher can be comparatively long, with the teacher only providing feedback after the performance is finished. This feedback must somehow be associated with the student's memory of their performance and the body positions used to achieve that particular part of the performance [22]. If the student succeeds in relating the feedback with performance, and then goes on to repeat the process as intended, the feedback loop still only iterates relatively slowly, with the appropriate length passage needing to be performed each time. Of course, real-time feedback loops iterate as quickly as the performer can perceive the information, and thus there is also little recall involved in associating body and muscle positions with sound results.

3. SONIFICATION

Visualisation systems are widespread; we are often systematically taught to understand numerical information in a visual manner within the primary and secondary school system. However, despite auditory methods for communicating information being shown to function well in various roles, they remain relatively unused. According to Walker and Kramer [21] techniques that use sound to convey information include the following:

Alerts and notifications: are simple sounds designed to alert a user to refocus their attention on some object or event.

Auditory Icons: which are the auditory equivalent of visual icons, and represent their target with sounds that the target produces.

Earcons: are sounds that represent a larger range of messages and meanings, and represent their target metaphorically, possibly with a melody or symbolic sound that is learnt over time.

Sonification: is the use of non-speech audio for information display. The data is 'mapped' to a parameter of sound and therefore the sound changes along a particular axis to represent changes in the data , thus 'sonifying' the data.

It is this final method that is most important for this research.

We are assisted in understanding these issues by the field of information theory. Moles provides an important primer regarding its application to sound and acoustic signals [10]. Moles has defined *channels* as a '...material system which conveys a message from transmitter to a receiver...' He goes on to define *natural channels* as being those channels related to a sense organ, which in our case is clearly the *auditory channel*. If we seek to transmit sonified information using the auditory channel, it seems strange to also attempt to receive information on the same channel. It is possible that the two forms of sound may interact in unwanted and arbitrary manners. This is a possible reason for avoidance of sonification in favour of visualisation. However, this is not a confounding problem when treated with a little finesse, and there are advantages of using sonification as a data communication method for this particular circumstance. For instance, it has been shown that instrumentally trained subjects are more attuned to the parameters often used for sonification [13]. Also, as sonification does not require the monitoring of a visual source, and only interjects aurally when a problem needs to be indicated, it is applicable in 'eyes busy' situations. One such situation is the practice of a piece of music. The most compelling reason for exploring sonification as a feedback mechanism is the possibility that musicians concentrate strongly on the sound they are producing when practising critically. Asking a musician to concentrate on both a visual source to identify auditory problems may disconnect the perception of the two, rendering a musician reliant on visual feedback to hear problems in their sound. Experiments that compare feedback transmitted using various natural channels may clarify whether this is the case.

3.1 Data Sources

For sonification to occur there needs to be data sources to sonify. This data comes primarily from the musician themselves, in the form of an audio waveform. The data is then processed using several acoustic analysis algorithms to provide relevant information about the musical sample. By developing these sonifications in Max/MSP [16] we are able to exploit Jehan's 'analyzer∼' [8] (which incorporates real-time pitch and loudness models), based on the 'fiddle∼' and 'bonk∼' algorithms developed by Puckette [15]. Primarily these algorithms provide us with:

- The time of attack of the note.
- The fundamental frequency of the note played.
- An estimation of the loudness of the note (using a basic psychoacoustical model).

It is by filtering and relating these data sources musically that we can provide more relevant information to sonify. Below we outline these filtering processes.

4. SONIFICATION 'STUDIES'

To teach the student innate musical and aural skills it seems most efficient to focus on one specific task. A reactive feedback system, responding to user's musical sound input with suggestions for correction of one *specific* musical parameter allows the student to focus on improvement. In a typical musical tuition framework we would term these programs 'studies' as they target a specific musical skill. These 'studies' respond intuitively to the note the instrumentalist is playing. They are designed to be either pleasurable or totally silent when the target parameter is within acceptable boundaries, and to intrude with sonified information when the user steps outside these boundaries. These sonifications are explained in further detail below.

4.1 Fine Pitch

Fine pitch control is the ability of the musician to play precisely in tune. It is a skill consisting of many sub-skills, such as noticing and estimating the pitch error, and the use of muscles to correct this pitch error. Over time players will develop strong skills at predicting or 'pitching' notes based on a flexible aural model which can attune itself to the external variables around it (key centre, environmental variables and other players with less precise tuning). Another knowledge base that is developed is an intimate understanding of the notes on the instrument, specifically the natural tuning of the instrument, and the difficulties associated with tuning these particular notes.

