University of Namur

MASTER IN COMPUTER SCIENCE

Sound as means for data representation for data mining

Warin, Cédric

Award date:
2002

Facultés Universitaires Notre-Dame de la Paix, Namur
Institut d'Informatique
Année Académique 2001-2002

Sound As Means for Data Representation for Data Mining

Cédric Warin

Mémoire présenté en vue de l'obtention du grade de Maître en Informatique

Abstract

Representing and transforming data into a form that facilitates discoveries is a key component in the overall process of knowledge discovery. This study shows a few techniques that use the possibilities of sound. Sound can be an alternative to the traditional approach based on visualization. These models which transform data into sounds can become an intrinsic support in the emerging field of data mining.

Through this Master's thesis, we will discover some recent models, principles and other sound support that are interesting for data mining. An Application using the Java Synthesizer called Jsyn is developed in order to implement some of the principles and to analyze a few features of the sound representation.

Keywords: Sound, models, principles, Jsyn

Résumé

Représenter and transformer des données dans une forme qui facilitent les découvertes est un composent clé dans le processus général de découverte de connaissance. Cette étude montre quelques techniques qui utilisent les possibilités du son. Le son peut être une alternative à l'approche traditionnelle basée sur la visualisation. Ces modèles qui transforme les données en sons peuvent devenir un support intrinsèque dans le champ émergeant du data mining.

A travers ce mémoire, nous découvrirons des modèles récents, des principes et autre support pour le son qui sont intéressants pour le data mining. Une application utilisant le synthétiseur Java appelé Jsyn est développé dans le but d'implémenter quelques-uns des principes et d'analyser quelques caractéristiques du son.

Mots-clés : Son, modèles, principes, Jsyn

Acknowledgements

First of all, I would like to thank my supervisor, Monique Noirhomme, for her help in all this study.

I would also like to thank my training course master, Simeon Simoff, for all his help during my four months stay at the Sydney University and more generally all the staff at the University of Technology in Sydney. People in Sydney were so friendly that I cannot forget anyone. Thanks to all for the help in the beginning of my stay, for the accommodation and the tools furnished during the four months.

Furthermore, I would like to thank Anne Hart, the Personnel Services Coordinator for my appointment as Visiting Scholar but also Michael Fry and his colleagues for their assistance and their information.

I would like to thank all the people who have helped during this study and during the making of this report. Also I will thank the people who have answered the different questions for the application and their interest and support. Their feedback was constructive for the application. Every answer has been taken into account and has contributed to the final results.

Particularly I would like to thank Donal, Martin, Oscar, Fabrice, and Anupam for all the support, feedback, assistance, interest. I sincerely thank them and hope they will recognize themselves.

Contents

List of Figures

Chapter 1

Introduction

The explosive growth of many business, government, and scientific databases has far out outpaced our ability to interpret and digest this data, creating a need for a new generation of tools and techniques for automated and intelligent database analysis. Advances in data storage technology, such as faster, higher capacity, and cheaper storage devices (e.g. magnetic desks, CD-ROMS), better database management systems, and data warehousing technology, have allowed us to transform this data deluge into "mountains" of stored data (e.g. Wal-Mart, a U.S retailer has to handle 20 million transactions a day).

Such volumes of data clearly overwhelm the traditional manual methods of data analysis such as spreadsheets and ad-hoc queries. Those methods can create informative reports from data, but cannot analyze the contents of those reports to focus on important knowledge. A significant need exists for a new generation of techniques and tools with the ability to assist intelligently and automatically humans in analyzing the mountains of data. These techniques and tools are the subject of the emerging field of knowledge discovery in databases.

Representing and transforming data into a form that facilitates discoveries is a key component in the overall process of knowledge discovery. For example, in visual data mining, the data is transformed by visualization models into a form that allows the extensive use of visual computing. The basic assumption is that large and normally incomprehensible amounts of data can be reduced to a form that can be understood and interpreted by a human through the use of visualization techniques. The design of visualization models for visual data mining, in broad sense, is the formal definition of the rules for translation of data into graphics. By analogy we can describe the design of audio models for data mining, i.e. models that transform data into sounds, by defining the rules of such translation. These models based on sounds can become an important support.

An approach for designing sounds that are useful information rather than "noise" is thus important. This Master's thesis is divided in three subprojects that represent three distinct areas. In a first step, we will identify current models that use sound in supporting explorative data mining and analysis. The investigation includes an identification of the features of the sound representation that can be useful for facilitating discovery processes but also the use of sound, sound models, sound supports, formal methods for data analysis and data mining. The second step will be to develop a model for mapping data into audio. This step includes a formal description of the metaphor where the sound forms are part of the metaphor, a description of the sound components and characteristics that will be used, the mappings between data and sound and the semantics of this mappings. The third step is the implementation of a prototype that uses Java Technologies and that is based on the metaphor described in the second step. This prototype will be experimented with data sets.

Annex A offers a theoretical overview of the concepts of sound and sonification. It introduces the features of the sound that will be used all along the study. It motivates the utilization of methods based on sound in data mining.

Chapter 2 focuses on sound supports, sound models and other principles that could be used in data mining. They form a list of information that could be used in a more formal method.

Chapter 3 offers an overview of the concepts of the TaDa method that is very useful and is the base of a formal method, both for description and utilization.

Chapter 4 describes the information requirements for the Task and Data analysis that we find in the TaDa method.

Chapter 5 focuses on principles for auditory information design. The user will learn to think in terms of principles in order to retrieve an everyday example by the information requirements like found in chapter 5.

Chapter 6 offers a more practical tool for designing sounds based on the information types in a database. The method is similar to a method for the color choice and is called the TBP prototype for Timbre, Brightness and Pitch.

Chapter 7 shows an example called the GeoViewer. It is described with the knowledge learnt in the three previous chapters and it shows the usefulness of all the methods seen before.

Chapter 8 describes a Java application that has been developed. This application is called SoundApplic. The chapter explains the metaphor that has been thought and describes the different sound component. It also explains which sound characteristics are important for each sound played. An evaluation on the sounds played has been realized. It gives the opportunity to test if people recognize the different sounds and are able to discern the modifications in the sounds. The results of this evaluation are found in Annex B.

Chapter 9 offers an analysis and criticism of the Tada method and especially the tools and methods that support the Tada method.

Conclusions of the present dissertation will be found in Chapter 10.

Chapter 2

Sound models and methods: previous approaches

2.1 Introduction

Approaches to auditory display have been classified into the semiotic types of lexical, syntactic and semantic that place different emphasis on learnability, organization, and discrimination.

The syntactic approach focuses on the organization 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 optimization of the sounds for a specific device.

Vocabulary:

Semiotics is a theory of signs and their meanings that has been used to analyze communication media. Some key terms and concepts in semiotics are introduced here.

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.

2.2 Syntactic approach

Syntactic principles bear on the way signs are organized to produce meanings.

The **earcon** [Bar97a] 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". Transformations, combinations, inheritance, and polyphony can organize the syntactic structure of an earcon.

Advantage:

- 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.

Disadvantage:

- 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. 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. The sounds modeled 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. The design of an auditory icon starts with an analysis of interactions between objects in the interface that would cause sounds in the physical world.

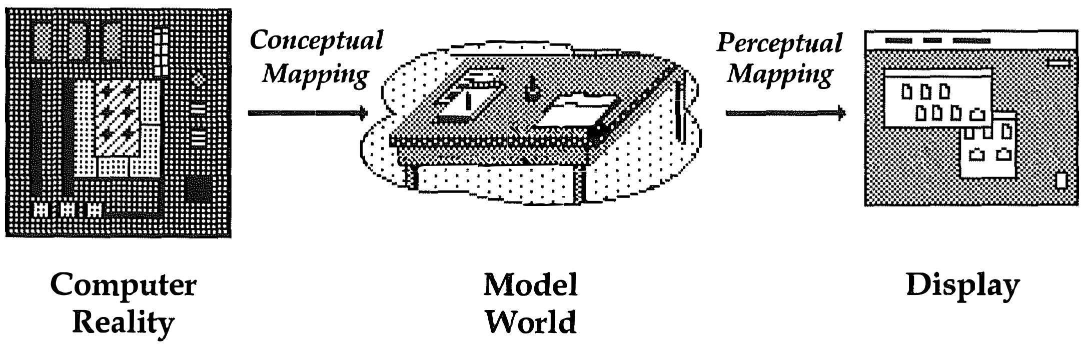


Figure 2.1: Mappings between computer reality, model world and display; Two mappings exist between the reality of the computer and the display to which the user has access. The first is a conceptual mapping between events in the computer and those in some model world (for instance, that of a desktop). The second is a perceptual mapping between events in this model world and their perceptible manifestations, be they visual or auditory [Gav].

For example, the SonicFinder [Gav] augments 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.

Certain experiments show that the speed and accuracy of identification of the source of a sound depends on the listener's expectations, context and experience.

2.4 Pragmatic approach

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

This is an example of study of auditory displays for aircraft by Deatherage [Bar97a]:
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 distances (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 of the aircraft that is flying.
  (c) Automatic information - recorded word signals as from automatic enunciators.
- 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 localize (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.

McCormick and Sanders [Sha97]

They provide some general guidelines that extend to other types of auditory displays outside the cockpit. 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 - selection of signals should exploit learned or natural relationships (e.g., high frequency associated with up or high or wailing sound of sirens for emergency vehicles) Remark: we should avoid extremes of auditory dimensions - high intensity signals, for example, can cause a startle response and actually disrupt performance.

- Approximation - Two-stage signals should be considered for complex information
  1) Stage 1 - alerting signal to attract attention
  2) Stage 2 - designation/message signal (once you have attracted the attention you can deliver the information)
- Dissociability - Auditory signals should be easily discernible from ongoing audio input (e.g. use two different frequencies if we listen to two or more channels)
- Parsimony - Input signals should not provide more information than is necessary
- Invariance - The same signal should designate the same information every time

Principles of installation of auditory displays

- Test signals to be used
- Avoid conflict with previously used signals
- Facilitate changeover from previous display - where auditory signals replace some other mode of presentation (e.g., visual), continue both modes for a while until new signal is suitably learned

Brewster, Wright and Edwards [BWE95]

They propose some guidelines for the creation of earcons. These are synthetic sounds that can be used in structured combinations to create sound messages. They try to answer the question "what sounds should be used at the user interface?" Their empirical observations are summarized as guidelines for designing earcons, shown in Table 2-1.

Timbre	Use musical instrument timbres, simple tones such as sinewaves or square waves are not effective. 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".
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 better.
Pitch	Do not use pitch on its own unless there are very big differences between those used (see register). Complex intra-earcon pitch structures are effective in differentiating earcons if used along with rhythm or another parameter. The maximum pitch used should be no higher than 5 kHz and no lower than 125 Hz-150 Hz. By this way, the sounds are not easily masked and are within the hearing range of most listeners.
Rhythm, duration and tempo	Make rhythms 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 0.0825 sec. Changing the tempo, speeding up or slowing down the sounds, is another effective method for differentiating earcons
Intensity	Listeners are not good at making absolute intensity judgements. Therefore, intensity should not be used on its own. Some suggested: 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. The major problem is the annoyance due to sound pollution.
Spatial location	This may be stereo position or full three-dimensions if extra spatialisation hardware is available. This is very useful for differentiating parallel earcons playing simultaneously.
Combinations	When playing earcons one after another, use a gap between them so those 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

Bergman [Hur91]

His theory proposes that acoustic elements are grouped into "streams" according to heuristics such as similarity, proximity, closure and familiarity. Streams are perceptual groups that form when sounds occur simultaneously and in sequences, as in everyday listening. Auditory streaming entails two complementary domains of study. How sounds cohere to form a sense of continuation is the subject of stream fusion. How concurrent activities retain their independent identities is the subject of stream segregation. Stream determining factors include: timbre, fundamental frequency (pitch), temporal proximity, harmonicity, intensity and spatial origin (however spatial origin seems in this theory a weak factor because the reverberation that can be used in spatial origin can indeed confound localization).

The two important distinctions in Bergman's theory are between primitive grouping and schema-based segregation.

- Primitive grouping - fast, bottom-up process, streams are parsed according to acoustic properties of the proximal stimuli.
- Schema segregation - slower, top-down process, 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.

- 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 that 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 [Wil]
Psychoacoustic observations describe the relation between acoustic variations and what a listener hears. This pragmatic method suggests that psychoacoustic measurements and theories can assist in the design of an auditory display.
Williams lists gestalt heuristics as principles for auditory display design as shown in Table 2-2. This approach may help in the design of concurrent and overlapping sounds.

