

Chapter 2

AN OVERVIEW OF SONIFICATION

The aim of this overview of data sonification is to provide the reader with an understanding of the origins and conceptual issues involved in this young, interdisciplinary and quickly evolving discipline. It begins by summarising different ways sonification has been defined, the types and classifications of data that it attempts to represent with sound, and how these representations perform under the pressure of real-world usage. The need for better tools for data sonification is raised and this leads to discussion of the value of collaborative research towards this end and a reflection on the relationship between music sound and science.

There are numerous reasons why sound might be the preferred representational medium for information in particular circumstances, including the known superiority of the hearing sense to discriminate certain kinds of structures. For example, it is easy to personally verify that a purely visual comparison of spatially separated representations requires high levels of concentration and is thus very prone to error, especially over extended periods of time, while listening to the reading of such representations is much easier. The presence of auditing (*hearing of accounts* from the Latin *auditus*) has been inferred from records of Mesopotamian civilizations going back as early as 3500 BCE. To ensure that the Pharaoh was not being cheated, auditors compared the *soundness* of strictly independently scribed accounts of commodities moving in, out and remaining in warehouses (Boyd 1905). In the alternating intoning of such lists, differences can be easily identified aurally. A faster and more secure method that eliminates any 'copy-cat' syndrome in such alternation, is to have the scribes read the records simultaneously—a type of modulation differencing technique. While we have no evidence that these techniques were practiced in ancient times, such a suggestion does not seem unreasonable, and would represent possibly the earliest form of data sonification.

2.1 Defining and classifying sonification

A primary distinction can be made between so called *audifications*, which entail the direct amplification or filtering of existing sounds, such as is accomplished with the esophageal stethoscope, and the use of sound to convey inherently silent abstractions such as *variables* and *data*. The non-chemical¹ definition of *sonification* has evolved over the last fifteen years as its use in auditory displays has developed. For the purpose of discussing multivariate data mappings, Bly (1994: 406) described sonification as *audio representation of multivariate data*. The sonification of univariate data is also possible, and thus Scaletti proposed a more formal working definition for her investigation of auditory data representation, as

a mapping of numerically represented relations in some domain under study to relations in an acoustic domain for the purposes of interpreting, understanding, or communicating relations in the domain under study (Scaletti 1994: 224).

In order to differentiate sonification from other uses of sound, Scaletti explicitly draws attention to two parts of her definition: a *technique* (mapping numerical data to sound) and an *intent* (to understand or communicate something about the world). Barrass, reworking Scaletti's definition, *en route* to a definition of auditory information design "the design of sounds to support an information processing activity", emphasises the idea of information (the content) over data (the medium):

a mapping of information to perceptual relations in the acoustic domain to meet the information requirements of an information processing activity (Barrass 1997: 29-30).

Both Scaletti's and Barrass' definitions can be read to mean both the process of representing, and the resulting sonic object. The *Sonification Report* (Kramer et al. 1999) was a major effort at summarising the field to date. Its focus is on sonification as a process:

¹ In biology, the term simply means *the production of sound waves* and is used to refer to a technique known as *sonication* [sic], in which *a suspension of cells is exposed to the disruptive effect of the energy of high-frequency sound waves* (Online Medical Dictionary: *sonification*) and (Biology Online: *sonication*). In chemistry, it refers to *the use of (often ultra-) sound waves to increase the rate of a reaction or to prepare vesicles in mixtures of surfactants and water* (ChemiCool: *sonication*). The term is also used for a process similar to the chemical one described that 'resonates' subsurface geological structures for oil extraction, See, for example, http://www.admin.mtu.edu/urel/news/media_relations/3/.

The use of non-speech audio to convey information. More specifically, sonification is the transformation of data relations into perceived relations in an acoustic signal for the purposes of facilitating communication or interpretation.

The first sentence of this 1999 definition appears to be the most succinct and widely used. Speech audio is specifically excluded, presumably so as to discriminate sonification techniques from speech-related practices, such as text-to-speech software. While speech research is an extensive research field in and of itself, there is no reason why speech audio should *necessarily* be excluded from the definition. As Hermann argued (2002: 23), speech does have attributes which, if data-driven, could be useful for sonification purposes, and some research has suggested speech-audio displays could be used to convey non-verbal information upon which people can make useful decisions (Nesbitt and Barrass 2002). While speech and non-speech sounds share such identifiable auditory characteristics as pitch, rhythm, articulation, and rugosity, as well as some larger gestalts such as phrase and prosody, the ability to simultaneously listen to music and talk or read without confusion is well known and easily demonstrated. This is supported by cognition research that finds that the auditory cortices in the two hemispheres of the brain are relatively specialised enough to be able to exploit the temporal and spectral differences between speech and musical sounds (Zatorre, Belin and Penhume 2002). Recent research found that speech that is temporally compressed until incomprehensible as speech could significantly improve menu navigation in PDA devices (Walker, Nance and Lindsay 2006). All these reasons emphasise the importance of not excluding speech-like sounds from the definition of sonification without a thorough evaluation of the information bandwidth that would be lost by doing so.

While *the representation of data relations in sound relations* (Anderson, Sanderson and Norris 2002) is likely the most succinct definition of sonification, it avoids the intent referenced in the 1999 definition: *for the purposes of facilitating communication or interpretation*. Nor does it quite capture distinctions between, for example, data sonification and data-driven music composition. While a purpose of music is the expression of musical knowledge and broader cultural considerations, whatever they may be, between composers, performers and listeners, the purpose of sonification, as the term was originally used, is to represent data in sound in such ways that

structural characteristics of the data become apparent to a listener. This distinction is emphasised by the new portmanteau expression *soniculation* (from *sonic articulation*). It is used occasionally in this thesis when such a distinction is important, without suggesting it be promoted to a term, as that would probably only add to the general confusion. In the early literature (Kramer 1994b), the term *sonification* is used as a shortened form of *data sonification* as a sub-category of *auditory display*, but the distinction, perhaps unfortunately, seems to have subsequently disappeared. All three terms are now used both nominally and verbally, as are the expressions *to sonify* and, though less common, *to audify*.

Terminology such as *auditory display* and *sonification* are best considered descriptions, rather than definitions within a strict taxonomy, because their meanings, driven by the need for finer distinctions and qualifications, have too great an inertia to be arrested by any current desires for semantic tidiness. So, this thesis uses the expression *data sonification* with the following interpretation, to clarify the object of the action and to lift the “non-speech audio” restriction:

Data sonification is the acoustic representation of data for relational interpretation by listeners, for the purpose of increasing their knowledge of the source from which the data was acquired.