In this context chromatic electronic tuners that use visual meters have been a useful form of feedback employed. However, there are also multiple sonification solutions to this problem:

- Playing the 'correct' note for a comparison.
- Playing a sound when the note is out of tune.
- Mixing the sound returned to the instrumentalist with the correct note and/or distortion to produce beat frequencies.

A mix of all three of these options provides the sonification system most likely to be useful to the musician. When the musician plays a note slightly lower or higher than the target note the system responds by playing the *closest correct note*. Amplitude modulation is exhibited at a rate equal to the difference between the two wave's frequency. The comparison note's gain is controlled by the deviation from the closest correct note. The amount of distortion employed is also related to this measurement. Guitarists often use distortion to increase the audibility of beats between two notes when tuning. In this situation, information appears from the interaction of the sonification and the sound source the information being sonified is extracted from.

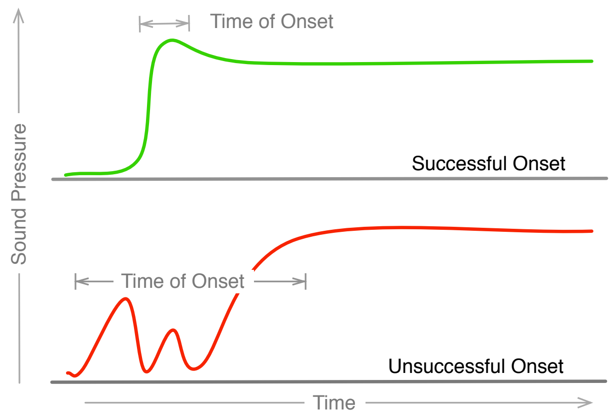

Figure 3: Successful and unsuccessful onsets can be distinguished by the time taken for the level to stabilise.

Of course this sonification design necessarily implies a method of determining correct notes, and for simplicity and flexibility we use the MIDI note scale. This is only relevant for rough tuning tasks, as the equal temperament system on which MIDI note names are based is not always used in instrumental performance. Musicians often gravitate towards the just intonation scheme if they are in a context that does not include an equal tempered instrument (e.g. a piano). A just intonation tuner would be a sophisticated tool for developing a flexible and rigorous understanding of fine pitch manipulation, however it would possibly require additional user input regarding the key the piece of music is being played in.

4.2 Note Onset

One important skill that can be much harder for certain instrumentalists than for others is achieving a clear note onset. Whilst pianists can be assured of a relatively precise onset, members of the woodwind and brass families (due to the acoustics of their instruments) need to practice carefully to achieve a successful onset. Furthermore their success also depends on the range and dynamic level of the note being attempted. It is useful to have some feedback as to how well they articulated a given note, in order to develop a high 'strike rate' for this task.

A clear onset of musical sound is a change from the lack of presence of a sound to the presence of sound in a short space of time. A simple algorithm for determining whether an onset is clear is to measure the time it takes to get from no sound to a stable sound pressure level (Figure 3). The influence of background noise would depend on the path length from the instrumental source to the microphone receiver, and also on the level at which the sound is stabilised at. The instrument acoustics usually do not cause a major problem when the dynamic is fairly loud and/or the instrument is playing in a comfortable range, only when there is a lower level of air pressure involved. Thus the average level used is very important in understanding the difficulty of achieving a satisfactory onset. However, accurate measurement of the amount of sound produced by an instrument requires a calibrated microphone and a constant distance between instrument and the microphone. If this precaution is taken we can provide a measurement of the sound pressure level at which the musician can consistently articulate.

Sonification algorithm presents a melodic sound representing either a successful or unsuccessful onset. A user-defined parameter would decide at what range the onset time is acceptable. A melody, based on the pentatonic scale closest to the pitch played by the musician, would rise by a number of scalar values (between 2 and 5) that indicates by how much the onset time was less than the acceptable onset time. Similarly it would fall related the amount the onset was greater than acceptable onset time. This allows the student to gauge their progress, whilst also controlling the standard they apply to themselves.

