Chapter 1

PROLOGUE

[S]tructure, which is the division of the whole into parts; method, which is the note-to-note procedure; form, which is the expressive content, the morphology of the continuity; and materials, the sounds and silences of the composition...(Cage 1961/1967: 36).

1.1 Motivations for the work

This thesis has its early origins in an attempt to explain a phenomenon frequently experienced when composing algorithmic music with computers. At certain times in the process, usually towards the end of a major section or the work as a whole, the algorithms are set aside and an 'adjusting' or 'tuning' is undertaken 'by ear' that might involve experimenting with rescaling or perhaps re-quantising the pitch gamut, adjusting rhythmic hierarchies, limiting or compressing the audio bandwidth, reducing reverberation in the tenor register, and so on. All these high-level 'global' actions are performed more–or–less intuitively, until the work *gels* and the internal acoustic data representations begin to function as part of a cohesive whole; in which the structure becomes secondary to the form–the morphology of the continuity. That is, an identifiable unfolding continuity.

Many questions arise in thinking about this process, which, if generalised, is not confined to computer music, or even just to music. What are the underlying principles that inform this practice and can some general methodology be drawn from them? Perhaps a new kind of music theory is needed that does not confine musical information to that defined in current didactic texts, both old and new, or in the results of the reductionist practices of laboratory-bound psychoacoustic research, interesting though they are. The embodiment is a somewhat delicate procedure: taken too far, the result is auditory mush[1]. Clearly a balance has to be orchestrated that creates a sense of cohesion but that does not blur the articulation of structural

---

[1] A soft or soggy mass, but also, an expression that is excessively sweet and sentimental.

features. Musical orchestration is often, though not always, concerned with the mixing of individual instrumental timbres into rich, cohesive complexes; and its principles are well documented and taught in music schools everywhere. While, in his ground-breaking overview, Albert Bregman (after Helmholtz) described the basic dimensions of analytic and synthetic listening in terms of auditory stream integration and segmentation (Bregman 1994: 395-453), there is yet to be written a generalised exposition appropriate for many sonification tasks: how to synthesise auditory cohesion while maintaining a clear articulation of the separate components.

For the purpose of the current research, it was decided to put aside those parts of the process that one could easily identify as stylistic, so as to concentrate on understanding the synthesis of perceptions that afford the transfer of information structures at the expense of the cultural–in full cognisance of the dangers of such dualisms. This does not imply that the style is forgotten, simply faded into the background so as to simplify the problem at hand. In fact, on playing examples of the capital market parameter-mapping sonifications discussed in Chapter 5, a composer colleague asked, "So, why do these sonifications sound so like computer music?"

Before introducing the content of each chapter individually, a broad summary of the context of the thesis as a whole may be beneficial. To date, the most common approach to sonifying multivariate datasets has been to apply a technique often used in computer-music composition, namely, to map data dimensions to acoustic or psychoacoustic parameters, in the hope that the information content of the data will be "revealed". However, it is now recognized that such an approach has not produced the sorts of results required of sonification, namely, clear and reliable perceptions of the information. The dilemma has become known colloquially in the field as "the mapping problem". The simplest explanation for the failure of the method is related to the non-orthogonal or co-dependent nature of aural perception as it is usually parameterized. For example, under certain circumstances, an increase in loudness is also perceived as a rise in pitch. The most common "solution" when using this technique is to test empirically which of a number of fine-tunings or "tweaks" of parameter space mappings is the least problematic; perhaps in the hope that eventually, over time, a generalised model may become evident.

Rather than proffering such a generalised solution supported by extensive empirical evidence, a much less ambitious aim of the research reported in the early chapters of this thesis is to try to better define the nature of the problem, for only then will the demands on the computational tools necessary to develop such solutions be understood. A major difficulty, and one that is frequently elided, is the distinction between the concepts of information and data. In fact, the frequently used expression "data sonification" promotes that elision and in doing so, implicitly supports the idea that information can automatically "pop-out" of a sonification once an optimal parameter-mapping of the dataset is found. This thesis argues that an understanding of the historically volatile nature of the differences between data, data-embedded information, "sense data" and perception can explain why such an expectation is unrealistic; that the early attempts by phenomenologists to define such purely mental constructs leads to a tautological reduction to Platonic Ideals and all the difficulties that they imply; that the search for these mental contructs is another example of the Cartesian disembodiment 'trap' and that there is still no known basis for the reliably robust formation of such abstract mental structures.