This, and other descriptions of data sonification, preserves the ‘mystery’ of the implications of the phrase *for relational interpretation by listeners*. Further, it does not intimate the fact that it is often the data *relations* that are sonified rather than the data itself. Relations are abstractions of, or from, the data and the sonification of them more explicitly implies sonifier intent. This thesis uses the expression *information sonification* to express that distinction. The term *information* adds a complexity of its own; one that has changed over time, as discussed more fully in the next chapter. When the distinction is not important to the argument, this thesis simply uses *sonification*, without a qualifier.

2.2 Classifying sonifications

Sonifications can be classified in a number of ways: distribution technology (medical, public arena, interactive) intended audience (anaesthetists, stock brokers, the visually impaired), data source (electrocardiograms, securities

markets, hearing aids) data type (analog, digital, real-time, spatial, temporal) and so on. Twenty or so years since concerted research began, data sources and target applications for sonification have now become too numerous to review them in this context. They are also the easiest to locate using keyword searches on the world wide web.² For the current purpose, we partially follow Kramer (1994:21-29) in using a description of the type of representation that is used to present information to the auditory system.

There are two broad representational distinctions at either end of an analogic–symbolic continuum. Analogy³ is a high-level cognitive process of transferring information from a particular subject (the analogue or Source) to another particular subject (the Target) or of establishing the relation between the Source and the Targets themselves. The purpose of analogic representation is to make the structure of the information better suited or simpler for the Target than it was in its original form. Analogic representation is more connotative than denotative. A good example of analogic sonification is a Geiger counter that produces clicks in a loudspeaker at a rate proportional to the strength of radiation in its vicinity.

By contrast, a symbolic⁴ representation is a categorical sign for what is being represented. It is more denotative than connotative and thus more abstracted from the source than an analogic representation. In symbolic representation, source data is aggregated and assigned to the elements of a schema (symbols) according to a set of rules that, as a result, implicitly makes the distinction between data and information. These symbols are then displayed to the Target, who processes them according to their knowledge of the schema used. By illustration, consider a stream of noise emanating from a source. On listening to the stream, a Target recognises that some segments of it are symbols in one schema that they know, some sounds are symbols in another known schema and there are still other sounds for which they do not have a schema (other than the ‘everything–I–do–not–have–a–schema–for’ schema). An example would be a sound stream of English speech, spoken by

² For example, in June 2008, an internet search of the phrase *stockmarket sonification* revealed over six thousand links (c.f. fifty million for “stockmarket”).

³ From Gk. *analogia* ‘proportion’, from *ana-* ‘upon, according to’ + *logos* ‘ratio’, also ‘word, speech, reckoning’. [OLED]

⁴ From Gk. *symbolon* syn- ‘together’ + stem of *ballein* ‘to throw’, evolves from the “throwing things together’ to ‘contrasting’ to ‘comparing’ to ‘token used in comparisons to determine if something is genuine’. Hence, the ‘outward sign’ of something. [OLED]

someone with influenza, who, at some point says something to a dog in a language that the Target doesn't recognise. The semantic meaning of the English is understood by the Target, as are the influenza symptoms, but the dog-directed sound contains no recognisable information: it is just data. Two important features of this symbolic process are (a) a variety of display schemata can be used to symbolically represent the same information, or different information by different discrete–element–aggregation of the data, and (b) the relationships between the symbols themselves do not reflect relationships between what is being represented. For example, there is no relationship between the ideas represented by the words *loss*, *boss* and *lass*, nor the numerals of the number 1055, even though the structures of the symbols are similar.

In data sonification, for which the Source is data and the Targets are listeners, the types of representations a sonifier might employ is dependent on the structure of the data, the kind of information needing to be extracted from it and how easily (efficiently and effectively) that information can be processed by the listeners' hearing systems: physiological, perceptual, cognitive and memoric. This description emphasises the importance, for sonifiers, of developing a thorough working knowledge of the ways hearing systems organise acoustic waves into discrete events using Targets' ability to perform such fundamental perceptual and cognitive processes as "auditory stream segregation and auditory stream integration" (Bregman 1994). Ongoing research in the fields of psychophysics and experimental psychology is directed at the empirical investigation of these processes. As is to be expected, most of this work is directed at understanding the segregation and integration of simultaneous, or near-simultaneous, components constituting auditory streams, rather than the larger or more complex scenarios of real-world situations.

2.3 Types of data representation

The data representation types described below, loosely based on de Campo's groupings (2007) are discrete, continuous and interactive. These types can be located at different points on the analogic-symbolic continuum discussed above. Discrete representations mostly function symbolically, continuous representations mostly analogically, and interactive representations, mostly

analogically but in a discontinuous manner. In practice, a particular sonification may use one or more of these types, even simultaneously, to satisfy the needs of the specific application.

2.3.1 Discrete data representations

Discrete data representations are representations in which every data point (datum) is sonified with an individual auditory event. Discrete data representations are strongly symbolic and can be used as signifiers when there is no single preferred ordering of the datum. User-defined subsets of the data can be formed at will and randomly iterated over, much as a visual scene can be scanned in whatever way the viewer chooses. This flexibility of access to the data assists the user to build up a mental representation of the data space and the position of each data point in it. Whether or not a representation *appears* discrete or continuous will often depend on the interplay between data scaling and perceptual thresholds. A sequence of discrete frequencies may appear as a continuous glissando if they are closely clustered, or the close examination a continuous representation may cause it to appear discrete when it is time-stretched. Such definitional conundrums emphasise that these explanations are more descriptive than strictly taxonomic. Another example is the way in which auditory beacons (Kramer 1994; §2.3.2.1) function within an otherwise continuous data representation adds further emphasis to this point. However, when individual isolated sounds are initiated as environmental events, such as in HCI, the distinction is clearer and it is these we consider next.

2.3.1.1 Auditory warnings: Alarms and Alerts

Many animals use sounds of various kinds to warn those in their vicinity to the presence of others. Alarm signals, known in animal communication sciences as anti-predator adaptations, can be species specific, and are particularly effective when imminent danger is sensed. Boisterous alarms tend to provoke a *fright or flight* response, which, in ongoing monitoring situations, is perhaps not subtle enough: the ensuing panic is more likely to lead to a silencing of the alarm, or when that isn't possible, adaptation by cognitive filtering.

The term *alert* indicates the possibility of a more considered, attentive approach, but the distinction is by no means common practice. Patterson (1982 and 1989) experimented on alerts for aircraft navigation and produced a set of guidelines covering all aspects of warning design. The warning signals he suggests are meant to be instantly recognisable by listeners and use quite low-level intensities and slower onsets/offsets times to avoid startling the pilot. The auditory warning must impress itself upon the consciousness of the operator, yet not be so insistent that it dominates cognitive function. Patterson categorized warning signals into three priority levels: *emergency*, *abnormal* and *advisory*.