It is also useful to sonify the loudness of the sound produced, so there is an extra note added after a short pause at the end of the melody. This uses a general categorisation of the loudness level into integers analogous to the familiar *fortissimo* to *piano*, and is sonified using scalar values of the pentatonic scale again, but with a different timbre and a longer envelope. These sources of information provide students with ample feedback to gauge their ability at producing clear note onsets at different dynamic levels.

4.3 Rhythm

The passage of time has been sonified for centuries, through the use of clocks, bells and other timekeeping devices. A metronome is a timekeeper designed to help students develop innate rhythmic abilities. However, a metronome does not allow much understanding about the magnitude of the correctness in time to be imparted, rather the student learns only whether they sound 'in' or 'out' of time. Indeed, if the student is in time, and the attack of their sound occurs simultaneously with the 'tick' of the metronome and it is much more difficult to hear, due to grouping or masking effects present when two sounds onsets occur simultaneously. Thus the student often alters their rhythm to be before or after the beat to be sure they are listening to the metronome and keeping a correct pace. It could be argued that the reliance on an external source demotes a musician's personal understanding of the beat.

An alternative method of sonifying the beat is to measure the difference between the time of the attack and the correct time as determined based on the beat. The difference is the data source we wish to sonify. Again taking the approach that the user only needs to know about problems they may wish to fix, we design an algorithm that is silent when there is no rhythmic mismatch. This promotes the internalisation of this musical attribute, rather than the reliance on an external source.

This sonification algorithm uses a continuously pulsed unpitched sound to denote inaccurate rhythmic placement. Pulsing increases perceptual prominence, which is important for sounds of such short intervals. This sound is created either; after the attack of a note that is early compared to the correct time; or after the correct time has passed until an attack is detected (as per Figure 4). The length and loudness of the sound is related on the amount of rhythmic inaccuracy, and serves to provide feedback that can be used for understanding not only the occurrence of inaccuracy, but also the extent.

4.4 Loudness Control

Crescendos and decrescendos (increase and decrease in loudness level respectively) are a seemingly simple, yet often quite crucial musical concept. The loudness of sound is of great importance in

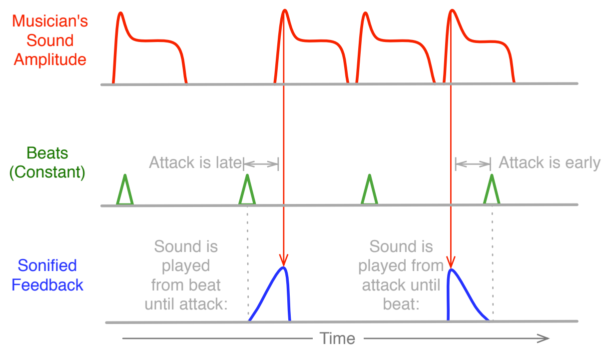

Figure 4: Method for sonification of rhythmic inaccuracy.

producing emotional responses to musical performances, as Schubert has shown [18]. Often teachers spend a lot of their time impressing on students the necessity for rapid changes in loudness to produce adequate excitement for the student's intended audience.

Increasing loudness at a defined rate is only simple before consideration of the difficulties inherent in estimating the likely perception of what 'loudness' means and how to measure it. Loudness as a perceptual sensation has been discussed by many authors [5, 19], and has also been modeled with a degree of usefulness [26, 12, 6, 8]. Furthermore, adaptation and fatigue often affect a listener in judgements of loudness [11], and loudness is strongly affected by stimulus frequency and spectral content [14]. It is generally reasonable to assume that by using a loudness model we more closely understand the percept that would be created for a listener, than if we use only sound pressure level.

Once we have decided on a scale for the loudness of a sound we can decide on what to measure regarding loudness. One of the most important abilities of a musician developing control over loudness is to achieve precise control of a single long note. This develops muscle control, respiration and endurance and is often a very important warm-up procedure for players of wind instruments. Another important ability is control of loudness whilst changing note, especially across registers. Following the mastering of static loudness levels, crescendos with constant (often quite slow) change of loudness level are often also attempted.

Thus sonification of data from a loudness model is an alternative method for developing an understanding of loudness. An algorithm that provides very clear representation is to play a short pulsing sound when the loudness exceeds boundaries around a specified loudness level. This pulsing sound represents the extent the loudness is exceeding the limit by the pulse frequency. A separate pulse sound is used for exceeding upper and lower loudness boundaries. This pulsing sound is clearly quite intrusive, but only occurs when the boundaries are overstepped, and thus the motivation to control loudness is increased.

