1 • Introduction

Sounds are distinguished from noises by being useful, informative, answering a question, and supporting a task or activity. In the kitchen the sounds of food frying, liquids boiling, timers going off, and pots clanging help the chef prepare a meal. On a construction site the workers use the hammering of nails, clattering of tiles and revving of engines to coordinate their tasks in a common project. On a farm a stockman may listen for cattle hidden in the bush, or diagnose a faulty pump that is making an erratic sound. Sounds are a natural consequence of activity in the physical environment, and although we are seldom aware of it, we are always listening and using sounds in many ways.

1.1 Motivation

The computer-based workplace is unnaturally quiet...and disquietingly unnatural...

Contrast the informative soundscape of the everyday world with the silence of the computer-based workplace. More and more people are spending more and more time carrying out their workday activities in this environment, and there is a suggestion that just as sounds help us in the everyday world they may also be helpful in computer-based activities too. This suggestion is sometimes met with the reservation that, although it might be fun, a "noisy" computer could become annoying. This concern highlights the need to design useful sounds, rather than amusing novelties or distracting noise. Just like the soundscape of the real world, a well-designed soundscape for the computer-based workplace will seldom be noticed, except through the marked effects of its absence.

Sound design is a well-developed practice in cinema. The first movies were silent, but the silent era did not last long. Within 30 years a soundtrack with church bells, train whistles and accompanying music, had become a part of the movie experience. Today the moviegoer is immersed in a multi-layered 3D soundscape of voices, sounds, music and special effects that is often essential to the understanding of the film. The simple bells and whistles of multimedia computer interfaces are reminiscent of the first movie soundtracks, and they too signal the end of a silent era. Although there is much to learn from existing practices, the design of sounds to support information-processing activities is a new challenge because of the types of information involved, and the need to communicate that information clearly and unambiguously. Some major obstacles that need to be addressed in auditory display were described by Smith in a panel on Human Perception and Visualisation [Smith S. (1990)]...

The first obstacle is the prevailing sonification model, which is simply to map data to sound parameters arbitrarily. The resulting sound is typically unpleasant and lacking any natural connection to the data represented (one intuitively feels that medical images, for example, ought to somehow sound different from demographic data or satellite imagery). Models of sonification more sensitive to the kinds of data presented must be developed.

The second major obstacle is the lack of suitable sound generation hardware. Sonification requires a general purpose real-time sound synthesis capability, preferably at an affordable price.

Finally, the third major obstacle is the nearly total absence of the kinds of models that allow design of computer graphics software systems that can run successfully on hardware made by many different manufacturers. The principle reasons for this situation are the *lack of a satisfying comprehensive theory of timbre perception and the lack of an agreed upon theory of timbre generation.*

These translate directly into the situation we observe today: multiple incompatible sound-generation devices, each accompanied by its own suite of non-standard application packages.

The need for a linkage between the characteristics of the data and the auditory perception of the data is reiterated in other surveys of the state-of-the-art. Kendall makes some suggestions about how this linkage might be made in an article titled Dream Machines [Kendall G.S. (1991)] ...

Some classifications of sound events tend to be categorical. Excitation functions are typically of discrete types such as hitting, scraping, blowing, vocal glottis etc. Some classifications of sounding objects are similarly categorical - metal, wood, hollow, solid, vocal tract etc. These simple categorical distinctions can potentially be exploited in auditory presentations to communicate important distinctions in the data.

Beyond these categorical distinctions, the essential goal is that perceptually continuous auditory attributes are scaled and mapped to data attributes in a way that is meaningful to the observer. Relevant changes in data should insure a change in what is perceived. Changes in what is perceived should signify meaningful changes in the data. The appropriate scaling functions will probably not exist a priori. Psychophysical scaling experiments may be needed in order to create perceptual scaling functions through which collections of auditory stimuli are mapped. This is made feasible only by utilizing a limited number of auditory tokens with well-understood perceptual properties. This suggests that sets of tokens be developed and scaled in advance.

Issues of perceptual representation are also highlighted by Frysinger [Frysinger S.P. (1990)]...

Some serious questions which must be addressed if we are to generate reliable and meaningful auditory displays. We must discover the set of truly useful auditory parameters and understand their perceptual transfer functions so that displays can be designed to take advantage of them. Likewise we need to understand which data analysis tasks can most benefit from Auditory Data Representation, and what types of displays apply to them.

Kramer found that the psychoacoustic interaction between acoustic parameters in his multidimensional sonifications had a significant influence on the comprehension of the display, and concluded that the realisation of a balanced auditory display with independently observable auditory dimensions may not be possible in practice. Like Frysinger he also notes the importance of a task-oriented approach [Kramer G. (ed) (1994b)] ...

Sonification system design should be heavily task dependent. Techniques that are applicable to one task, e.g. real time monitoring, may not be as effective on other tasks, e.g. exploration of a high-dimensional data set. Practical use of the techniques (and of sonification in general) will prove or disprove their utility.

The need for demonstrations of practice, and tools to support practice, is raised by Scaletti in her list of open questions in sonification from ICAD'92 [Scaletti C. (1994)]...

Broad categories for further study include:

Applications: further and more sophisticated examples of sonification applied to specific problems

Sonification science: studies in the perception, cognition and neurophysiology of data-driven sound

Systems: Needs for future hardware and software include: integrated sonification/visualisation languages, tools for getting from an imagined sound to a realised sound, the integration of sonification tools into mass market software like spreadsheets or statistical analysis packages, and more and better tools for exploratory sonification.

This thesis is motivated by the need address obstacles identified by previous researchers as important for progress in the field of auditory display, which I have listed as

- usefulness of the sounds in an activity
- faithful representation of data relations
- semantic linkage to the application domain
- psychoacoustic control
- device-independent display specification
- computer aided tools for auditory information design
- demonstrations of practical application

1.2 Thesis

My thesis is an approach for designing sounds to support information processing activities. The approach focuses design on an auditory representation to meet the information requirements of the activity. This focus is at the core of a system of methods that address the array of issues raised by previous researchers. A task-analysis addresses the issue of usefulness. A data characterisation addresses the issue of faithful representation. A case-based design method addresses semantic linkage with the application domain. A rule-based method addresses psychoacoustic control. A perceptually linearised sound space addresses device-independent specification. This multifaceted system is flexible to cope with the obstacles encountered in design practice. Most of these methods have never been applied in auditory display before, and each has been adapted specially for this design domain.

1.3 Layout and Overview

The layout of the thesis chapters is organised to reflect the TaDa design process. An introduction to each Chapter is given here to give you an overview of what is to come.

Chapter 2: Previous approaches

Approaches to auditory display have been classified into the semiotic types of lexical, syntactic and semantic [Blattner M.M. Papp A.L. and Glinert E.P. (1994)] which place different emphasis on learnability, organisation, and discrimination. However there are other approaches to the design of information displays that raise other issues about usefulness, usability, social value and the realisation of the design on a device. This chapter begins with an introduction to semiotic terms, followed by descriptions of approaches to design that emphasise syntactic, semantic, pragmatic, perceptual, task, connotation and device issues. The collection of a broad range of issues may help us to develop a broad and flexible approach to auditory information design

Chapter 3: Designing useful sounds

This chapter proposes an approach for designing useful sounds. The approach builds on Scaletti's working definition of sonification, which is analysed to have two parts - one part has to do with information requirements and the other with information representations. The requirements part addresses issues of usefulness in a task and the selection of useful data relations to display. The representation part addresses the need to ensure that people can hear the required information in the display. These parts are shaped into a framework that focuses on the design of an information representation to meet the information requirements of a task. The phrase *auditory information design* indicates this focus on *useful information* which is at the core of this approach.

Chapter 4: TaDa: information requirements analysis

The purpose of the TaDa approach is to design sounds that carry useful information. The core of the approach is a meeting of information requirements with an information representation. This chapter describes the methods used to elicit the information requirements of a design problem. The first section introduces scenario analysis as a technique for capturing key features of the problem. The next sections describe the particular flavour of task analysis and data characterisation used to decompose the problem, and give a detailed account of the parts of analysis.

Chapter 5: EarBenders: case-based design from stories

Designers often base a new design on a previous version that has proven successful in similar problems. A previous solution can be a way to quickly come to grips with a design problem, and provides a starting point for a top-down process of iteration. This chapter introduces the case-based method of design by example, and describes how it has been adapted for auditory information design. The case-based method relies on a rich source of examples to be effective, but as yet there are not many examples of auditory display design to draw upon. An alternative resource was developed by collecting stories about everyday listening experiences into a database, which I call EarBenders. The information requirements of a design problem can be used to search this case-base for everyday examples which share a similar task, data and information structure with the problem. The ability to do this search required each story to be analysed with the TaDa Information Requirements Analysis developed in the previous chapter. In addition an auditory characterisation was developed to describe the sounds in each story, and provide a footing for auditory design. The sound characterisation also provides an opportunity to extract principles of sound design from regularities between auditory structure and information structure in the example cases. The case-based design of auditory information is demonstrated on a problem in a geological visualisation interface, called the GeoViewer.

Chapter 6: Hearsay: principles for auditory design

This chapter proposes the Hearsay principles to help a designer to meet the information requirements specified by the TaDa analysis. Hearsay integrates principles for information design with observations about auditory perception. Each Hearsay principle was investigated by generating a simple auditory demonstration to confirm that characteristic properties can be heard. The demonstrations show that the required information can be represented by auditory relations, and that the Hearsay principles are applicable in practice. The principles were tried out in a design of an auditory display for Bly's 'dirt and gold' scenario. The display enables a listener to quickly answer the question "is there gold in this pile of dirt?". The effectiveness of this display indicates that the principles are helpful in practice.

Chapter 7: Information-Sound Space: a cognitive artifact

The previous chapter introduced the Hearsay principles for auditory information design, that summarise some knowledge that can help in the design process. Although they are helpful, principles and guidelines can be unwieldy in practice because of the need to keep referring back to them. Principles cannot be simply applied by rote, they have to be learnt and understood. This chapter describes an alternative representation of the Hearsay rules in the form of an Information-Sound Space (ISS). The ISS is a three dimensional spatial organisation of auditory relations that embodies the Hearsay principles. This tool bridges the gap from theory to practice by changing the way the designer can think about and manipulate relations between sounds. Rather than having to follow written principles the designer is able to think in terms of simple spatial structures that represent information relations. The following sections describe the development of the ISS, which is based on the HSL colour space that has been applied in many areas of design including scientific visualisation of data sets. The feasibility of implementing an ISS is investigated in several experiments that draw upon psychoacoustic observations made by Von Bismarck, Slawson, Grey, and Bregman.

Chapter 8: GreyMUMS: an Information-Sound Space

This chapter describes the realisation of an Information-Sound Space (ISS). The raw material for the construction is the McGill University Master Samples (MUMS) reference palette of musical samples that is specifically intended for research into musical timbre. The ISS was constructed in 4 stages - the pedestal, the frame, the grating and the plasticine. The pedestal is 8 equally spaced timbre steps organised in a circle by similarity. The frame is an arrangement of brightness profiles which define the limits of dynamic range in pitch and brightness for each timbre. The grating consists of grids of equal differences in brightness and pitch for each timbre. The grating grids are joined on a central axis and radiate outward like segments of a mandarin. The plasticine is a continuous medium moulded to the grating and frame to model the behaviour of the overall space. The resulting sculpture has the property that there is a relationship between distance and the strength of perceptual grouping between points in the space. A vertical line is a pitch scale, and may be used to represent continuous data. A radial line is a scale of equal brightness increments for a timbre, and may also be used for continuous data. A circle of constant radius is a contour of constant brightness across the range of timbres which can be used to represent categorical data. These properties are a rich area for further experiment with data mappings.

Chapter 9: Personify: computer-aided design tool

Multi-modal interfaces are becoming increasingly important, and designing sounds for the human-computer interface is something more people are going to want to do. This chapter describes the Personify tool that can assist in the design of useful and effective sounds. The tool integrates a principled design approach with a direct manipulation interface. The guidance provided by the tool makes it quick and easy to use, and improves the likelihood of producing an effective display. The chapter begins with an overview of interfaces for handling audio material found in musical tools. This is followed by an overview of tools that are specialised for auditory display design, with attention to the way these tools allow you to handle sounds. The Personify tool is then described in two parts - the Requirements part and the Representation part. The meshing of the parts as a system is demonstrated in a design scenario that involves resource monitoring by satellite data.

Chapter 10: TaDa! demonstrations of auditory design

This chapter describes the design of 4 auditory displays for information processing scenarios drawn from mining exploration, resource planning and climatology applications.

The RiverAndRain scenario is about the siting of a new sewerage treatment works to minimise the environmental impact on a river system. The PopRock scenario involves the assessment of risk in digging a mineshaft. The cOcktail scenario is about modelling climate change from measurements of oxygen isotopes in sea-bed drill-core sites. The LostInSpace scenario is about navigating back to an interesting place through irregular structures in a 3D visualisation of geology. Experiences with the multimedia interfaces that were implemented shows that the sounds can provide information that is difficult to obtain visually, and can improve the usefulness of the display. Besides showing ways that sounds can be useful, the demonstrations also show how the TaDa approach to design works in practice.

2 • Previous approaches

Approaches to auditory display have been classified into the semiotic types of lexical, syntactic and semantic [Blattner M.M. Papp A.L. and Glinert E.P. (1994)] which place different emphasis on learnability, organisation, and discrimination. However there are other approaches to the design of information displays that raise other issues about usefulness, usability, social value and the realisation of the design on a device. This chapter begins with an introduction to semiotic terms, followed by descriptions of approaches to design that emphasise syntactic, semantic, pragmatic, perceptual, task, connotation and device issues. The collection of a broad range of issues may help us to develop a broad and flexible approach to auditory information design.

2.1 Semiotics

Semiotics is a theory of signs and their meanings that has been used to analyse communication media. Some key terms and concepts in semiotics are introduced here (based on Chandler's tutorial [Chandler D.C. (1997)]).

A "sign" is anything from which a meaning may be generated - words, sounds, photographs, clothing etc. A sign has two parts - a "signifier" which is the form that the sign takes, and the "signified" which is what it represents to the person who perceives it. Semiotic principles are commonly divided into 3 kinds - syntactic, pragmatic, and semantic.

- Syntactic principles bear on the way signs are organised to produce meanings. Signs may be organised in parallel (paradigm) or serial (syntagym). Rules for organising signs are called grammars.
- Pragmatic principles bear on the material form of the signifier in the sign. It is usually important for different signs to be perceptually distinct, recognisable and memorable. A set of signifiers that have a special purpose, such as the letters of the alphabet, are called a lexicon, or sometimes a palette.
- Semantic principles bear on what is signified by the sign. The signified is a concept in the mind of an observer, not a material "thing". The signified concept of a ball contains many different objects of different materials, sizes, colours and shapes. Concepts can be learnt from other people, or from experiences. The association between a signifier and the concept that is signified is commonly classified into 3 kinds
    - Symbolic the signifier does not resemble the signified e.g. the spoken word "ball" doesn't sound like a ball, and the written word "ball" doesn't look like a ball.
    - Indexical the signified is causally connected to the signifier e.g. hitting a tennis ball makes a characteristic whacking sound, bouncing a basketball makes a bouncing sound.
    - Iconic the signifier resembles the signified e.g the sampled sound of a cricket chirping in the soundtrack of a movie signifies "real" crickets chirping, and

the photographic image of a person signifies a “real” person.

Signs can signify more than one thing at the same time. The concept that the sign stands for is called the “denotation” and additional signifieds are called “connotations”. Cultural associations generate connotations by metonym and metaphor. A metonym invokes an idea or object by some detail or part of the whole - a picture of a horseshoe may be a metonym for a horse. A metaphor expresses the unfamiliar in terms of the familiar - a picture of a tree may be a metaphor for a genealogy. A metonym is considered more “natural” than a metaphor because it does not require such a leap of imagination (transposition). Connotations may transform the meaning of a sign through emotional overtones, subjective interpretations, sociocultural values and ideological assumptions.

Auditory display techniques were classified as syntactic, semantic and lexical by Blattner, Papp and Glinert [Blattner M.M. Papp A.L. and Glinert E.P. (1994)]. An example of a syntactic approach is the earcon, which is a short musical motif with structure modelled on pictographic writing [Blattner M. Sumikawa D. and Greenberg R. (1989)]. An example of a semantic approach is the auditory icon, which is modelled on the everyday sounds caused by interactions between material objects [Gaver W.W. (1986)]. Lexical approaches map data variations to acoustic variations, and examples include Bly’s multivariate mappings [Bly S. (1994)], Kramer’s “parameter nesting” [Kramer G. (1994a)] and the granular texture technique described by Smith, Pickett and Williams [Smith S. Picket R.M. and Williams M.G. (1994)].

2.2 Syntactic approach

The emphasis in the syntactic approach is on the organisation between auditory signs. Morse code is a syntactic approach where variation in duration and rhythm of a non-speech sound can communicate coded text messages. The earcon is a syntactic method for designing non-speech sounds to represent information in human-computer interfaces. An earcon is built from components that may vary in rhythm, pitch, timbre, register, and dynamics. Each earcon has a unique meaning that must be learnt - for example a tone X with pitch 440 Hz may mean “file”, and tone Y with pitch 600 Hz may mean “deleted”. These earcons can be combined to communicate more complex messages - for example playing X and Y in series produces a rising XY earcon that means “file deleted” [Blattner M. Sumikawa D. and Greenberg R. (1989)]. The syntactic structure of an earcon can be organised by transformations, combinations, inheritance, and polyphony. Blattner et al. suggest that earcons have the advantages of

- *Ease of production: earcons can be easily constructed and produced on almost any computer with tools that already exist for music and audio manipulation.*
- *Abstract representation: earcon sounds do not have to correspond to the objects they represent, so objects that either make no sound or an unpleasant sound can still be represented.*

Earcons were added to a map of a research facility to provide extra information that could be heard by pointing or selecting a region with the mouse [Blattner M.M. Papp A.L. and Glinert E.P. (1994)]. Access privileges of a building were heard by a knocking pattern of a tom-tom drum. Higher restrictions are heard by a faster knock and higher pitch. The presence of computers in a building was heard as a four note flute tune. The presence of an administrative unit in a building was heard by a 3 note saxophone tune. This demonstration of earcons raises an important issue - symbols are categorical and the categorical representation of ordered information requires a decoding phase. The need for more direct representations of ordered values is recognised in the speeding up of rate to represent the higher restriction levels in the access privileges earcon. However this type of representation is not an explicit part of the earcon method.

A major problem with earcons is learnability. Novices are able to learn 4-6 symbolic sounds within minutes, but further learning of up to 10 signals can take hours. Beyond 10, the process is prolonged and some listeners may never learn the catalogue completely [Patterson R.D. (1982)]. There is no standard syntax or lexicon of earcons, and the investment of time and effort in learning a new set may be too great for many applications.

2.3 Semantic approach

The emphasis in the semantic approach is on what is signified by a sound. The semantic method for sounds in user interfaces is called the auditory icon. The auditory icon method is to map what is signified by a familiar everyday sound to objects and events in the user interface [Gaver W.W. (1986)]. Gaver suggests that sounds modelled on real world acoustics are likely to be learnable and easy to understand because humans are adapted to hear information in these kinds of sounds. This suggestion is based on a theory of perception that says that physical variations in acoustic energy generated by interactions between material objects are innately perceived and intuitively understood [Gibson J.J. (1966)].

The design of an auditory icon starts with an analysis of interactions between objects in the interface which would cause sounds in the physical world. For example, moving a file in the Apple desktop GUI involves dragging it between windows. An auditory icon for this event is based on the sound caused by dragging a real file across a real desktop. Gaver demonstrated auditory icons in the SonicFinder, which augmented the Apple desktop GUI with auditory icons for selecting, dragging and copying files, opening and closing folders, selecting scrolling and resizing windows, and dropping files into and emptying the trash can [Gaver W.W. (1994)]. He comments that although many people found the auditory cues useful, others found them irritating or merely entertaining - providing an indication of the challenges facing the designer in real-world practice. In another demonstration, called the ARKola experiment, the auditory icons were designed to provide information that was not visible in the interface [Gaver W.W. Smith R.B. and O'Shea T. (1991)]. Participants collaborated in pairs to control a simulated cola bottling factory made up of 9 interconnected machines, such as a heater and a bottler, with on/off and rate controls. Each participant could only see half of the factory, but could hear a characteristic sound for each machine, which varied in a predictable manner with the rate of operation. Each sound was designed to have a semantic association with the machine it represented, for example the heater made a whooshing sound like a blowtorch, and the bottle dispenser sounded like bottles clattering along a conveyer belt. The participants were able to quickly learn and remember the meaning of the auditory icons in the process of using them. The sounds helped them to track ongoing processes, and to monitor individual machines as well as the overall condition of the factory. The participants could refer to the sounds to discuss unseen elements, and the sounds increased the enjoyment of the activity. Some problems in the discrimination of overlapping sounds were observed, which Gaver et al. suggest could be addressed by a more systematic approach to shaping the psychoacoustics of the sounds.