Extant guidelines for ergonomic and public safety applications of auditory warnings (McCormick and Sanders 1984) can be adapted for both general sonification purposes (Warin 2002) and more specific Human Computer Interaction (HCI) tasks, the initial purpose for which research into auditory icons and earcons was undertaken.

2.3.1.2 Auditory icons

The first detailed studies of the use of the sound capabilities of the newly emerged personal computers in the 1980s was as interface tools to operations of the machine itself. The desktop metaphor, first widely available on the Apple Macintosh computers (c. 1984), was recognized as a paradigm shift and quickly adopted by all personal computer manufacturers. In summarizing his work going back to 1988, Gaver (1994) outlines how sounds that were modelled after real world acoustics and mapped to computer events, could be used to enhance this desktop metaphor. He called these sounds *auditory icons*, for conceptual compatibility with Apple's computer Desktop Icon and Finder metaphors (folders, rubbish bin, menus etc). Gaver's reasons for using real-world sounds were based on an interpretation of James Gibson's theories of direct perception in which information about events in the world is perceived directly through patterns of energy, and understood innately, rather than through inferential representational structures in the mind. (Gibson 1966).

Building on the earlier work, Gaver developed, with others, various Finder-related tools: SonicFinder, Alternative Reality Kits (ARKs) and Environmental Audio Reminders (EARs). They extend auditory icons from sampled recordings of everyday sounds to synthesized abstract models of

them that could more easily be manipulated (by pitch change, filtering etc) to represent qualitative aspects of the objects being represented. He also suggests extending the range of applicability to remote machines, such as to indicate whether a printer in another room is functioning and the frequency of the page output.

For blind users, Mynatt developed a library of auditory icons, called *Mercator*, to complement the existing widget hierarchy of an existing graphical user interface. She reasoned that “auditory icons offer the most promise for producing discriminable, intuitive mappings” on the same ‘direct perception’ basis as Gaver and thus argues that “[w]hat this realization means to an interface designer is that we can use sounds to remind the user of an object or concept from the user’s everyday world” (Mynatt 1994).

Real-world sounds easily convey simple messages that are cognitively processed without much learning if the sounds are easy to identify. Yet sound in an interface that “seems cute and clever at first may grow tiresome after a few exposures” (Gaver and Smith 1990). In evaluating *Mercator*, Mynatt also observed a number of limitations, including the difficulty users had in identifying cues, especially those of short duration; the need for high (spectral) quality of sound and the need to evaluate auditory icons in a complete set for maximum dissimilarity while maintaining ease of identification. Both Gaver and Mynatt raise the question of whether or not the concept of auditory icons breaks down with virtual objects that do not have a clear counterpoint in the everyday world. Such objects are known phenomenologically as *mental* or *immanent* objects (Husserl 1927)⁵. Mynatt concludes, at the same time as cautioning that the participant sample size was too small for formal statistical analysis, that while non-sighted users could successfully navigate the system, “the overall test results indicated that the goal of designing intuitive auditory icons was not satisfied in the prototype interface.” There were suggestions, by Blattner, Sumikawa and Greenberg (1989), that the problems in memorizing a large number of such icons was due to there being no structural links between them. In addition, auditory icons can be homonymic - different physical events give rise to similar sounds. Later research (Sikora, Roberts and Murray 1995; Roberts and Sikora

⁵ This issue is addressed more extensively in the next chapter.

1997; Bussemakers and de Haan 2000) shows that, when compared with earcons and visual icons, test participants found that auditory icons were the least pleasant and least appropriate for use in computer interfaces.

2.3.1.3 Earcons

One alternative to using real-world sounds to reflect various computer activities, is to use structured sequences of synthetic tones called *earcons*, which are “non-verbal audio messages that are used in the computer/user interface to provide information to the user about some computer object, operation or interaction” (Blattner, Sumikawa and Greenberg 1989). Earcons are made by transforming a tone’s psychophysical parameters—pitch, loudness, duration and timbre—into structured, non-verbal ‘message’ combinations. Blattner, Papp and Glinert. (1994) found the use of musical timbres proved more effective than simple tones, however gross differences were needed for effective distinction to occur.

Brewster, Wright and Edwards (1994) undertook a number of experiments investigating the comprehensibility of earcons under various parametric transformations. They created motifs for a set of simple operations, such as *open*, *close*, *file* and *program*. A compound earcon was then created that gave a sound for *open file* or *close program* by simply concatenating the two motifs. They experimentally tested the recall and recognition of different types of earcons and reported that 80% recall accuracy could be achieved with careful design of the sounds and that earcons “are better for presenting information than unstructured bursts of sound ...and high levels of recognition can be achieved by careful use of pitch, rhythm and timbre. “

One of the most powerful features of earcons is that they can be combined to produce compound messages. So, symbolic sound elements (motifs) and their transformations can be used hierarchically to represent events or objects. A drawback, however, is that the deeper the structural hierarchy being represented, the longer a sequentially structured earcon takes to play, thus interfering with the speed at which a user can interact with the computer. Playing earcons more rapidly would risk recognition errors, though this could presumably be controlled for. Brewster, Wright and Edwards (1993) experimented with a different approach, namely to play the earcons at the same rate but present the information in parallel, taking the

time of only a single earcon to do so. Their results indicated that parallel and serial earcons were recognised equally as well. This is an important finding because it indicated that cognitive processing of an auditory display for known informational content may provide greater immutability than psychoacoustic conflict might otherwise indicate⁶. In both cases, recognition rates improved with repeated usage, which is the case in most HCI applications. Interestingly, more *training* produced greater improvements for parallel earcons than serial ones, 90% recognition rates being easily achieved. Furthermore, given their symbolic structures are similar to the semantic aspects of music, the hypothesis that musicians would perform better than non-musicians was not verified: there were no differences in performance between the two groups, except that some of the non-musicians took slightly longer to learn the system. Results from several experimenters had confirmed that correct identification of earcons decreased markedly if they occurred concurrently. McGookin (2004) undertook a detailed investigation of this phenomenon and produced a set of design principles including the use of spatial separation and a 300 ms onset-time offset. His findings indicate that, even when using these guidelines, the accuracy of concurrent earcon identification still decreased from 90% to 30% when the number of earcons increased from one to four. However these guidelines did improve the accuracy with which individual earcons were identified. Overall, earcon studies indicate that they can be an effective means of communicating hierarchical structures but that the number of them that can be usefully identified is quite limited when they are used concurrently.

2.3.1.4 Speech noises

NoiseSpeech is made by digitally processing sounds so that they have some of the acoustic properties of speech (Dean 2005). It is made either by applying the formant structures of speech to noise or other sounds, or by distorting speech sounds such that they no longer form identifiable phoneme sequences. The 'hybrid' sounds that results from this process encapsulate some of the affective qualities of human speech, while removing the semantic content. Empirical experimental evidence suggests that most listeners cluster the sonic

⁶ This suggestion was is tested in Experiment 5 in chapter 5.

characteristics of NoiseSpeech with speech rather than those musical instrument or environmental sounds (Dean and Bailes 2006; 2009).