This algorithm can be applied equally to stationary loudness levels or changing loudness levels. All that is needed to apply this algorithm to changing loudness is to calculate the change in loudness over time and compare this with a user-defined change speed. Then the algorithm needs a user-defined start point, most easily provided by using the loudness calculated at the start of a long note, the attack of which is detected using an algorithm like the previously mentioned 'bonk∼'. In this way, students can easily and immediately perceive an accurate description of their ability at controlling both constant and changing loudness.

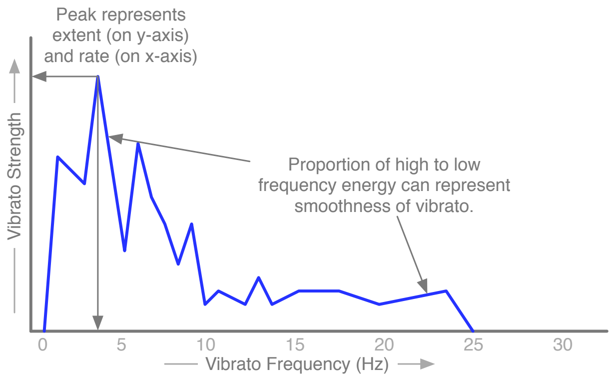

Figure 5: Examining the Fourier transform of a train of pitch values provides information about characteristics of the musician's vibrato.

4.5 Loudness and Legato

Legato, the musical impression of smoothness is highly prized in musical performance. Many factors can affect legato, but at least in woodwind playing one of the chief problems is synchronising depressing multiple fingers upon keys, and maintaining adequate air pressure. Practising musicians need to learn how listen for these gaps, and how to gauge which gaps are the most problematic. Some type of measurement of the extent of this gap may aid the user.

The use of the loudness sonification algorithm described previously is obviously an option, excepting that the parameters need to be redesigned. This situation is obviously different; the aim is not to maintain constant loudness, but only to not let loudness drop suddenly to low levels for short periods of time. Devising a strategy based on user-definition of the period of time that qualifies as a 'gap' allows the user to set a boundary condition that will trigger an earcon sound to be played.

4.6 Vibrato

Vibrato can be subdivided into three further attributes. Vibrato *Rate*, *Extent* and *Smoothness* can be defined in terms based on the vibrato spectrum. We may define the vibrato spectrum as the fourier transform of the pitch values sampled over a period of time, commonly one second or thereabouts.

In a 'vibrato spectrum' we may expect a peak at the fundamental frequency of the vibrato, somewhere in the range between 2-10 Hz, dependent on the singer, the note and the part of the note being sung. Vibrato rate changes dynamically and expressively, and Bretos and Sundberg have described how the vibrato rate tends to rise exponentially towards the end of important notes in classical arias [1]. The vibrato 'smoothness' can be estimated from the comparison of high frequency with low frequency energy in the vibrato spectrum (see Figure 5). This is based on the assumption that the smoothness of vibrato is related to sinusoidal (as opposed to harmonically complex) modulation at the vibrato rate. Whilst it is possible to use a ratio of the sum of energy below and above an arbitrary midpoint, a more appropriate method is to calculate the spectral centroid. The centroid gives us a frequency which can be considered the 'centre of gravity' of the spectrum, and can be calculated as follows:

$$C = \frac{\sum_{n=1}^{n} f_n a_n}{\sum_{n=1}^{n} a_n} \quad (1)$$

where f is the centre frequency of a bin, a the amplitude, and n the number of frequency bins in the spectrum. The vibrato spectral centroid is mapped to the amplitude of a sine tone an octave below the current pitch. As the vibrato spectrum centroid decreases towards the fundamental frequency, and the 'smoothness' of the vibrato increases, the amplitude of the sine tone increases. This may give the impression of the sound being 'reinforced', and perhaps also an increase in the size in the sound [2], and will thereby provide user feedback regarding this performance aspect.