The auditory icons in the ARKola experiment were able to signify amounts, such as rate or quantity, directly and without the need to refer to a legend. This is an important demonstration that sounds can convey quantitative information, and support decision making based on quantitative data. The fact that the participants learnt the meaning of the sounds very quickly seems to indicate that the semantic design method is intuitive to understand, as Gaver suggested. However the theory that auditory icons are more learnable than other sonic signs has not been supported by empirical investigations. Lucas found no significant difference between auditory icons and earcons in the amount of time taken to learn associated meanings, in the number of errors made in the process of learning, or in the improvement in these factors over two trials. It took significantly less time to learn speech associations, and these were consistently error free. The factor that most influenced the accuracy of recognition of the auditory icons and earcons was an explanation of the rationale behind the sound design [Lucas P.A (1994)].

This result does not accord with the ecological theory, but rather indicates that intermediate mental structures do have a role in listening. This observation is supported by Ballass' experiment which found that the speed and accuracy of identification of the source of a sound depends on the listeners expectations, context and experience. The experiment tested the ambiguity of the relation between a sound and its perceived source by embedding subjectively similar sounds, such as a fuse burning and leaves crackling, in contextual sequences which were biased toward one or the other interpretation. An inappropriate context had a clear negative effect, although an appropriate context did not seem to have a symmetrical positive effect [Ballass J.A. (1994)].

Truax comments that if an audio signal is perceived or measured to have "fidelity" to the original source it is thought to have been successfully reproduced - but this assumption causes a schizophrenic fracture between the original source and its later out-of-context usage [Truax B. (1992a)]. Beck describes how most of the sounds in movies, radio and theatre are not sampled directly from the real world but are deliberately constructed to foster an impression in the mind of the listener, through expectations and cultural experience, as well as physical experience [Beck M. (1996)]. The subjective nature of auditory experience was highlighted by the inconsistency that Mynatt found in her experiment on associations between sampled sounds and interface concepts. [Mynatt E.D. (1994)]. The experiment tested associations between 28 sounds, such as a zipper, and 15 interface concepts, such as a pull-down menu, from a sound to a concept and from a concept to a sound. The same sound was chosen for many different concepts, and there was rarely any convergence on a single sound for a single concept. Mynatt concluded that the design of sounds for the human-computer interface is difficult and painstaking process which depends on the skill of gifted individuals. She summarised her findings as guidelines that may assist other designers

- *Choose short sounds that have a wide bandwidth, and where length, intensity and sound quality are roughly equal.*
- *Evaluate the identifiability of the auditory cues using free form answers.*
- *Evaluate the learnability of the auditory cues that are not readily identified.*
- *After deciding on a collection of sounds that are easily identified or learned, evaluate possible uses of the auditory cues in the interface being designed.*
- *Evaluate possible sets of auditory icons for potential problems with masking, discriminability and conflicting mappings.*
- *Finally the complete interface should undergo usability evaluations with subjects or*

users conducting specified tasks, as well as long-term evaluations that monitor the performance over a long period of use.

The studies of learnability and identification with auditory icons suggest that what is signified depends on listening experience and context. Although the method is based on a theory of innate perception that should ensure consistent interpretation the experimental results show significant individual differences.

2.4 Pragmatic approach

The pragmatic approach emphasises the form of the signifier. The set of signifiers in a lexicon need to be discriminably different to represent different signifieds, and to prevent ambiguous combinations.

Early work on the pragmatics of sounds was prompted by the need to ensure that a pilot could hear warnings and alarms against a background of engine noise and radio conversations. A study of sounds for aircraft display systems was carried out by Patterson. Although his guidelines were intended for sounds in cockpits, they may be helpful in computer interfaces too [Patterson R.D. (1982)].

- *Overall level - The lowest intensity level of a warning sound should be 15dB above the threshold imposed by background noise. The upper limit is 25dB above threshold.*
- *Temporal - Component pulses of a warning sound should have onsets and offsets 20-30 ms in duration. These avoid a startle response in the listener. Pulses should be 100-150 ms in duration with a 150 ms inter-pulse gap for urgent sounds and a 300 ms gap for non-urgent sounds. Distinctive rhythms of 5 or more pulses should be used.*
- *Spectral - The fundamental frequency of a warning should be within the range 150-1000 Hz. There should be four or more component harmonics to help avoid maskings. The overall spectral range of warnings should be 500-5000 Hz.*
- *Ergonomics - Manual volume control should be avoided and automatic control restricted to a range of 10-15 dB variation. There should be no more than six immediate action warnings.*
- *Voice warnings - These should be brief and use a keyword format. They should not be repeated in a background version of the warning. Voice warnings used as immediate awareness warnings should use a full-phrase format and be repeated after a short pause.*

In another study of auditory displays for aircraft, Deatherage not only investigated the design of sounds as alarms but also as aids for spatial information and orientation [Deatherage B.H. (1972)].

Alarm and warning signals:

- *At a minimum, use sounds having frequencies between 200 and 5000 Hz, and if possible, between 500 and 3000 Hz, because the human ear is most sensitive to this middle range.*
- *Use sounds having frequencies below 1000 Hz when signals must travel long dis-*

tances (over 1000 ft.) because high frequencies are absorbed in passage and hence cannot travel as far.

- *Use frequencies below 500 Hz when signals must bend around obstacles or pass through partitions.*
- *In noise, signal frequencies different from those most intense frequencies of the noise are best in order to reduce masking of the signal.*
- *Use a modulated signal to demand attention. Intermittent beeps repeated at rates of one to eight beeps per second, or warbling sounds that rise and fall in pitch are seldom encountered, and are therefore different enough to get immediate attention. If speech is necessary during an alarm, use an intermittent, pure-tone signal of relatively high frequency.*
- *Use complex tones rather than pure sinusoidal waves, because few pure tones can be positively identified but each complex sound is noticeably different from other sounds.*

Spatial information:

- *Use auditory displays to relieve the eyes. Although the eye is better for spatial discrimination, it can look in only one direction at a time. In general, auditory spatial displays are recommended only when the eyes are fully engaged and additional spatial information is needed.*
- *Use auditory displays (other than speech) to present restricted information, such as the following:*
  (a) "Yes-no" information and indications of amount or degree. Auditory displays can represent error or deviation from a course, speed, attitude, or other "normal" condition.
  (b) Continuous information. For example, radio-range signals present practically continuous information about one kind of event - the course the aircraft is flying.
  (c) Automatic information - recorded word signals as from an automatic annunciator.
- *Use auditory displays of tonal or noise signals when speech channels are already fully employed. Most of the auditory displays that utilize tonal signals can be heard through speech, and, conversely, speech can be understood while hearing the tonal signals over the same receiving system.*

Spatial orientation:

- *Confine representation to a single dimension; multidimensional displays are less effective than their visual counterparts.*
- *Provide a standard stimulus to represent the "normal" then make abrupt changes to indicate departures from the normal. Human listeners are sensitive to frequency or intensity changes but poor at identifying a unique signal.*
- *Provide changes in intensity rather than frequency as a spatial cue. Because everyone with normal hearing can detect changes in intensity, it is easier to control these changes.*
- *Use intermittent to repeated changes in a signal rather than a single change followed by a continuous signal. The ear is much more likely to detect changes in a signal*

occurring every second or two than at longer intervals.

- *If absolute identification is required, limit the number of signal categories to four because listeners cannot identify correctly more than a few different intensities, pitches, or interruption rates.*
- *The following "natural" relationships between auditory signals and the dimensions they represent are quickly learned or are perceived with little training:*
  (a) Binaural intensity differences serve to localise (in bearing) the direction of a sound.
  (b) Pitch differences naturally represent up and down (high and low pitch). To indicate climb or "upward pointing" raise the pitch. Combined with binaural changes in pitch from one ear to the other, "left wing high" for instance, can be represented.
  (c) A slow interruption rate is a natural indication of speed - an increase or decrease in interruption rate is immediately perceived as a change in speed (or rate) of interruption.

These recommendations were brought together with other observations, by McCormick and Sanders, to provide some general guidelines that extend to other types of auditory displays outside the cockpit [McCormick E.J. and Sanders M.S. (1983)]. They begin with a list of circumstances where an auditory display is preferable to a visual display:

- *When the origin of the signal is itself a sound.*
- *When the message is simple and short.*
- *When the message will not be referred to later.*
- *When the message refers to events in time.*
- *When sending warnings or when the message calls for immediate action.*
- *When presenting continuously changing information of some type, such as aircraft, radio range, or flight path information.*
- *When the visual system is overburdened.*
- *When speech channels are fully employed (in which case auditory signals such as tones should be clearly detectable from the speech).*
- *When illumination limits vision.*
- *When the receiver moves from one place to another.*

General principles

- *Compatibility - where feasible, the selection of signal dimensions and their encoding should exploit learned or natural relationships of the users, such as high frequencies associated with up or high, and wailing signals with emergency.*
- *Approximation - two-stage signals should be considered when complex information is to be presented, these stages consist of 1) Attention-demanding signal: to attract attention and identify a general category of information 2) Designation signal: to follow the attention demanding signal and designate the precise information within the general class indicated above.*
- *Dissociability - auditory signals should be easily discernible from any ongoing audio*

input (be it meaningful input or noise). For example if a person is to listen to two or more channels the frequencies of the channels should be different if it is possible to make them so.

- *Parsimony - Input signals to the operator should not provide more information than is necessary.*
- *Invariance - The same signal should designate the same information at all times.*

Principles of presentation

- *Avoid extremes of auditory dimensions - high intensity signals, for example, can cause a startle response and actually disrupt performance.*
- *Establish intensity relative to the ambient noise level - this is simply saying that the intensity level should be set so that it is not masked by the ambient noise level.*
- *Use interrupted or variable signals - where feasible, avoid steady-state signals and, rather, use interrupted or variable signals. This will tend to minimise perceptual adaption.*
- *Don't overload the auditory channel - only a few displays should be used in any given situation. Too many displays can be confusing and will overload the operator. (For example, during the Three Mile Island nuclear crisis, over 60 different auditory warning displays were activated).*

Principles of installation of auditory displays

- *Test signals to be used - such tests should be made with a representative sample of the potential user population, to be sure the signals can be detected by them.*
- *Avoid conflict with previously used signals - any newly installed signals should not be contradictory in meaning to any somewhat similar signals used in existing or earlier systems.*
- *Facilitate change-over from previous display - where auditory signals replace some other mode of presentation (e.g. visual), preferably continue both modes for a while, to help people become accustomed to the new auditory signals.*

The pragmatic design of sounds to support human-computer interaction was investigated by Brewster, Wright and Edwards who sought to answer the question “what sounds should be used at the user interface?” Their empirical observations are summarised as guidelines for designing earcons, shown in Table 2-1 [Brewster S.A. Wright P.C. and Edwards A.D.N. (1994)].

Timbre	*Use musical instrument timbres. Where possible use timbres with multiple harmonics. This helps perception and avoids masking. Timbres should be used that are subjectively easy to tell apart e.g. use “brass” and “organ” rather than “brass1” and “brass2”.*
Pitch	*Do not use pitch on its own unless there are very big differences between those used (see register below). Complex intra-earcon pitch structures are effective in differentiating earcons if used along with rhythm. Some suggested ranges for pitch are: maximum 5 kHz (four octaves above middle C) and minimum 125 Hz to 150 Hz (an octave below middle C).*
Register	*If this alone is to be used to differentiate earcons which are otherwise the same, then large differences should be used. Two or three octaves difference give good rates of recognition.*
Rhythm	*Make them as different as possible. Putting different numbers of notes in each rhythm was very effective. Patterson says that sounds are likely to be confused if rhythms are similar even if there are large spectral differences. Small note lengths might not be noticed so do not use notes less than eighth notes or quavers. In the experiments described here these lasted 0.125 seconds.*
Intensity	*Although intensity was not examined in this test, some suggested ranges from Patterson are: maximum 20dB above threshold and minimum 10 dB above threshold. Care should be taken in the use of intensity. The overall sound level will be under the control of the user of the system. Earcons should all be kept within a close range so that if the user changes the volume of the system, no sound will be lost.*
Combinations	*When playing earcons one after another, use a gap between them so that users can tell where one finishes and the other starts. A delay of 0.1 seconds is adequate. If the above guidelines are followed for each of the earcons to be combined, then the recognition rates should be sufficient.*

Table 2-1: Guidelines for the pragmatic design of earcons

Auditory representation techniques have also been evaluated from a pragmatic perspective. A system for psychometric evaluation of auditory displays was implemented by Smith, Levkowitz, Pickett, and Torpey [Smith S, Levkowitz H., Pickett R.M. and Torpey M. (1994)]. They classify perceptual discrimination into 3 types

- Detection - the discernment of a shift (or just noticeable difference)
- Recognition - the discernment of a shift and discrimination between different kinds of shifts
- Scaling - the discernment, discrimination and characterisation of a shift

The threshold of shift detection for an auditory parameter is measured by a three-alternative, forced-choice, up/down technique. Evaluations are made against a statistically parameterised data set, in which the mean, standard deviation and other features can be systematically controlled. The system was demonstrated in an investigation of an asynchronous granular synthesis (AGS) algorithm with ten parameters. The results show complex interactions between the parameters, perceptual non-linearities in response to linear changes in the data, and wide variation in the dynamic range of various perceptual effects. Smith et al. concluded that their AGS algorithm is less than ideal for auditory data representation. Some of the issues they raise are that sounds that are interesting and lively in musical and artistic compositions can become irritating when presented repeatedly in an experiment. Sound synthesis algorithms have idiosyncratic parameters that are not aligned with perceptual variations, and which cannot be generalised. They comment that the development of a device-independent and easy to use way to specify timbres is critical before auditory data representation can make a further leap forward.

The types of discriminations measured by Smith et al. are called “capabilities” by Watson and Kidd. They propose a principle of “proclivity” to explain their findings that performance on very familiar sets of stimuli may not be predictable from performance on the discrimination of unfamiliar sounds [Watson C.S. and Kidd G.R. (1994)]. Proclivities reflect responses that cannot be scored as right or wrong, such as judgements of sound quality, of the value of a sound on a perceptual scale, and of the similarity of a sound to another sound. Psychophysical measures of acuity are not sufficient to explain the large individual differences in performance, and central “cognitive” factors of learning, memory and attention may have to be considered if we are to predict the human operator’s performance with arrays of auditory signals. Their observations are summarised by some guidelines

- *Auditory capabilities determined under minimal uncertainty are both valid descriptions and misleading guides.*
- *Auditory capabilities determined under high levels of stimulus uncertainty or under “ecologically valid” conditions can be both invalid descriptions and useful guides.*
- *Performance on auditory identification tasks can be greatly improved through intense training, but pilots, cardiologists, and nuclear power station operators aren't likely to devote the required hours in learning an arbitrary auditory catalogue of more than 7 or 8 discriminable sounds.*
- *One way to convey more information is to use the auditory signal only to indicate (a) that a message has arrived and (b) the priority of that message. The actual message then can be conveyed by voice or by alphanumeric display.*
- *Establishing larger catalogues of auditory signals may be possible, but may benefit from:*
  (a) exploiting pre-existing sound-event associations,
  (b) designing sounds with maximal pairwise discriminability, based on psychoacoustic principles,
  (c) establishing a grammar, or other form of sound organisation.

The perception of overlapping auditory elements has been explained by Bregman in his theory of auditory scene analysis [Bregman A.S. (1990)]. This theory proposes that acoustic elements are grouped into “streams” according to heuristics such as similarity, proximity, closure and familiarity. The heuristics help to predict the outcomes of a collaboration and competition between two auditory processes:

- Primitive grouping - fast, innate grouping by acoustic properties of the proximal stimuli.
- Schema segregation - slower, conscious selection of elements from groups that have been formed by the primitive process, and active restructuring of groups by effort of will.

The primitive level explains auditory illusions and why sounds can be hidden or camouflaged due to the acoustic interactions. The schema level explains why the motivation and experience of the listener can make such a difference to what they hear. Some properties of streams that may be relevant in display design are [Bregman A.S. (1990)]

- Only material grouped into the same stream can obscure or camouflage a target.
- Simple tasks such as counting the number of tones are more accurate if the tones are in the same stream.
- Temporal relations are difficult to make across streams - for example it is very difficult to judge the order of elements in separate streams, or compare the rates of cyclic sequences which have segregated.
- An element in a stream may be captured by another stream with elements that are similar
- A rhythm tends to be defined by sounds that fall in the same stream. The allocation of sounds to different streams affects what rhythms may be heard

Williams says that a knowledge of the potential auditory streams that may arise from a particular acoustic signal is essential in order to predict the possible interpretations of that signal, and lists gestalt heuristics as principles for auditory display design as shown in Table 2-2 [Williams S.M. (1994)]. This approach may help in the design of concurrent and overlapping sounds. However it doesn't tell us how to map data relations to auditory relations.

Similarity	*components which share attributes are perceived as related.*
Proximity	*the closer two components are, the more likely that they belong together.*
Good continuation	*components that display smooth transitions from one state to another are perceived as related.*
Familiarity	*recognition of well-known configurations among possible subcomponents leads to these subcomponents being grouped together.*
Belonginess	*a component can only form part of one object at a time and its percept is relative to the rest of the figure-ground organisation to which it belongs.*
Common fate	*components which experience the same kinds of changes at the same time are perceived as related.*
Closure	*incomplete figures tend to be completed.*
Stability	*having achieved an interpretation that interpretation will remain fixed throughout slowly changing parameters until no longer appropriate.*
Articulation	*the separation of a figure from ground requires energy.*

Table 2-2: Gestalt heuristics

Psychoacoustic observations describe the relation between acoustic variations and what is heard by a listener. This pragmatic method works well for simple displays by ensuring the discrimination of categories and linearity of scales, suggesting that psychoacoustic measurements and theories can assist in the design of an auditory display. However complex multidimensional displays are not easy to measure, and the predictions made from these measurements are not always valid outside the laboratory situation. Perhaps the biggest problem with this approach is that psychoacoustics does not provide guidance about the relation between perception and information.

2.5 Perceptual approach

Graphs show information in sets of abstract numbers. Graphic information is not contained in individual signs, but in the perceptual relations between signifiers. Bertin proposed that graphic relations were a new form of semiology that is distinctly different from other sign systems in the way the signified is perceived [Bertin J. (1981)].

- *To perceive a pictograph, a road sign for example, requires a single stage of perception: what does the sign signify? Stop! All the useful information is perceived. The aim of pictography is to define a set or concept.*
- *To perceive a graphic requires two stages of perception: 1st: What are the elements in question? 2nd: What are the relationships among those elements?*

The signifieds in a graphic are resemblance, order and proportion, transcribed by visual variables that have signifying characteristics shown in Table 2-3.

Perceptual Variable	*Quantitative*	*Ordered*	*Differential*	*Visibility*
X,Y spatial dimensions	*quantitative*	*ordered*	*selective*	*constant (associative)*
size	*quantitative*	*ordered*	*selective*	*variable (dissociative)*
value (lightness)		*ordered*	*selective*	*variable (dissociative)*
texture		*ordered to some extent*	*selective*	*constant (associative)*
colour			*selective*	*constant (associative)*
orientation			*selective*	*constant (associative)*
shape				*constant (associative)*

Table 2-3: Bertin's signifying properties of the visual variables

A representation is designed by mapping data relations onto visual relations with similar characteristics. Quantitative visual relations signify quantitative data relations, qualitative visual relations signify qualitative data relations. The point at which a visual variable disappears signifies the zero in a data variable. The effectiveness of Bertin's scheme for visually representing quantitative data was empirically validated by Cleveland [Cleveland W.S. (1985)]. This perceptual scheme is the basis for MacKinlay's rule-based tool which can automatically generate a graph from a characterisation of the data to be represented [MacKinlay J. (1986)].