Less abstracted than NoiseSpeech, *spearcons* (speech earcons) are spoken phrases that have been time-compressed until they are not recognizable as speech (Walker, Nance and Lindsay 2006). They were designed to enhance hierarchical menu navigation for mobile and screen-limited devices, in which they can be created automatically by converting menu item text (e.g., *Export File*) to synthetic speech via text-to-speech software. Keeping pitch invariant, this synthetic speech is then time-compressed, rendering it incomprehensible. Whilst each spearcon in an application is unique, phonetic similarity is maintained. For example, the initial phonemes are invariant in *Open File*, *Open Location...*, and *Open Address Book* as are the endings of *Open File*, *Export File*. The relative lengths of the text strings are maintained by their spearcons and this assists the listener to learn and identify the mappings.

2.3.1.5 Discrete data representations compared

The postulation that, because of the connection to fundamental percepts, auditory icons should be easier to learn than earcons, was questioned by Lucas (1994) who found it took significantly less time to learn spoken messages than it did to learn either earcons or auditory icons, and unlike earcons and auditory icons, these spoken messages were consistently interpreted error-free. Further, he found no significant differences in the amount of time needed to learn the meanings associated with earcons or auditory icons. This result is at odds with a simple interpretation of ecological theory (Gibson 1966; 1979) and Ballas (1994) suggests that the activity of listening contains intermediate mental processes that take account of a listener's expectations and experience and the context in which the auditing occurs. One could speculate that, because the listener is also immersed in a real auditory scene as well as the computer's virtual one, a differentiation between virtual and real elements requires greater cognitive processing than if the two scenes were not superimposed. Bussemakers and de Haan (2000) undertook a comparative study of earcons and auditory icons in a multimedia environment and found that having sound with a visual task does not always lead to faster reaction times. Although reaction times are slower in the

experiment with earcons, it seems that users are able to extract information from these sounds and use it. Furthermore, users find real-life sounds annoying when they hear them frequently.

Earcons are constructed from lexical elements that are ordered categorical symbols. So, while, like musical themes, their meaning is multiplicative under symmetry group transformation (transposition, retrogradation, inversion etc), the information needs to go through a decoding phase. As there is no standard syntax or lexicon, their meanings have to be learnt (Blattner, Sumikawa and Greenberg 1989), requiring a high initial cognitive load. Moreover, without absolute standardisation across all software and hardware, disarray rather than greater clarity seems the more likely outcome.

Palladino and Walker (2007) conducted a study comparing menu navigation performance with earcons, auditory icons, and spearcons. Their results indicate that faster and more accurate menu navigation was achieved with spearcons than with speech-only, hierarchical earcons and auditory icons (the slowest). They suggest that one reason for the speed and accuracy of spearcons is that, because they retain the relative lengths of their sources, these different lengths provide a “guide to the ear” while scanning down through a menu, just as the ragged right edge of items in a visual menu aids in visual searching. Because spearcons can be created and stored as part of a software upgrade or initialisation process, access times would be no longer than earcons and auditory icons. Furthermore, language or dialect-specific issues can be managed by the operating system’s standard internationalisation procedures.

Since the mapping between spearcons and their menu item is non-arbitrary, there is less learning required than would be the case for a purely arbitrary mapping. Unlike earcon menus, spearcons menus can be rearranged, sorted, and have items inserted or deleted, without changing the mapping of the various sounds to menu items. Spearcons may not be as effective at communicating their menu location as hierarchical earcons. However, spearcons would still provide more direct mappings between sound and menu item than earcons, and cover more content domains, more flexibly, than auditory icons.

Whilst auditory icons and earcons clearly have lexical properties (pitch, duration etc) they are used as auditory cues or messages; as signifiers of stable gestalts, i.e. as clear denoters of the presence of known, separable or discrete information states. The design of sets of auditory icons, earcons, spearcons etc is concerned with finding ways to effectively indicate these states under auditory segregation pressure. That is, the ability with which they remain immutable in human audition, in the presence of each other and within the perceptual environment in which they exist. These audio cues (or *messages*) function semantically in that they convey information about the state of the system of which they are a part and so can be used to reduce visual workload or function as a way of monitoring the system when visual or other means are not available or appropriate.

A weakness of the primarily ontological classification of discrete data representations as earcons, icons, etc, is that an emphasis on obtusely veridical denotations rather than often subtle, meaningful qualities, can have the effect of discouraging the user's active listening and so significantly discourage their engagement with the auditory display altogether. Such subtlety is more likely to be achieved by the cohesive use of careful mixtures of real-world audio-graphic and syntactic constructions whose symbolic power comes not from their concreteness or abstractness but from their intuitively understood affect. The design challenge is how to do this without recourse to distractional splendiferous phantasmagorias and in ways that support the user's need to segregate their computed and physical environments with a minimum cognitive load.

2.3.2 Continuous data representations

Continuous data representations treat data as analogically continuous. They rely on two preconditions: an equally-spaced metric in at least one dimension and sufficient data to afford a high enough sampling rate for aural interpolation between data points. Continuous data representations are most commonly used for exploring data in order to learn more about the system that produced it. Their applications range from monitoring the real-time operation of machines, capital-market trading, geographic and demographic features, weather and the environment, and so on, so as to discover new regularities and to assisting those with visual impairment to gain access to

information normally presented graphically. There are three types of continuous data representation that, for consistency, are labelled *Parametric Mapping Sonification*, *Direct Data Audification* and *Homomorphic Modulation Sonification*.

2.3.2.1 Parametric mapping sonification

Parameter mapping is the most widely used sonification technique for representing high-dimensional data as sound. Parameter mapping sonifications are sometimes referred to as sonic scatter plots (Flowers, Buhman and Turnage 1997; Flowers 2005) or n^{th} -order parameter mappings (Scaletti 1994). Typically, data dimensions are mapped to sound parameters: either to physical (frequency, amplitude), psychophysical (pitch, loudness) or perceptually coherent complexes (timbre, rhythm). Parameter mapping sonifications can have both analogical and symbolic components. Analogic variations in the sound can result when mapping from a large data domain into a small perceptual range or when data is specifically mapped to acoustic modifiers such as frequency or amplitude modulators. Parametric mapping sonification is sometimes referred to as *multivariate data mapping*, in which multiple variables are mapped to a single sound. Scaletti describes one way of implementing it by “mapping of each component of a multidimensional data point to a coefficient of a polynomial and then using that polynomial as the transfer function for a sinusoidal input” (1994). Within an overall analogic mapping, symbolic representations such as auditory beacons (Kramer 1994) can be used to highlight features such as new maxima and minima, or absolute reference points in a sonification such as ticks to indicate the regular passing of time.