Similarity	Components, which share attributes, are related.
Proximity	Components close together are grouped.
Good continuation	Components that display smooth transitions from one state to another are related.
Habit or Familiarity	Recognition of well-known configurations among possible sub-components leads to these sub-components being grouped together.
Belonginess	A component can only form part of one object at a time.
Common fate	Components, which experience the same kinds of changes at the same time, are related.
Closure	Incomplete figures tend to be completed.
Stability	Having achieved one interpretation of an acoustic signal, that interpretation will remain fixed through slowly changing parameters until the original interpretation is no longer appropriate.

Table 2-2: Gestalt heuristics

The techniques for evaluating perceptual processes are generally focused on pitch, timbre, rhythm, and localization, resulting from the frequency, duration, and intensity attributes of the acoustic signal. It seems that 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. It does not tell us how to map data relations to auditory relations.

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 [Bar97a] has 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.

- 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 signified in a graphic are resemblance, order and proportion, transcribed by visual variables that have signifying characteristics shown in Table 2-3 [Sto00].

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 (with the granularity)	Selective	constant(associative)
Color			Selective	Constant(associative)
Orientation			Selective	Constant(associative)
Shape				Constant(associative)

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

This is the definition of the properties:

- Quantitative: in quantitative perception, the viewer must determine the amount of difference between two ordered values. With a quantitative variable, the user does not need to refer to an index or key to determine how much more of quantity is represented by a given mark.
- Ordered: in ordered perception, the viewer must determine the relative ordering of values along a perceptual dimension. Given any two visual elements, a natural ordering must be clearly apparent.
- Differential: in selective perception the viewer attempts to isolate all instances of a given category and perceptually groups them into a single image. A visual variable is selective, only if the grouping is immediate and effortless.
- Visibility: can be constant or variable. To give an example, size can limit the visibility of a display.

A representation is designed by mapping data relations into 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 [Bar97a].

Scaletti [Bar97a] gives a similar approach in sound. He proposes an appropriate synthesis technique based on data characteristics. It is 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

Flowers, Buhman and Turnage [FWG01] have implemented an application based on the sonification of daily weather. They have evaluated the effectiveness of a simple pitch variation over time as a means to represent abstract data relations. The sonification of daily weather observations provides an example of representing multivariate time series data by sound using the timbre feature. A general scheme could be applied to a variety of other time series data. 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 sonification has potential to introduce auditory display principles to a wide audience.

Kendall [Bar97a] 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. This method would focus on a faithful mapping between auditory relations and data relations.

Madhyastha [Mad90] has proposed a mapping where data sample corresponds to a single note or chord. But he adds that is only a simply suggestion that can inspire more experimentation. A mapping that uses the entire musical phrase is possible. For example, tempo or pitch can change with the data while another sonification is being played in the foreground. Furthermore, Madhyastha and Reed [Bar97a] have proposed a method that 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. 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. Same, the auditory display may be perceived as an annoying noise if the information in it is not useful.

2.6 Task-oriented approach

Wittgenstein [Bar97a] observed that meaning of a sign is not the object it signifies, but the way it is used. Signs can be used in many ways to generate families of meanings. Usefulness is one of the criteria often used to evaluate an interface design.

Kramer [Bar97a] suggests that two broad types of tasks are important in auditory display

- 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.

Frysinger [Bar97a]
Frysinger recognized the need to design sounds that provide information relevant to a task in his proposal of a taxonomy of tasks and data types as a foundation for choosing auditory representations. 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 parameterized 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 analyzing trends in a data set.

Erickson [Bar97a]
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. 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 analyzes.

ESM method [BWE94]
The method is called the Event, Status, Mode method (ESM) and is a semi-formal method for graphical user interface (GUI) design that was extended by Brewster to integrate auditory and graphical information at the interface.

An event is a discrete point in time. The user (e.g. a mouse click) or the system (e.g. mail arriving) can initiate it. Events depend on context, for example clicking the mouse in one window may select an icon, and while in another it may position a cursor. Events can be inputs events (mouse clicks) or outputs events (a beep indicating an error). Events can be hidden because the system does not display them, or because the user does not perceive them.

Status is the information about the state of the system that is perceivable by the user. 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. Status information can be rendered in different ways, graphically, through sound or a combination of both.

A mode is a system context that alters the interpretation placed on events. In one mode typed characters may appear on the screen, while in another they may be interpreted as commands. Mode errors occur when the status information is hidden. Once the event, status and mode information has been described four dimensions of feedback must characterize it.

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 key pressed that has activated a 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 while 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 2-5: Brewster's ESM method

This informal, structured technique can be used to investigate problems with interactions what might be causing the mistakes. 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. From this analysis we can summarize 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. The mental workload required to perform the task was significantly less when sound was used and overall preference was for the auditory scrollbar. All these results indicate the addition of sound was successful and the ESM model proven to be effective.

Alty [Alt93]
He lists reasons why multimedia interfaces may be advantageous for control applications.

- Telepresence - Multimedia options allow us to regain the natural link between the operator and observable. Avoids events by proxy. Preserves implicit cues. Video and audio enable us to some extent to be present at the scene of operations, we are watching and hearing.
- 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 characterizing knowledge. Need a characterization of Media and their knowledge carrying capabilities. Need to match the two.
- Complexity - Alty's group have tried to answer the question 'Does a multimedia approach help in understanding a complex situation.' They found that it was subject-dependent and user-dependent.
- Apparent Redundancy - Humans prefer redundancy of information even if redundant information is useful but not strictly necessary. Users like to have different forms of presentation to choose from. Also it is useful when information quality deteriorates.
- Operator Choice - Operators have hidden goals and implicit knowledge. It is important to let the user choose the information medium as much as possible.
- Intrusion - Some media are more powerful than others are. Sound can be used to break in when there is danger or an emergency, or to reinforce a point in learning. A switch of medium can have the same effect. The user's attention has to be refocused after the intrusion.
- 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).
- Synchronization - Media when used together and synchronized are very powerful, e.g. Video and sound or animations with synchronized text or sound overlay work well.

Alty tested the validity of these principles in an experiment that compared the performance of subjects using different display media. The task was to control Crossman's Waterbath that involves balancing in-flow, out-flow and heater temperature to prevent error conditions. Combinations of graphics, text, speech, and sound showed information about flow rates, temperature, and water level. The results show that differences in the display media 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 [Bar97a] 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 characterized?

The goal of the user is described in a process flow diagram that allows flexible connections and iterations between tasks. There are four types of tasks.

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 stabilized. Activities include Identify, Compare, Derive, Guess, Reason.
Predicting	Identifying potential consequences of plant deviations in order to prevent them. Activities include Model, Simulate, Run.
Controlling	Direct impact on the operations of the processing system. Activities include Record, Create, Delete, Edit, Alter, Enter, Move, Load, Save.

Table 2-6: Task types

The knowledge required by each task is characterized in terms of three 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 organization 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 2-7.

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	10 to 300

	values	
Ordering	Does the variable have an ordering	Ascending
Units	The units of measurement	Celsius
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 2-7: Knowledge characterization

The Knowledge characterization is matched against a characterization of a display medium, shown in Table 2-8. This characterization is similar to Bertin's visual variables, but the concrete characterization 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 2-8: Media characterization

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. Alty comments that the impact of the multimedia interface was not as great as it could have been because 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. 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 users receive it, and may be especially important in commercial applications. Positive connotations may encourage users to experiment and learn the display. Different examples could be cited to show the effects of sonification on interface or environment. It seems that people reacted more favorably to an interface when the sounds were switched on, and that they were often influenced by the sounds even when they said they had not noticed them.

2.8 Device approach

A major problem in audio applications is the variation in the characteristics of audio devices (each device behaves in subtly different ways). These characteristics include the parameters for adjusting overall volume, equalization, 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 [Tka96] describes an approach for balancing the sounds in computer games across various FM and Wave Table synthesizers on different devices and operating systems. He had to manually edit and evaluate the global controls on each target system (define the appropriate settings for each device driver), 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.

2.9 Conclusion

Methods using sound can be classified in six different approaches that are the syntactic approach, the semantic approach, the pragmatic approach, the perceptual approach, the task oriented approach, the connotation approach and the device approach.

Clearly there are many ways to go about designing sounds, and many factors to consider. A comprehensive method will need to include the psychoacoustic properties of the signifier, the perceptual relations between signifiers, the organization of signifiers into structures, the learnability of the signified, signification through use, connotations and social values, the need for transportability and reproduction.

Chapter 3

The TaDa Method

3.1 Introduction

The previous chapter describes a lot of principles, methods that help to understand what sound is and how the listeners react to it. This chapter tries to introduce to a formal method that will help the user to find a model to transform data into sounds. We present an overview of a formal method called the TaDa Method. The next chapters will describe in more details the TaDa method and its different tools.

3.2 Overview of the TaDa Method

TaDa stands for a Task-oriented Data-sensitive method for auditory information design. The TaDa method is divided in different parts: the Task analysis (Ta) and Data characterization (Da) specify the information requirements (Irequire) and the 'Person' and 'Display' parts specify the Information representation (Irepresent). The 'Person' part point to the need to consider how people hear information in sounds, and the way the display device influences the range of available sounds for the 'Display part' [Bar97g].

The method has four phases:

1.scenario description
2.requirements analysis
3.representation design, and
4.realization.

The **scenario description** is a short story about the information processing activity that the display is being designed to support. The story describes the purpose of the activity, and the interaction and organization of information elements. These key features are extracted by recasting the story as a question, in accordance with the observation that "useful information is the answer to a question".

A task analysis of the question, an information analysis of answers to that question and a data characterization of the elements involved derive the requirements. These requirements specify a **representation** that is useful to the task and true to the data. Computer-aided design tools have been built to support the TaDa method through

- Case-based synthesis
- Rule-based mapping
- Interactive refinement.

In **case-based synthesis**, the requirements are used to look-up examples from the Earbenders database of 150 stories about everyday listening. A similarity match on TaDa analysis fields can retrieve the cases from a database. The information requirements of a design problem can be used to search this case-based for everyday examples that share a similar task, data and information structure with the problem. An auditory characterization is needed to describe the sounds in each story, and provide a footing for auditory design. However, the small number of examples of auditory display limits the potential for case-based design, the observation that people use sounds in everyday activities opens the door to an alternative source of examples. The stories capture the way sounds are useful, the ways that people hear information in sounds, and the organization of the sounds. The stories contain semantic, syntactic and pragmatic elements that are important for designing symbols. [Bar97c] [Bar97f] [Bar97g]

Advantages of case-based design [Bar97c]**:**

Cases are **top down**: keyword searches with the key features of the problem can be used to retrieve other stories about that domain. Most real problems are large and complex, and the existence 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 synthesize a new design. The cases that are retrieved can be a source of useful metaphors. 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 do not normally make sounds, so are likely candidates for a metaphorical design.

A case-based 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 that 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 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. A pattern method can be set. 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. 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.

Computer tools can assist a case-based method by making it easy to store and retrieve design cases from a database. The retrieval of relevant cases depends critically on a representation that captures features of the cases that are important in a solution.

The **rule-based** is used to look-up a principled mapping of information relations to acoustic relations. Sarah Bly [KWT+97] drew attention to the lack of a systematic approach for designing the data-to-sound mapping as a "gaping hole" impeding progress in this field of design practice. A taxonomy of mappings has been generalized from the literature of psychoacoustics and data visualization. These mappings can improve the veridical perception of information in an auditory representation. A similar rule-based has been developed for color visualization. [Bar97f] [Bar97g] [Bar97d]

Advantages of design principles [Bar97d]:

Design principles can make specialist expertise easier to learn and apply in practice. Some advantages of design principles are listed below.

Making expert knowledge accessible: Principles, guidelines and rules are a way to capture and formalize knowledge. This knowledge can then be used to produce effective designs, without the need for every designer to have in-depth expertise in every aspect of the design problem, or have to design from first principles every time.

Generality: Principles are general observations that have many uses. For example principles of aerodynamics influence the design of bicycles, windmills, bridges and space shuttles.

Constrained guidance: Principles can help you to home in on a good solution more quickly. A rule system can detect contradictions and exceptions during the design process. However there is the problem that formal methods may impair innovation by forcing all designs into the same mould.

Computer assistance: Rules can be programmed into a computer to deduce the outcome of a design decision, and support interactive simulations and feedback about a design.

Once the representation has been perceptually organized, the designer may listen and make refinements (**Interactive refinement**).

3.3 Conclusion

This short introduction has begun with the presentation of the TaDa Method. The TaDa method allows to specify the Information requirements and the Information representation for a specific problem. The chapter 4 describes in more details the requirements analysis that is the base of the TaDa method. Two interesting computer-aided design tools support the TaDa method.