At the ICAD'92 conference Bly pointed out that progress in auditory display is impeded by the lack of a principled approach similar to that which has been developed for graphic representation [Bly S. (1994)]. She showed that different auditory representations are not equally effective by challenging researchers to design a display to support the classification of a 6 dimensional multivariate data. The results for the 3 displays that were evaluated ranged from chance to significant. The variation in the results led Bly to conclude that a knowledge of the data structure is crucial for the design of an effective auditory display. This observation is echoed by Scaletti's suggestions of appropriate synthesis techniques based on data characteristics as shown in Table 2-4. These suggestions capture experience and expert knowledge in the design of displays to represent these types of data, and are helpful for other designers faced with similar data.

Data Characteristics	*Suggested synthesis techniques*
Oscillating between states	*timbral interpolation or morphing*
Axes and grids	*resonators, fixed tones*
Comparison	*sums, products, differences, correlation*
Textures and tendencies	*granular synthesis, FM, waveshaping, sonic histogram*
Periodicity detection	*data as samples, autocorrelation*
Virtual objects in VR space	*physical models, sampled sounds*
Data with an attitude	*instrumental sounds, musical scales, sampled sounds*

Table 2-4: Scaletti's suggested synthesis techniques

Hayward similarly characterises data in the seismic domain, as shown in Table 2-5 and Table 2-6. As well as characterising the data and the interpretation task, Hayward also provides a rationale for applying the auralisation technique to seismic data.

Sample rates	*250 to 4000 Hz*
Bandwidth	*usually less than 3 octaves* *10-100 Hz common* *20-500 Hz high resolution* *20-2000 Hz special studies*
Dataset size	*12,000 to several billion samples* *0.2 to 5 seconds/record* *usually < 4000 samples/record* *multiple shots per reflection point*
Interpretation	*Every peak is significant and waveform details (width or phase) are often also interpreted*
Dynamic range	*8 - 24 bits commonly used*

Table 2-5: Characteristics of exploration seismic signals

Sample rates	*1 to 120 Hz* *multiple sample rate streams are common*
Bandwidth	*1 hour period to 40 Hz (17 octaves)*
Dataset size	*Continuous recordings 3-100 Megabytes/day* *Triggered recordings < 1 Megabytes/day*
Interpretation	*Precise timing and identification of phase (P,S etc.) is critical. Much interpretation is based on the envelope and spectra of the event*
Dynamic range	*24 bit recording is common* *Most signals easily represented by 16 bits*

Table 2-6: Characteristics of planetary seismic signals

Hayward says that the direct mapping (auralisation) of seismic data works well because the data is physically constrained by the elastic wave equation which is common to both acoustic and seismic vibration. The resemblance between the acoustic signifier and seismic signified is a basis for everyday interpretations - for example a series of echoes followed by an explosion is recognised as physically ridiculous. Loudness is related to energy, pitch implies harmonic structure, and intervals between echoes infer the size of an echo chamber. He comments that auralisation of arbitrary time-series data, such as stock prices, is often unsuccessful because the data relations are physically unconstrained and generate acoustics that are not matched to the perceptual capabilities and expectations of the listener. This is a similar argument to that given by Gaver for the intuitiveness of auditory icons.

The systematic acoustic relation between the signifier and the signified is clearly a useful way to design sounds, but it doesn't help us to signify symbolic relations in abstract data, like stock prices. The effectiveness of a simple pitch variation over time was evaluated as a means to represent abstract data relations by Flowers, Buhman and Turnage. They found that normal users could quickly and easily understand the basic properties of simple functions, distribution properties of data samples, and patterns of covariation between two variables, and that the auditory display was as effective as a visual graph [Flowers J.H. Buhman D.C. and Turnage K.D. (1996)]. This is an important starting point for a method of signifying properties of auditory elements modelled on Bertin's signifying properties of visual marks. Kendall has suggested such a method when he observed that auditory representations of categorical data should sound categorical, that continuous data should sound continuous, and that uniform steps along the continuum should sound uniform [Kendall G.S. (1991)]. This method would focus on a faithful mapping between auditory relations and data relations. A step in this direction has been taken by Madhyastha and Reed, who proposed a matching of the importance of a data variable to the perceptual weighting of an acoustic variable heard in the display. This method was designed to handle the problem that extreme values of some acoustic parameters can cause others to lose their effect - for example it can be difficult to hear the pitch of a short sound, or the timbre of a high pitched sound [Madhyastha T.M and Reed D.A. (1992)]. The perceptual ranking is informal, i.e. "pitch and rhythm are the most distinguishing characteristics of a melody, and thus, can be considered more significant than, say, volume", but it is an example of an approach to auditory design based on perceptual signification of abstract relations.

The matching of perceptual properties of sounds with the properties of the information to be conveyed may improve the directness and correctness of a display. However the display is a nuisance if the information is not useful. The viewer can simply not look at a useless visual display, but auditory displays are not so easy to ignore and prone to become annoying when the information in them is not useful. The auditory display may be perceived as an annoying noise if the information in it is not useful.

2.6 Task-oriented approach

Semiotic methods have been criticised because they are static in time, and do not include motivation as a factor in either the construction or reading of signs. Wittgenstein observed that meaning of a sign is not the object it signifies, but the way it is used [Wittgenstein L. (1953)]. Signs can be used in many ways to generate families of meanings. Signs can be useful without having a material signified at all - for example the greeting “hello”.

Usefulness is one of the criteria often used in HCI to evaluate an interface design. Casner says that different presentations of the same information best support different tasks, and that the usefulness of any information presentation is a function of the task it is being used to support. He built an automatic tool to deduce a graph for a specific task and data [Casner S. (1992)]. Kramer suggests that two broad types of tasks are important in auditory display [Kramer G. (ed) (1994b)]

- *Analysis - tasks where the user cannot anticipate what will be heard and is listening for “pop-out” effects, patterns, similarities and anomalies which indicate structural features and interesting relationships in the data.*
- *Monitoring - a “listening search” for familiar patterns in a limited and unambiguous set of sounds.*

The need to design sounds that provide information relevant to a task was recognised by Frysinger in his proposal of a taxonomy of tasks and data types as a foundation for choosing auditory representations [Frysinger S.P. (1990)]. He suggests that the effectiveness of different auditory display techniques could be evaluated against standard tasks and sets of data. However the definition of a corpus of tasks and parameterised data sets is not trivial. Frysinger points out that the identification of task types is difficult because often the analyst is not able to describe what they are doing very precisely, and a task may consist of a combination or compound of simpler tasks. Sometimes a task is considered to be a small closed action, like pushing a button, other times it can be something bigger, like filling in a form, or something more complex like analysing trends in a data set. Storytelling has been suggested as a way to describe a task by Erickson, who uses a collection of stories to come to grips with the user requirements of an interface [Erickson T. (1996)]. The users are asked to tell stories about their activities which are not expected to be formal or complete or even particularly accurate. A collection of stories contain information about what the users like, and dislike, what works well, what users who are expert in the activity think will work well, and concrete explanations of real problems. Stories are informal, unconstrained and easy to remember. The stories of different users may overlap and provide snapshots of an activity from different perspectives and in different situations. The collection is a rich source of contextual, experiential and concrete knowledge about the problem that can provide a basis for more formal analyses.

A semi-formal method for graphical user interface (GUI) design was extended by Brewster to design useful sounds for the human-computer interface. The method is called the Event, Status, Mode method (ESM). An event marks something that happens at a discrete point in time, due to some action by the user (e.g. a mouse click) or the system (mail arriving). Events depend on context, for example clicking the mouse in one window may select an icon, whilst in another it may position a cursor. Events can be hidden because the system does not display them, or because the user does not perceive them. Status information is similar to an event, except that it is persistent, for example the location of a mouse cursor. Status information can become hidden even if it is visually present, due to visual fixation on other parts of the screen, and because it is static and may fade from attention. A mode is a system context that alters the interpretation placed on events. In one mode typed characters may appear on the screen, whilst in another they may be interpreted as commands. Mode errors occur when the status information is hidden. Information about the events, status and modes in an interaction is characterised by 4 dimensions of feedback.

Information	*Description*	*Sound*
Action dependent/ independent	*Does the feedback depend on a user or system action? Events are action dependent, status and modes are action independent.*	*A keypress activated beep is an action dependent sound.* *A constant tone indicating mode is action independent sound.*
Transient/ sustained	*Is the feedback sustained throughout a particular mode? Events are transient, status is sustained, and modes may be either.*	*A short beep to indicate an error is a transient sound* *Sustained sounds can be habituated, and will be perceived only when it changes in some way, or by conscious attention.*
Demanding/ avoidable	*Can the user avoid perceiving the feedback? Events and modes should be demanding.*	*Sound is attention grabbing so is good for demanding feedback, whereas graphic displays are often missed. Care must be taken in designing avoidable sounds so that they are not demanding by mistake.*
Static/ dynamic	*Does the feedback change whilst it is presented or is it constant? Events are static, status can be static or dynamic.*	*A constant tone is static, music is dynamic*

Table 2-7: Brewster's ESM method

The method begins with the identification of all the modes that are present in an interaction. The events and status information required in that mode are listed and compared with the event and status information that is available in the interface. Discrepancies identify the missing or hidden information that may be causing errors and reducing task performance. The characteristics of the required feedback are then represented by sounds with appropriate characteristics. Brewster demonstrates the method by analysing 7 interface widgets, including a caps-lock key, dialogue boxes, buttons, menus, scrollbars and windows. In every case the source of errors was identified as the avoidability of status information about modes. From this analysis we can surmise that the most useful way that sounds can provide information in an interface is by alerting the user to transient events that may be missed due to visual attention being elsewhere at the time. Brewster tested the method by designing some earcons for the scrollbars, buttons and windows, and evaluating user performance with and without the sounds switched on. He found that the earcons decreased the amount of time to carry out a task, and the amount of time taken to recover from errors. The users preferred the sound enhanced interfaces over visual-only interfaces, and Brewster comments that this was because the sounds provided information that was needed, they were not gimmicks.

Alty lists reasons why multimedia interfaces may be advantageous for control applications [Alty J. (1995)]

- *Telepresence - Multimedia options allow us to regain the natural link between the operator and observables. Avoids events by proxy. Preserves implicit cues.*
- *Measurable Media Differences - Match the medium carrying capabilities with the knowledge output requirements. Differences must be discernable.*
- *Goal - Main factor in choice of a medium is how the presented information will be used. Need a taxonomy for characterising knowledgeNeed a characterisation of Media and their knowledge carrying capabilities. Need to match the two.*
- *Complexity - The medium matters in more complex situations.*
- *Redundancy - For humans more is better. Humans prefer redundancy of information. Multimedia interfaces exploit the whole mind. Useful when information quality deteriorates.*
- *Operator Choice - Operators have hidden goals and implicit knowledge. Often they do not know what information they use. Give them an element of choice. Feedback gives useful pointers to media usage.*
- *Intrusion - Some media such as sound are intrusive. Such media are often useful as context switchers. Intrusive media can obliterate lower level goals, so the system should remember "interrupted" goals and remind operators when the intrusion is over. Media switches can also aid problem solving.*
- *Metaphor - Presenting information in a different medium can be illuminating i.e. music is a visual medium (height and distance metaphors for notes and time).*
- *Synchronisation - Media when used together and synchronised are very powerful. Minimising overload on short term memory.*

Although these seem like very good reasons to use multimedia, Alty comments that in fact there is little empirical evidence to support the view that multimedia interfaces are any better. He tested the validity of these principles in an experiment which compared the performance of subjects using different display media. The task was to control Crossman's Waterbath which involves balancing in-flow, out-flow and heater temperature to prevent error conditions. Information about flowrates, temperature, and water level were shown by combinations of graphics, text, speech, and sound. The results show that differences in the display media do influence task performance, though it is hard to know why or how. Alty observes that the different media have different syntactic, pragmatic and semantic properties, and some are better suited to represent different types of knowledge. He concludes that is it difficult to produce good interfaces by chance or ad-hoc techniques, and that a formal method of multimedia design is necessary.

The approach he proposes has the premise that the goals of the user should determine what information is required and how it should be rendered, and is summed up by a series of questions

- *what is the goal?*
- *what task is needed to achieve it?*
- *what knowledge is required?*
- *how is the knowledge characterised?*

The goal of the user is described in a process flow diagram that allows flexible connections and iterations between tasks. There are 4 types of tasks, shown in Table 2-8.

Task type	*Description*
monitoring	*checking on critical variables to spot deviations as soon as possible, in order to maintain optimal running conditions and plant safety. Activities include Identify, Search, Browse, Instantiate, Scan, Check.*
diagnosing	*identifying causes of deviations so that the plant conditions can be stabilised. Identify, Compare, Derive, Guess, Reason.*
predicting	*identifying potential consequences of plant deviations in order to prevent them. Model, Simulate, Run.*
controlling	*direct impact on the operations of the processing system. Record, Create, Delete, Edit, Alter, Enter, Move, Load, Save.*

Table 2-8: Task types

The knowledge required by each task is characterised in terms of 3 kinds of variables - primitive, derived, or complex. A derived variable is a new variable calculated from a primitive variable(s). Complex variables relate a variable with an organisation in space or time. For example a set of Temperature Variables in Time (i.e. Temperature history) is described as {(Temperature, Time)}. Most variables are complex. The description of a variable is shown in Table 2-9.

Variable	*Description*	*Example*
Name	*Conventional name for the variable*	*Temperature*
Type	*Nominal, Ordinal or Quantitative*	*Quantitative*
Cardinality	*Single-Values, Fixed-Multiple Valued or Variable-Multiple-Valued*	*Single-Valued*
Accuracy	*The accuracy of representation*	*0.01*
Range	*The range of possible values*	*-10 to 300*
Ordering	*Does the variable have an ordering*	*Ascending*
Units	*The units of measurement*	*Celsius*

Table 2-9: Knowledge characterisation

Variable	*Description*	*Example*
Stability	*Static or Dynamic*	*Dynamic/every 5 secs*
Continuity	*Continuous or Discrete*	*Continuous*
Directionality	*Scalar or Vector*	*Scalar*
Derived from	*Variables list*	*[]*
Derivation	*How derived*	*Primitive*

Table 2-9: Knowledge characterisation

The Knowledge characterisation is matched against a characterisation of a display medium, shown in Table 2-10. This characterisation is similar to Bertin's visual variables, but the concrete characterisation for auditory media is not described.

Medium	*Description*	*Values*
Name	*Conventional name for the variable*	
Number	*the variables involved*	
For Each Variable		
	Type	*Nominal, Ordinal, Quantitative*
	Ordering	*Does the variable have an ordering*
	Continuity	*Continuous or Discrete*
Processing	*Possible processing options*	
Carrier Details		
Resource Needs	*Audio or Visual +formula for Resource requirement*	

Table 2-10: Media characterisation

The method was applied to design a multimedia control system for a nuclear power plant. The operators were asked to try out the new multimedia display for a trial period, and were asked to evaluate the display. They generally agreed that the sounds made handling alarms quicker and easier, and helped avoid mistakes related to the analysis of alarms. However the sound was disliked as there were too many sounds in the plant. Alty comments that the impact of the multimedia interface was not as great as it could have been. All operators found the system complex and difficult to use, and needed much more training to make full use of the system. One found it fun to experiment with the new opportunities. Another found the new system too tedious to work with. A number of operators drew attention to the poor quality of the voice output.

Task-oriented design methods help to focus the design on a display that is useful. This is particularly important in auditory display because the sounds cannot be easily ignored. However the usefulness of the sound only becomes apparent with use. The first impression that a listener has of an auditory display is aesthetic, and the usefulness of the display can only become evident if it is used.

2.7 Connotation approach

Most people are used to hearing high quality sounds in music CD's, movie soundtracks and computer games. Composers and sound designers are concerned with aesthetic, stylistic and affective connotations of the sounds. These connotations reflect the value that society places on both the signifier and the signified in a sign. The connotations of the sounds in an auditory display are likely to influence how well it is received by users, and may be especially important in commercial applications. Positive connotations may encourage users to experiment and learn the display. An example of how sounds can influence credibility is provided by Tkaczevski's sound design to imply an air of fraud and humour in a computer game [Tkaczevski A. (1996)]. The humorous connotations that arise from the juxtaposition of high tech graphic objects with low tech sounds is surely a lesson to auditory display designers about the importance of auditory connotations! The need to satisfy expectations of audio quality is described by Dougherty in his account of the sound design for the Taligent Operating System [Dougherty T. (1996)]. Engineers on the project had experienced the 8 bit sound quality of the SonicFinder demonstration (which was constrained by the 8 bit sound card that was available at the time), and were concerned that the sounds should be of compact disc quality, and should be very low volume. Dougherty shaped high quality sampled sounds so they could be discriminated at low volumes. He found that people reacted more favourably to the interface when the sounds were switched on, and that they were often influenced by the sounds even when they said they hadn't noticed them.

2.8 Device approach

A major problem in audio applications is the variation in the characteristics of audio devices. These characteristics include the parameters for adjusting overall volume, equalisation, stereo position, reverberation etc., and the ranges of these parameters. Frequency response, speaker locations, ambient noise and many other factors can influence the display design. Tkaczevski describes an approach for balancing the sounds in computer games across various FM and Wave Table synthesisers on different devices and operating systems. He had to manually edit and evaluate the global controls on each target system, tweaking them to reach a compromise that worked across the board. Although a compromise is a practical solution, it can only work when the designer knows what sounds will be heard and when. It also compromises the capabilities of the better quality devices. These problems have been addressed by the Porsonify toolkit developed by Madhyastha and Reed to support auditory displays [Madhyastha T.M and Reed D.A. (1992)]. Porsonify provides a uniform network interface to sound devices through table driven servers that encapsulate device specific parameters without hiding unique hardware functions. A sonification client can query a device daemon to obtain a description of available parameters and their ranges. This description can be used to automatically constrain the mapping to sounds that can be realised on that device. As part of this work, Madhyastha and Reed raise the need to introduce a data characterisation, and rules for mapping data into sounds in a device-sensitive manner. However this approach does not address the problem that different devices with the same interfaces produce perceptually different sounds, or Smith's comments that the need for a device-independent method for specifying timbre is a major impediment to progress in auditory display [Smith S. (1990)].

Transportability has also been an issue in satellite imagery and visualisation of data. Different devices produce colours in different ways, and an image that looks good on one device can look terrible on another, even similar, device. The need for a transportable way to specify colours for these applications was addressed by Robertson using the device-independent, perceptually uniform CIE colour system [Robertson P.K. and O'Callaghan J.F. (1986)]. The behaviour of a display device was modelled from perceptual measurements at regular points in the device parameter space and could be used to optimise colour sequences for the dynamic range of a device. This approach was developed in an interactive tool for designing colour images to represent data-sets. The designer is able to tailor the colour selection by manipulating graphic widgets overlaid on slices through the colour space that show the limits and behaviour of the display device [Robertson P.K. Hutchins M. Stevenson D. Barrass S. Gunn C. and Smith D. (1994)].

The approach taken in the design of device independent colour displays may also help in the design of transportable auditory displays. Plainly the characteristics of the output device are a major factor to consider in display design.

2.9 Summary

This chapter described previous approaches to the design of sounds to support information processing activities. The approaches were divided into types - syntactic, semantic, pragmatic, perceptual, task-oriented, connotative, and device-oriented. The syntactic approach focuses on the organisation of auditory elements into more complex messages. The semantic method focuses on the metaphorical meaning of the sound. The pragmatic method focuses on the psychoacoustic discrimination of the sounds. The perceptual method focuses on the significance of the relations between the sounds. The task-oriented method designs the sounds for a particular purpose. The connotative method is concerned with the cultural and aesthetic implications of the sound. The device-oriented method focuses on the transportability of the sounds between different devices, and the optimisation of the sounds for a specific device. Clearly there are many ways to go about designing sounds, and many factors to consider. A comprehensive method will need to address issues that include the psychoacoustic properties of the signifier, the perceptual relations between signifiers, the organisation of signifiers into structures, the learnability of the signified, signification through use, connotations and social values, the need for transportability and reproduction.