Parametric mapping sonification has a number of positive aspects, which Scaletti (1994) outlines in some detail. Many data dimensions can be listened to simultaneously. It is very flexible and the mappings can be easily changed, allowing different aural perspectives of the same data. In addition, acoustic production can be assigned to sophisticated tools originally developed for computer music synthesis. These are readily available and permit many quite sophisticated parameter mappings to be synthesised in real-time.

The main limitation of the technique is the lack linear independence or orthogonality⁷ in the psychophysical parameter space: loudness can affect pitch perception, for example. Though conceptually simple, in practice, parameter mapping requires a working knowledge of how the parameters interact with each other perceptually. Linear changes in one domain produce non-linear auditory effects, and the range of the variation can differ considerably with different parameters and synthesis techniques. These perceptual interactions, caused by coupled perceptual parameters, can obscure data relations and confuse the listener. Flowers, an experienced multivariate data sonifier, observed that while “the claim that submitting the entire contents of ‘dense and complex’ datasets to sonification will lead to the ‘emergence’ of critical relationships continues to be made, I have yet to see it ‘work’” (Flowers 2005). However, although a truly balanced multivariate auditory display may not be possible in practice (Kramer 1994), given powerful enough tools, it may be possible to heuristically test mappings to within acceptable limits for any given application. Frysinger (2005) provide a useful overview of the history of the technique, and Flowers (2005) highlights some of its pitfalls and possible future directions.

2.3.2.2 Direct Data Audification

Direct data audification is a technique for translating data directly into sound. Kramer (1994: 186) used the unqualified *audification*, which he describes as “a direct translation of a data waveform to the audible domain for the purposes of monitoring and comprehension.” Direct data audification may be applicable as a sonification technique for datasets that have an equally-spaced metric in at least one dimension. It is most easily applied to those that exhibit oscillatory time-series characteristics, though this is not a requirement.

⁷ A simple description of orthogonality is that of vectors that are perpendicular, such as the X and Y axes of two-dimensional geometry. The etymological origins are from the Greek ὀρθός (orthos), meaning "straight", and γωνία (gonia), meaning "angle" (OLED). The term is used here in its more general vector space definition: Two vectors x and y in an inner product space V are orthogonal if their inner product is zero. Formally, a linear transformation $T: V \rightarrow V$ is called an orthogonal linear transformation if it preserves the inner product. That is, for all pairs x and y in the inner product space V , $\langle Tx, Ty \rangle = \langle x, y \rangle$. Our concern here is to develop sonification spaces with an equally-spaced metric, that is, that preserve the psychophysical inner-product between a data vector and a linear transformation of it in that space. For a more formal definition of orthogonality, see Behnke, Bachmann, Fladt and Süß (1983: 273). There *are* subtle differences between orthogonality, dependence and correlation (Rogers, Nicewander and Toothmaker 1984) that need not particularly concern us here.

Because of the integrative capabilities of the ear, audification is useful as a technique for very large numerical datasets whose datum can be logically arranged as a time sequence of audio samples. These samples can be either stored in an audio file for delayed audition or streamed directly through the computer's audio hardware in real-time. On playback, any number of standard audio-signal processing techniques such as filtering, frequency shifting, sample interpolation, and time-and-amplitude compression can be applied, perhaps under user control, to enhance information detection. The inclusion of these techniques indicates that *direct* is used as a general descriptive, rather than taxonomic, classifier. So, while direct data audification allows for signal-processing techniques, the defining characteristic is that there are no sound-generating or acoustical models used.

Direct data audification has been shown to be effective in situations where the data is voluminous, such as that produced by monitoring physical systems such as seismology. Speeth (1961) for example, audified seismic data for the 90% successful differentiation of events caused by bomb blasts from those caused by earthquakes. Speeth's experimental results were remarkable because the task is apparently very difficult to achieve using visual plots of the data (Fryinger 2005). By time-compressing the signals to bring them into audio range, analysts could review twenty-four hours worth of data in a few minutes. Hayward (1994), in describing the use of audification techniques on seismic data, found the technique very useful, but stressed that proper evaluation and comparisons with visual methods are needed. In summarising a body of work on earthquake data, Dombois (2002) remarks that

eyes and ears give access to different aspects of the phenomenon of earthquakes. Free oscillations are relatively easy to recognize as the ear is all too familiar with many kinds of resonance. On the other hand synthetic seismograms, which are calculated for fitting as good as possible the curve of a measured seismogram, show low significance in the auditory display. This is less astonishing if one remembers that the power of calculation routines has been judged only by the visual correspondence of measured and synthetic seismogram.

In monitoring the operation of a complex machine, Pauletto and Hunt (2005) determined that key set of attributes (noise, repetitive elements, regular oscillations, discontinuities, and signal power) in helicopter flight data were equally discernable via an audification or a visual spectrogram. Krishnan et al. (2001) undertook a comparative study of direct data audification and other sonification techniques to represent data related to the rubbing of knee-joint

surfaces and did not find that audification was the best technique for displaying the difference between normal and abnormal signals.

In summary, direct data audification with variable sample-rate playback can be useful for data ‘dredging’ large datasets at high speed, for bringing sub-audio information into an audible range, and for realtime monitoring by allowing buffered time-lapse playback of the most recent data. Because of the ways these types of audification appeal directly to a listener’s low level pre-cognitively auditory stimulus-processing faculties, such as those described by the Gestalt psychologists and J. J. Gibson (1966, 1979, §Appendix 2), this technique is useful for monitoring global features of large time-series and in situations requiring extended passive auditing.

2.3.2.3 Homomorphic Modulation Sonification

A homomorphic mapping is one in which the changes in a dimension of the auditory space tracks changes in a variable in the dataset, with only as few mediating translations as are necessary for comprehension (Kramer 2004b: 26). This section describes a narrow interpretation of the term; here named Homomorphic Modulation Sonification. There is a subtle but important distinction between the mapping described here and the parametric mapping approach, described above, in which each datum is played as, or contributes to a separate tone with its own amplitude envelope. In the separate-tones case, the audio-amplitude profile of the resulting audible stream fluctuates from-and-to zero while with modulation, a single continuous pulsed waveform results. In the case of frequency modulation, there is the opportunity for the amplitude formant to be held relatively constant. In research reported in Chapter 5, this appears to assist the listener’s auditory system to respond to it as a single modulating tone rather than a sequence of auditory objects individuated by rapid onset transients. This difference is illustrated in Figure 2.1 and may result in lower perceptual loading, especially for extended listening periods. Patterson’s warning design guidelines study of navigation alerts mentioned earlier (§2.3.1.1), support this suggestion in recommending slower onsets/offsets times to avoid startling the auditor, to impress itself upon their consciousness without it dominating cognitive function.