The first one is the Case-based synthesis where the requirements specified in the TaDa analysis are used to look-up examples from a database called Earbenders.

The second one is the Rule-based mappings that will help to think in terms of principles and these principles can be the source to develop a mapping from data into audio. The chapter 5 describes the principles for auditory information design.

As seen the two methods support a list of advantages that we have to keep in mind.

Chapter 4

TaDa: task and data analysis of information requirements and auditory characterization

4.1 Introduction

This section describes the scenario description, the Tada analysis and the auditory characterization. It forms the Tada method and it can be used in the Earbenders database. The method is mentioned by S. Barrass [Bar97b] [Bar96j]. This chapter describes the methods used to deliver the information requirements. The information requirements of a design problem are used in the Earbenders database to search for everyday examples that share a similar task, data and information structure with the problem. After having introduced the scenario as a technique for capturing key features, the task, the information and the data analysis are used to decompose the problem and furnished the requirements that give details for the different parts of the analysis. Also, a description of the sound forms the base of the auditory characterization [Bar97c] [Bar96j] and can accompany the requirements. Once the requirements have been done, they can be stored in the Earbenders database.

4.2 Describing the problem with a story

The method of problem description adopted 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.

4.3 A bridge from story to requirements

Once the problem has been described it can be analyzed 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 analyzes may be carried out with the same problem description. The observation that "useful information is the reply to a question" 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. [Bar97b]

4.4. Requirements

The TaDa requirements analysis draws on task analysis and data characterization methods that have been established in Human Computer Interaction (HCI) and visualization. There are three areas of analysis - Task, Information and Data. Each area analyzes a different key feature from the Scenario Description. The Task section analyzes the Question key, the Information section analyzes the Answers key, and the Data section analyzes the Subject key.

4.4.1 Task analysis

Task analysis is an established technique in Human Computer Interaction (HCI) and visualization design, and the fields proposed here are an amalgam of components borrowed from several different analyzes. The authors cited in [Bar97b] are Wurman R.S., Norman D.A., Robertson P.K., Kaplan B. and MacCuish J., Rogowitz B.E. and Treinish L.A., Lewis C. and Rieman J. (1994), Alty J. These components have been selected for their relevance to designing information in 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 [Bar97b] 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 that requires local information is about a single element. A question that 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.

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 organized ? 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 of a task is identified in most task analyzes. A set of Purposes was obtained by interpreting the purposes of the generic questions, as shown in Table 4.2. The ten purposes proposed here are an amalgam of the search and compute types that have been defined in graphic display and interactive confirmation, navigation, and alert types found in interface design methods. Several extra Purposes were added after analyzing the uses made of sounds in everyday listening experiences. 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 misdialled my parents phone number by a change in the familiar tone-dialing 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 organized ? 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 ?	alert	highlight or draw attention

has anything changed ?		to an element or subset
	relax	mask unwanted noise, de-emphasize 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 to say 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 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, and be ongoing or very short.

Discrete/procedural	Tasks are initiated by a single event, linear, defined closure, short, seldom overlapping
Continuous/tracking	Tasks require constant monitoring, constant feedback is used to iteratively make refinements, often overlapping other tasks, undefined closure.
Branching/decision	Tasks affect the subsequent course of action

Table 4.4: Task event type

Style

Researchers in both sonification and visualization have recognized two quite different styles of information processing tasks. The first is the exploration of data sets for interesting and has yet unknown features, which requires a faithful, veridical or isomorphic representation that preserves structural relationships. The second is the presentation of known features that may involve an intentional transformation of structure to draw attention or highlight or exaggerate.

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

4.4.2. Information

The Answers to the Question key contain the information needed to carry out the activity. A characterization 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	related to a single element
Information	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 that are not readily swayed by the opinions of others. 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 is 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. In data analysis the data relations are characterized as 4 main types - nominal, ordinal, interval and ratio. 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)
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 answers to the Generic Question contain the range. The number of answers can have a significant effect on the design of a display.

Organization
The information in a display is contained in the interplay of relations between the elements. These relations can be organized in a variety of ways.

Category	Pertains to organization of goods or types. Category can mean different models, types or questions. This mode lends itself to organizing items of similar importance. Category is well reinforced by color 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 does not always have to refer to a geographical site. Doctors use locations in the body as groupings to study medicine.
Alphabet	Lends itself to organizing 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 organization works when another form such as category or location may not.
Continuum	Organizes 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 organization

4.4.3. Data
The Subject key of the Scenario Description identifies the phenomenon of interest. The characterization of this phenomenon can help the designer to 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 that 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 that may be needed in the mapping to the display representation.

Organization

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

Category	organization by difference
Time	organization by time
Location	organization by spatial position
Mnemonic	organization by mnemonic, e.g. alphabet
Continuum	organization by continuous order

Table 4.11: Phenomenal organization

4.4.4. The auditory characterization

The auditory characterization is both a description of sounds in the cases, and a specification of sounds for the design scenario. The Sounds Key that has been added to the Earbenders Scenario Analysis to capture the auditory features of the story heads the characterization. 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 characterization of the sounds, which may assist, in the pragmatic and syntactic aspects of the sound design. There is no all-encompassing way to characterize 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 is {nature, level, streams, occurrence, pattern, movement, type, compound, descriptors}.

Nature

The first characteristic is about the nature of the sound.

Everyday	acoustic events in the physical environment e.g. knocks, scrapes, rumbles, crashes
Musical	musical sounds generated by instruments specialized 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	recognizable words, singing

Table 4.12: Nature of the sound

Level

Analytic and holistic listening are musical terms for different listening styles.

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

Table 4.13: 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 [Hur91] of auditory scene analysis. Streams are perceptual groups that form when sounds occur simultaneously and in sequences, as in everyday listening. 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 4.14: Streams

Occurrence

Sounds can be one-off or ongoing, just as tasks may be discrete or continuous. It takes at least 4 seconds for the primitive stream grouping process to stabilize, and once an interpretation of the number of sources has occurred it does not matter if one or other of them briefly disappears for a second or two.

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

Table 4.15: Occurrence

Pattern
Sounds can vary in many ways over time.

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 4.16: Pattern

Movement
The movement of a sound is relative to the listener.

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 4.17: Movement

Type
The type is a characterization 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 organization. Continuous is divided in metathetic and prothetic types. 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 organizations 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 4.18: Type

Compound
Sometimes it is possible to separate aspects of the sound from its overall identity and other sounds are more integral and it can be hard to hear them as anything but a whole.

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 4.19: Compound

Descriptors
The descriptors are a list of the words used to describe the sounds. Shared descriptors may indicate consistent variation or similarity relations. Multi-word descriptions may indicate the separability and attention to particular aspects in the sounds.

4.5 Conclusion

After having described the problem with a story, the key features of the requirements analysis are furnished by the triplet (Question, Answers, Subject).

The Task section analyzes the Question key, the Information section analyzes the Answers key and the Data Section analyzes the Subject key. The auditory characterization is a description of sounds and a specification for the design scenario.

The Task analysis is articulated around five fields {Generic question, Purpose, Mode, Type, Style}. The Information analysis is articulated around five fields {Level, Reading, Type, Range, Organization}. The Data analysis is articulated around three fields {Type, Range, Organization}. The auditory characterization is articulated around nine fields {Nature, Level, Streams, Occurrence, Pattern, Movement, Type, Compound, Descriptors}

All these requirements represent the heart of the TaDa method and can be presented in an Access^{TM} database that is the Earbenders database. This database is a tool for a case-based method that allows to store and retrieve some knowledge about a specific problem.

You will find a specification using the TaDa analysis in the GeoViewer example in the chapter 7. The Java application found in the chapter 8 is also described with the TaDa analysis.

Chapter 5

Principles for auditory information design

5.1 Introduction

This section gives the principles for the rule-based method and is a part of Stephen Barrass's ideas [Bar97d] [Barh]. The five fields from the information requirements in the TaDa method {Reading, Type, Level, Organization, Range} can be used in helpful principles. With this method, the designer meets the information requirements specified by the TaDa analysis. This method gives an approach for designing the data-to-sound mappings. The principles have been investigated by generating a simple auditory demonstration to confirm that characteristic properties can be heard. All the sounds and examples shown in [Bar97d] to test the principles were made with Csound [Cso] and can be tested in [Bar97g]. The principles are also called Hearsay principles.

5.2 Principles for information design

This part develops some principles for designing useful sounds. The TaDa analysis characterizes the information required from the display. The information characteristics are the starting point for design. Five fields: reading, type, level, organization and range characterize the information. These fields can serve as anchor points for principles that couple the requirements to the representation. In the quest for a principled approach to auditory display we can look to methods of graphic display that involve similar issues of representation. There has been a great deal of effort put into understanding how graphs can best show different types of information. This effort has resulted in the development of principles for graphic information design that have been broadly applied and found to be effective in practice. This approach to design has progressed to the point where rule-based computer tools can automatically construct a display from descriptions of the task and the data.

The principles of information display developed for graphic design may also be helpful in auditory information design. Some principles that have been consistently identified, and broadly applied, are linked with each of the TaDa information characteristics in Table 5.1.

Reading	The most direct representation is the one with the shortest psychological description
Type	An appropriate representation provides the information required by the task: neither more nor less
Level	The power of a graphical display is that it allows us to summarize general behavior and at the same time to examine details
Organization	Useful information involves regrouping. The interactive reorganization of relations between elements can uncover information in the interplay of the data

Range	Any undetectable element is useless. Utilize the entire range of variation

Table 5.1: Some principles for information design

5.3 Principle of reading or directness

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. 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. These symbols are slow to read (several per second), and people can only keep about seven discrete items in short term memory, which may limit the operations that can be performed.

5.4 Principle of type

If the task requires qualitative information then use a qualitative representation. If the task requires quantitative information then use a quantitative representation. For example coloring the countries in qualitatively different hues can appropriately support the task of finding a country on a globe. If the task is to find the country with the highest rainfall then hues would make this difficult because large differences in hue do not look ordered and cannot be compared. The design of an appropriate representation requires a description of the information to be represented.

If we look at the table 4.8, we can see that the information types are characterized by elementary relations of difference, order, metric and zero that are the building blocks of more complex information structures. Order is a directed difference, which might be expressed as more or less, or low and high. Metric is an equal unit of difference that is consistent, for example a 1-degree rise in temperature is the same no matter what the current temperature is. Zero is a point of correspondence between all scales independent of unit, so for example zero rainfall is the same whether your rain gauge is in mm or inches.

Difference	qualitative or quantitative
Order	directed difference
Metric	equal units of difference throughout the range of variation
Zero	a point of correspondence between all scales independent of unit

Table 5.2: Elementary information relations

The information building blocks can be aligned with perceptual building blocks that have similar properties.

Difference	all perceptual elements are detectably different
Order	the perceptual elements have a discernable order
Metric	there is a unit of equal perceptual difference
Zero	an absolute point of reference for variations at any scale

Table 5.3: Elementary perceptual relations

The following examples demonstrate that the perceptual building blocks can be heard in auditory relations, as required by this process. See in [Bar97g] to hear the different sounds.

Difference

Plugging in parameters to a synthesis instrument can easily generate qualitatively different sounds. The demonstration is a FM instrument and a score that generates a sequence of three different sounds. You can hear that the sounds are different by listening to the sequence in a loop. When the sounds are not very different you can hear a double sound in the loop.

Order

Changes in a sound are sometimes described by words like buzziness, or squelchiness or heaviness that indicate a degree of order in the sounds. The demonstration is a vibrato at three different rates. When you listen to the sequence in a loop it is easy to hear an ordered change in the sound.

Metric

A metric variation has a unit of equal perceptual difference. This can be heard by a unit step in difference no matter where the step occurs. Examples of units that are available are semitones and decibels. The example demonstrates equal steps in pitch. The size of the steps can be heard by listening to the sequence in a loop. The difference between the middle sound and those on either side should seem equal.

Zero

A perceptual zero can be detected no matter where it occurs in a sequence, and no matter what the scale of variation. There are three types of zero that can be listened for

- a natural zero where the sound disappears altogether
- an original zero where an observable aspect of the sound disappears
- a conventional zero, such as middle c, which may be compared against a reference, or perhaps learnt

The original zero can be demonstrated with the vibrato instrument. The sound is played in three times: a sound is produced supplied by a vibrato effect, the vibrato disappears although the sound remains (the sound is played without the vibrato), the vibrato reappears.