3 • Designing useful sounds

Ask not the meaning, but the use. [Wittgenstein L. (1953)]

Shirley Robertson: In terms of utilising hearing in the world, when somebody's going through the [blindness] training program, what we really try and look at is focusing on sounds that are useful and filtering out the sounds that are not useful. So sounds that are useful are sounds that can help you maintain a straight line, and help you to orientate so that you know where you are. Sounds that are not useful in that situation, are sounds that are moveable for example. So sounds of people talking, they're useful in terms of knowing that there's somebody in your way, but they're not useful in terms of an orientation clue. [Swan N. (1996)]

This chapter proposes an approach for designing useful sounds. The approach builds on Scaletti's working definition of sonification, which is analysed to have two parts - one part has to do with information requirements and the other with information representations. The requirements part addresses issues of usefulness in a task and the selection of useful data relations to display. The representation part addresses the need to ensure that people can hear the required information in the display. These parts are shaped into a framework that focuses on the design of an information representation to meet the information requirements of a task. The phrase "auditory information design" indicates the focus on useful information which is at the core of this approach.

3.1 Scaletti's definition of sonification

Scaletti proposed a working definition of sonification as ...
a mapping of numerically represented relations in some domain under study to relations in an acoustic domain for the purpose of interpreting, understanding, or communicating relations in the domain under study [Scaletti C. (1994)].

There are three different relations in this definition - numerical relations, acoustic relations, and domain relations. The observation that information is contained in relations between elements, rather than in the elements of themselves, has been made many times. Gibson says that people and animals obtain information about the environment from the perceptions of energy proportions and ratios [Gibson J.J. (1966)]. Bertin says that graphic information is contained in the relations between visual marks that signify resemblance, order and proportions [Bertin J. (1981)]. Sebba found that consistent pairings of paintings and music were influenced by perceptions of contrast, order and ratios [Sebba R. (1991)].

These observations support a substitution of the word "information" for the phrase *relations in the domain under study* in Scaletti's definition. This substitution allows us to separate the definition into 2 parts that are connected by **information** - the requirements part, and the representation part. The requirements part is about specifying information that is useful in some purpose. The representation part is about displaying information as sounds.

3.1.1 The requirements part of the definition

The requirements part of Scaletti's definition is the phrase *for the purpose of interpreting,* *understanding, or communicating relations in the domain under study*. The purpose has been classified into 3 types of information processing activities - interpreting, understanding and communicating. This focus on particular activities can greatly simplify the design problem, particularly when the domain is large or complex. Rather than designing a generic display to represent everything about a domain, the display can be focused on a subset of information that is useful and relevant to the task at hand. From this perspective the requirements part of the definition could be rephrased *to meet the information requirements of an information processing activity*.

3.1.2 The representation part of the definition

The representation part of Scaletti's definition is the phrase *a mapping of numerically represented relations in some domain under study to relations in an acoustic domain*. This is a mapping from one representation to another. Numerically represented relations are contained in measurements of some phenomenon, be it a physical phenomenon such as sea surface temperatures, or an abstract phenomenon like share trading prices. However it may be difficult for a person to understand much about the phenomenon by reading a list of numbers. A perceptual representation, such as a graph or a sonification, is another way to represent information that can allow a person to more quickly obtain a better understanding of what is being represented.

Although the definition outlines the basic process of representation there are issues of faithful perception and faithful reproduction that are not addressed. Numeric relations in data sets are usually classified into 4 types - nominal, ordinal, interval and ratio. A faithful perception of this information requires that the listener be able to hear these types of relations in the auditory display. If the information in the display is to be understood correctly then the mapping is first a mapping to auditory perception, followed by a mapping to acoustic variations produced by a display device. In this view sounds exist only in the mind of the listener, whilst acoustic vibrations exist even when a listener is not present. From these observations we may rephrase the representation part of Scaletti's definition to *a mapping of* ***information*** *to perceptual relations in the acoustic domain*.

3.2 A definition of auditory information design

If we put the modified parts of Scaletti's working definition of sonification back together we obtain ... *a mapping of information to perceptual relations in the acoustic domain to meet the information requirements of an information processing activity*.

This phrase can be made more succinct by substituting "sounds" for the phrase *perceptual relations in the acoustic domain*. The definition for auditory information design becomes

the design of **sounds** *to support an* **information** *processing activity*.

3.3 An approach to auditory information design

The approach to auditory information design proposed here builds on Scaletti's definition of sonification. The approach focuses on the design of sounds to support an information processing activity. The approach has two parts that hinge on **information**

1. Requirements: analysis of the information requirements of an activity
2. Representation: design of an auditory representation of the information requirements

3.3.1 Requirements

There are a variety of methods for analysing information. Task analysis is a method developed in Human-Computer Interaction (HCI) design to analyse information required to manipulate events, modes, objects and other aspects of user interfaces [Kaplan B. and Goodsen J. (1995)]. This form of analysis is particularly concerned with actions that occur in sequence and parallel, and the feedback of the current state of the interface. Data characterisation is a method developed in scientific visualisation to describe the relations contained in data sets [Robertson P.K. (1991)]. This analysis addresses concerns about the validity and faithfulness of a representation that is to be re-represented in some other form. A combined task analysis and data characterisation can define a mapping from data relations to information that is useful in a task. This mapping may involve transforming or selecting parts of the data set, for example highlighting a region. The combination of task analysis and data characterisation has been demonstrated in a system for designing colour representations called PRAVDA [Bergman L.D. Rogowitz B.E. and Treinish L.A. (1995)], and in an automatic graphing tool called BOZ [Casner S. (1992)]. These tools operate on the task and data descriptors with a rule base that selects a representation scheme. However there is a problem that the addition of a new task to the system requires new rules to be formulated for every type of data. As the number of tasks and data types increases there will be a combinatorial explosion of rules to cope with each special case. This problem can be addressed by introducing an explicit description of information requirements that separates the task and data, but is a function of these influences. When a new task type is added to the system it is only necessary to add a new rule to map that task type to the information requirements. The information requirements are the pivotal point of contact between the analysis and the realisation of the design. I call this the TaDa approach because the design is focused on information that is useful in a task and true to the data.

3.3.2 Representation

Once the information requirements of the activity have been analysed we are in a position to design a representation of that information. This is the stage where the designer maps the required information into sounds in which a person can hear that information. Listening is a complex process which has been described in terms of "innate primitive" and "learnt schema" levels of perception in Bregman's theory of auditory scene analysis. The TaDa approach addresses these two levels with different design methods - a case-based method for schema design, and a rule-based method for primitive design. The case-based method is a source of familiar auditory scenes that have an information structure that can represent the required information. For example the sounds of a kettle coming to the boil may be a familiar schema for an auditory display of temperature levels in a turbine boiler.

The rule-based method aligns auditory structure with information structure in accordance to principles from graphic design and models from psychoacoustics. Examples of these rules are that if the information is ordered then the sounds should be ordered, if the information is categorical then the sounds should be categorical. Equal differences along a continuum should sound equal to the listener, and zero values should sound like they are zero.

The final stage in the design process is to produce specified sounds on an auditory display device. It is called an auditory display device because the sound specifications are perceptual (auditory), rather than device specific (acoustic), so the display can be transported to other devices. The display device may be a compound of hardware, software, synthesis algorithms, samples and audio components such as amplifiers, speakers or headphones. The reproduction of perceptual specifications on a device requires a measurement of the mapping from perceptual coordinates to control parameters, device capabilities and audio ranges. There is no point specifying sounds that cannot be produced by the display so a knowledge of the device characteristics is vital in the design process.

3.3.3 TaDa tiles

The TaDa approach is summarised by an arrangement of six tiles organised into two trapezoidal shapes shown in Figure 3-1. The upper trapezoid is the requirements part of the approach. The lower trapezoid is the representation part of the approach. The component tiles are facets of the design process, and common edges are connections between these facets.

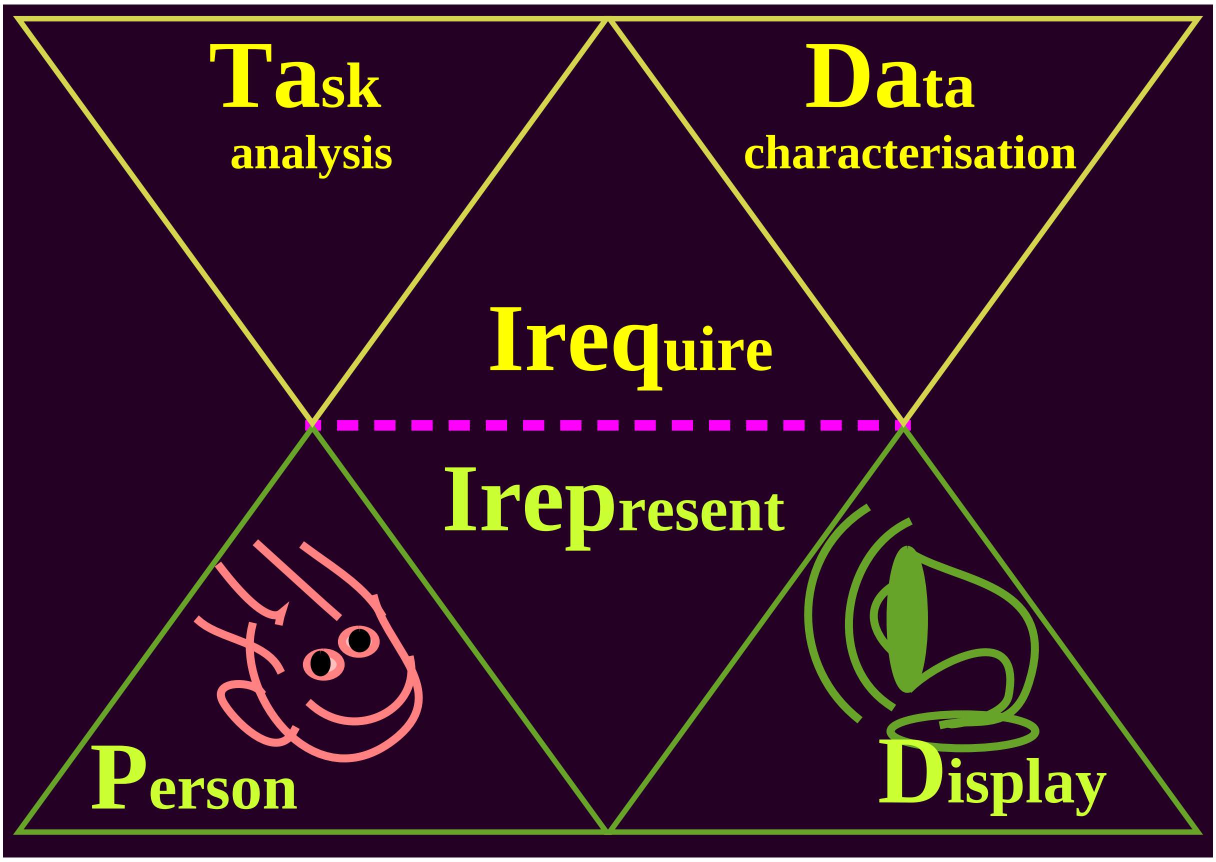

Figure 3-1: Facets of the TaDa approach to auditory information design

The requirements trapezoid consists of the Task analysis (Ta) tile, Information requirements (Ireq) tile, and the Data characterisation (Da) tile. The Ta and Da tile sit upon the Ireq tile to indicate that both of these facets are necessary for the analysis of information requirements. The representation trapezoid consists of the Person (P) tile, Information representation (Irep) tile, and Display device (D) tiles. The Irep tile sits upon the P and D tiles to show that the design depends critically on the types of information relations that can be heard by a human listener, and the types of sounds that can be produced by a particular auditory display device. The core of the TaDa approach is the central diamond where the trapezoids connect. This is where the information requirements (Ireq) are met by the information representation (Irep). The TaDa approach moves through the phases of requirements analysis, design, and representation as shown in Figure 3-2. The phases generally move from left to right, top to bottom through the TaDa tile arrangement. However it is expected that certain facets and connections will be revisited and improved during the process, which is not intended to be strictly linear.

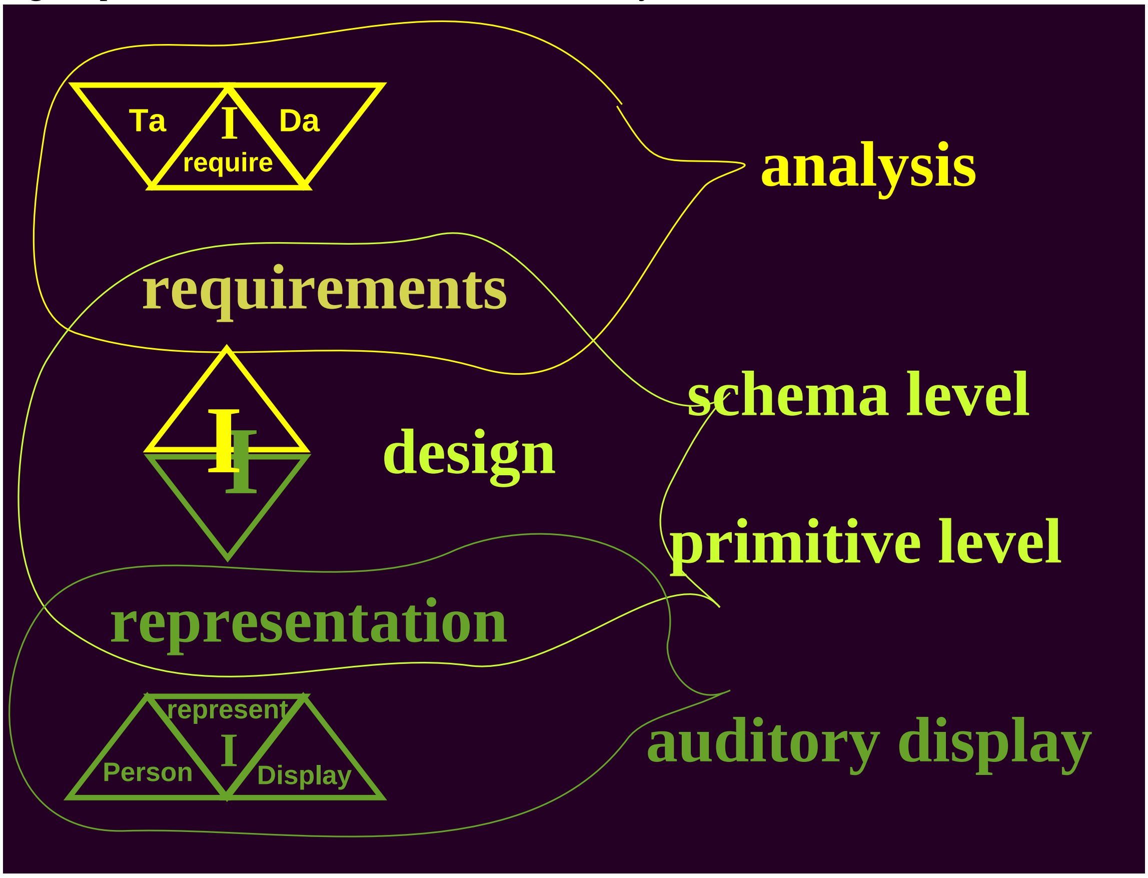

Figure 3-2: TaDa process

3.4 Summary

This chapter proposed the TaDa approach to auditory information design. The approach builds on Scaletti's working definition of sonification, which has two parts - a representation part and a requirements part. Some aspects of the definition were modified based on issues raised in previous approaches to auditory display (see Chapter 2). These issues included the need to analyse useful information using techniques such as task analysis and data characterisation, an explicit differentiation between auditory and acoustic domains, and the need to consider the characteristics of the display device in the design process. A modified version of the definition was proposed as a definition for auditory information design... *the design of* **sounds** *to support* **information** *processing activities.* This definition is the core of the TaDa design approach, which integrates task analysis, data characterisation, perceptual factors, and device characterisation into a multifaceted framework of methods. The framework is described by triangular tiles organised into a requirements trapezoid sitting on a representation trapezoid. The design process is intended to be iterative and the hybrid process of methods is intended to allow the designer scope to be flexible in the process. A multifaceted approach may be particularly beneficial in real world design practice where problems are messy and involve many complex issues.

4 • TaDa: task and data analysis of information requirements

Information is a difference that makes a difference. [Bateson G. (1972)]

Useful information is the answer to a question. [Bertin J. (1981)]

The purpose of the TaDa approach is to design sounds that carry useful information. The core of the approach is a meeting of information requirements with an information representation. This chapter describes the methods used to elicit the information requirements of a design problem. The first section introduces scenario analysis as a technique for capturing key features of the problem. The next sections describe the particular flavour of task analysis and data characterisation used to decompose the problem, and give a detailed account of the parts of analysis. The TaDa analysis is supported by a form interface to an Access™ database.

4.1 Describing the problem with a story

The design of useful information display must involve a description of how the display will be used. This may best be obtained from someone who is involved in that activity. User-centred design has become a prominent topic in Human-Computer Interaction. This method involves the user in the design process, and communication between the designer and the user. Techniques, such as storyboards, scenarios, interviews, and case studies, have developed as a means of information exchange between technical and non-technical groups.

Storyboards are effective for discussing and presenting design situations within collaborative design groups, and between designers and their clients [Lewis C. and Rieman J. (1994)]. Scenario analysis incorporates the knowledge and experience of the user into the design process through the recounting of instances of interaction with the system, which vary from real episodes to more or less constructed stories [Klausen T. and Aboulafia A. (1995)]. Interviews are often used to evaluate an existing system, as in the analysis of a decision support system for the St. John Ambulance service [Wong, Sallis and O'Hare (1995)]. Use-case analysis is based on descriptions of interactions which are written by the designer from the perspective of the system [Kaplan B. and Goodsen J. (1995)]. These techniques force the designer to be specific about the features of the real system and user, rather than addressing abstract issues which may have little real impact on the situation of use [Lewis C. and Rieman J. (1994)].

The method of problem description adopted in this thesis is a text description of an activity written or spoken by the user involved in that activity. The intent is to obtain short story-like descriptions in very general terms at any level of the problem. The word activity is used, rather than task, to indicate a loose association between elements, rather than a structured account. Below, and shown in Figure 4-1, is an example of a problem scenario transcribed from a verbal description and demonstration of the problem by Chris G., a software engineer working on a geophysical visualisation interface at CSIRO.

The GeoViewer is a 3D interactive view of rock strata for mine planning and other geological applications. You can move about anywhere in the space, and see the rock layers. The rock- type of the layers is shown by colour and texture, and a mouse click can popup a text description. You can see more by turning on the transparent view, and speed up the interaction with wireframe views. One problem is that it can be hard to see what type of rock it is when it is transparent, or wireframe. Also the popup text can get in the way of what you are looking at.