Pre-attentive perceptual faculties are also applicable when the data is used to frequency- or amplitude-modulate a simple carrier signal. This is a relatively unexplored territory but initial results (Worrall 2004; de Campo 2007) are encouraging: in situations where audification may be an appropriate technique, but the data needs to be ‘massaged’ beforehand, or when there is not enough data to sustain it. For example, it is possible to apply signal processing techniques such as granulated time-stretching (vocoding), shifting the pitch while keeping time invariant, using the data to amplitude- or frequency-modulate a carrier signal. De Campo also suggests modulating frequencies of an array of sines for detection of polarity and time alignments in multiple channels of EEG data.

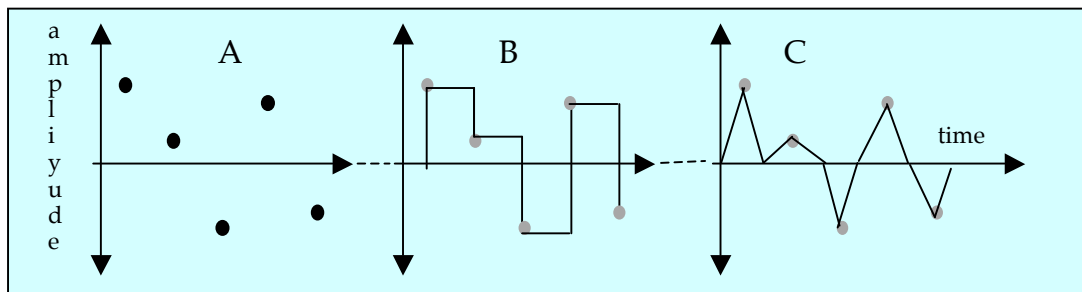


Figure 2.1. The different amplitude profiles of (A) the samples (B) being realised with amplitude modulation, and (C) individually enveloped events.

This possible difference in perceptual loading is emphasised by the noticeable difference between the resultant effect of passing parameter mapped and homomorphic modulation sonification of signals through post-synthesis signal ‘enhancing’ modification such as reverberation. In the case of parametric mapping, the by-products appear to enhance the original, making it more embodied, while for the homomorphic mappings, they appear as distractional, noisy artifacts, as is discussed in the homomorphic modulation sonification study reported in Chapter 5 (§5.7).

The observation that (sonic) object continuity enhances auditory attention has received research attention in the wider context of object continuity in the attentional loading of perception (Best, et al. 2008). The acoustic startle reflex is a gross version of the effect described here. This effect can be summarised as follows:

The minimum stimulus required to elicit a response in normal rats is about 80-90 dB of 50 msec. duration, provided that the stimulus reaches full potential within about 12 msec. initial onset. In mammals, the amplitude of the startle response is

proportional to the duration of the stimulus... The startle reflex has been observed in all mammals tested (Ladd, Plotsky and Davis 2000).

While there does not yet seem to be any empirical studies of the relative perceptual loadings of mono-modal stimuli with different spectral characteristics, there is a growing recognition of the potential for auditory displays to *inform* rather than to just alert (Vincente 2002; Watson and Sanderson 2007) and this is an important consideration for the design of interfaces to continuous sonifications.

2.3.3 Interactive data representations

2.3.3.1 Sound graphs

The term *auditory graph* is used in a variety of ways, often simply meaning the output of a multivariate data sonification. In order to provide a restricted meaning I use the term *sound graph*, to refer to a sonic representation of a visual graph (Stockman, Hind and Frauenberger 2005; Harrar and Stockman 2007). Other names by which it is known are *tone graph*, *auditory graph*, *tree-graph* and *auditory box plot*. Its function is to provide a sonified interface to a discrete dataset so that the relationships between the datapoints can be investigated interactively and asynchronously. The addition of auditory tick-marks, axes, and labels to add context is not uncommon (Stockman, Nickerson and Hind 2005).

Although the original impetus for the sound graph was to provide visually impaired people with access to line graphs and spreadsheets, it clearly has wider application, including as a design space for parameter mapping sonifications. A slightly different interpretation is provided by Vickers (2005) whose *CAITLIN*, based on design principles similar to auditory icons, is used to aurally identify and locate bugs in computer programs.

2.3.3.2 Model-based sonifications

Model-based sonification is a relatively new technique for sonification, in which parameters of a virtual physical model function in same way as the coordinates of a visual display. There are two basic kinds of models available: those based on the known physical properties of acoustic resonators and those based on the structure of the data to be sonified. Digitally modelling the resonating components of a musical instrument (called *physical modelling*) is a

relatively common practice in computer music. For such dynamic models, the temporal behaviour of the components of the model are determined ahead of time by detailing their resonant properties and the way they are connected together to form a single resonator. The instrument is then excited with virtual bows, drumsticks, scrapers etc and virtual contact-microphones placed at strategic places on the 'body' of the instrument to capture its resonance. (Pearson 1996).

In *Model-Based Sonification* (Hermann and Ritter 1999; Hermann 2002), a variable of the dataset to be sonified is assigned to some structural properties of a component (elasticity, hardness etc) of the model. A user interacting with this model via 'messages'—virtual beaters, scrapers etc—causes it to resonate. The resulting sound is thus determined by the way the data integrates through the model. By virtually beating, plucking, blowing and scraping the model, the characteristics of the dataset are available to the listener in the same way that the material and structural characteristics of a physical object is available to a listener who beats, plucks, blows or scrapes it.

Model-based sound synthesis has proved effective in producing complex, 'natural' sounding acoustic events (see Pearson, op.cit. for an extensive bibliography). While the building of such models requires considerable skill, this would not be a barrier for their use in data sonification if templates were readily available. And, although they can be computationally expensive, such models tend to lend themselves to parallel computing. The application of these models to sonification is relatively new and the results of empirical testing against other sonification methods will, in time, determine the degree to which they become generally accepted techniques. Before they become more generally used, it will be necessary to make intuitive implementation tools more widely available.

One reason for exploring such techniques is to try to improve the dimensional orthogogonality of a sonification by integrating a dataset in, or through a quasi-physical medium⁸. Attempts to produce acoustically-dimensioned linear psychophysical spaces, such as with equally-distanced timbres (Barrass 1997: 88-216), so as to differentiate multiple orthogonal dimensions, on which parameter-mapping techniques rely, have not been yet

⁸ This matter is discussed more fully in Chapter 3.

been very successful. The representational dimensions of physical systems, on the other hand, *appear* to be more promising as they are of a higher order than acoustic parameter-maps, involve the redundant use of acoustic parameters, and are structurally closer to those in ecological use in every day listening.