An elementary characterization of some sounds

There are a many auditory variations that we might harness in an auditory display.
These include everyday sounds, musical sounds, synthetic sounds, vocal sounds, and verbal sounds. Some sounds of each of these types have been characterized in terms of the elementary relations of difference, order, metric and zero. The table shows how some sounds can be described in these terms, but is not meant to be definitive or complete. The characterization of sounds in this way can help select an appropriate representation for a display element.

Sound relation	Difference qualitative/ quantitative	Order 1D, 2D, 3D, ND	Metric ratio/ difference unit	Zero natural original conventional
Door knocks	Qualitative	-	-	-
object material	Qualitative	-	-	-
event type	Qualitative	-	-	-
Rhythm	Qualitative	-	-	-
harmonicity	Qualitative	-	-	-
Tune	Qualitative	-	-	-
musical key	Qualitative	-	-	-
phasor pattern	Qualitative	-	-	-
binaural cohesion	Qualitative	-	-	-
temporal order hiss, tone, buzz, `ee'	Qualitative	-	-	-
vowels a,e,i,o,u	Qualitative	-	-	-
animals moo, woof, meow, baa	Qualitative	-	-	-
Formants	Qualitative	ND	-	-
Timbre	Qualitative	ND	Difference MDS	-
Squeakiness	Qualitative	1D	-	Natural
Flapping	Qualitative	1D	-	Natural
Popcorn Popping	Qualitative	1D	-	Natural
Music tempo	Qualitative	1D	-	Natural
Machine rate	Qualitative	1D	-	Natural
Machine work	Qualitative	1D	-	Conventional
Pitch class	Qualitative	1D	Difference Semitone	Conventional
Event force	Quantitative	1D	-	Natural
Drum stretch	Quantitative	1D	-	Original
Fuzz level	Quantitative	1D	-	Original
Reverb wetness	Quantitative	1D	-	Original
Vibrato depth	Quantitative	1D	-	Original
Vibrato rate	Quantitative	1D	-	Original
Tremolo rate	Quantitative	1D	-	Original
Tremolo depth	Quantitative	1D	-	Original
Phasor depth	Quantitative	1D	-	Original
Phasor rate	Quantitative	1D	-	Original
Brightness	Quantitative	1D	Ratio Acum	Original
Object size	Quantitative	1D	-	Conventional

Filling a bottle	Quantitative	1D	-	Conventional
Rolling marble	Quantitative	2D	-	Conventional
Granular density	Quantitative	1D	-	Conventional
Pitch scale	Quantitative	1D	Difference Semitone ratio Mel	Conventional
Repetition rate	Quantitative	1D	Ratio B =1.0	natural
White noise duration	Quantitative	1D	Ratio B =1.1	natural
Binaural loudness	Quantitative	1D	Ratio B =0.6	natural
Monaural loudness	Quantitative	1D	Ratio B =0.54	natural

Table 5.4: Elementary characterization of some sounds

5.5 Principle of level

Higher level information is contained in the groupings, clusters, trends, correlations, outliers and other relations between data elements. The level of question that can be immediately answered from the display can determine the level of the display.

Local	Can answer questions about a single element
Intermediate	Can answer questions about subsets and groups of elements
Global	Can answer questions about the entire set of elements as a whole

Table 5.5: Levels of information

According to Bregman's theory [Hur91], there are two levels of listening processes- a global level of overall analysis, and a local level of attention to details. These processes group and segregate acoustic elements into coherent sounds or “streams”. Levels of information may be linked with levels of auditory scene analysis.

Local	Answered by listening to an element within a stream
Intermediate	Answered by listening to a stream
Global	Answered by listening to an auditory scene

Table 5.6: Levels of auditory information

A simple experiment can show the importance of the segregation. A looped sequence of three different sounds is played at the slow rate of 1 sound per second. We can easily write down the order of the three sounds. Now speed it up to 10 sounds per second by changing the tempo from 60 beats per minute to 600 beats per minute. If we play the loop and try to write down the order again - this time it will be very difficult to tell which sound comes after which. This is because the sounds have segregated into different auditory streams. The segregation of elements into streams can make simple tasks like counting much harder. Some consequences of streaming for auditory display are

- streams are categorical and exclusive
- judgements involving elements in the same stream are easy
- judgements involving elements in different streams are difficult

An understanding of the factors that influence streams can guide the design of a higher level display. As mentioned earlier, there is a global level and a local level. The global level, or primitive process, is a default bottom-up grouping by acoustic factors such as local can answer questions about a single element intermediate can answer questions about subsets and groups of elements global can answer questions about the entire set of elements as a whole spectral similarity. The local level, or schema process, allows the listener to alter the default grouping by mental effort. Mental schemas detect familiar acoustic patterns and draw attention to them. Schemas are a top-down process that explains why what we hear depends so much on attention and previous listening experience. In this view, the characteristics of sounds are calculated from streams, not directly from the acoustic array. This is very different from a straightforward signal-processing model of auditory perception.

Primitive grouping in auditory displays

The factors that influence the primitive process group operate sequentially in time, and simultaneously across the spectrum. The perception of a new sound depends on the streams that exist when it is introduced. Parts that are acoustically similar to an existing stream will be grouped with it, leaving the residue to be heard as the new sound. A listing of factors in order of influence can be made from results of experiments that have placed various factors in competition. The sequential factors are toward the top of the list. This listing may provide a basis for controlling the primitive grouping of elements in a higher level auditory display.

- the difference between spectral centroids
- difference in fundamental frequency in the range 4-13 semitones
- binaural harmonic correlation
- correlated frequency modulations
- correlated amplitude modulations
- harmonic relations
- parallel spectral movement
- synchronous onsets (part of a syllable before the vowel ex.: 2 letters starts: blob, blot,...)

Schemas in auditory displays

Mental schemas detect familiar patterns and draw attention to them. They are important in auditory design because attention and previous learning have a marked influence on what is heard. We can take advantage of familiar patterns to improve the detection of information elements, and to improve the coherence of information in a mixture. The semantics of familiar sounds can also be used to improve the interpretation of the display in a particular task - for example rain sounds can be easily related to rainfall records. Some consequences of schemas in auditory display include

- improved coherence and separation of figure from ground
- the selection of streams and material from streams
- recognition of familiar patterns
- restoration of hidden material

The effects of a schema can be demonstrated by the restoration of a damaged tune. Despite all the noise and interference, the tune can be identified.

5.6 Principle of organization

Useful information involves regrouping. Bertin [Bar97d] demonstrated the way regrouping alters the information in a display by moving cards with simple marks, such as spots of different sizes, around on a table. In one example he transcribes the length of stay of hotel guests on these cards then physically reorganizes them on the table to show different information about peak booking periods which was not evident in the original graph. The useful information does not correspond with the values of the individual elements but with the structures formed by the interplay of these elements with each other as a whole. Only the spatial organization of the elements was permuted, not any of the other visual variables such as lightness or size. This is because you can only see two distinct cards if they have different positions on the table, or are in the same place at different times. This is why space and time are called the ``indispensable" dimensions of a visual display. Elements that use-up an indispensable dimension constrict the options for permutation. For example a time-series plot uses-up the horizontal dimension, leaving only the vertical for permutation. A map uses-up both the horizontal and vertical dimensions and so cannot be permuted. Streaming experiments have shown two sounds can occupy the same space and time but still be heard as separate identities when they occupy different parts of the spectrum. It seems that the auditory display designer has a great deal of freedom to organize display elements. The degree of freedom depends on the capabilities of the display device that may employ multiple synthesis parameters to reorganize spectral relations. Real-time input sensors support interactive permutation and exploration. A limiting factor is the amount of computation required to generate the sounds in real time. Any apparent lag in reaction can compromise the usability of an interactive display. One way to address this problem is to design the synthesis to be as computationally simple as possible. Another way is to take advantage of fast hardware for audio synthesis.

5.7 Principle of range

The number of elements that can be differentiated in a display depends on the range of perceptual variation available on the display. Most people cannot hear the pitch of frequencies below about 80 Hz in which case human hearing is the limiting factor. Some devices cannot play frequencies above 4 kHz, in which case the device is the limiting factor. The range of perceptual variation on a device is called the display gamut. The knowledge of a gamut allows the designer to optimize the display for the device. A transportable display must be designed to fit in the intersection of the gamuts of the target devices. The orchestra and score in example demonstrate the effect of available range. The sequence is four levels of rainfall {none, light, medium, heavy} mapped to loudnesses (0, 40,60,80} dB. This demonstration assumes that you can easily change the loudness setting of your audio equipment. Generate the sequence and turn the loudness knob down low to avoid the risk of an uncomfortably loud sound. A display that relies on loudness will need to be calibrated to ensure that all the elements are discriminable.

5.8 Conclusion

In conclusion, using the principles seen in this section can develop useful sounds. The principles are articulated around five fields that are reading, type, level, organization and range. These five fields characterizes the Information requirements as seen in the TaDa Analysis and are the starting point for design and useful principles.

Reading will focus on the most direct representation, Type on an appropriate representation neither more nor less, Level will allow to give information at a general level and at a detail level, Organization will involve regrouping and Range will tend to utilize the entire range of variation.

The integration of these principles may provide a systematic and useful approach for designing sounds.

The Java application found in chapter 8 has tried to use these principles.

Chapter 6

Information-Perception Space

6.1 Introduction

The simple categorical distinctions of sound events can potentially be exploited in auditory presentations to communicate important distinctions in the data. The goal is that perceptually continuous auditory attributes are scaled and mapped to data attributes in a way that is meaningful to the observer.
The Hearsay principles summarize 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. That is why an alternative representation of the Hearsay principles in the form of an Information-Sound Space is presented according to the experiments made by Stephen Barrass in [Bar97e] and [Bar94k]. The ISS is a three dimensional spatial organization of auditory relations that bridges the gap from theory to practice by changing the way a 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. Written rules are not the only way that principles can be represented. A more direct and accessible form of representation may better support design practice. The approach used is similar to the HSS model (Hue, Saturation and Skill) that is commonly used to assist in the choice of color scheme. It gives a motivation for a similar method for understanding and specifying auditory relations. The Information-Perception Space makes the Hearsay principles more direct and easy to apply in auditory design practice.

6.2 Description of the Information-Perception Space

The Information-Perception Space is a cylindrical polar organization that has a cyclic dimension of 8 categories, a radial of 8 equal, ordered differences, and a vertical axle of 100 equal, ordered differences. Each principle is addressed by the organization of the IPS as follows:

Reading
The IPS focuses the design on direct perceptual relations between elements (timbre, pitch,...) that is a very different from the usual design of conventional symbols that must be learnt and read from the display. It can also help to select conventional symbols with prescribed perceptual properties. An example might be to select some symbols, which will be perceived as distinctly separate in a conventional display, by selecting perceptual points that are equally spaced around the categorical circle.

Type

The combination of a circle of globally unordered difference, an axis with order, an axis with a metric and a zero covers all of the elementary perceptual relations {difference, order, metric, zero}. This enables the IPS to support the TaDa information types {boolean, nominal, ordinal, ordinal-with-zero, ordinal-bilateral, interval, ratio} that have been defined in terms of these relations.

The pedestal of categories

The other properties of the IPS rest upon the pedestal of categories. The circle can be divided into 8 regular categories to accord with the limits of short-term memory. The equally spaced pedestal has the following properties

- differences do not have an observable order
- adjacent points are subjectively equally different
- the circle has a conventional zero, so that cycles are perceptually seamless

The disc of radial spokes

The disc of radial spokes is an extension of the pedestal of categories by a radial variation within each category. It is important that this radial component does not cause a perceptual change in category. In the color solid the radial saturation of a hue can vary without the hue changing, for example pink can change to red. The original zero is an anchor for absolute judgments along any radius. For example Grey is a zero in the color space, independent of hue or lightness. The radial component has the following characteristics:

- observable variation throughout each category
- a perceptual metric
- an original zero which is a common point of origin, independent of category

The vertical axle

Transfixing the disc of radial spokes on a vertical axle completes the IPS. The variation in this dimension must be observable and ordered everywhere in the space. This dimension has a natural zero, which marks the absence of a perceptual element. The axle contains all the original zero points of the radial scales. In color models it is sometimes called the "Grey" axis because all the desaturated points lie along it, stretching from the dark point to the light point.

The vertical axle has the following characteristics:

- observable variation throughout each category
- perceptual independence from the radial dimension
- a perceptual metric
- a natural zero for absolute judgements

Level

The possibility to combine the three different types of perceptual axes into an orthogonal basis provides the opportunity to construct bivariate and trivariate representations. The axes are scaled so that Euclidean distance corresponds with perceptual difference. To be a truly uniform the scaling needs to account for the perceptual interactions between dimensions - for example the color space is scaled by just noticeable differences in each dimension at each point.