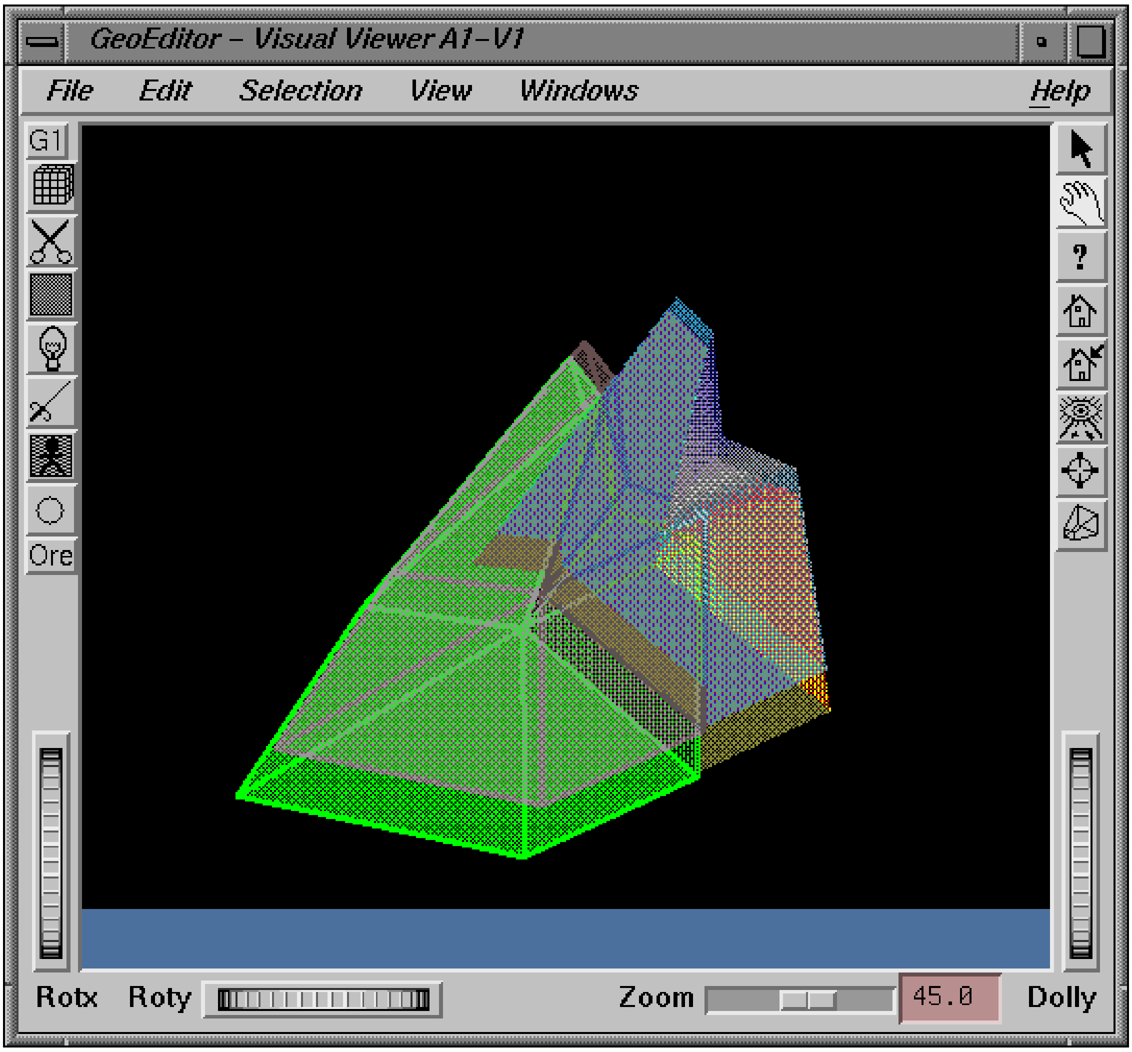

Figure 4-1: The GeoViewer

4.2 A bridge from story to requirements

Once the problem has been described it can be analysed to find key features and requirements of a solution. The analysis identifies key features of the problem description that are relevant in the design domain, and different analyses may be carried out with the same problem description.

The observation that “useful information is the reply to a question” [Bertin J. (1981)] provides a way to identify features that are relevant in the design of an information display. The problem captured by the scenario story is recast as a question, and the range of possible answers is identified. The questions extract the information required to carry out the activity, the relationships between the answers further specify the information type, and the subject of the question is the phenomenon of interest in this activity. The key features (Question, Answers, Subject) are a bridge from the story to the Requirements Analysis. They are an important first step toward the representation of the problem in a standard format that provides a framework for systematic design.

The Scenario Description for the GeoViewer, is shown in Figure 4-2. The story was recast as a Question “what type of rock is this?”. Chris agreed that this question adequately summarised the problem. However he pointed out that there is no definitive answer to this question - there are as many answers as there are types of rocks in geological data sets which can be visualised with this software. The GeoViewer handles this by allowing the user to arbitrarily associate rock types with colours in an editing panel. A similar panel could be made for sound allocations from a palette. Rather than consider all possible answers, a representative subset was chosen to focus the design on a useful implementation. The Answers {coal, sandstone, granite and marble} were selected as typical. The Subject of the Question in this example is the rock-type.

Scenario

	Title	GeoViewer
	Storyteller	C.G.
STORY	The GeoViewer is a 3D interactive view of rock strata for mine planning and other geological applications. The rock- type of the layers is shown by colour and texture, and a mouse click can pop-up a text description. You can see more by turning on the transparent view, and speed up the interaction with wireframe views. A problem is that it can be hard to tell the rock-type when it is transparent, or wireframe. Also the popup text can get in the way	
KEYS	*Question*	what rock-type is it?
	Answers	coal, sandstone, granite, marble
	Subject	rock-type
	Sounds	???

Figure 4-2: Scenario Description for the GeoViewer

4.3 Requirements

The TaDa requirements analysis draws on task analysis and data characterisation methods that have been established in HCI and visualisation. TaDa amalgamates variants of these methods to produce an Information Requirements specification that drives the design of an Information Representation. There are three areas of analysis - Task, Information and Data. Each area analyses a different key feature from the Scenario Description. The Task section analyses the Question key, the Information section analyses the Answers key, and the Data section analyses the Subject key.

4.3.1 Task analysis

Task analysis is an established technique in HCI and visualisation design, and the fields proposed here are an amalgam of components borrowed from several different analyses [Wurman R.S. (1989)], [Norman D.A. (1991)], [Robertson P.K. (1991)], [Kaplan B. and MacCuish J. (1993)], [Rogowitz B.E. and Treinish L.A. (1993a).], [Lewis C. and Rieman J. (1994)], [Alty J. (1995)].

These components have been selected for their relevance to designing information in a temporal medium like sound. The task analysis is rooted in the Question key from the Scenario Description.

Generic question

Although there are an unlimited number of questions which may be asked, Bertin proposed that they may all be classified in terms of three levels of information. The subject of the question can be used to make this classification. A question which requires local information is about a single element. A question which requires intermediate information is about a subset of elements. The global question is about all the elements as a whole. By replacing the subject of a question with a generic tag, such as “it”, or “they”, a range of generic questions is proposed as a classification. New questions can be added to the classification scheme, shown in Table 4-1, as they are encountered.

Local Questions Subject {it}	Intermediate Questions Subject {they,which,what}	Global Questions Subject {everything, anything}
who is it ? what is it ? where is it ? is it ready ? is it time ? is it ok ? how good/bad is it ? how much is it ? what is wrong with it ? is it organised ? what was that ? where did it go ? what does it remind me of ?	where are they ? are they the same ? are they similar ? which is more ? which are the same ? which are similar ? which are different ? what is over there ? where am I ?	is anything here ? what is happening ? is everything ok ? has anything changed ? where am I ?

Table 4-1: Generic questions by information level/subject

Purpose

The purpose, or goal, or aim of a task is identified in most task analyses. A set of Purposes was obtained by interpreting the purposes of the generic questions, as shown in Table 4-2. The 10 Purposes proposed here are an amalgam of the search and compute types that have been defined in graphic display [Casner S. (1992)], and interactive confirmation, navigation, and alert types found in interface design methods [Brewster S.A. (1994)]. Several extra Purposes were added after analysing the uses made of sounds in everyday listening experiences (see Chapter 5). *Relax* is a purpose that captures the way people sometimes use low-level background noise or music to block out unwanted auditory disturbances which interfere with sleep, or mental attention in some activity. *Remember* has been included because sounds can be used to remember things - for example I recently noticed that I had mis-dialled my parents phone number by a change in the familiar tone-dialling tune. Sounds attract attention and are an important part of the engagement of interest in entertainments of all kinds. *Engagement* has been included because it is an important role that sounds have in movies and computer games, and there is potential to use them this way to improve workplace activities too.

Question	Purpose	Description
are they the same ? are they similar ? which are similar ? which is different ? which are different ? what is over there ? what is here ? what is happening ? has anything changed ?	analysis	observe relationships, groupings, trends, outliers, patterns
is it ok ? is it ready ? is it time ? are they the same ? is everything ok ?	confirm	absolute boolean confirmation
who is it ? what is it ? what is wrong ? what state is it ? what is over there ?	identify	absolute identification from a familiar set
how good/bad ? how much is it ? is it organised ? is everything ok ?	judge	absolute classification from a familiar ordered set
which is more ? which are same ? which are similar ?	compare	relative comparison of ordered properties
where are they ? what is here ? where am I ?	navigate	interactive movement through an ordered space
where did it go ? where are they ?	track	track an object through an ordered space
what was that ? has anything changed ?	alert	highlight or draw attention to an element or subset
	relax	mask unwanted noise, de-emphasise highlights
what does it remind me of? has anything changed ?	remember	remember places, times, people, information
what is over there ? has anything changed ?	engage	attract, entertain, maintain interest

Table 4-2: Generic question by purpose

Mode

The distribution of attention between overlapping tasks may help capture the ebb and flow between what is information at one moment and noise the next. An interactive task requires full attention, a monitoring task requires focused attention, and a background task can continue while something else is the focus of attention. The task mode field is directly borrowed from Kaplan and Goodsen, and is shown in Table 4-3.

interactive	manipulation with feedback e.g. tuning a radio
focus	conscious attention to an element e.g. conversation in a noisy room
background	attention focus is elsewhere e.g. watching television while babysitting

Table 4-3: Task attention mode

Type

Sounds can overlap, form patterns and cycles, be ongoing or very short. The analysis of overlapping tasks are a part of the HCI task analysis proposed by Kaplan and Goodsen [Kaplan B. and Goodsen J. (1995)]. Discrete/Procedural tasks are initiated by a single event, linear, short and seldom overlapping. Continuous/Tracking tasks are interactive, overlapping and have undefined closure. Branching/Decision tasks affect the subsequent course of action. The task event type field is borrowed from Kaplan and Goodsen, and is shown in Table 4-4.

discrete/procedural	initiated by a single event, linear, defined closure, seldom overlapping
continuous/tracking	require constant monitoring, constant feedback is used to iteratively make refinements, often overlapping other tasks, fuzzy closure.
branching/decision	affects the subsequent course of action

Table 4-4: Task event type

Style

Researchers in both sonification and visualisation have recognised two quite different styles of information processing tasks. The first is the exploration of data sets for interesting and as yet unknown features, which requires a faithful, veridical or isomorphic representation that preserves structural relationships. The second is the presentation of known features which may involve an intentional transformation of structure to draw attention or highlight or exaggerate. The task style field, shown in Table 4-5, is similar Kramer's exploration and presentation tasks [Kramer G. (ed) (1994b)], Cleveland's analysis and communication tasks [Cleveland W.S. (1985)] and the type I and type II tasks defined by Bergman et al. [Bergman L.D. Rogowitz B.E. and Treinish L.A. (1995)]

exploration	veridical representation that preserves information relations
presentation	intentionally transform the structure to draw attention to particular features, exaggerate details, or segment into regions.

Table 4-5: Task style

4.3.2 Information

The Answers to the Question key contain the information needed to carry out the activity. A characterisation of these answers can specify the information requirements of a display to support that activity.

Level

The level of information describes whether it concerns a single element (local), a group of elements (intermediate) or all of the elements as a whole (global). The level of information in the design scenario is obtained from a classification of the Generic Questions in these terms (see the previous section), for example "what is it?" is a local question.

local	information related to a single element
intermediate	information related to a subset of elements
global	information about all of the elements as a whole

Table 4-6: Information level

Reading

A direct representation can be understood with little training, can be understood almost immediately, and allows judgements which are not readily swayed by the opinions of others [Ware C. (1993)]. Some examples of direct representations are scatterplots, satellite images, and geiger counters. Conventional symbols, on the other hand, depend on learning or a legend to be understood. However they have the advantage that they may carry complex concepts built on layers of reference. Some examples of conventional representations are traffic signs, morse code, and hand gestures.

conventional	learnt, cultural, varies between individuals
direct	little training, immediate, resists bias, cross-cultural

Table 4-7: Information directness

Type

The set of Answers to some questions are qualitative, and others may be quantitative. For example a set of Answers, such as "coal" or "sandstone" that identify materials as categorical types is qualitative. Answers such as "twice as much coal as sandstone", or "no shale at all" provide information that involves some form of measurement. In data analysis the data relations are characterised as 4 main types - nominal, ordinal, interval and ratio. This characterisation may also be used for the information relations contained in the Answers. Some specialisation of the types occurred when this scheme was used to design the demonstration scenarios described in later chapters. The additional types are boolean, ordinal-with-zero, and ordinal-bilateral.

none	no information is involved
boolean	2 different categories (e.g. yes, no)
nominal	difference without order (banana, apple, orange)
ordinal	difference and order (e.g. low, med, high)
ordinal-with-zero	difference, order and a natural zero (e.g. none, some, lots)

Table 4-8: Information type

ordinal-bilateral	difference, order, central zero (e.g. less, same, more)
interval	difference, order and metric (e.g. Celsius temperature scale)
ratio	difference, order, metric, and natural zero (e.g. mm of rainfall)
unknown	information type is unknown

Table 4-8: Information type

Range

The number of Answers can have a significant effect on the design of a display. Large catalogues may need to be organised and represented differently from small sets of 3 or 4. In the case of continuous information the range of variation is recorded in this field. In exploratory tasks or unknown domains where the range of the answers is unknown the Range is tagged as *unknown*.

Organisation

The information in a display is contained in the interplay of relations between the elements. These relations can be organised in a variety of ways. A characterisation of information organisations developed for graphic design has been borrowed directly from Wurman [Wurman R.S. (1989)]. This scheme may be used to characterise the organisation of the Answers, for example {coal, sandstone, shale, limestone} are a *category* organisation

category	Pertains to organisation of goods or types. Category can mean different models, types or questions. This mode lends itself to organising items of similar importance. Category is well reinforced by colour as opposed to numbers which have inherent value.
time	Works best for events that happen over fixed durations. It is an easily understandable framework from which changes can be observed and comparisons made
location	The natural choice for examining and comparing information from diverse sources or locales. If you were examining an industry you may want to know how it is distributed around the world. Location doesn't always have to refer to a geographical site. Doctors use locations in the body as groupings to study medicine.
alphabet	Lends itself to organising extraordinarily large bodies of information, such as words in a dictionary or names in a telephone directory. As most of us know the alphabet this organisation works when another form such as category or location may not.
continuum	Organises items by magnitude from small to large, least expensive to most expensive, order of importance etc. It assigns value of weight. Which department has the highest rate of absenteeism? What is the smallest company engaged in a certain business? Unlike category, magnitude can be illustrated with numbers or units

Table 4-9: Information organisation

4.3.3 Data

The Subjects key of the Scenario Description identifies the phenomenon of interest. The characterisation of this phenomenon can help the designer select a representational mapping that provides useful information about the relevant aspects of the phenomenon.

Type

In data analysis the general data types (nominal, ordinal, interval, ratio) have been developed to represent different phenomena. For example the Subject "type of rock" is a nominal phenomenon, whereas "percentage of a rock type in a rock sample" is a ratio phenomenon which can be measured and has a zero.

none	
nominal	difference without order (banana, apple, orange)
ordinal	difference and order (green, crisp, ripe)
interval	difference, order and metric (temperature)
ratio	difference, order, metric, and natural zero (rainfall)

Table 4-10: Phenomenal type

Range

The number of categories in a nominal or ordinal phenomenon can have a great bearing on the representational mapping of that phenomenon in a display. The display may need to show some or all of those categories, depending on the information required. In a continuous phenomenon the range of variation is recorded in this field as a basis for scaling transformations which may be needed in the mapping to the display representation.

Organisation

Physical phenomena such as temperature are organised in the physical continua of space, time and energy. Abstract phenomena such as stock prices may be organised by alphabet or category.

category	organisation by difference
time	organisation by time
location	organisation by spatial position
mnemonic	organisation by mnemonic, e.g. alphabet
continuum	organisation by continuous order

Table 4-11: Phenomenal organisation

4.4 Requirements of the GeoViewer

This section demonstrates TaDa Information Requirements Analysis by analysing the GeoViewer geological exploration interface.

Task analysis of GeoViewer

The Question Key is the bridge to the Task Analysis. The Question {what type of rock is it?} from the GeoViewer is transformed to the Generic {what is it?} by removing the subject, and referring to Table 4-1. The Purpose {identify} is looked up from Table 4-2 with the Generic question. The Mode is {interactive} because the question is made through a mouse click which is a manual operation. The Type is {discrete} because the question only occurs at discrete times during the activity. The Style is {exploration} because the user does not know the rock structure before using the GeoViewer.

Information analysis

The Answers Key is the bridge to the Information Analysis. The Answers for the GeoViewer are {coal, sandstone, shale, limestone}. The Reading is {direct} because the diversion of visual attention by popup text is a problem in the visual display design. The Type is {nominal} because the Answers are different but have no ordering. The Level is {local} because each Answer is about a particular rock strata. The Organisation is {category} because the relationships between Answers are not ordered in space or time, but are simply different. The Range is {4} which is indicative of the number of rock types that might be expected.

TaDa Analysis

Generic	what is it?	
Purpose	identify	TASK
Mode	interactive	TASK
Type	discrete	TASK
Style	exploration	
Reading	direct	
Type	nominal	INFO
Level	local	INFO
Organisation	category	INFO
Range	4	
Type	nominal	DATA
Range	4	DATA
Organisation	category, space	DATA

Figure 4-3: Example of TaDa requirements analysis

Data analysis

The Subjects Key is the bridge to the Data Analysis. The Subject in the GeoViewer is {type of rock}. The Type is {nominal} because types of rocks have difference but no order. The Range is {4} because there are only 4 types of rocks in the region of interest. The Organisation is {category, space} because the rock structure is both spatial and material.

4.5 Computer aided support

The TaDa Information Requirements Analysis has been implemented as a graphical form-like interface in an Access™ database. This tool makes it easy to store and recall design problems. The options for each field can be selected from a popup menu, which aids the transcription and analysis of a new problem.

4.6 Summary

This chapter presented a scenario-based method for analysing the information requirements of a design problem. The method recasts the problem as a Question and some Answers about the Subject of the Scenario. These Keys are the bridge between the informal story that describes the problem, and the formal analysis. The analysis is an amalgam of visualisation and HCI task analyses and data characterisation methods specifically selected to support the design of auditory information. There are three parts to the analysis - the Task analysis of the Question, the Information analysis of the Answers, and a Data characterisation of the Subject. The rationale for choosing each analysis field was described, and examples were used to illustrate each field. The method was demonstrated in an analysis of the GeoViewer tool for interactive exploration of rock structures. A form-based interface was implemented in an Access™ database to support the method.

5•EarBenders: case-based design from stories about listening

The first obstacle is the prevailing sonification model, which is simply to map data to sound parameters arbitrarily. The resulting sound is typically unpleasant and lacking any natural connection to the data represented (one intuitively feels that medical images, for example, ought to somehow sound different from demographic data or satellite imagery). Models of sonification more sensitive to the kinds of data presented must be developed.[Smith S. (1990)]

Designers often base a new design on a previous version that has proven successful in similar problems. A previous solution can be a way to quickly come to grips with a design problem, and provides a starting point for a top-down process of iteration. This chapter introduces the case-based method of design by example, and describes how it has been adapted for auditory information design. The case-based method relies on a rich source of examples to be effective, but as yet there are not many examples of auditory display design to draw upon. An alternative resource was developed by collecting stories about everyday listening experiences into a database, which I call EarBenders. The information requirements of a design problem can be used to search this case-base for everyday examples which share a similar task, data and information structure with the problem. The ability to do this search required each story to be analysed with the TaDa Information Requirements Analysis developed in the previous chapter. In addition an auditory characterisation was developed to describe the sounds in each story, and provide a footing for auditory design. The sound characterisation also provides an opportunity to extract principles of sound design from regularities between auditory structure and information structure in the example cases. The case-based design of auditory information is demonstrated on a problem in a geological visualisation interface, called the GeoViewer.

There are two appendices for this chapter. Appendix 5-1 lists the EarBenders case-base. Appendix 5-2 lists the software programs that retrieve relevant cases from the database, and synthesise a sound design from them according to the method.

5.1 Advantages of case-based design

The stories in many professional design journals and magazines are case-studies which convey information about a good design in a way that can be easily understood and assimilated. The popularity of these magazines attest to the value of case-studies in design practice. The designer can use good examples as a starting point, or borrow elements from them. In HCI the intelligent borrowing of good ideas from existing interfaces is promoted as an effective design technique which is less risky, quicker and easier to implement than a totally original design, and has the advantage that users will already be familiar with the borrowed elements [Lewis C. and Rieman J. (1994)]. In a similar vein, comparability analysis uses the precedent system, and lessons learnt from it, as the starting point for the design of the replacement system [Carlow International Inc. (1992)]. The synthesis of a design from examples is suited to design domains where there is little theory and few principles that can provide guidance. A critical survey and comparison of designs can identify new principles of design from examples, as demonstrated by Tufte in his collection and critique of graphs [Tufte E.R. (1983)]. Some advantages of the case-based method are listed below.