However, I go on to argue, that the work of Polanyi and others on tacit and embodied knowledge may prove a fruitful path to explore, particularly as this approach is currently being pursued by interdisciplinary teams of philosophers and cognitive scientists, following the generally recognised failure of abstract computational models to solve "the hard problem" in machine learning research. An argument is thus advanced that the cognitive stability of aural structures such as melodies, which were of intense interest to the early phenomenologists because they are examples of apparently abstract mental structures, may be related to their origins in body actions; of speaking, singing and playing. The implication of this is that if software is to be capable of contributing to the translation of information in data into more reliable perceptual objects, it may have to be capable of simulating embodiment; a different, perhaps more difficult task, than the production of sound from acoustic or psychoacoustic parameter-mappings. With that task in mind, the thesis then proposes and reports on the development and testing of a software framework, called *SoniPy*, which affords such research.

Chapter overview

The thesis can be broadly divided into two themes. Chapters 1–3 provide an historical overview and theoretical context, Chapter 4 is a short, somewhat speculative link to the more practically–oriented design of a software framework powerful enough to create and undertake research into sonification and information together with some technical experiments that test parts of the framework on the sonification of capital market trading data.

1.1.2   Chapter 2

Empirical research in the synthesis of auditory designs for the pragmatic communication of non-musical and non-speech acoustic representations began to emerge in the 1990. Chapter 2 reviews the field as a whole, first by examining some descriptive definitions of sonification and suggesting some small improvements. The use of discrete sounds for alerts and alarms present designers primarily with differentiation problems: between the sounds themselves and between the sounds and the environment in which they function. Though related in subtle ways, these discrete audifications do not address an opposite issue (the "mapping problem"): how can data *relations* be represented acoustically for interpretation by listeners, for the purpose of increasing their knowledge of the source from which the data was acquired. That problem can be recast as the task of creating mental 'objects' for active contemplation, rather than how to correctly elicit a timely response to a well-differentiated auditory stimulus. Somewhat between these two is the task of continuous monitoring of production and environmental processes, and so forth.

An informal browse through a number of other theses in the field was one reason behind the decision not to include another cursory overview of the physics or psychophysics of sound in this thesis. Another was the ability to reference personal material previously published that more fully covers the material. However, the most important reason was a sense that the discussion needed to move on. The physics or psychophysics is important from an analytic perspective but for it to be useful for synthesis, it needs to be in the form of inverse filters, such as that for Fletcher-Munson as informally applied in Chapter 6 (§6.9) of this thesis. There is some peripheral work currently being undertaken (Cabrera Ferguson and Schubert

2007) and it would be useful if were to be generalised. The concern of this thesis, however, was to look further forward, to try to find a basis for better mental instantiations of multivariate datasets using sonification.

The term *sonification* has passed ('been appropriated' would probably be a more accurate description) into creative practice fora, possibly in order to avoid some of the associations the term *composition* engenders in funding bodies and the public at large. This thesis attempts to maintain the distinction, not in order to promote territorial disputes, but because having such a distinction makes it easier to compare and contrast motivations and results. Research in music can be very beneficial to research in sonification, however one of the disciplines of the latter, or so it seems this author, is the need to use music's tools and findings without being seduced by the aims and functions of music itself. So Chapter 2 ends with a comparison of data sonification and data music, and iterates the principle reason why there is a frequently–expressed need for a new generation of software for sonification; tools that integrate flexible sound-synthesis engines with those for data acquisition, analysis and manipulation, in ways that afford both experiments in cognition and lucid, interpretive *soniculations* (that is, sonic articulations).