A factor not included in the color space is the control of perceptual grouping. This capability is important for designing higher level intermediate and global information displays that depend on perceptions of grouping and segregation. The selection of perceptual attributes for each axis may be made from factors that have an influence on grouping. The categorical factor should be particularly strong to maintain cohesiveness while other aspects vary.

Organization
Audio and video sequencers are common tools for organizing sounds and pictures in time. These aspects are not a part of the IPS but may be organized with another tool, called Personify [Bar95i] [Por].

Range
The dynamic range of each perceptual axis in the IPS constrains how representation schemes can use the space. The range of each dimension can be set in many ways, but the color solid is the framework we are starting with, and so we will set the ranges accordingly.

The Information-Perception Space is shown in the following figure.

- the pedestal has 8 categories
- the radial spokes have 8 steps
- the vertical axle has 100 steps

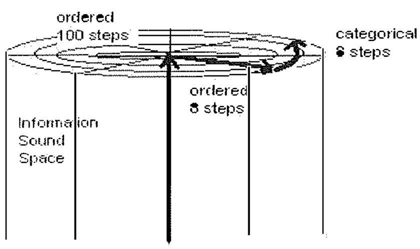


Figure 6.1: The Information-Perception Space

6.3 Representations in IPS

Paths through the space that have particular properties due to the organization of the space can specify the design of representations in the IPS. The angular axis is qualitative difference, the radial axis is quantitative difference with order, metric and a zero, and the vertical axis is quantitative difference with order and a metric. The elementary information relations in the IPS are listed in Table 6.1.

Elementary relation/ ISS dimension	Difference	Order	Metric	Zero
Angle	Qualitative		8 steps	Conventional
Radius	Quantitative	Yes	8 steps	Natural
Vertical	Quantitative	Yes	100 steps	Natural

Table 6.1: elementary information relations in the IPS

A direct representation preserves the characteristics of the information when it is heard in the display. The rules for choosing a direct representation in the IPS are

- Rule 1 - if the Information Type has qualitative difference vary the qualitative Angular axis
- Rule 2 - if the Information Type has quantitative difference vary the quantitative Vertical axis
- Rule 3 - if the Information type has a zero vary the Radial axis starting from the origin

These rules specify paths through the IPS that directly represent different types of information relations, as shown in Table 6.2. As seen in the table 4.8, the information types {Boolean, nominal, ordinal, ordinal and zero, ordinal bilateral, interval, and ratio} have been described with the properties of order, metric, zero and difference. It allows to describe the information types with each dimension.

Info Type	Angle category	Radial zero	Vertical metric	IPS mapping
Boolean	X			Opposite angles
Nominal	X			Circle
Ordinal	X		X	Coil
Ordinal and zero	X	X	X	Spiral
Ordinal bilateral	X	X	X	Sloped Line
Interval			X	Vertical Line
Ratio		X	X	Radial Line

Table 6.2: information relations

6.4 Specialization

Assigning auditory dimensions to each dimension in accordance with the Hearsay principles makes the specialization. The result is a 3 dimensional space of auditory relations that can represent information relations, called the Information-Sound Space (ISS). For each dimension, a pilot study has been made to evaluate some candidates. A set of criteria that must be satisfied to realize a dimension are proposed. A null hypothesis for each criterion is given. A criterion is supported if its null hypothesis is rejected. Listening to auditory sequences having prescribed relations in terms of Zero, Difference, Order, Metric, Level and Range carried out the tests. The tests can be found in [Bar96g] in the ISS page.

The Pedestal criteria

The starting point of the ISS is the pedestal of auditory categories.

Zero - categories do not have a perceptual zero
Null hypothesis: the sequence has a perceptually singular point.
Zero sequence: a repeating cycle of points regularly spaced around the circle.
The task is to indicate the start of the sequence in a repeating cycles of the sequence. The null hypothesis is accepted if a starting point is consistently identified.

Difference - each category sounds different
Null hypothesis: two or more elements of the sequence are identical
Difference sequences: complementary and adjacent triplets
The task is to find two points that sound the same. The search through all pairwise comparisons is very large. A limited analysis can be obtained by listening to all complementary and adjacent triplets. Complementary triplets are sets of three elements chosen at equal intervals around the circle. Adjacent triplets are three points at successive positions around the circle. The null hypothesis is supported by the consistent identification of two identical elements in a triplet.

Order - categories are not heard to have a simple order
Null hypothesis: subsets in the sequence have a simple order.
Order sequence: complementary triplets
The listener hears each complementary triplet in a repeating cycle. The null hypothesis is accepted if a repeating triplet is consistently heard to have a simple unidimensional variation.

Metric - the difference between categories is regular
Null hypothesis: adjacent elements do not have regular spacings
Metric sequence: adjacent triplets
The listener hears sets of adjacent triplets from around the circle. The task is to identify the most similar pair in each triplet. The null hypothesis is accepted if there is a consistent pairing that indicates irregular spacing in a triplet.

Level - categories segregate into different auditory streams
Null hypothesis: the categories do not segregate into different streams
Level sequence: adjacent pair in Van-Noorden's XOX-XOX galloping sequence, at rates from 500 ms to 50 ms
The listener hears a pair of categories in a sequence of the form XOX-XOX where X is one category and O is an adjacent category, and - is silence. This is the sequence that Van Noorden used to measure the temporal coherence of sounds. The task is to hold each triplet together as a single unit, and indicate the rate at which this can no longer be done. To reject the hypothesis, the segregation has to be heard. The point of segregation is signaled by a galloping rhythm where the X-X-X-X is heard in one stream and O---O---O is heard in the other. The typical cohesion threshold is between 50 ms for very similar sounds to 150 ms for dissimilar sounds. The null hypothesis is supported if the triplet is very cohesive as indicated by segregation only occurring at fast rates with onsets of less than 100 ms.

Range - there are 8 discriminable steps
Null hypothesis: two of the categories are the same
Range sequence: complementary and adjacent triplets

This test is identical to the difference test. If the difference test fails then there are less than the requisite 8 categories, and a listener may hear two different categories as the same.

The experiment has been made on 4 different pedestals. This is the results:

Pedestal	Zero	Difference	Order	Metric	Level	Range
Pitch	Ok	Ok	Fail(4)	Ok	Fail	Ok
Formant	Ok	Ok	Fail(2)	Ok	Fail	Ok
Static Timbre	Ok	Ok	Fail(4)	Fail(3)	Ok	Ok
Timbre	Ok	Ok	Fail(1)	Ok	Ok	Ok

Table 6.3: Results for the pedestal; The values between parenthesis for the Order and the Metric represent the test that have failed. For the timbre, the test 1 is unordered for example. For the level, the test has failed because one example has show that no segregation could be heard.

Timbre seems the best candidate and has been chose like the pedestal.

The radial spokes

The pedestal provides a platform for the rest of the Information-Sound Space to sit on. The next stage of investigation is to fill in the pedestal with a disc of radial variation. This variation can be thought of as spokes that radiate from the center of the pedestal out to the categorical node on the perimeter of each segment. The pilot study was to investigate the radial spokes. The investigation is framed by the theory that a radial variation can represent quantities for comparison, without altering the perceptual category. The radial dimension has the characteristics of difference, order, metric and an original zero. The spokes can only exist if they can be observed throughout every category in the pedestal. Therefore the radial variation needs to be observable across a general range of timbres, or else highly specialized to the particular timbres in the pedestal. The straightness of the spokes relies on the selection of an aspect of variation that does not cause a change in category as it traverses its range. The choice is constrained by the common point of origin and the need to support smooth transitions through the origin.

Criteria of the radial spokes

A set of criteria that must be satisfied to realize the radial spokes of a Timbre Disc are proposed.

Zero - radial zero is a common point of origin across categories.
Null hypothesis: the radial zeros in each category are not similar
Zero sequence: the zero from each category + random 50% point
The listener hears repeating cycles of the sequence. The task is to identify the most dissimilar point. The null hypothesis is accepted if there is consistent choice of the point with 50% brightness.

Difference - different elements sound different
Null hypothesis: two or more elements of the sequence are identical
Difference sequences: repeating cycles of 8 steps along the radial.
The listener hears repeating cycles of the radial sequence. The task is to listen for level regions or turning points in the sequence. These points are where different

values sound the same. The null hypothesis is supported by the consistent identification of a level point or a turning point.

Order - ordered subsets sound ordered
Null hypothesis: ordered subsets do not sound ordered
Order sequence: repeating cycles of 8 steps along the radial.
The listener hears repeating cycles of the radial sequence. The task is the same as the difference task - to listen for level regions or turning points in the sequence. These regions indicate the possibility that an ordered subset will not be heard as ordered. The null hypothesis is supported by the consistent identification of a level point or a turning point.

Metric - regular intervals sound regular
Null hypothesis: regular intervals along the radius do not sound regular
Metric sequence: ordered triplets with regular spacing 25%, 50%, 75%
The listener hears sets of ordered triplets with regular spacings. The task is to identify the most similar pair in each triplet. The null hypothesis is accepted if there is a consistent pairing that indicates irregular spacing in a triplet.

Level - difference influences sequential grouping strength.
Null hypothesis: radial difference does not influence sequential grouping strength
Level sequence: pairs of tones in the XOX-XOX galloping sequence.
Three test sequences are generated at 100 ms rate. The first has maximum difference, the second is 50% maximum difference and the third is 5% maximum difference. The listener is asked to try to hear the sequences as single repeating sounds. The task is to answer whether there is more than one sound perceived in each sequence. The null hypothesis is supported if the answers to all three tasks are consistent- either they all had only one sound, or all had more than one sound.

Range - there are 8 discriminable steps
Null hypothesis: there are not even 3 discriminable steps
Range sequence: triplet with regular spacing 25%, 50%, 75%
The listener hears a slowly repeating triplet with regular spacings. The task is to identify whether there are one, two or three different sounds. The null hypothesis is accepted if less than 3 different sounds are consistently heard.

Brightness is parameter implemented in Csound [Cso] by linearly adjusting the cut-off frequency of a first orderlow-past filter. To give a difference of brightness, compare Rockmore to Hoffman, and you will hear differences in brightness. Whatever, the transformation made allows to increase the brightness of a sound. With this transform the sound is louder and it is made much more acoustically realistic than increasing the gain.

Results on brightness:

Zero	Difference	Order	Metric	Level	Range
Ok	Fail	Fail	Fail	Ok	Fail

Table 6.4: Results for the radial spokes

The brightness radius failed on the criteria of Order, Difference and Metric. The difference and order criteria may be satisfied by scaling of the brightness variation to ensure equal steps in brightness along the radius. The range criteria could be addressed by providing information about the number of equal steps available at each pitch of each sample. The designer could then ensure the necessary range is available for a particular sound. By this way, the problems of saturation of the brightness in different timbres are solved.

The vertical axle

Fixing the disc of radial spokes on a vertical axle completes the IPS. The variation in this dimension must be observable and ordered everywhere in the space. This axle also contains all the original zero points of the radial scales. In color models it is sometimes called the "Grey" axis because all the desaturated points lie along it stretching from the dark point to the light point. The vertical axle has the following requirements:

- observable separability throughout each category
- perceptual orthogonality to the radial dimension
- a perceptually scaled metric
- a natural zero

Criteria of the vertical axle

Zero - the variation has a natural zero
Null hypothesis: the vertical variation does not have a natural zero.
Zero sequence: repeating cycles of the vertical variation.
The listener hears repeating cycles of a vertical sequence. The task is to identify the point at which the sequence cannot be heard. This is a natural zero that is an absolute anchor point for all vertical sequences. The null hypothesis is accepted if the point of disappearance cannot be consistently identified.

Difference - difference is heard as difference
Null hypothesis: two or more elements of the sequence are identical
Difference sequences: repeating cycles of a vertical sequence.
The listener hears repeating cycles of the vertical sequence. The task is to listen for level regions or turning points that indicate points of repetition. The null hypothesis is supported by the consistent identification of a level point or a turning point.

Order - ordered subsets sound ordered
Null hypothesis: ordered subsets from the vertical do not sound ordered
Order sequence: repeating cycles of a vertical.
The listener hears repeating cycles of the vertical sequence. The task is the same as the difference task - to listen for level regions or turning points in the sequence. These regions indicate the possibility that an ordered subset will not be heard as ordered. The null hypothesis is supported by the consistent identification of a level point or a turning point.