Top down

Cases are top down - example cases address the key features of the problem, and provide an account of those features. Most real problems are large and complex, and the reuse of a previously successful design can more quickly lead to a better quality solution. The recall of several previous solutions provides an opportunity to integrate features to synthesise a new design.

Open ended

A case-base is an open-ended store of knowledge. New cases can always be added to improve the depth and breadth of examples. Many different solutions to the same problem can be included, providing a store of variation which can illuminate the features of the design space. The database can change over time to reflect a particular perspective or style as cases are added.

A source of principles

A systematic analysis of cases may lead to the discovery of general design principles. These principles capture regularities in the relations between features of the problem space and features of the solution.

Computer tools

Computer tools can assist a case-based method by making it easy to store and retrieve design cases from a database. Case-based tools have been implemented for designing buildings, computer interfaces, fire engines, and CAD drawings [Maher M.L. Balachandran M.B. and Zhang D.M. (1995)]. The method depends on the retrieval of relevant cases. The retrieval of relevant cases depends critically on a representation that captures features of the cases that are important in a solution.

5.2 Everyday stories as examples

The case-based method depends on a rich source of examples. The small number of examples of auditory display limits the potential for case-based design in this field. However the observation that people use sounds in everyday activities opens the door to an alternative source of examples. I investigated the idea that everyday uses of sounds might serve as design examples in auditory display by collecting stories about listening experiences into a database. The collection was started by an email message to all staff at CSIRO Mathematics and Information Sciences that included some examples of everyday uses of sound, and a request for similar stories.

During the next few weeks if you notice yourself using your hearing to help you, say, find a lost pin, or search for hollow spaces in a wall, or notice something wrong with *your car, or tell whether the dog is hungry, anything at all, then please email me with a brief story about the occasion.*

Nearly 200 stories have since been collected. The initial round of stories were received by email, but many more have come from conversations, newspapers, books, television, movies, and ongoing general observation. Each story is written down and stored in a computer database, along with the name of the storyteller and a title. The database is called EarBenders because “to bend your ear” is an Australian colloquialism to do with telling a story or having a conversation. The stories are about sports, cooking, camping, domestic chores, games, car maintenance and many other activities. Some example stories are shown in Figure 5-1

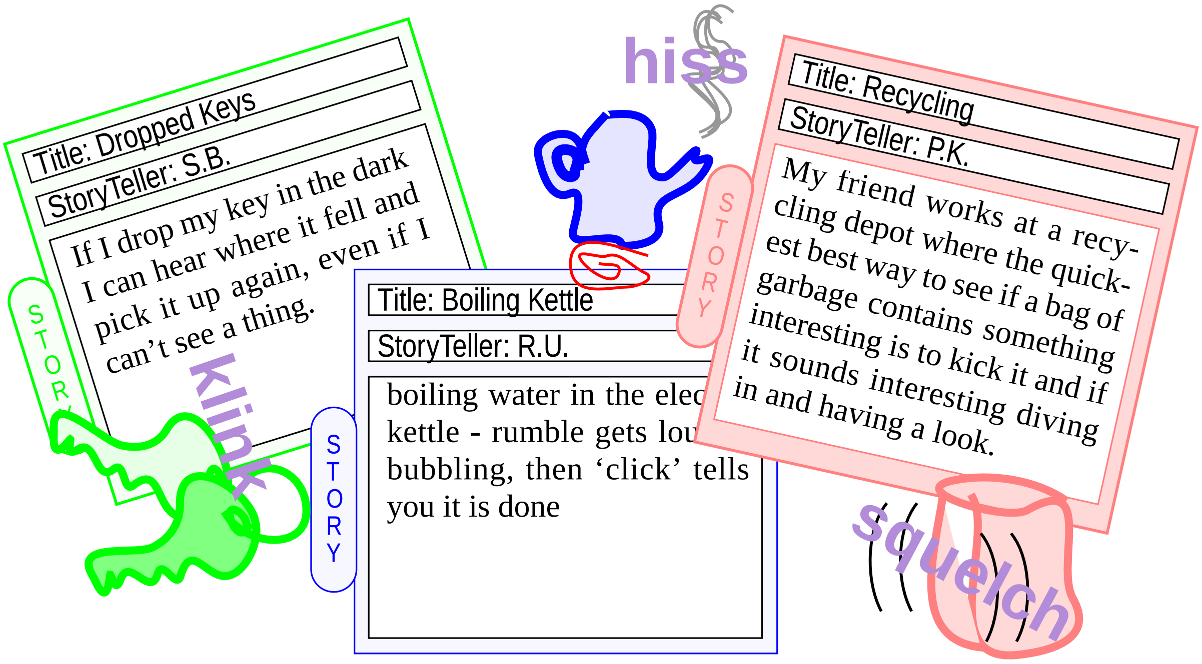

Figure 5-1: EarBenders stories

The stories capture the way sounds are useful, the ways that people hear information in sounds, and the organisation of the sounds. The use of the sounds can be found by recasting each story as a question. In the examples the questions are “where did the key drop?”, “has the kettle boiled?” and “what is in this rubbish bag?”. The answers to these questions are sounds described in the stories. Where the sounds are not explicitly described the objects and events that cause them are. The organisation of the answers helps to identify the syntax of events in the story. If the {rumble, bubbling, click} sequence in the Boiling Kettle story did not occur in the right order then you would know something was wrong. These observations indicate that the stories contain semantic, syntactic and pragmatic elements that are important for designing symbols.

Implementation details

The EarBenders database is currently implemented in Microsoft Access™ 6.0. This tool lets you design a form as the interface to the database records. Keyword searches can be carried out on individual fields, or on all the fields at once. Partial matches can be specified with wildcard characters. An example of the interface for entering an EarBenders story is shown in Figure 5-2.

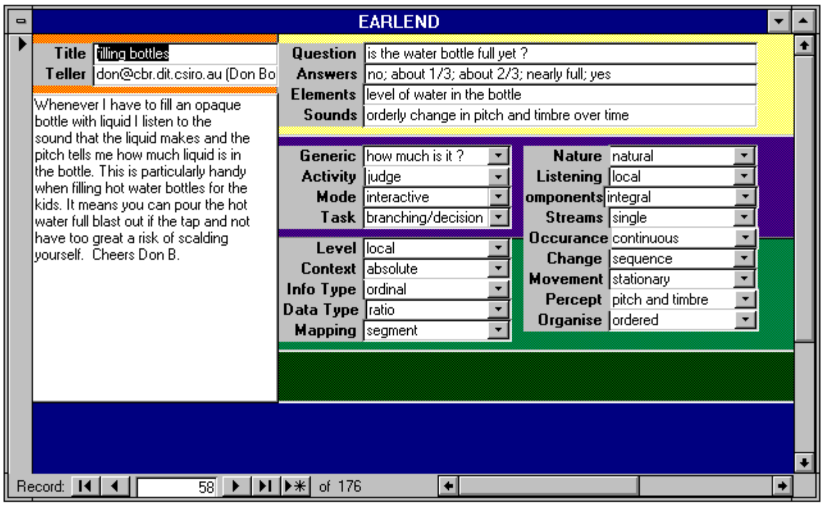

Figure 5-2: EarBenders database form

5.3 Three case-based methods

The EarBenders stories describe situations where sounds have been useful. Collecting and reading the stories was an educational experience which has changed how I think about the uses of sounds, and made me more aware of the many ways that people use them all the time. Entering the stories into a database provides the extra capability to search and retrieve examples that transforms EarBenders from a collection of stories into a resource for case-based design. There is more than one way that EarBenders may assist in auditory display design. Three methods that I have explored are called the metonymic method, the metaphoric method, and the pattern method, each of which is described in the following subsections.

5.3.1 Metonymic method

A metonym connotes the whole by a part of that whole - for example the sound “woof” may be a metonym for a dog. The metonymic method uses sounds that are a part of the design scenario. The sounds in a medical scenario might be heartbeats, breathing, or instrument beeps. Any sounds mentioned in a design scenario may be the beginnings of a metonymic palette. If there aren’t any sounds explicitly mentioned, then objects and events that are described may be a basis for sound design. This method is modelled on Gaver’s auditory icon method where he maps events and actions in the interface to events and actions that cause sounds in the everyday world [Gaver W.W. (1994)].

Keyword searches of the EarBenders database can assist a metonymic design. Keywords from the scenario domain can retrieve other stories about that domain which may provide a broader design perspective. For example a search for {medicine} may retrieve other medical stories that describe uses of medical sounds. If sounds are explicitly described in a scenario then you may search for other stories which have the same sounds in them. For example you might retrieve all stories with the word “crunchy” in them. Keyword searches for objects and events in a scenario may retrieve stories that also include those objects and events. For example, the scenario might involve opening a file, so you might search for “open” or “file” to find stories that have sounds that maybe related to these keywords. Two stories retrieved by the keyword {open} are shown in Figure 5-3. Door-muffle describes how you can hear the sounds from a room become louder when someone opens the door. Window-open describes the whistling that happens when you are driving along with a car window partly open.

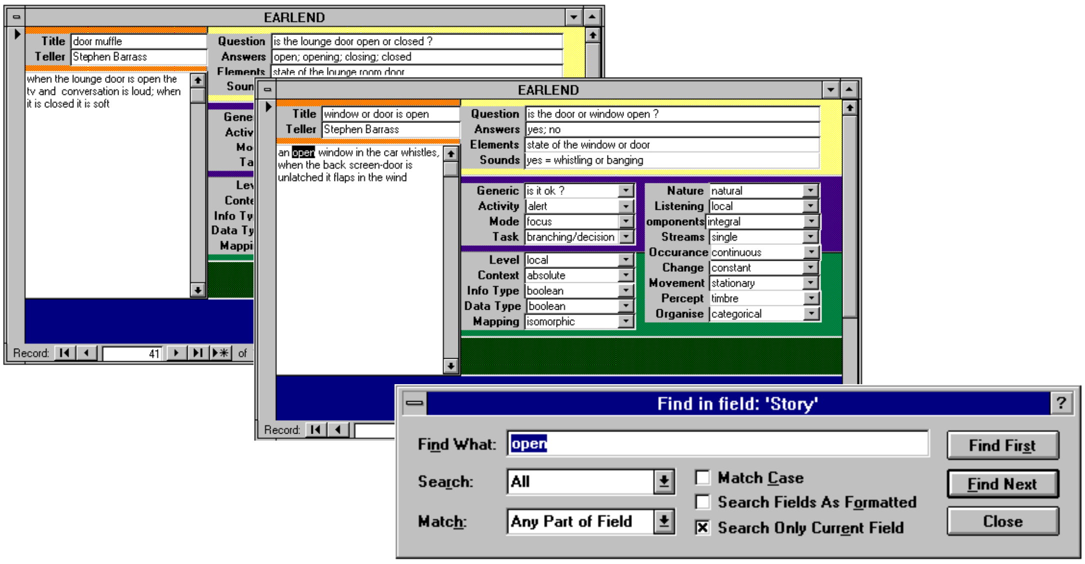

Figure 5-3: Retrieving stories by keyword search

5.3.2 Metaphoric method

A metaphor expresses the unfamiliar in terms of the familiar, for example a tree may be a metaphor for a filing system. A metaphoric design can help when sounds are not a natural part of the design scenario, which is often the case in computer-based applications. Stock prices and internet traffic don't normally make sounds, so are likely candidates for a metaphorical design. But what metaphor to choose? The metaphor needs to be easy to understand and relate to the phenomenon it represents. The information in the metaphor must be information about the phenomenon, or else it is useless. A relevant metaphor can be found by searching EarBenders for a story that can support the information requirements of the design scenario. These requirements are described by a TaDa analysis. The facility to search on TaDa fields was added to EarBenders by appending a TaDa analysis to each story. You can retrieve stories by entering the TaDa requirements of a design scenario into a query form. The cases that are retrieved are a source of useful metaphors. After some trials I found that it was quite rare to find cases that match the query in every way. The solution was to change the retrieval process from an exact match to a search by similarity. The measure of similarity is the number of fields that match. The query returns a ranked list where the top-most cases are most similar in TaDa structure to the design scenario.

TaDa analysis for an EarBenders story

The TaDa analysis of the EarBenders stories is the key to the metaphoric method that allows the retrieval of relevant metaphors. The process of analysis is briefly outlined in this subsection, and is summarised in Figure 5-4. The story we will analyse is the Recycling story introduced earlier. The Question is {what is inside this bag?} The Answers to the Question are {household garbage, bottles, crockery, nappies}. Probably there are many more answers than were given, but a comprehensive list may not be necessary to analyse the information relationships. The Subject of the Question is {inside this bag}. The Sounds that give each Answer are {plasticy, clink, clatter, squelch}. The identification of the Scenario Keys leads to the next stage of TaDa Analysis. First the Question {what is inside this bag?} is analysed in the Task section. The Generic question {what is it?} is obtained by removing the subject of the question, and referring to Table 4-1. The Purpose of the Generic question is {identify} which is found in Table 4-2. The Mode is {interactive} because the question is asked by kicking a bag. The Type is {discrete} because the question has an immediate response. The Style is {exploration} because the answer to the question is unknown. The Answers Key {household garbage, bottles, crockery, nappies} is analysed in the Information section. The Reading is {direct} because the identification can be made without reference or comparison. A new employee will probably need to look inside the bag but quickly learns to make a direct auditory identification to avoid an unpleasant experience. The Type is {nominal} because the Answers are different but have no order. The Range of Answers is {4} in this example. The Level is {local} because the Answer only pertains to the bag that is being kicked. The Organisation is {time}, because the answers are separated from each other in time. The Subjects Key {inside this bag} is analysed in the Data section of the TaDa analysis. The Type is {nominal} because the phenomena are mixtures of garbage objects that have characteristic differences in content but no natural physical order. The Range is {unlimited} to indicate that each garbage bag has a unique content. The Organisation is {location} because the garbage bags are separated by their locations.

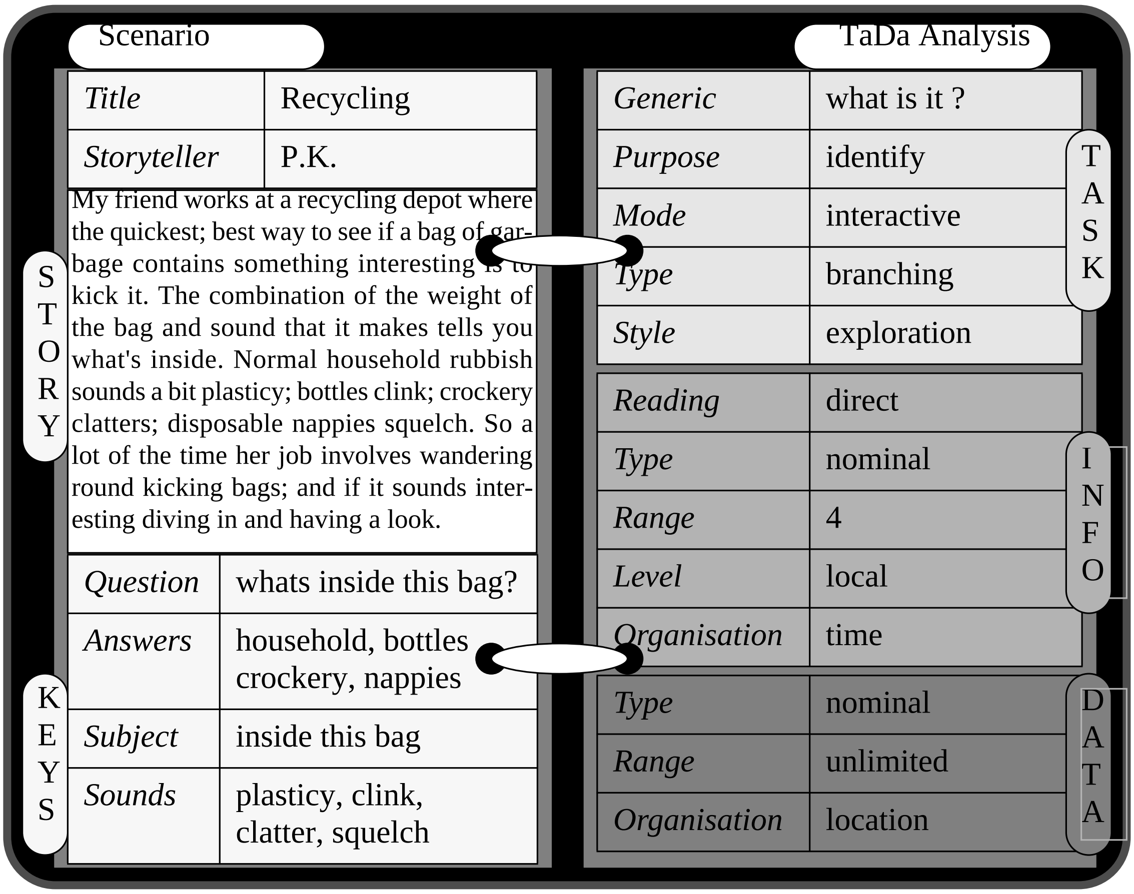

Figure 5-4: TaDa analysis for the Recycling story

Implementation details

The Access™ database can only retrieve records by exact match on all the query fields. This matching procedure does not allow the retrieval of cases by similarity that is necessary in the metaphoric method. The lookup of cases ranked by similarity to a TaDa query was implemented in a Perl program [Wall L. and Schwartz R. (1991)] called lookup.prl (see Appendix 5-2). The lookup process involves saving the Access™ database to a file as ascii text with a special “|” character as a field delimiter. The lookup.prl program accepts the text file as input, and generates a reorganised version ranked by similarity with the query. Similarity is measured by the number of fields that match. The best matches can be taken from the bottom of the generated file, or the file may be re-imported to Access™ and browsed with the forms interface. This process is shown in Figure 5-5.

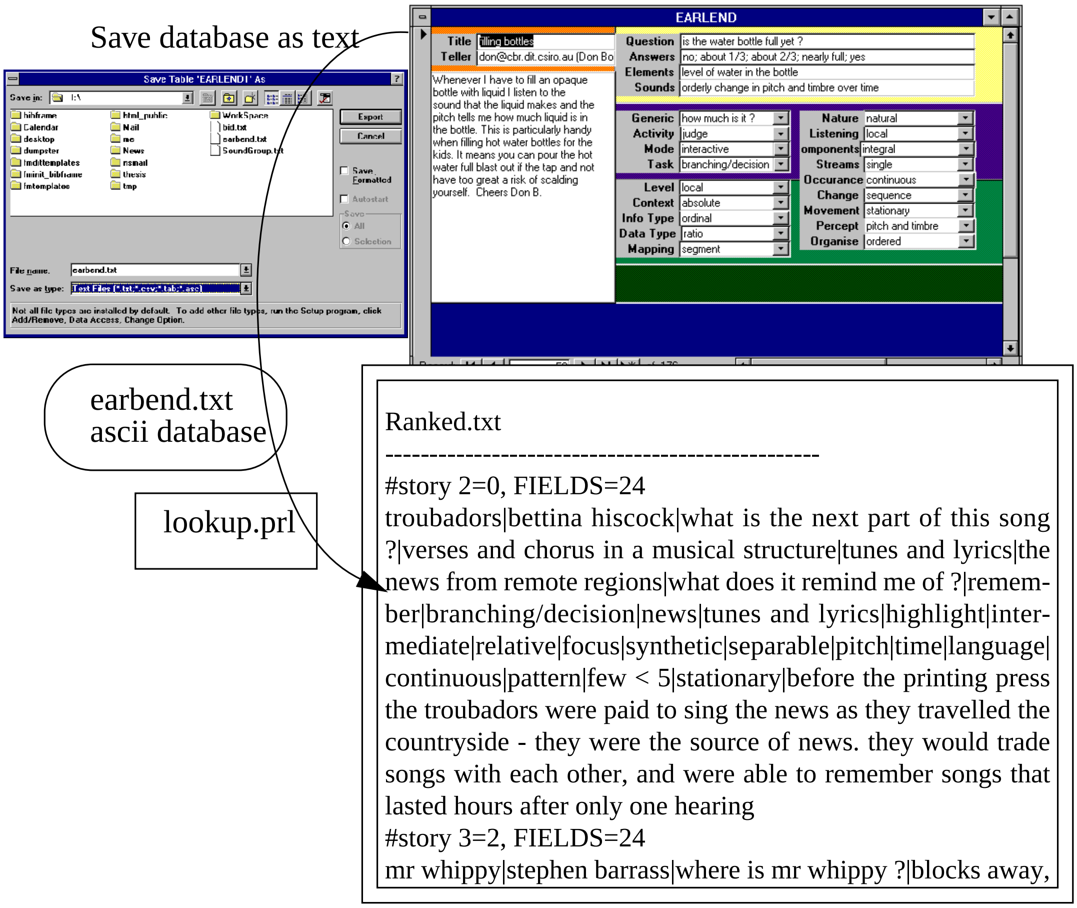

Figure 5-5: Looking up stories by similarity to a TaDa query

5.3.3 Pattern method

A pattern is a regularity in the mapping between the problem and solution domains. The pattern method begins by identifying features shared by different solutions to the same problem that may capture patterns in the mapping between domains. The pattern method was enabled by appending a description of auditory characteristics to each EarBenders case. When a TaDa query retrieves cases with similar information structure it also retrieves the auditory characterisation of each case. The characterisations are scanned for regularities identified by the majority trend in each auditory characteristic. The weighting of each trend is calculated from the number of cases which follow the trend. The result is a synthesis of the auditory characteristics of the retrieved cases. These cases have an information structure which is similar to the information requirements of the design scenario. Regularities in the auditory characteristics of the retrieved cases may capture patterns in the mapping between information and sounds. Hence the design of an auditory display from these regularities may support the information requirements of the design scenario.