The vibrato rate is sonified by using a mapping to the harmonic series. A categorical binning of the range between 2 and 10 Hz (avoiding the DC component and its rolloff) is first undertaken, mapping the vibrato rate to a series of integers, which are then multiplied by the current pitch, and then played back to the musician using a low-amplitude sine tone. A cross-fade is used for changes between these pitches to avoid arbitrary boundaries, and to avoid startling the user.

For different instruments vibrato and tremelo are often terms that are interchangeable, and refer to both cyclical pitch modulation and cyclical amplitude modulation. This sonification and analysis system can be applied to either pitch or amplitude trains.

5. IMPLEMENTATION, PRACTICALITIES

Use of the feedback system requires a computer, microphone and headphones. The software is reliant on the *Max/MSP* environment [16], and also requires a small number of freely available digital signal processing extensions developed by Jehan [8].

Acoustically, a room with low background noise, a medium to large volume and a reasonably low reverberation time is optimal. The avoidance of unwanted acoustic effects is important, as they could strongly affect some of the algorithms, especially those that deal with time. An optimal microphone position would be quite close to the instrument, but must be certain not to preference particular notes on instruments whose radiation pattern depends on the note played (e.g. woodwind).

6. EVALUATION

It is clear both theoretically [22, 9], and from qualitative studies [3] that there is great promise in visual feedback systems for musicians. However, little research has been carried out into acoustic analysis feedback systems that do not employ visual displays, and as far as the author knows this is the first example of an interactive sonification feedback system for this purpose. The existence of an alternative method of feedback allows research into comparison between methods.

Any feedback communication system is only as good as the feedback it is intending to give. If there are errors in the feedback presented to students, they will either learn bad practices or start to ignore the unreliable feedback being presented. Whilst this implementation intuitively attempts to only display information that is useful for musical instrument students, some method of evaluating the relevance of particular acoustic information would be helpful to maximise the applicability of acoustic analysis to musical practice.

This system introduces another layer additional to traditional sonification algorithms; the use of a single sensory channel for both data source and sonified feedback requires that interaction be controlled to avoid arbitrary interaction. This possible interaction brings the possibility of planned or unplanned emergent sonification.

This system for feedback can be taken at face value, as a system that seeks to 'lay down groundrules' to train a musician towards ideals of technical ability. However, it can also be approached with a less orthodox viewpoint. The sonification methods, if subverted, could lead to interesting and original results. Conversely, the algorithms used to derive the data sources could be altered towards usefulness for various purposes. As they are based within traditional musical practice they may be less abstract than other data sources sometimes used for sonification and may lend themselves to interactive or improvisational musical performances.

Whilst feedback does seem to be useful, it is important to investigate the usage of these systems in a tuition situation with user studies and quantitative methods. This will give some idea of the applicability of this mode of tuition.

7. CONCLUSION, RESEARCH AGENDA

We have described an interactive sonification system for supplying feedback to a musician to guide their instrumental practice and tuition. The system acts to provide the musician knowledge of their results, in turn associating their actions directly with their musical results and thereby developing muscle memory and technical musical skills.

Further research could focus on optimising this approach to music education, as well as a demonstration of its efficacy (or otherwise) in a teaching situation. Comparison and partnership with visual methods would be interesting for its multi-modal implications. Alternatively, exploration of the creative reuse and alteration of both sonification algorithm and data source derivation could be attempted. Most importantly though, methods for accessing musical parameters that musicians respond to and manipulate need to be collected for feedback systems to be effective.

8. ACKNOWLEDGEMENTS

This research was carried out with the assistance of an Australian Postgraduate Award and a Departmental Top-up Scholarship. The author acknowledges discussions with Andrew Johnston, Shigeki Amitani, Andrew Vande Moere, Densil Cabrera and his instrumental teachers Sharman Pretty, Mark Walton, and Josef Hanic.

9. REFERENCES

[1] J. Bretos and J. Sundberg. Measurements of vibrato parameters in long sustained crescendo notes as sung by ten sopranos. *Journal of Voice*, 17(3):343–352, 2003.

[2] D. Cabrera and S. Tilley. Vertical localization and image size effects in loudspeaker reproduction. In *Proceedings of the AES 24th International Conference on Multichannel Audio*, Banff, Canada, 2003.

[3] J. Callaghan, W. Thorpe, and J. v. Doorn. The science of singing and seeing. In F. Zimmer, editor, *Proceedings of the Conference of Interdisciplinary Musicology*, Graz, Austria, 2004.