Metric - regular intervals sound regular
Null hypothesis: regular intervals up the axle do not sound regular
Metric sequence: ordered triplets with regular spacing

The listener hears sets of ordered triplets with regular spacings. The task is to identify the most similar pair in each triplet. The null hypothesis is accepted if there is a consistent pairing that indicates irregular spacing in a triplet.

Level - difference influences sequential grouping strength.
Null hypothesis: vertical difference does not influence sequential grouping strength
Level sequence: pairs of tones in the XOX-XOX galloping sequence.
Three test sequences are generated at 100 ms rate. The first has maximum difference, the second is 50% maximum difference and the third is 5% maximum difference. The listener is asked to try to hear the sequences as single repeating sounds. The task is to answer whether there is more than one sound in each sequence. The null hypothesis is supported if the answers to all three tasks are consistent- either they all had only one sound, or all had more than one sound.

Range - there are of the order of 100 discriminable steps
Null hypothesis: there are less than 20 discriminable steps
Metric sequence: A triplet with 5% spacing
The listener hears a slowly repeating triplet with regular spacings. The task is to identify the whether there are one, two or three different sounds. The null hypothesis is accepted if only one or two sounds are consistently heard.

Vertical Axle	Zero	Difference	Order	Metric	Level	Range
Loudness	Ok	Fail	Fail	Fail	Ok	Fail
Duration	Ok	Ok	Ok	Fail	Ok	Fail
Pitch	Fail	Ok	Ok	Ok	Ok	Ok

Table 6.5: Results for the vertical axle

6.5 A prototype of an Information-Sound Space

A prototype can be formed using the candidates tested. This ISS will be organized to have perceptual characteristics of difference, order, metric, zero, level and range that are necessary to represent the TaDa information types.
The Timbre Circle was the most effective of the four Pedestals that were tested. The Timbre Circle consists of a subset of subjectively equally spaced musical instrument timbres. It provides a platform for data mappings that preserve the unordered difference between elements mapped to sounds, so that it may veridically represent categorical data.

The radial axis can be used if we scale the brightness dimension. Three candidates were tested for the vertical axis - loudness, duration and pitch. The final test was of pitch, which does not have a natural zero, but does have a good range. The pitch axis may be an acceptable compromise because the brightness radius, which has an original zero, provides the zero that is a necessary characteristic of the IPS.
The Timbre Circle, Brightness Radius, and Pitch Axle can be combined to form a prototype Information-Sound Space. The TBP model is a polar cylindrical space that revolves around a pedestal of 8 equally different timbres. The radial axis is 8 equal steps in brightness, and the vertical axle is 100 equal steps in pitch. There are other permutations that may also satisfy the ISS criteria, but the TBP basis is an initial point

for further investigation. Note that the motivation for the space is not to describe hearing perception but to support a method of data-sensitive auditory display.

The TBP prototype of an Information-Sound Space :

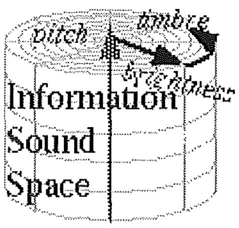


Figure 6.2: The TBP prototype

The TBP sound model was developed to have properties similar to those of the HSL color model. The advantages of the TBP model are:

- Natural specification, comparison and matching - Timbre, Brightness, and Pitch are perceptually separable attributes of sounds.
- Natural order - the Timbre Circle is ordered by an underlying perceptually orthogonal basis, which arranges complementary timbres diametrically opposite each other.
  The Brightness and Pitch axes both have a natural order.
- Independent control of perceptually aligned parameters - Timbre, Brightness, and Pitch can be changed independently.
- Geometric interface - the 3D sound solid provides the opportunity for spatial interaction with sounds.
- Transportability - the TBP model may be used to specify sounds in natural terms rather than device coordinates.

6.6 Representational Mappings in TBP ISS

The various information paths in the Information-Perception Space become auditory representations when they are mapped to the TBP Information-Sound Space. The angular axis is qualitative timbre difference, the radial axis is quantitative brightness difference with order, metric and a zero, and the vertical axis is quantitative pitch difference with order and a metric. The elementary auditory relations in the TBP ISS are shown in Table 6.6.

Info Type	ISS Mapping	TBP description
Boolean	Opposite angles	2 very different timbres
Nominal	Circle	Up to 8 categorically different timbres
Ordinal	Coil	Categorically different timbres ordered by pitch
Ordinal and zero	Spiral	Categorically different timbres, ordered by pitch, with dull zero
Ordinal bilateral	Sloped Line	Dull central zero, -ve and +ve category timbres, ordered by pitch
Interval	Vertical Line	Ordered change in pitch

Ratio	Radial Line	Ordered change in brightness, starting from a dull zero

Table 6.6: auditory relations in the TBP ISS

6.7 Conclusion

The Hearsay principles summarize some knowledge that can be useful in the design process but these principles need to be learnt and understood to be applied. This section has fulfilled the gap between theory and practice by testing some sound features in a three-dimension model. The user can now follow the method and ideas given by the Information-Sound Space (ISS) articulated around three dimensions that are Timbre, Brightness and Pitch. The TBP ISS describes some auditory relations in terms of sound characteristics (timbre, brightness and pitch) according to the information type. The goal is reached because the user does not have to keep in mind the Hearsay principles but he can directly play with the sound characteristics.

Chapter 7

A Complete example: the GeoViewer

7.1 Introduction

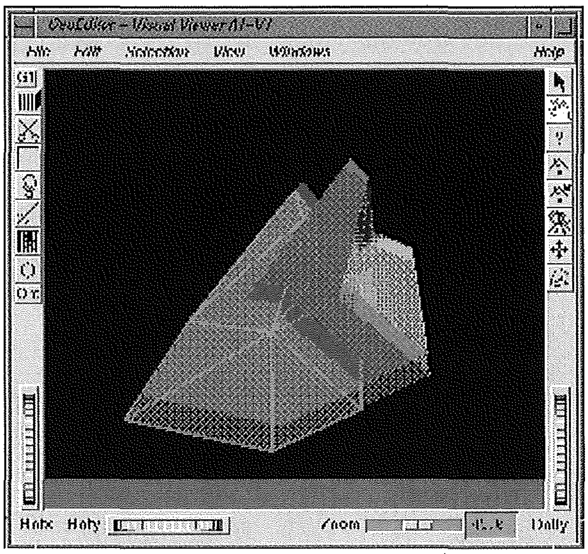
This section shows a complete a complete example that is the GeoViewer. The GeoViewer is analyzed with the scenario description, Task analysis, Information analysis, Data analysis and Auditory Characterization according to the TaDa method.

7.2 Description

GeoViewer is a software that allows to gather data from a variety of sources into one interactive view. GeoViewer offers a single, coherent view of large amounts of information, located and presented quickly and efficiently. The geographical information system (GIS) provides access to multiple databases and presents them as if they came from a single database. This is the description for the GeoViewer.

Title	GeoViewer
Storyteller	C.G.
Story	The GeoViewer is a 3D interactive view of rock strata for mine planning and other geological applications. Color and texture show the rock- type of the layers 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
Question	What rock-type is it?
Answers	coal, sandstone, granite, marble
Subject	rock-type
Sounds	???

Table 7.1: Scenario description of the GeoViewer [Bar97f]



Picture 7.1: Example of rock strata viewed by the GeoViewer [Bar97f]

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 Geo Viewer 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 Organization 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.

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 Organization is {category, space} because the rock structure is both spatial and material.

Auditory characterization

The TaDa requirements were used to look-up everyday examples with similar task and data structure from the Earbenders database. The three best matches were shaking cereal containers to determine the contents, kicking garbage bags at a recycling depot to sort them, and listening to the weather outside the tent to decide whether to sleep-in. If the semantics of the sound design are not obvious, then one of these examples may suggest a suitable everyday scheme. Perceptual aspects that are common to all the examples were copied straight into the sound design, and are shown with an asterisk in the next table.

Requirements Analysis		Sound design	
Generic	What is it?	Answers	Different types (*)
Purpose	Identify	Nature	Non-verbal (*)
Mode	Interactive	Level	Local (*)
Type	Exploration	Streams	Single (*)
Reading	Direct	Occurrence	Isolated
Type	Nominal	Pattern	Discrete
Level	Local	Movement	Stationary
Organization	Category	Type	Categorical (*)
Range	4	Compound	Integral (*)
Type	Nominal	Aspect	Timbre (*)
Range	4		
Organization	Category space		

Table 7.2 : Requirements and sound design of the GeoViewer [Bar97f]

The rule-based suggests that perceptually-equally-different timbres can represent categorical information because they do not imply a spurious ordering of the elements. Personify [Bar95i] [Por] (a sound-editing tool) was used to choose four equally different timbres. Each rock was arbitrarily assigned an instrument timbre at constant pitch and brightness: granite = cello, limestone = tenor sax, shale = English horn, marble = trombone. The GeoViewer plays a one-second sample of an identifying timbre when the user taps on a rock with the mouse. A rock strata that is difficult to visually identify can be heard, without having to divert visual attention to a text. A sound can be heard in order to an expensive change of viewpoint operation. A development could provide information about the number and material of overlapping hidden layers.

7.3 Conclusion

This section has shown the utility of the Tada method. The Tada requirements were used to look-up in the Earbenders database to find some similar example.

A mapping based on the timbre was chosen to differentiate the different sounds produced in the GeoViewer application.

Chapter 8

Application

8.1 Introduction

This section describes a Java [Jav99a] [Jav99b] [Jav00] application that tries to illustrate the principles seen in the previous chapters. We introduce it by a specification of the application by using the concepts seen in the Tada Method. A more detailed explanations of the database used, a description of the metaphor and the goals we wanted to reach is presented. Also a description of the sounds and how there are produced is presented. Also, you will find a description of this application in terms of principles as seen in the previous chapters. An experimentation has been realized to test the different sounds and to know how people react when they hear it.

Furthermore, we would like to mention that the application implemented during my stage was divided in two parts. The first part was a visual part presenting data in a visual form, the second part presented data into audio. Only this part is presented in this master's thesis according to the topic of this study.

8.2 Scenario description

This section illustrates the scenario description that is a part of the Tada Method. As seen before, the method adopted is a text description in general terms written by the user involved in that activity. The problem captured is recast as a question and the range of possible answers is identified.

Title	SoundApplic
Storyteller	Cedric Warin
Story	SoundApplic transforms a huge data set into an easier way that will give an idea on the weather during the time considered. The parameters of temperature, precipitation and wind can simply be heard by the sound produced by this application. The information is thus divided into an easy way of understanding for the listener.
Question	How is the weather?
Answers	Cold, moderate, warm, no rain, light rain, moderate rain, heavy rain, not windy, windy, very windy, storm and a mix of this basic concepts

Subject	Atmospheric data
Sounds	???

Table 8.1: Scenario Description for SoundApplic

8.3 TaDa Analysis

The requirements analysis is given by the TaDa method (task analysis and data characterization). This method divides the analysis in three areas: the Task section analyzes the Question key, the Information section analyzes the Answers key and the Data section analyzes the Subject key. Also the sounds will be described in the auditory characterization.

Task analysis
The Question key is the bridge to the Task analysis. The Question {How is the weather?} from the Weather is transformed to the Generic question {How is it?} by removing the subject. According to table 4.2 with the Generic question, the purpose is {identify} because we need to identify the weather from a source of data. According to table 4.3, the mode is {interactive} because the question is made through a mouse click, which is a manual operation. According to table 4.4, the type is {continuous/tracking} because the tasks are interactive, overlapping and they have undefined closure. According to table 4.5, the style is {exploration} because preserve the relations and do not exaggerate details, we explore the data sets to find some interesting and unknown features.

Information analysis
The Answers Key is the bridge to the Information Analysis. The Answers for the Weather are { cold, moderate, warm, no rain, light rain, moderate rain, heavy rain, not windy, windy, very windy, storm and a mix of this basic concepts}. According to table 4.7, the reading is {direct} because the understanding is immediate, need a little training (The sounds used are from everyday life). According to table 4.8, the type is {ordinal} because the answers give a difference and order. For example, cold, moderate and warm are like low, medium and high. According to table 4.6, the level is {information} because the answer concerns a subset of elements. For example, the weather is cold, windy and rainy. According to table 4.9, the organization is {category} because the answers pertain to organization of types. The range is {48}, which is indicative of the different weather that might be expected.

Data analysis
The Subject Key is the bridge to the Data Analysis. The Subject in the Weather is {atmospheric data}. According to table 4.10, the type is {interval} because the data have differences, order and metric. The wind is characterized by a speed in m/s and a direction, the rain by the precipitation in mm and the temperature is measured in degree centigrade. The range is {48} because there are 48 different weathers in the region of interest. According to table 4.11, the organization is {space, time, energy} because the data correspond to some physical phenomena's.