The auditory characterisation

The auditory characterisation is both a description of sounds in the cases, and a specification of sounds for the design scenario. The characterisation is headed by the Sounds Key which has been added to the EarBenders Scenario Analysis to capture the auditory features of the story. This Key is a high level description of the sounds that give each Answer in the scenario. However a verbal description does not specify how to produce the sound on an output device. The Sounds Key is a bridge to a more detailed characterisation of the sounds which may assist in the pragmatic and syntactic aspects of the sound design. There is no all-encompassing way to characterise sound, so several different perspectives from music, psychoacoustics, perceptual psychology, and auditory display have been taken up. The characteristics are not necessarily orthogonal or independent, and only through practice will it be possible to determine which, if any, describe important features for auditory design. The current set of characteristics are {nature, level, streams, occurrence, pattern, movement, type, compound, descriptors}. Each is described in the following subsections.

Nature

Everyday sounds are acoustic events generated by physical interactions between objects rubbing and colliding in the environment [Gibson J.J. (1966)]. Everyday sounds also describe artificially generated sounds that have an environmental basis, for example digital recordings of acoustic events, or auditory icon algorithms [Gaver W.W. (1986)]. Musical sounds are generated by specially engineered instruments that shape the acoustics to engage a human listener. These instruments allow control over specific perceptual aspects, such as pitch, rhythm, timbre and loudness. Synthetic sounds are artificially generated sounds that do not have an acoustic basis. The sounds can be generated by electronic circuits or computer algorithms. Vocal sounds are moans, croaks, or other expressive but non-verbal sounds made by people and animals, including clicks and hums that insects make by means other than a vocal tract. Verbal sounds are recognisable words. Auditory Display researchers sometimes distinguish their field as the use of non-speech sound, but there are many ways that words can be used that are not the same as speech.

everyday	acoustic events in the physical environment e.g. knocks, scrapes, rumbles, crashes
musical	musical sounds generated by instruments specialised to shape acoustics to engage musical perceptions e.g. pitch, rhythm, loudness, timbre, etc.
synthetic	synthetic sounds with no acoustic basis e.g. car alarm, keycard beep, computer error quack
vocal	animal communications formants, phonemes, moans, grunts and sighs humming, whistling
verbal	recognisable words, singing

Table 5-1: Nature of the sound

Level

Analytic and holistic listening are musical terms for different listening styles. Picking the flute part in a symphony is an example of analytic listening. This attention to a single element is called local information in Bertin's graphical method. Holistic listening to the overall sound of the orchestra is analogous to the attention to global information. A good display allows attention to information at more than one level [Bertin J. (1981)].

local	analytic listening to a single element e.g. violin in orchestra
global	holistic listening to many elements e.g. whole orchestra

Table 5-2: Listening level

Streams

The ability to segregate the flute or clarinet from the rest of the orchestra is explained in terms of perceptual streams in Bregman's theory of auditory scene analysis [Bregman A.S. (1990)]. Streams are perceptual groups that form when sounds occur simultaneously and in sequences, as is usual in everyday listening. Bregman's theory explains some of the factors that influence the grouping of acoustic events into perceptually cohesive sounds in the mind of the listener. This grouping enables the listener to consciously select between different sound streams as they become interesting or useful - an ability sometimes called the cocktail party effect. The number of streams may be estimated from the number of sound sources that can be consciously identified in an auditory scene.

single	only a single stream is involved e.g. a voice talking
pair	a pair of streams are involved e.g. bass line and melody
few < 5	a few streams are involved e.g. shaking a box of muesli
some <10	5-9 streams are involved e.g. car sounds while driving
many 10	more than 10 simultaneous streams e.g. the aural scene during a picnic lunch in the park

Table 5-3: Streams

Occurrence

Sounds can be one-off or ongoing, just as tasks may be discrete or continuous [Kaplan B. and Goodsen J. (1995)]. It takes at least 4 seconds for the primitive stream grouping process to stabilise, and once an interpretation of the number of sources has occurred it doesn't matter if one or other of them briefly disappears for a second or two [Bregman A.S. (1990)]. This 4 second hysteresis is found in other perceptions, and provides a basis for classifying the occurrence of sounds. A continuous sound doesn't have a definite beginning or end, and interruptions last for periods less than the 4 second hysteresis of perceptual continuity. A regular sound repeats at predictable intervals that are greater than 4 seconds apart. A sporadic sound repeats at frequent but unpredictable intervals of more than 4 seconds. An isolated sound does not occur often at all.

continuous	an ongoing sound in which breaks are < 4 seconds e.g. a waterfall
regular	a sound that repeats at intervals > 4 seconds e.g. a dripping tap
sporadic	unpredictable repetition at intervals > 4 seconds e.g. wind-chimes in light breeze
isolated	a one-off sound e.g. a dropped key

Pattern

Sounds can vary in many ways over time. A discrete sound may have a typical duration of only about a second or two, and so may be too short to hear a pattern in. A constant sound does not change over its duration. An unpredictable sound changes in unpredictable ways but maintains its essential identity. A cycle is a cyclic variation than can be predicted after it has been heard. A sequence is a sound that has a predictable directed variation.

discrete	a short sound with definite start and end e.g. a hand clap
constant	does not change much e.g. air-conditioner hum
unpredictable	unpredictable variation e.g. wind-chimes, pop-corn popping
cycle	predictable cyclic variation e.g. a squeaky wheel
sequence	predictable directed variation e.g. water bottle filling

Table 5-4: Pattern

Movement

The movement of a sound is relative to the listener. The stationary sound stays in roughly the same location. A change in distance occurs in movements when other cues for spatial location are not available. Jumping sounds shift location in a discontinuous fashion. The location of a smoothly moving sound can be tracked and predicted. A texture sound does not have an identifiable location or movement.

stationary	the sound is fairly stationary e.g. rattling mudguard as you ride a bike
distance	the distance of the sound is changing e.g. walking to a surf beach
jumping	the location of the sound jumps about in space e.g. flying grasshopper wing clicks
smooth	the sound moves smoothly through space e.g. a plane flying overhead
texture	the sound has no identifiable location or movement e.g. the rain

Table 5-5: Movement

Type

The type is a characterisation of the perceptual relations between the sounds. Perceptual psychologists typically classify perceptions as categorical or continuous. A categorical perception has difference but no order. A continuous perception has a unidimensional organisation. Continuous perceptions were further divided by Stevens into metathetic and prothetic types [Stevens S.S. (1966)]. A metathetic perception is not additive, for example the simultaneous occurrence of two sounds of the same pitch is not heard as a sound with an increased pitch. On the other hand a prothetic perception is additive, so for example when two sounds of the same loudness are heard together there is an increase in overall loudness. All of these types of relations have been scaled to create organisations of equal perceptual difference.

categorical	difference e.g. piano note, engine rev, dog bark
metathetic	difference and order e.g. pitch, brightness
prothetic	difference, order, and natural zero e.g. loudness, duration

Table 5-6: Type

Compound

Sometimes it is possible to separate aspects of the sound from its overall identity, for example you may listen specifically to the rate of vibrato, or the depth of tremolo. Other sounds are more integral and it can be hard to hear them as anything but a whole - for example the crunching of gravel, the popping of a cork or a hand clap.

separable	aspects of the sound can be readily heard e.g. vibrato, tremolo, pitch
integral	it is difficult to separate aspects of the sound from the sound itself e.g. crunching gravel, popping cork, hand clap

Table 5-7: Compound

Descriptors

The descriptors are a list of the words used to describe the sounds. Verbal rating scales have been used to measure the primary dimensions of perceived variation between sounds in multidimensional scaling experiments [Von Bismarck G. (1974a)]. Shared descriptors may indicate a consistent variation or similarity relations. Multi-word descriptions may indicate the separability and attention to particular aspects in the sounds.

Auditory characterisation of an EarBenders story

The Sounds Key for the Recycling case is {plasticy, clink, clatter, squelch}. The Sounds Characterisation, shown in Figure 5-6, is a more detailed description of the sounds that may provide a basis for organising the perceptual and psychoacoustic aspects of an auditory display. The Nature is {everyday} because these sounds are generated by everyday physical interactions between materials rubbing and colliding. The Level is {local} because the sounds emanate from the particular bag that is being kicked. The Streams are {few} because, for the most part, each kick forms a cohesive auditory entity. However it may be possible to also hear separate objects, perhaps a plate breaking or a can crushing, as highlights in the overall mass. The Occurrence is {isolated} because the sound only occurs when the bag is kicked. The Pattern is {discrete} because it only occurs once per kick. The Movement is {stationary} because

Character

Sounds	plasticy, clink, clatter, squelch
Nature	everyday
Level	local
Streams	few
Occurrence	isolated
Pattern	discrete
Movement	stationary
Type	category
Compound	integral
Descriptors	plasticy, clink, clatter, squelch

SOUND

Figure 5-6: Auditory characterisation of the Recycling story the sound doesn’t move (unless you really send the bag flying!). The Type is {category} because the relations between the sounds are unordered, although some sounds may be more similar to each other than others. The Compound is {integral} because the sound is an overall or holistic identity. The Descriptors are a list of adjectives that were used in the description of the sounds. Shared descriptors may indicate principal dimensions of perceptual variation. In this case of integral sounds the descriptors are purely categorical and are copies of the Sounds themselves.

Implementation details

The pattern method is supported by a Perl program, called synthesis.prl (see Appendix 5-3), which automatically generates an auditory specification from trends in the auditory characteristics of EarBenders cases. This program accepts a file of cases in ascii text format, as output by lookup.prl. The best cases at the bottom of the file are scanned for trends. The most common value for each field is recorded, along with the count of cases that had that value. An auditory design is generated from the majority value of each characteristic, and printed to the standard output. The complete process, shown in Figure 5-7, is tied together by a csh script called casedesign.csh (see Appendix 5-4).

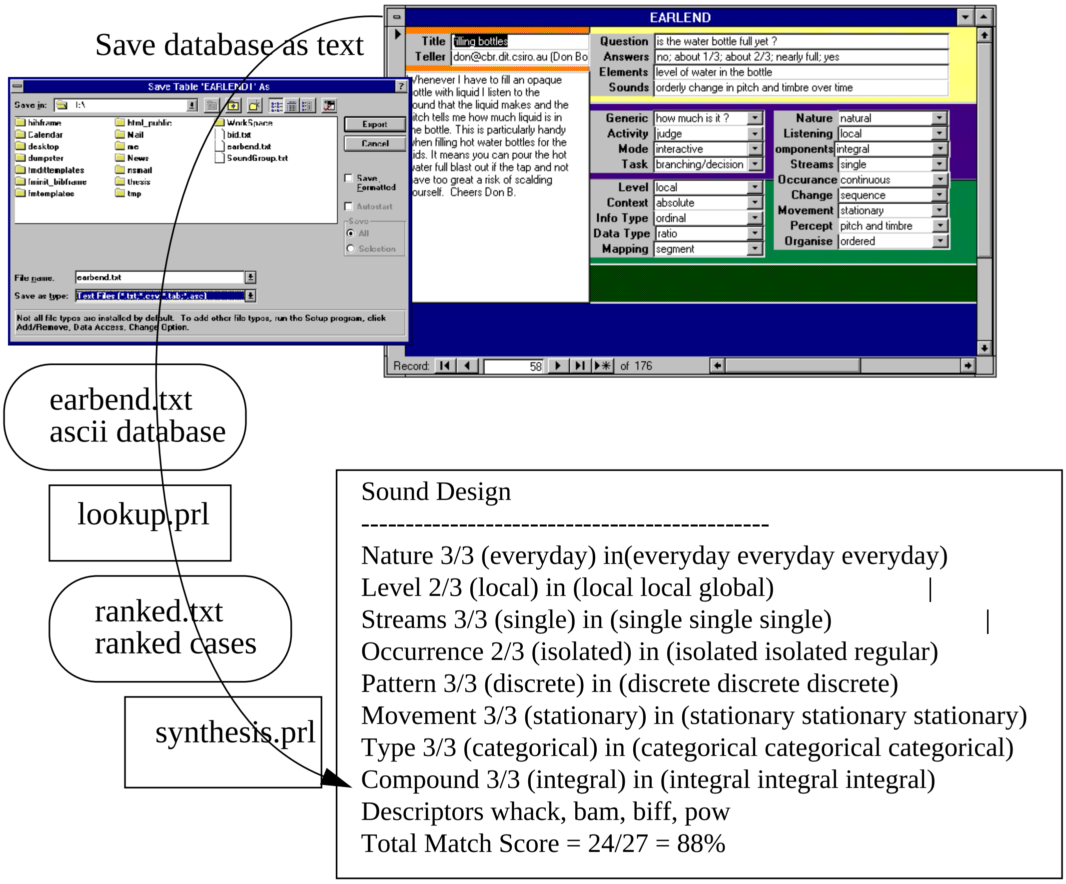

Figure 5-7: From TaDa query to sound design

5.4 Demonstration of EarBenders design

This section demonstrates EarBenders by designing sounds for the GeoViewer scenario that was introduced in the previous chapter. The metonymic, metaphoric and pattern methods are each demonstrated in the following sections. The GeoViewer scenario and analysis are summarised in Figure 5-8

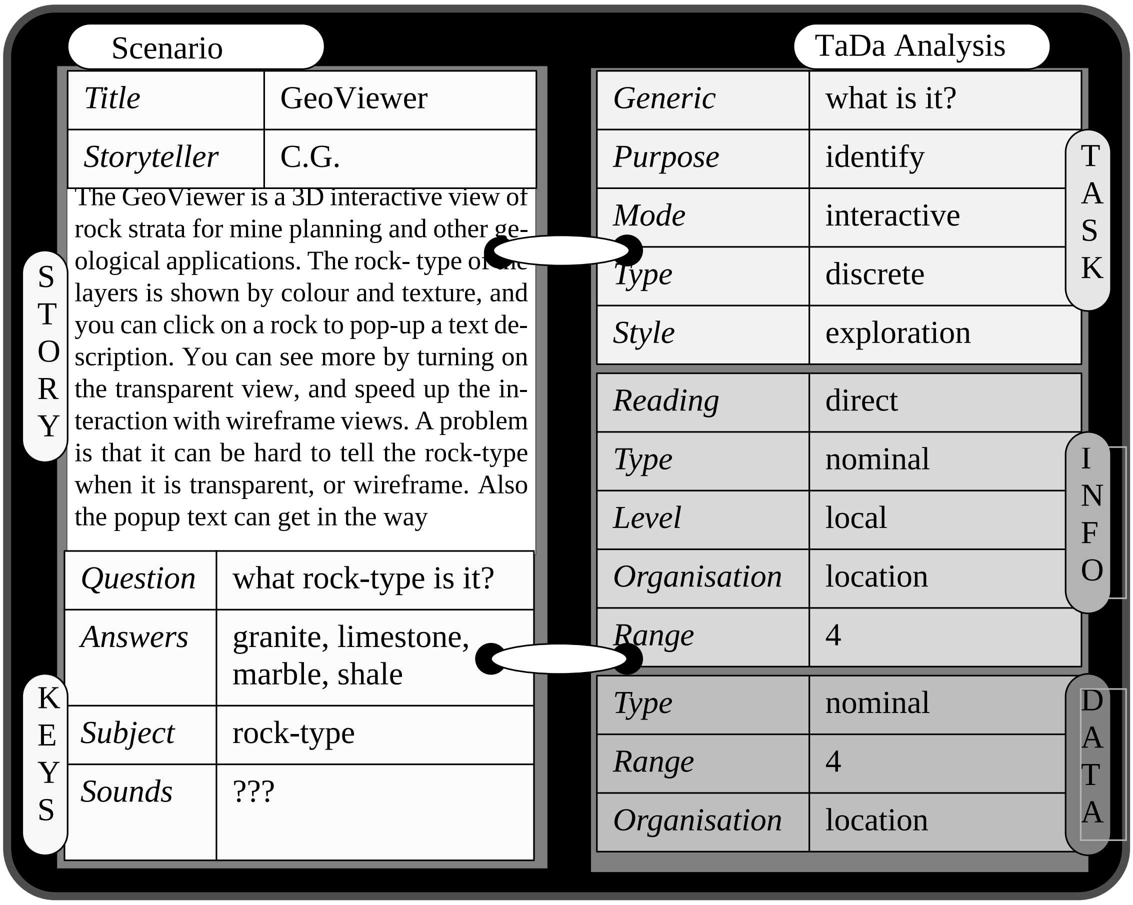

Figure 5-8: TaDa Analysis for the GeoViewer

5.4.1 Metonymic method

In the metonymic method we are looking for sounds that are a part of the design scenario. The GeoViewer is from the domain of mining exploration, and some keywords that may find other examples from similar scenarios are {mining, exploration, geophysics}. I tried searching EarBenders with these fields but did not turn up any cases. The story in the GeoViewer scenario, shown in Figure 5-8, does not mention any sounds so there are no leads there. The subject of the scenario is the rock-type, which is the object of interest. The actions that involve the rock are - clicking on it to find out what type it is, and moving it to see what shape it is. We can search EarBenders for object keywords like {rock, strata, layer}. This returned one story about dropping a rock down a hole to find out how deep it is, which did not seem relevant. The event keywords are {click, tap, knock, impact}. Click turned up a lot of stories which had clicking sounds in them but were otherwise irrelevant. Some possibilities were turned up by {tap} and {knock}which retrieved 4 stories - ripe fruit, gas cylinder, hammering a nail, and finding studs, shown in Figure 5-9 and Figure 5-10.