1.1.3   Chapter 3 and Appendix 1

A goal of data sonification is to use sounds to aid listeners' acquisition of knowledge about a phenomenon, so it is logical to suppose that an understanding of the essential characteristics of that acquisition process, the extraction of information, may influence the design of the software used to compose and render sonifications. Such software will need to afford the exploration of the cognitive and psychological aspects of the perception of mental objects formed through the sonification of datasets that have no analogue in the material world, and the purpose of Chapter 3 is to explore the epistemological dimensions of that task.

It is rare to find references to philosophical inquiry when reading scientific literature that reports on the results of empirical experimentation–a trend probably with its origins in the Gestaltist's desire to separate their experimental motivations from those of the 'pure' psychologists and understandable because a discussion of the validity or otherwise of the empirical techniques is more appropriate in philosophy of science arenas. Discussion of a philosophical nature is more common

in fundamental science, especially on either side of a paradigm shift such as occurred in quantum physics and is currently occurring in cognitivism. Currently, sonification research is hardly settled and there are references in the literature, some, unfortunately, not very informed. Whilst not as dire, the same can be said for much published work on new media. The empiricist John Locke (1632-1704) seems particularly favoured when some degree of philosophical respectability is called for, probably, apart from his empirical leanings, because he wrote in English. Reference to Hume's refutation of some of Locke's work is as rare as Immanuel Kant's resolution. Occasional mentions of the intention to write an overview was met with enthusiasm so, having some previous experience in the field it was decided to attempt to lay out the philosophical framework as succinctly as possible.

Clearly, a complete philosophical and psychological overview is outside the scope of the current thesis, however if sonification software is to access complexly structured data, support informational enquiry, presentation and retention, in a perceptually and cognitively efficient manner, a thorough understanding of the dimensions of the problem and the contribution of others from the past, should be empowering. The approach is to use primary sources (at least English translations of them) as much as possible in order to maintain the flavour of the original enquiry, and to use sound-related examples when examples are called for–something that the original texts rarely do, and secondary sources, almost never.

The chapter begins with a discussion of some meanings of the term information, and Appendix 1, a pragmatic summary of different modes of knowledge acquisition, functions to support these definitions. Considered in this way, the trans*form*ation of in*form*ation into knowledge is an internal process–whether to an individual, a group or a community, and while there may be sonification techniques to enhance that such processes[2], they lie outside the scope of the current thesis. Most of the contents of Appendix 1 are widely understood, however it was included because such an inclusive yet succinct summary was not found elsewhere. In addition to the various forms of inference, and embodied knowledge, the inclusion of Reliablism, so apt an epistemological description of current scholarly practice, will add a less–well–known flavour.

---

[2]   Such as those used to enhance learning by entraining brain beta frequencies.

The relationship between our sensing of a variegated world and the mental models we use to represent it has been a major theme in Western philosophy and the remainder of Chapter 3 provides a reasonably thorough introduction.

1.1.4   Chapter 4

Chapter 4 is a short *intermezzo* between the theoretical orientation of the epistemology of Chapter 3 and the more practical orientation of software in chapter 5. Its purpose is to address, in a discursive way, the question: "If sonification software is to meet and even anticipate the needs of sonifiers in the future, what sorts of problems will it be required to address?" In some ways Chapter 4 is the underling enquiry of the thesis as a whole but a way was not found to express it properly without reference to the epistemology of Chapter 3, which points very strongly to the inadequacy of a purely mind-oriented solution to the problem of how to sustain abstract immanent phenomenal objects of multivariate datasets for cognitive enquiry and reflection. If this inadequacy is a reality it is probably more effective if it is considered a 'design feature' of the human condition, rather than as a 'bug'.