Generic	How is it?
Purpose	Identify
Mode	Interactive
Type	Continuous/tracking
Style	Exploration
Reading	Direct
Type	Ordinal
Level	Information
Organization	Category
Range	48
Type	Interval
Range	48
Organization	Space, time, energy

Table 8.2: Requirements of the SoundApplic

Auditory characterization

According to table 4.12, the nature is {everyday} because we face to acoustic events in the physical environment. According to table 4.13, the level is {global} because we are listening to many events. {A few} streams are involved, table 4.14. According to table 4.15, the occurrence is {continuous} because the sounds are ongoing and can change every second. According to table 4.16, the pattern is {unpredictable} because we cannot predict the variations. The movement in table 4.17 is {texture} because the sound has no identifiable location or movement. The type in table 4.18 is {prothetic} because we have difference and order in the sound by frequency, amplitude,... and almost one natural zero (when the sound disappears for the wind or the precipitation). According to table 4.19, the compound is {separable} because aspects of the sound can be easily heard.

Nature	Everyday
Level	Global
Stream	A few < 5
Occurrence	Continuous
Pattern	Unpredictable
Movement	Texture
Type	Prothetic
Compound	Separable

Table 8.3: Auditory characterization of the SoundApplic

This section 8.3 has allowed to gather some knowledge on the TaDa analysis requirements (the Task, the Data and the Information) and the Auditory characterization. This set of information is an interesting source of specification and definition. These specifications could be added to the Earbenders database with a description of the mapping used. Other examples that share a similar task, data and information structure could match the SoundApplic application and use the same mapping. Furthermore, the database could store the knowledge about this application. Many solutions of this problem could be included in the database or the first solution could be improved.

8.4 Variable definition and example of data used

This part shows the description of the database and the format associated with each variable. Only three parameters are important in this study: the temperature, the wind speed and the precipitation. Initially, the others parameters were shown in a visual way but accordingly with the topic of this study only the three parameters that use the possibilities of sound will be presented and described.

Date
The date of the observation represented by height-digits in the format MM-DD-YYYY where MM is months, DD is days and YYYY is years.

Date
10-08-2001

Hour or **Time of Observation**
Time of observation on the 24-hour clock represented by four-digits in the format HHMM, where HH is hours and MM is minutes. The following parameters will always be an average value in the time considered.

ObsTime
00:15

Precipitation
Prec = The amount of precipitation (in mm)

Prec
0
0
0
1
1

Humidity
Hum = Humidity (in percent)

Hum
30
32
34

Temperature
Temp = Temperature (in degree centigrade)

Temp
10
12
14

14

Wind

The format is SS for speed and DD for direction, where DD is direction from which the wind is blowing with direction noted to 8 points of the compass (N, NE, E, SE, S, SW, W, NW) and SS is speed in miles per hour.

WindSpd = Speed of the wind (in mps)
WindDir = Direction of the wind (N, NE, E, SE, S, SW, W, NW)

WindSpd	WindDir
0	N
5	N
6	NE
7	E
20	E
40	E

Parameters together

Date mm-dd-yyyy	ObsTime hh:mm	Prec (mm)	Hum (%)	Temp (c)	WindSpd (mps)	WindDir (dir)
10-08-2001	20:00	8	33	2	107	E
10-08-2001	20:15	8	33	2	106	E
10-08-2001	20:30	8	33	3	107	NE
10-08-2001	20:45	8	33	2	107	E
10-08-2001	21:00	8	33	2	107	E
10-08-2001	21:15	8	33	2	107	E
10-08-2001	21:30	8.5	33	2	107	E
10-08-2001	21:45	8	33	2	107	E
10-08-2001	22:00	8	33	2	107	E
10-08-2001	22:15	8	33	2	107	E
10-08-2001	22:30	8	33	2	107	E

This example shows a simulation of a storm (see the WindSpd) with heavy rain (see the Prec) at a moderate temperature (see the Temp). It occurs the 10-08-2001 from 20:00 to 22:30.

8.5 Description of the metaphor

The model used tries to base on everyday sounds. Three kinds of data presented in a file are manipulated to produce particular sounds. The file has been described in the previous section. Data are taken from meteorological data and the important data form the triplet temperature, precipitation and wind. By using the Java Synthesizer called Jsyn [Jsy97], each data alone produce a sound that we usually hear in a normal world. The reason is that a common sound should be easier to recognize than a strange sound or a noise. So the sounds come from the real world acoustics and are by this way easily learnable and understandable because humans are adapted to hear information in these kinds of sounds. If we look at the three kinds of data, the description will be the following.

For the precipitation, the sound produced will be the same than the noise of the rain that falls on the floor. With data of higher values the sound of the rain will be louder and we will hear more impacts. With data of lower values we will hear the opposite.

For the wind, the idea is similar. The sound produced is the same that when we hear the wind blowing. Again, with higher values, the sound heard will be louder, changing and faster. It will be the opposite with lower values.

The sound produced for the temperature is a bit different because it is produced only when there is a change of category. The temperature is divided in three categories: below 0, from 0 to 15 degree and above 15 degree. To give an example, a sound will be played if the temperature goes from -2 degree to 0 degree or from 14 degree to 16 degree. The sound played is the sound of a bird singing. The idea is that in a cold weather, the bird is not likely to sing and produce only a few different sounds. In a warm weather, the situation will be the opposite. The bird will be happy to sing a lot and will produce a lot of different sounds.

The justification why using a metaphor is the following: the idea is not to develop a general tool that allows to map data into audio and to analyze if this mapping is correct (this kind of experiment has been made in other thesis, especially in [DS01] and [Anr99]) but to develop a tool that allows to analyze the features of sound by manipulating some of these characteristics. Some of these features could be interesting and could be included in a more formal method. Nevertheless, this application could be used to analyze the usefulness and correctness of such a mapping or to analyze some other data than meteorological data but it was not the main goal.

Furthermore, the temperature and the wind behaves like a stream, we mean by that way that a sound for these two data is always produced except if there is no wind and no precipitation. The reason given is that according to the recent analysis, humans have sometimes some difficulties to recognize a sound amongst a list of sounds. Thus, a stream seems better to recognize a change in the data that means a change in the parameters of the stream. To give an example, a stream for the wind is played from 20 to 60 mps, above 60 mps the same stream will be played but with different parameter (different frequency, amplitude...). By this way, it is easy to distinguish a trend in the data.

The sound produced for the temperature is more usual because only produced when there is a change in the different categories for the temperature. The reason for that choice was that two different streams seem already enough for the understanding of the sounds presented. Another stream could cause a sound pollution and it will be difficult to discern all the information presented in the three streams. Furthermore, it seems more difficult to associate a stream based on a usual sound with the temperature because this parameter is not "noisy".

All the signals do not provide more information than is necessary (principle of parsimony) and the same signal designates the same information every time (principle of invariance).

8.6 Mappings between data and sound

This part explains the sound components, the characteristics used and the mappings between data and sound. Each sound is described with the classification that is used, a description of the units and parameters used in Jsyn and the mappings.

8.6.1 Stream for the wind

A. Classification:

Data for the wind can be divided in four classes according to a usual classification of this parameter. The four classes are {No wind, Windy, Very windy, Storm}.

This is the classification with the intervals:
- 0 mps <=No wind < 20 mps
- 20 mps <= Windy < 60 mps
- 60 mps <= Very windy < 100 mps
- 100 mps <= Storm

B. Synthesized Sound

Using Jsyn Api that is a Java Synthesizer creates the sound of the wind. The sound of the wind uses a circuit. Using this technique, very complex sounds could be constructed from simple units.

There are dozens of unit generators in JSyn that provide all sorts of sound generating and sound modifying functions. These can all be connected together and we can develop a large library of sounds by connecting various units using the class SynthCircuit. A SynthCircuit can contain multiple JSyn units connected together. When you start a SynthCircuit, all of its sub units are started simultaneously. This can be important when you want all the parts of a sound to be synchronized.

The wind sound produced is the result of a white noise through a low pass filter with random modulation. It uses the following units:

Units:

- WhiteNoise: This unit controls a white noise.
- StateVariableFilter: This filter is convenient because its frequency and resonance can each be controlled by a single value.
- RedNoise: It generates a waveform with linear ramps between random values. This unit interpolates straight-line segments between pseudo-random numbers to produce "red" noise. It is a grittier alternative to the white generator Noise_White. It is also useful as a slowly changing random control generator for natural sounds. Frequency port controls the number of times per second that a new random number is chosen.
- MultiplyAddUnit: This unit is used to combine some input together in that way: Output = (inputA * inputB) + inputC; The RedNoise (inputA) is applied with a rate (inputB) and a depth (inputC) that fix the values for the modulation.

Remark: The Applet tj_seeosc found in [Jsyn] allows hearing the different waveforms that can be used. By testing it, you can hear for the red noise for example something like blowing.

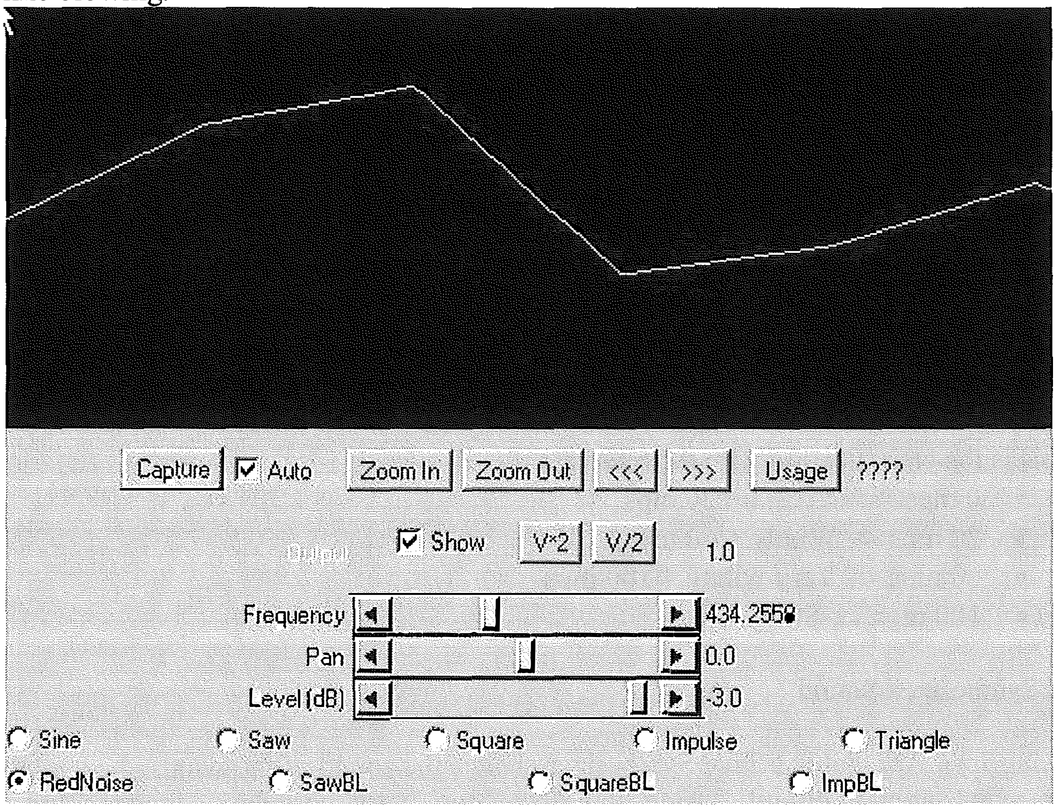


Figure 8.1: tj_osc applet in [Jsyn]; The waveform selected is the rednoise.

This is a description of the parameters that we can manipulate according to the units used. There are six parameters:

- ModRate and ModDepth are used to fix the modulation rate and the modulation depth for Red Noise. The rate is the frequency with which the modulation is calculated and the depth is the difference in the random values generated randomly.
- CutOff is a regulator used with the Red Noise, in other words it rules the frequency.
- NoiseAmp is the amplitude of the White Noise.
- The parameters of resonance and amplitude are used with the filter.

C. Mapping:

The wind from the data set is "separeted" in six parameters as seen before.

Wind →(noiseAmp, modRate, modDepth, cutoff, resonance, amplitude)

With these six parameters we can affect the sound that we want to hear. The range of these 6 parameters is able to render a sound like "No Wind, Windy, Very Windy and Storm".