Scenario

Title	ripe fruit
Storyteller	D.B.
STORY	You can tell whether a water melon is ripe by tapping it and listening - this also works for apples - it must be something to do with the damping of the sound when the fruit isn't ripe, because it sort of rings when it is ripe.
Question	is this fruit ripe?
Answers	yes, no
Subject	the fruit
Sounds	yes=ringing no = damped, dull

Scenario

Title	gas cylinder
Storyteller	S.B.
STORY	which gas cylinder has more in it? you tap them and the fuller one has a low dull sound while the emptier one is higher and has a brighter hollower tone
Question	which cylinder has more gas in it?
Answers	this one, that one
Subject	the cylinders
Sounds	more gas = lower, duller tone

Figure 5-9: Stories retrieved by the keyword {tap}

Scenario

Title	hammering a nail
Storyteller	R.U.
STORY	A solid sounding knock, then a softer slightly clangy sound means that the nail is bent. A solid sounding knock, then overly solid (higher pitch) means that the nail has hit something like a pipe or concrete.
Question	is the nail ok?
Answers	yes, no
Subject	the nail
Sounds	yes=solid knock no = softer slightly clangy, overly solid high pitch

Scenario

Title	finding studs
Storyteller	S.B.
STORY	you can find wall studs by knocking on the walls and listening - where the stud is sounds solid and dull compared to the hollow sound of the rest of the wall.
Question	where is the wall stud ?
Answers	here, not here
Subject	the wall stud
Sounds	here = dull solid not here = hollow

Figure 5-10: Stories retrieved by the keyword {knock}

The knocks and taps in these stories produce sounds that provide quite different information. The sounds in the stories do not seem connected to mining or rocks. However the idea of knocking on something to hear information does suggest tapping on a rock with a hammer. This idea may be the starting point for a design. My first investigation of this design was to go outside and tap some rocks with a hammer. I tapped on marble, granite, sandstone, and a conglomerate. They didn't sound very different. I closed my eyes and tried to learn the sound each rock made. I could hear how hard I was tapping, and that it was rock that I was tapping, but could not identify what type of rock it was. I couldn't hear any correspondence between the size of chunks of granite, and the tapping sound. The variations in the physical properties of the rocks don't seem to have perceptible affects on the sounds generated by tapping them. Tapping metal and wood did produce discriminably different sounds. I could correctly identify rock, wood and metal by tapping them. My experience is that the identification of materials by tapping on them requires familiarity, and the discrimination between materials requires that they have quite different material properties. These observations raised the possibility of a metonymic palette comprised of clearly discriminable and identifiable impact sounds. The impact-like nature of the sound could suggest tapping and connote the scenario.

A metonymic design for the GeoViewer was investigated by assigning an impact sound to each rock type as follows
granite = tennis serve
limestone = door slam
marble = glass break
shale = bell strike

The resulting sounds were clearly different and it was easy to answer the question "is this the same rock type as that?" straight away. After using the interface for a couple of minutes I was able to consistently answer the question "what rock type is it?" from the sounds. The sounds demonstrably do provide the information required by the design scenario.

The sounds also allowed answers to other questions that weren't in the design scenario. The intersections between rock surfaces can be visually ambiguous in a 3D view, and answering the question "which surface is in front?" involves a change in viewpoint which can be distracting and slow. The sounds let you answer this question by tapping at the intersection and listening for the fore-most surface. A development could provide information about the number and material of overlapping hidden layers. The prospect of unanticipated affordances of the sounds is very encouraging.

The investigation also turned up some problems. Firstly, there is an inconsistency in the degree of the impacts that does not correspond with the consistent nature of tapping with the mouse button. Secondly, the initial impression is that the breaking glass and the tennis serve are out of context because they clearly connote their origins. Nevertheless, after using the interface for a while I found that the sounds began to become a part of the tool, and their use in this context overtook other associations. All the same the initial impression of the tool depends critically on the connotations of the sounds in the context of the other symbols. The final problem was the need to learn the sounds, which is counter to the requirement for a direct and immediate interface. The metonymic method aims at a direct display by taking advantage of sounds that are part of the design scenario. Although the sounds in this design could immediately answer the question "is this the same rock type as that?" it took several minutes of use to learn to answer the question "what rock type is this?". It is difficult to know how to design a more direct display for categorical information. Auditory icons have a premise of directness, but it is not clear how this can be realised with this scenario involving rock-types. You can specify different materials such as wood, metal and glass in auditory icon algorithms, but these materials are no more a part of the problem domain than the other impact sounds that were tried, and so entail the same need for learning. Ballass found that up to 35 different sources may be attributed to the same everyday sound. He also found that the identification of discriminable but ambiguous everyday sounds depends critically on the expectancy and context of the listener [Ballass J.A. (1994)].

The problem that the answer associated with each sound takes a few minutes to learn can be overcome with verbal sounds. An experiment that compared the learnability of auditory icons, earcons and spoken words found that the time taken to associate interface actions with auditory icons and earcons was similar, but that speech feedback was significantly faster to learn and more reliable [Lucas P.A (1994)]. A verbal design was implemented by sampling a person saying each of the 4 rock-types, and storing the words as audio files.

granite = “granite”
limestone = “limestone”
marble = “marble”
shale = “shale”

Tapping the rocks in the interface produces the spoken response. The “what is it?” question is answered immediately, without any need for learning at all. It also answers the other questions just as well as the non-verbal design. New rock types can be added without introducing ambiguity because words are inherently discriminable. A large catalogue of rock-types could be represented this way, without the need for a long training period that is necessary for systems of more than 6 or 7 symbols [Patterson R.D. (1982)]. Although the verbal design is direct, it also has its problems. The verbal design is probably best suited to local information, since it may be difficult to understand more than a couple of words at the same time. Spoken words may interfere with other speech in the environment, where non-verbal sounds may not. There is some incongruity in that a talking rock requires a leap of imagination (transposition) that is expected in metaphors, but there are no other symbols in the interface to support this metaphor. Metonyms can seem more natural than metaphors because they are a part of the evoked object (contiguous). The words are semantically a part of the design scenario, but not the interaction in the interface. The words typically take about a second to say, and can slow the interaction with the interface down a lot. However speech can be time compressed by a factor of about 2 and still maintain intelligibility [Arons B. (1994)]. Time compression techniques could be used to equalise the durations of the words, and to set the durations to < 0.5s which is comparable with a short impact sound.

5.4.2 Metaphoric method

The metaphoric method is a way to represent the unfamiliar in terms of the familiar. The method works by looking up some EarBenders cases that are similar to the design scenarios in their information structure. The TaDa analysis for the GeoViewer, shown in Figure 5-8, was used to retrieve 3 EarBenders cases - Recycling, Walking and Kitchen, shown in Figure 5-11, Figure 5-12, Figure 5-13. Although the subject matter of each story is very different, they share a similar TaDa structure with the design scenario. The information required in the GeoViewer is the type of a rock strata that is being explored. The Recycling case describes how workers at a recycling depot classify garbage bags by listening to the sounds they make when kicked. The Walking case describes different sounds that footfalls make on different walking surfaces. The Kitchen story describes the identification of different cereals in opaque containers by shaking them.

Scenario

Title	Recycling
Storyteller	P.K.

STORY

My friend works at a recycling depot where the quickest; best way to see if a bag of garbage contains something interesting is to kick it. The combination of the weight of the bag and sound that it makes tells you what's inside. Normal household rubbish sounds a bit plasticy; bottles clink; crockery clatters; disposable nappies squelch. So a lot of the time her job involves wandering round kicking bags; and if it sounds interesting diving in and having a look.

KEYS

Question	whats inside this ?
Answers	household, bottles crockery, nappies
Subject	inside this
Sounds	plasticy, clink, clatter, squelch

TaDa Analysis

TASK	*Generic*	what is it ?
	Purpose	identify
	Mode	interactive
	Type	branching
	Style	exploration
INFO	*Reading*	direct
	Type	nominal
	Level	local
	Range	4
	Organisation	time
DATA	*Type*	nominal
	Range	unlimited
	Organisation	location

Figure 5-11: Recycling case

Scenario

Title	Walking
Storyteller	S.B.

STORY

walking home in the dark
footsteps make different sounds
as I walk over different surfaces
leaves are crackly
concrete is quiet thud
grass is muffled drag
gravel is crunchy

KEYS

Question	what surface am I walking on ?
Answers	leaves, concrete, grass, gravel
Subject	surface
Sounds	rustle, quiet thud, muffled drag, crunchy

TaDa Analysis

TASK	*Generic*	what is it ?
	Purpose	identify
	Mode	focus
	Type	discrete
	Style	exploration
INFO	*Reading*	direct
	Type	nominal
	Level	local
	Range	7+-2
	Organisation	time
DATA	*Type*	nominal
	Range	unlimited
	Organisation	location

Figure 5-12: Walking case

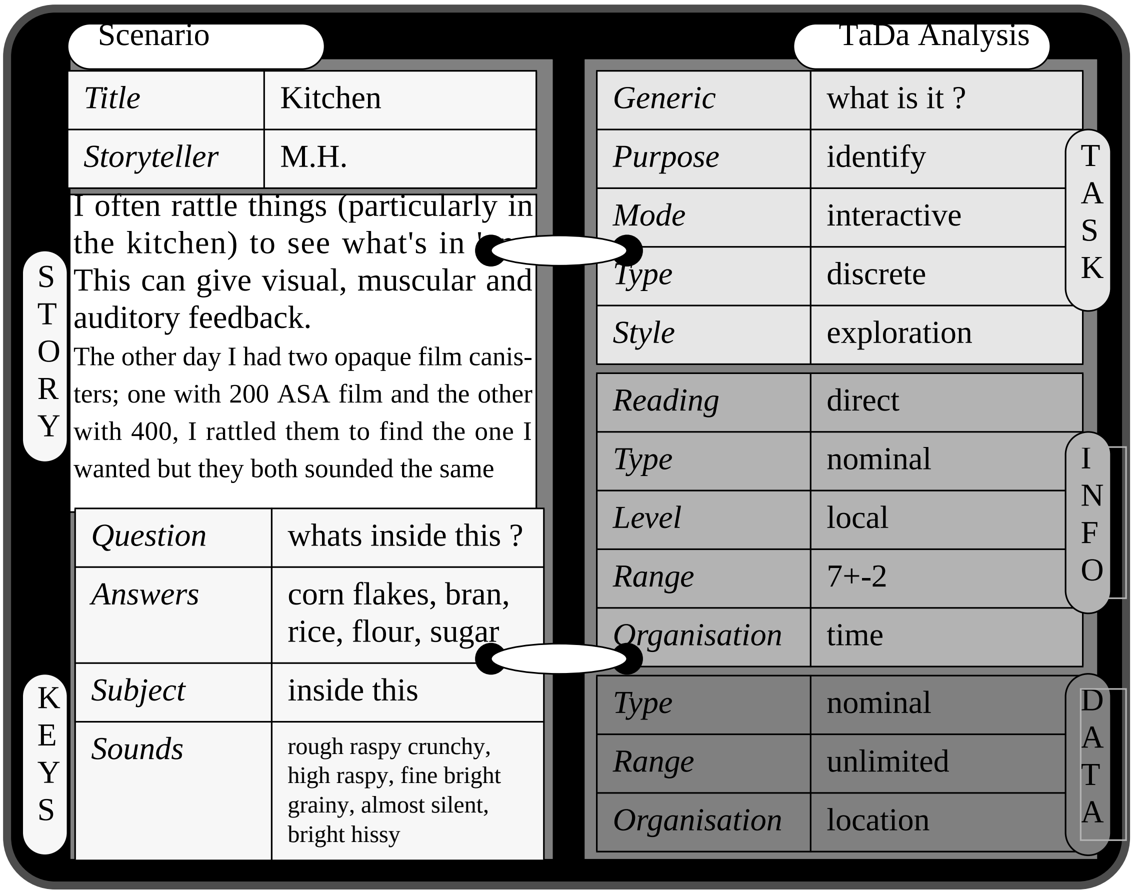

Figure 5-13: Kitchen case

Although the stories provide some interesting metaphors there is a problem. The metaphoric method works best when the suggestion of sounds that are part of the design scenario is weak or non-existent, allowing a leap in imagination. In the GeoViewer scenario the tapping of rocks is an easy sound to imagine, and it is hard to replace this with a metaphorical sound that has other connotations. The semantics of the metaphorical design need to be carefully considered in the context of the design scenario - for example garbage sounds might evoke pollution where none exists. Shaking cereals suggests a movement of the rocks which are stationary. Perhaps the footfalls on different surfaces has some potential to suggest walking over the rocks that could be a basis for a metaphorical design.

5.4.3 Pattern method

The pattern method is a search for trends in a set of solutions to similar problems that may capture consistencies in the mapping from the problem domain to the solution domain. The TaDa query from the GeoViewer scenario retrieved the cases shown in the previous section. The pattern method continues by scanning the auditory characterisations of these relevant cases for trends. The cases shared 6 of the 8 characteristics, shown by a score of 3 in the fields in Figure 5-14 - {everyday, local, single stream, discrete, stationary, category}. The other 2 fields had a 2/3 majority trend - {isolated and integral}. The auditory characterisations are very similar. However the actual sounds and descriptions of the sounds are quite different in each case. The Descriptors field contains 18 words that describe the sounds in the 3 cases. Only “crunchy” appears more than once - to describe both walking on gravel and shaking a container of cornflakes in the kitchen. The variety of descriptions reinforces the notion that it is the relations between sounds rather than the sounds themselves that are important in conveying information relations. The auditory design proceeds by selecting a set of sounds according to the characterisation synthesised from the trends. The actual choice of sounds has not been filled in Figure 5-14. This choice might be made with the metonymic or the metaphoric methods described in the previous sections. These methods consider the connotations of the auditory symbols with respect to the design problem. An alternative to these semantic approaches might be a structured palette of “abstract” sounds that do not have an everyday connotation that will interact with the connotations of the rest of the interface.

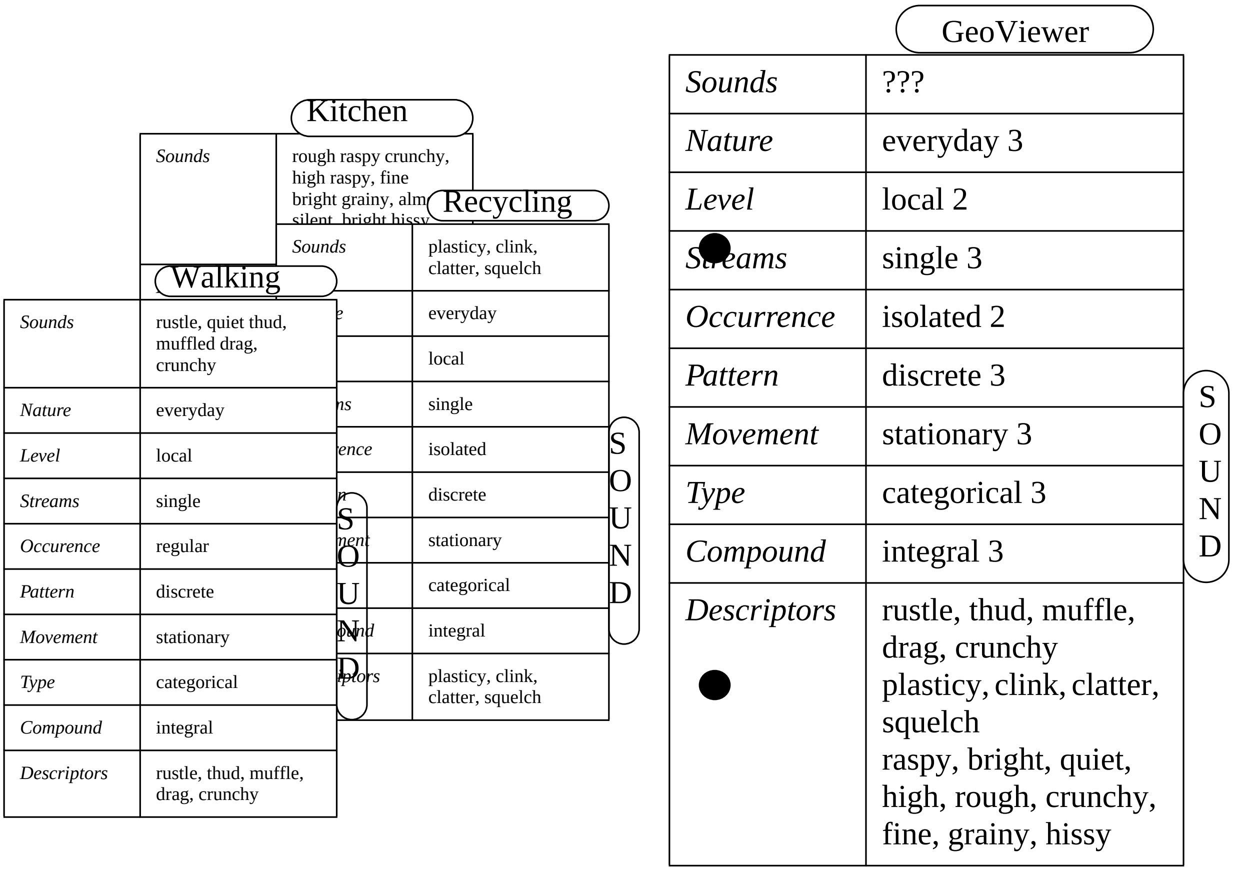

Figure 5-14: Pattern method for the GeoViewer

An example of an abstract palette might be the musical sounds from classical instruments. Although these obviously do have connotations, they are not usually associated with everyday tasks, and so may be more transportable between application scenarios. A palette of musical instrument samples was chosen in accordance with the auditory characterisation from the GeoViewer. However the Nature of the sound was changed from {everyday} to {musical}. The timbres of musical instruments are usually considered to be categorical, so 4 instruments with discriminable timbres were chosen. The pitch and duration of the sounds are the same, to prevent a misleading impression of order between the categories.

granite = tenor sax, pitch A3, duration 0.5s
limestone = cello, pitch A3, duration 0.5s
marble = English horn, pitch A3, duration 0.5s
shale = trombone, pitch A3, duration 0.5s

As with the metonymic designs, the answer to the question “is this the same rock type as that?” was immediate. However learning to answer the question “what rock-type is it?” took longer than it did with the metonymic palette of familiar impact sounds. Perhaps this was because the musical instrument timbres were much more similar to each other than the highly recognisable impacts, and I had to learn the timbral features of each instrument as well as the association with an answer. The musical palette has the advantage that the sounds can be much shorter than either the impact or the word designs. Just how short depends on the instrument. The instrumental sounds in Grey’s study of timbre were 350 ms in duration. Hammered and plucked instruments can be identified from durations of less than 100 ms [Grey J.M. (1975)]. Many everyday sounds, such as a tap drip or clock tick, can be identified from durations between 50 and 100 ms [Ballass J.A. (1994)]. These sounds are an order of magnitude shorter than words.

5.5 Summary

HCI designers, bridge engineers and home renovators all use previous examples as a starting point for new designs. Case-based design is a top down method that can help in complex problems and areas of practice where few predictive principles exist. Auditory display is a developing field where case-based design can be helpful. However there are not enough examples of auditory display to draw upon as a case-base. In this chapter it was proposed that everyday uses of sounds could provide a pool of examples that could make case-based design practical in auditory display. A collection of nearly 200 stories about everyday listening experiences was gathered by email and conversation, and entered into a database called EarBenders. The stories were found to contain semiotic elements that could help in the design of an auditory display and 3 methods of case-based design were developed to take advantage of this resource - the metonymic method, the metaphoric method, and the pattern method.

The metonymic method searches the case-base for keywords that are a part of the design scenario. The aim is to design a palette of sounds that are a familiar part of the application domain, and may evoke that context for a listener. The metaphoric method represents the unfamiliar with the familiar. This is suited to scenarios where sounds do not naturally occur, which is often the case in computer-based information processing activities. EarBenders stories can be retrieved by similarity to the information structure of the design scenario, as described by a TaDa analysis. The retrieved stories are a source of metaphors that can support the information requirements of the design scenario. The pattern method captures regularities in the mapping from the design domain to the solution domain. EarBenders stories have an auditory characterisation that can describe the characteristics of an auditory display. The trends in the characterisations of stories retrieved with the TaDa characteristics of the design scenario were used to automatically synthesise an auditory specification. This specification describes the relations between sounds, but not the sounds themselves. The actual choice of sounds to meet the specification might use a metonymic or metaphoric palette, or perhaps an abstract palette of musical sounds that do not have everyday connotations.

Each case-based method of auditory information design was demonstrated in a mining visualisation scenario. The GeoViewer is a 3D interactive view of rock strata for mine planning and other geological applications. A problem is that it can be hard to tell what type of rock you are looking at, so an auditory display was designed to provide information about the rock-type as well. A rock strata that is difficult to visually identify can be heard, without having to divert visual attention to a text. An unexpected advantage became apparent when the interface was used. The sounds allow the front most surface at an intersection of strata to be disambiguated by tapping there, saving on a distracting and computationally expensive change of viewpoint operation. A development could provide information about the number and material of overlapping hidden layers. The demonstration showed that the case-based method is functional in practice, and can provide a good starting point for the design of a useful auditory display.

5.6 Further work

There is great potential to use the Internet to gather many more stories for the EarBenders case-base. Examples from different cultures could allow the possibility of culturally specific metaphors. For example the sounds of cooking in a kitchen in Australia may be quite different from the sounds in a kitchen in Japan.