2.4 The need for better tools

The *Sonification Report* considered the field of sonification research to consist of three principal components: (1) Research in perception and cognition, (2) Development of research and application tools, and (3) Design and applications using the sonification techniques. To use these components in consort requires the integration of “concepts from human perception, acoustics, design, the arts, and engineering” (Kramer et al. 1999). Hermann amplifies this by considering knowledge of data availability and meaning, as well as domain-specific dependencies, to be sufficient for simple-purpose sonifications such as for auditory alarms and the enhancement of graphic user interfaces. However in sonifications of higher-dimensional data, he considers expertise in other techniques is also necessary: statistics and data-mining, HCI, computer science, acoustics, sound engineering, physiology and neurobiology, psychology, psychoacoustics, musicology, and cognition (Hermann 2002: 24-25).

The need for better software tools, including some general proposals for adapting sound synthesis software to the needs of sonification research, has been a constant theme of auditory display researchers and was highlighted in the *Sonification Report*. Several attempts to provide generic tools are either in hibernation or extinct but there is evidence that the issues continue to be addressed (de Campo, Frauenberger and Höldrich 2004). As exploratory research continues, and software development iterates, it becomes clearer that the tools needed are not simple adaptations of existing sound synthesis software, fine though much of it is, but more robust models which can better integrate current expertise and the software expression of it from the variety of fields, that need to meet in order to advance the research into how data might be better displayed to our auditory systems. The problems raised in the *Report* prompted the work undertaken for this thesis, as first reported discussed in Worrall et al. (2007) and now more fully developed in Chapter 4.

2.5 Music and sonification: some relationships

In everyday life, sounds are produced when physical objects touch or impact upon each other. It takes no special training to be able to infer certain material qualities from the sound of impacting objects: their density, hardness and perhaps shape for example, as well as the manner of impact: collision and recoil, scraping, plucking, blowing etc. All this information is available and perceived by listeners almost instantaneously and effortlessly millions of times every day, even when they are not looking at the objects, and often even when asleep. Not having earlids, the listener's hearing constantly monitors the surrounding world, and in doing so, directs their visual and kinesthetic attention. The observation that our hearing leads our vision in making sense of the world is amply demonstrated by the importance of Foley (sound effects) in film; the imitation of the sound of a horse's hooves, over stones, through mud etc, by knocking coconut halves together, is much more convincing than an unenhanced recording of the sound of the hooves themselves. An important aspect of Foley is that its effectiveness is in remaining hidden, that is, its functions not to distinguish itself, but to enhance the realism of the audio-visual scene as a whole (Alten 2002: 364–368; Chion 1994).

2.5.1 Acoustic generation

The ability for digital computers to generate, transform and present sound to users in a timely manner is fundamental to sonification research. The vast majority of the tools and techniques used for the computer synthesis of sound have been developed by composers and engineers engaged in the task of making new music, leading John Chowning to remark, "With the use of computers and digital devices, the processes of music composition and its production have become intertwined with the scientific and technical resources of society to greater extent than ever before" (Chowning in Roads 1996: ix). Because, as he goes on to say, "a loudspeaker controlled by a computer is the most general synthesis medium in existence", these researchers were quick to begin exploring adaptations of, and eventually alternatives to, the static models of musical instrument tones as described in the literature of the time (Olson 1967: 201-241) and experience which analog

synthesis had already revealed to be deficient of the lively qualities of the sounds of acoustic instruments.

The original periodic synthesis techniques include additive, subtractive and modulation models (Mathews 1967), which were augmented by dynamic (non-periodic) techniques such as microphone-sampled waveforms, stochastic function and granular synthesis-generated timbres (Xenakis 1971: 242-254; Truax 1988; Roads 2001). As computing hardware became faster and object-oriented software tools became more prevalent, software models of physical resonance systems were developed, ranging from difference equations, mass-spring, modal, non-linear excitation, waveguide, and formant (after speech) synthesis to the highly individual Karplus-Strong techniques for plucked-string and drum-like sounds. Roads provides a brief introduction to these techniques (1996: 261-315), most of which are still under active development. Their usefulness to sonification is that they provide appropriate coupling mechanisms between data structures and the synthesis models being used to express them. As Hermann (2002) concludes, apart from the often-considerable time involved in setting up such systems, their major disadvantages are the time and expertise needed to set up the models and the computing resources necessary to implement them. As experience grows and computation speeds increase, neither of these impediments is likely to prove permanent. Computing science and software engineering continues to play important roles in the development of the theoretical models and efficient algorithmic processes within which these sound synthesis techniques can be implemented.

2.5.2 Sonification research

The systematic study of using computers to *display* information to the human auditory system is a relatively young discipline. The first International Conference on Auditory Display (ICAD) was held in 1992 and Kramer's extensive introduction to the *Proceedings* (1994) provides an overview of interesting precedents, the then current state of the field as well as some possible futures. Before the computing resources with which to process large amounts of data were readily available, sonification was restricted to relatively simple tasks, such as increasing the frequency of a metal detector's audio oscillator in proportion its proximity to metals and the timely

production of attention-capturing beeps and blurps. As the size and availability of datasets have grown and the complexity of communication tasks increased, so also have the expectations of sonification, requiring sonifiers to draw on skills and research knowledge from a variety of disciplines.

In such interdisciplinary enquiries, researchers from various fields come together to share and integrate their discipline-specific knowledge and in doing so reinvigorate themselves with new questions and new methods. For example, when introducing the concept of intelligent help using knowledge-based systems, Pilkington (1992: vii) notes that different disciplines often adopt different forms of evidence. While psychologists prefer experiments, computer scientists prefer programs, composers prefer sound activities, and linguists prefer rules, for example, this can lead to communication failures when they try framing a theory that will satisfy each discipline's criteria of proof. Moore (1991: 23) observes that in interdisciplinary research, the challenge is to do justice to several points of view simultaneously because, "An awareness of [a variety of] points of view is important not so much because ignorance of any one of them makes it impossible to do anything but because what may be done will eventually be limited by that lack of awareness." As experience continually reaffirms, knowledge is not value free, so when these differing points of view come from disciplines with divergent aims, a lack of such awareness may not be obvious. The opportunities to more closely examine disciplinary assumptions that collaborative work provides to researchers can lead through apparent impasses and strengthen the foundations of domain knowledge.

In the current context, it may be worthwhile to speculate on the kinds of knowledge and experience a cultural activity such as music composition can bring to the data sonification process, and inversely, of what use are the results of experimental psychology to music composition. As numerous historical examples attest—from the ancient Greeks, who believed in a direct relation between the laws of nature and the harmony of musical sounds, through the numerical symbolism of the Middle Ages⁹ and beyond—musicians

⁹ One example is Dufay's 1436 motet *Nuper rosarum flores*, which is based on the proportions of the Florence Cathedral for whose consecration it was composed.

were sonifying numerical data long before computers and multivariate datasets occupied the Western mind. While their approaches were perhaps more heuristic and teleological, surviving artifacts from past eras are more likely than not to have been either workable solutions to practical problems, or sufficiently novel to have induced a new way of thinking music. However, except of the very recent past, no sound recordings are extant, and the function of the musical score as something more than an aide-mémoire is only a relatively recent one.