Any attempt to define the basis on which a paradigm that exploited this 'bug' that could be constructed for the translation of information contained in datsets into mental models that were more sustainable than those used in Parameter Mapping sonifications would require more empirical research than was appropriate in the context of the current work. The two closest models known to work, for different reasons and with different types of information are speech and music. These are powerful models. However, to function as a medium of information transfer, speech requires language and that requires community adoption. Esperanto does not seem to have been accepted, and the modernised talking drum, Morse code, probably could not sustain a 'come-back', enchanting though it would be. The seeming universality of music and the increasing acceptance, as evidenced in the popularity of 'world music' (Taylor 1997: 1), of broader range of musical paradigms than considered by Deryck (1959), are positive aspects. Because of its experimental nature, music can lead the way. The serialists exposed cognitive limitations to all but the highly trained, such as the lack of recognition of outside time temporal

transformations (retrogradation for example), and imitations in the recognition of pitch inversion.

Chapter 4 begins by returning to the Greeks again; for perspective and for inspiration. While music has its limitations, Chapter 4 outlines a somewhat speculative case for the need for software to be able to address issue of embodiment, whatever that may mean as a possible way forward. While this is not taken-up in any major way in the remainder of the thesis, it was an underlying motivation for developing the framework approach to software design that is discussed in Chapter 5, as mentioned earlier.

1.1.5   Chapter 5

The need for better software tools for sonification was highlighted in the *Sonification Report*'s comprehensive review of the field (Kramer et al. 1997). Their review included some general proposals for adapting sound synthesis software to the needs of sonification research. However, over a decade later, it is evident that the current demands being made of sonifications, especially those with large or multidimensional datasets, are much greater than the capabilities afforded by music the composition and sound synthesis software that is currently in use. Chapter 5 addresses some of the technical reasons this problem exists and discusses some major contributions towards achieving the *Report*'s proposals and current sonification demands.

The chapter outlines a broader and more robust framework model that can integrate other software developer's prior work and expertise, including that which has no direct connection to sonification, by using a public-domain community-development approach. Named *SoniPy*, it integrates various already existing independent components such as those for data acquisition, storage and analysis, cognitive and perceptual mappings as well as sound synthesis and control, by encapsulating them, or control of them, as *Python* Modules within the framework. In contemporary computer science the term *framework* has a specific meaning, and that is the meaning applied here.

A website has been created that outlines the various components of *SoniPy*. It functions as a first port–of–call for sonification-related activities using the *Python* programming language, and provides an introduction to modules that have passed

selection criteria testing for their use in undertaking various sonification-related tasks. While the site (at http://www.sonification.com.au/sonipy) is continually evolving at, a version is available on disk for off-line browsing.

1.1.6   Chapter 6 and Appendices 2 & 3 & 4

Chapter 6 details some experiments with capital markets data using the *SoniPy* framework. The sonification techniques employed include a new approach to the direct audification, using twenty-two years of an historical dataset, and the psychoacoustic parameter-mapping of information 'mined' from a high-frequency trading engine data. The latter work required the development of considerable data-handling capabilities in order to test the initial hypotheses. The practical experiments are preceded by a literature review of audification and prior work undertaken by others in sonifying economic, market and trading data, together with an overview of how a generic public market operates.

The sound rendering models used are as simple as possible for two reasons. Firstly, the aim is clarity not comfort, and secondly, experience has taught that hundreds of hours can be consumed trying to adjust one or more of a multitude of parameters in order to approximate a fuzzy target, only to find that the mind has adapted to the extent that what began as a clarinet-like sound ends up sounding more like a French horn, but in the interim the mind has convinced itself that it does in fact, sound more like a clarinet than it did before the whole exercise was begun.

Appendix 2 provides a succinct 'refresher' outline of some key statistical principles to make comprehension of the main text easier. Appendix 3 is the metadata specification of the high-frequency trading engine dataset and Appendix 4 contains various code listings, as detailed in the text; all of which are available, together with the sound examples, on the accompanying disk.

1.1.7   Chapter 7

Chapter 7 summarises the principal ideas of the thesis, draws some conclusions on what worked well, what not so well, and makes some suggestions for future similarly-motivated work as well as that which can build on the work undertaken here.

THIS PAGE HAS INTENTIONALLY BEEN LEFT BLANK