[4] S. Ferguson, A. Vande Moere, and D. Cabrera. Seeing sound: Real-time sound visualisation in visual feedback loops used for training musicians. In *Proceedings of the Tenth Symposium on Information Visualisation*, London, UK, 2005. IEEE.

[5] H. Fletcher and W. Munson. Loudness, its definition, measurement, and calculation. *Journal of the Acoustical Society of America*, 5:82–108, 1933.

[6] B. R. Glasberg and B. C. J. Moore. A model of loudness applicable to time-varying sounds . *Journal of the Audio Engineering Society*, 50(5):331–342, 2002.

[7] D. M. Howard, G. F. Welch, J. Brereton, E. Himonides, M. DeCosta, J. Williams, and A. W. Howard. WinSingAD: A real-time display for the singing studio . *Logopedics Phoniatrics Vocology*, 29(3):135–144, 2004.

[8] T. Jehan. *Perceptual Synthesis Engine: An Audio-Driven Timbre Generator*. Masters Thesis, Massachusetts Institute Of Technology, 2001.

[9] A. E. Kemp. *The Musical Temperament. Psychology and Personality of Musicians.* . Oxford University Press, Oxford, 1996.

[10] A. A. Moles. *Information Theory and Esthetic Perception*. University of Illinois Press, Urbana, Ill., 1966.

[11] B. C. J. Moore. *An Introduction to the Psychology of Hearing*. Academic Press, San Diego, California; London, 1997.

[12] B. C. J. Moore, B. R. Glasberg, and T. Baer. A model for the prediction of thresholds, loudness, and partial loudness. *Journal of the Audio Engineering Society*, 45(4):224–240, 1997.

[13] J. G. Neuhoff, R. Knight, and J. Wayand. Pitch change, sonification, and musical expertise: Which way is up? In *Proceedings of the International Conference on Auditory Display*, Kyoto, Japan, 2002.

[14] J. G. Neuhoff, G. Kramer, and J. Wayand. Pitch and loudness interact in auditory displays: Can the data get lost in the map? *Journal of Experimental Psychology*, 8(1):17–25, 2002.

[15] M. Puckette, T. Apel, and D. Zicarelli. Real-time audio analysis tools for PD and MSP. In *Proceedings of the International Computer Music Conference*, University of Michigan, Ann Arbor, 1998.

[16] M. Puckette and D. Zicarelli. *Max/MSP*. Cycling 74/IRCAM, version 4.5, 1990-2006.

[17] A.-L. Rostvall and T. West. Analysis of interaction and learning in instrumental teaching. *Music Education Research*, 5(3):213–226, 2003.

[18] E. Schubert. Modeling perceived emotion with continuous musical features. *Music Perception*, 21(4):561–585, 2004.

[19] S. S. Stevens. The measurement of loudness. *Journal of the Acoustical Society of America*, 27(5):815–829, 1955.

[20] C. W. Thorpe, J. Callaghan, and J. van Doorn. Visual feedback of acoustic voice features: New tools for the teaching of singing. *Australian Voice*, 5:32–39, 1999.

[21] B. N. Walker and G. Kramer. Ecological psychoacoustics and auditory displays: Hearing, grouping and meaning making. In J. G. Neuhoff, editor, *Ecological Psychoacoustics*. Academic Press, New york, 2004.

[22] G. F. Welch. A schema theory of how children learn to sing in tune. *Psychology of Music*, 13(1):3–18, 1985.

[23] G. F. Welch, E. Himonides, D. M. Howard, and J. Brereton. VOXed: Technology as a meaningful teaching aid in the singing studio. In *Proceedings of the Conference of Interdisciplinary Musicology*, Graz, Austria, 2004.

[24] G. F. Welch, D. M. Howard, E. Himonides, and J. Brereton. Real-time feedback in the singing studio: An innovatory action-research project using new voice technology. *Music Education Research*, 7(2):225–249, 2005.

[25] G. F. Welch, C. Rush, and D. M. Howard. The SINGAD (SINGing Assessment and Development) system: First applications in the classroom. *Proceedings of the Institute of Acoustics*, 10(2):179–185, 1988.

[26] E. Zwicker and H. Fastl. *Psychoacoustics: Facts and Models*. Springer, Berlin; New York, 1999.