So, although musicologists have described many of the un-notated performance practice and 'style' characteristics of the music from contemporaneous non-musical sources, such descriptions are rarely couched in the phenomenal language of modern science. From the contemporary design perspective, the appropriation of scripted musical gestures of the historical past, together with some a-historical harmony book's 'rules of composition' which omit reference to the pitch or temperament system employed, or the timbral distinctiveness of the instruments of the period, fails to grasp the cultural imperative: music is not science, and its aims are not always for clarity of line, or purpose. The sounds of music are social sounds and as society changes so the ways its components interact are determined as much by cultural forces as by psychoacoustics (Attali 1985: 46-86).

The parametric analysis of tone according to physical characteristics (frequency, loudness, etc) and the availability of computational tools for additive synthesis from these parameters, belies the psychoacoustic evidence that, for all but the simplest sonifications, the results of design by simply assigning data dimensions to such physical parameters quickly leads to unsegregated, difficult-to-interpret clumping (Flowers 2005). Whether stream segregation is maintained during a temporal simultaneity of two or more separate spectra as components of textures or 'chords', or whether integration, occlusion or emergence results (Bregman 1994: 456-528) is dependent upon a combination of characteristics, including the system of tuning in use, the spectral density profiles (timbres) involved, the duration of the simultaneity, the previous and following directional profiles of the spectral contributors, their relative amplitudes, and so on. Furthermore, hearing is very sensitive to spectral congruence, as illustrated by the account auditing example at the beginning of this chapter. Sequences of spectral

events that are highly congruent, such as occurs when using General MIDI instruments, or auditory icons for example, quickly pall on the ear. The extra cognitive load needed to work within such soundscapes is not insignificant, and possibly contributes to the annoyance commonly reported by listeners when using such sonifications for extended periods of time. There are ways out of this impasse, as illustrated by the introduction to sound synthesis techniques in common use in computer music, such as physical modelling (including voice), granulation and stochastics, as well as careful attention to second-order features such as reverberation and spatialisation; these latter being environmental ‘embodying’ of sound sources. The importance of embodiment in information sonification, is discussed more fully in Part B of Chapter 3.

Flowers (2005) comments that most of the problems that arise in data sonification stem from lack of adequate understanding about key properties of auditory perception and attention, and from inappropriate generalisations of existing data visualisation practices. Such generalisations arise more as a result of the overwhelming dominance of vision over audition in the perceptual psychology literature than experimental evidence (Bregman 1994: 1-3). They can be overcome with more sustained applied research in auditory perception, including the rôle played by silence, and a deeper understanding of the relationship between speech and non-speech auditory processing (Slevc, Rosenberg and Patel, 2008). For sonification design to advance, *sonification designing* needs to be practiced and critiqued as designs to be interpreted, not just listened to as *musique e d’ameublement*¹⁰. To that end, task and data analysis of information requirements (Barrass 1997: 35-45) together with generally available datasets and interesting mapping templates in software environments that combine sonic complexity with experimentation flexibility will encourage the much-needed community-wide sharing of examples towards a catalogue of best practice.

2.5.3 Data music

The advent of the Internet and the percolation of computing devices into ordinary, everyday objects and activities, means that most of us exist in an

¹⁰ Literally “furniture music”, as so called by the composer Eric Satie.

increasingly datarised cultural environment. One characteristic of music is that it is continually being recontextualised, both technically and culturally as it is influenced by, and influences, the fluctuating *zeitgeist*. So, from a cultural, experimental composition perspective, data sonification is of significant interest.

One way of classifying data music (data sonification music) is according to where it can be placed on a perceptual continuum, with works using representational data mapping on one end and those using free data transformation on the other. Closer to the 'free data' end would be data music that use arbitrarily formatted digital documents as control data for some sound synthesis routines as arbitrarily determined by the sonifier (Whitelaw 2004). In such circumstances, it is likely that the (possibly imagined) content of the documents would play a role in culturally contextualising the resulting soundscape, in-keeping with postmodernist or other mannerist aesthetics.

At the representational end is data music that employs the techniques used in pragmatic information interpretation for decision-making, system regulation, vision substitution etc.; techniques that are derived from a developing knowledge of psychoacoustics, and the cognitive sciences more generally. For composers who are interested in engaging with listeners as interpreters rather than the struggling receivers of obscure messages, this research offers a stronger technical basis for their practice, not necessarily for overtly programmatic narratives but to better control the overall dramaturgy of a composition (Landy 2007: 36-38). For example, although the techniques for making adherent tone complexes have formed the basis of studies in orchestration for at least two centuries, in order to discover how to disassemble the ensemble, to compose coherent sounds that maintain their perceptual segmentation, what is needed is 'unorchestration' studies that include psychoacoustic techniques for producing disaggregated complex sounds in situational environments (Bregman 1994: 458) and a developing understanding of auditory cognition.

Computer-based composition tools have their origins in the desire to explore and express algorithmically-generated structures at both micro-(sound synthesis) and macro-(gestural) compositional levels. As a cultural activity, making new music is not just about exploring new ways to put new noises together but about providing others with the opportunities to engage,

both actively and passively, in that exploration. Once digital musical instruments became available, MIDI was quickly adopted as a computer protocol suitable for capturing and transmitting data from human performance gesture and for coordination with other technologies, such as lighting controllers. However, not all interesting sources of data have MIDI interfaces or are related to human corporeality and proprioception. So, at a time in human history when there is an imperative for us to take responsibility for our relationships with both natural and man-made environments, the extension of sound synthesis software to aurally explore such datascares opens the possibility of compositional access to, and thus cultural dialogue about, a broader spectrum of phenomena: interspecies, planetary and interstellar.

Notwithstanding the particular application, whether for cultural spectacles or for more pragmatic reasons, there is a general need for a new generation of software; tools that integrate flexible sound-synthesis engines with those for data acquisition, analysis and manipulation in ways which afford both experiments in cognition and lucid, interpretive *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 abstract multidimensional datasets that have no analogue in the material world.

The formation and support of such embodied 'in-forma-tion' is a difficult task and remains a long-term goal. It will be aided by active cross-fertilisation between psychophysics, the relatively new field of cognitive science, and the older disciplines of the philosophy and psychology of perception. The next chapter and its appendices outline the epistemological and ontological dimensions of that study; the cultural 'background radiation' against which new, more effective, sonification software can be engineered.

THIS PAGE HAS INTENTIONALLY BEEN LEFT BLANK