This is the values that has been chosen for each category:

No Wind → (0.0,0.0,0.0,0.0,0.0,0.0)
Windy → (0.1,0.5,200,500,0.066,0.2)
Very Windy → (0.3,1.0,300,625,0.066,0.5)
Storm → (0.4,1.5,300,750,0.066,0.99)

The noiseAmp of the white noise filter and the general amplitude increase logically to give more volume to the sound heard. The modRate increases and is reflected like a bigger change in the sound, it means that the sound changes more if the modRate is high. The modRate gives more speed to the sound. The modDepth is difficult to use because it offers the possibility to fluctuate more from low value to high value. It is not so interesting because you can be in storm and hear almost nothing because the modDepth is high and can fluctuate to very low value, value that you hear less. Nevertheless, with the values chose in modRate and modDepth, we should be able to discern more fluctuations in storm than in very windy and in very windy than in windy.

The cut-off increases to give the sensation of more speed in the sound. We have the feeling that it blows faster. The resonance is as well difficult to use, the value of 0.66 seems a good value because it allows to look like a wind sound. In this case, the resonance seems not a good parameter because there is only one value used and the others values are not very useful to distinguish some different sounds.

8.6.2 Temperature

A. Classification:

Data for the temperature can be divided in three classes according to a usual classification of this parameter. The three classes are {Cold, Moderate, Warm}.

This is the classification with the intervals:

- Cold: less than 0 degree
- Moderate: 0 to 15 degree
- Warm: more than 15 degree

B. Synthesized sound

Another synthesized sound can be used in that way: the sound is played in the important event like in the beginning of a data set or when a change of category occurs. For example a change from moderate to warm weather.

Manipulating some units in Jsyn again plays the sound. This is the units that are used:

Units:

- TriangleOscillator: This unit produces a wave that looks like: /\/\/\ .
- SawtoothOscillator: This unit produces a wave that looks like: /|/|/|.
- ExponentialLag: Output approaches Input exponentially. This unit provides a slowly changing value that approaches its Input value exponentially. The equation is: Output = Output + Rate*(Input - Output); The Rate is calculated

internally based on the value on the halfLife port. Rate is generally just slightly less than 1.0.
This unit is used in this example to control the amplitude.

- WaveShapper: Look up a value by interpolating between values in a Table. When Input is -1.0, the value from the table is Table[0]. When Input is +1.0, then the value from the table is Table[NumValues-1]. The output is the interpolated table value times the Amplitude. The name "wave shaping" refers to the synthesis technique that involves connecting an oscillator such as a SineOscillator or SawtoothOscillator to the input of this unit. By modifying the contents of the table, and the amplitude of the input, interesting output waveforms can be created.
  Used like this in the example: it creates a waveshaper that converts the sawtooth input to an amplitude envelope. When the sawtooth goes / the waveshaper will go /\. Thus the amplitude will be 1.0 when the freq offset is zero, and will be 0.0 when the freq offset is -1.0 and +1.0.

Synthetic bird sounds are produced by randomly varying the depth and rate of modulation. The modulation is played several times with this changing parameters.

The units used gives the opportunity to manipulate the following parameters:

- Frequency of the SawtoothOscillator
- Modulation Depth
- Time of a modulation (in ticks)
- Number of modulation
- Top of the frequency

C. Mapping:

The temperature is “separeted” in 5 parameters that can affect the sound that we want to hear.

Temperature → (freq, depth, time, numb, topfreq)

The range of these 5 parameters is able to render a sound like “cold, moderate, warm”.

This is the values that has been chosen for each category:

- Moderate → (50.0,500.0,130,5,2500.0)
- Cold → (25.0,1.0,300,2,1000.0)
  With these parameters the sound produced seems slower (frequency decreases and the time of a modulation is bigger) and smaller than the previous one (only two different modulations). Furthermore, the top of the frequency and the modulation depth are lower than before that is why we should hear less difference in the sound. “The bird sings only a few sounds”.
- Warm → (200.0,650.0,70,12,4000.0)
  With these parameters the sound produced is faster and longer that the others with a lot a modulations and more possibilities to fluctuate. “The bird seems agitated and produces a lot of different sound”.

We notice here that a simply way to differentiate the different sounds is to base on the number of modulation. Simply, the number of modulation (the number of times a modulated sound is played) is useful because 2, 5 and 12 modulations give an order in the stream; it is like to count 2, 5 and 12. The others parameters can be used to accentuate the difference in the sounds played (longer, more range in the frequency, more difference in the modulation).

8.6.3 Stream for the rain

A. Classification:

Data for the precipitation can be divided in four classes according to a usual classification of this parameter. The three classes are {No rain, Light rain, Moderate Rain, Heavy rain}.

This is the classification with the intervals:

- No rain: 0 mm/h
- 0 mm/h < Light Rain <= 5 mm/h
- 5 mm/h < Moderate Rain <= 30 mm/h
- 30 mm/h < Heavy Rain

B. Synthesized Sound

We use again a synthesized sound that can reflect the different "rains" (no rain, rainy and very rainy).
This sound uses the Parabolic Envelope (an envelope describes the contour or shape of a parameter) that generates a very short arc. When the envelope is triggered (by the unit Poisson Trigger), it latches a random value for a sine oscillator using a LatchUnit. The sine wave is multiplied by the parabolic envelope to generate a very short burst of a sinewave similar to a wavelet. The sound on the left channel is the original grains. The sound on the right channel is the output of a MultiTapDelay that is used to make it sound like there are more grains playing at once.

This is the units that are used:

- ParabolicEnvelope unit: Output goes from zero to amplitude then back to zero in a parabolic arc. It generates a short parabolic envelope. The output starts at zero, peaks at the value of amplitude then returns to zero. This unit has two states, IDLE and RUNNING. If a trigger is received when IDLE, the envelope is started and another trigger is sent out the triggerOutput port. This triggerOutput can be used to latch values for the synthesis of a grain. If a trigger is received when RUNNING, then it is ignored and passed out the triggerPass port. The triggerPass can be connected to the triggerInput of another ParabolicEnvelope. Thus you can implement a simple grain allocation scheme by daisy chaining the triggers of ParabolicEnvelopes. The envelope is generated by a double integrator method so it uses relatively little CPU time.
- ParabolicGrain: ParabolicEnvelope modulating a sine wave.

- Unit Poisson Trigger: Generate triggers with a Poisson distribution. Port "probability" controls the probability of a pulse happening on any given sample frame. The average rate of the pulses is:
  AverageRate = (probability / Synth.getFrameRate())

The units used gives the opportunity to manipulate the following parameters:

- Probability - controls the likelyhood of triggering a parabolic envelope on any given sample. Triggers are generated by a PoissonTrigger circuit that outputs a trigger whenever a white noise source exceeds a threshold determined by "Probability".
- Frequency - determines the centre frequency of the sine wave oscillators.
- Spread - determines the maximum random deviation of the sine wave frequency.
- GrainSpeed - determines how quickly the ParabolicEnvelope moves from zero to maximum amplitude then back to zero. If a ParabolicEnvelope were retriggered continuously then it would generate this many arcs per second.
- Feedback - determines how much of the delayed sound is mixed back into the delay line. This determines how long it takes for the sound to decay.

C. Mapping:

The precipitation is "separated" in 5 parameters that can affect the sound that we want to hear.

Prec →(proba, freq, spread, grainspd, ampl, fdk)

The range of these 5 parameters is able to render a sound like "No rain, Light rain, Moderate rain, Heavy rain".

This is the values that has been chosen for each category:

No Rain →(0.0,0.0,0.0,0.0,0.0,0.0)
Light Rain →(0.001,60.0,60.0,100.0,0.9,0.4)
Moderate Rain →(0.002,180.0,60.0,120.0,0.9,0.4)
Heavy Rain →(0.003,300.0,60.0,140.0,0.9,0.4)

The parameters of proba, freq and grainspd are respectively used to give the feeling of more impacts, a sharper sound and less time between two impacts (more speed). The amplitude and the spread are not very useful because the amplitude is used to set a general volume and the spread do not really give a perceptual difference in the sounds played.

8.7 Description of the application

SoundApplic is divided in four parts that are the general panel, the tutorial panel, the oscillators panel and the database visualization.

General panel

> The first panel is a general panel where you can load data from a file. Once the data has been loaded, you can start the engine of the sound process just by clicking on the start button and hear the sounds associated with this database. Furthermore, you can stop the process whenever you want by clicking on the stop button. A description presents the topic of this thesis and the author.

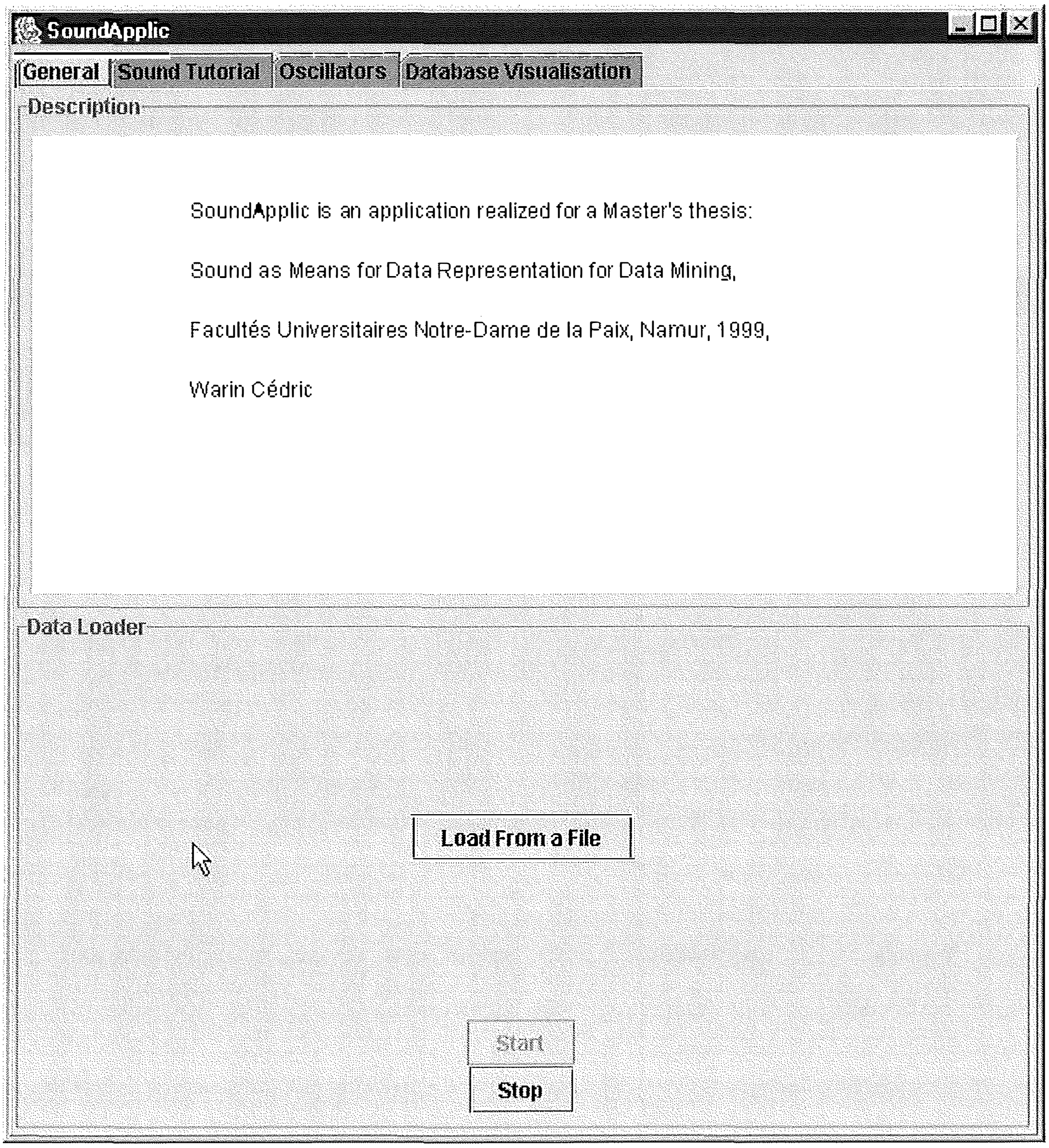


Figure 8.2: SoundApplic, General panel

Sound Tutorial

This panel is very important for a person that does not know anything about this application. It can be used to learn the different sounds produced and to ease the ears at these new sounds. As mentioned, each sound produced lasts four seconds expect the last one that is a combination of the three previous sounds. To give an example, the sounds for Light, Rainy and Very Rainy are played one after another in "all". By this way, we can learn easily the trend in each stream that mean a trend in the data associated. Just clicking one of the three buttons and selecting a category in the radio buttons you can play each sound.

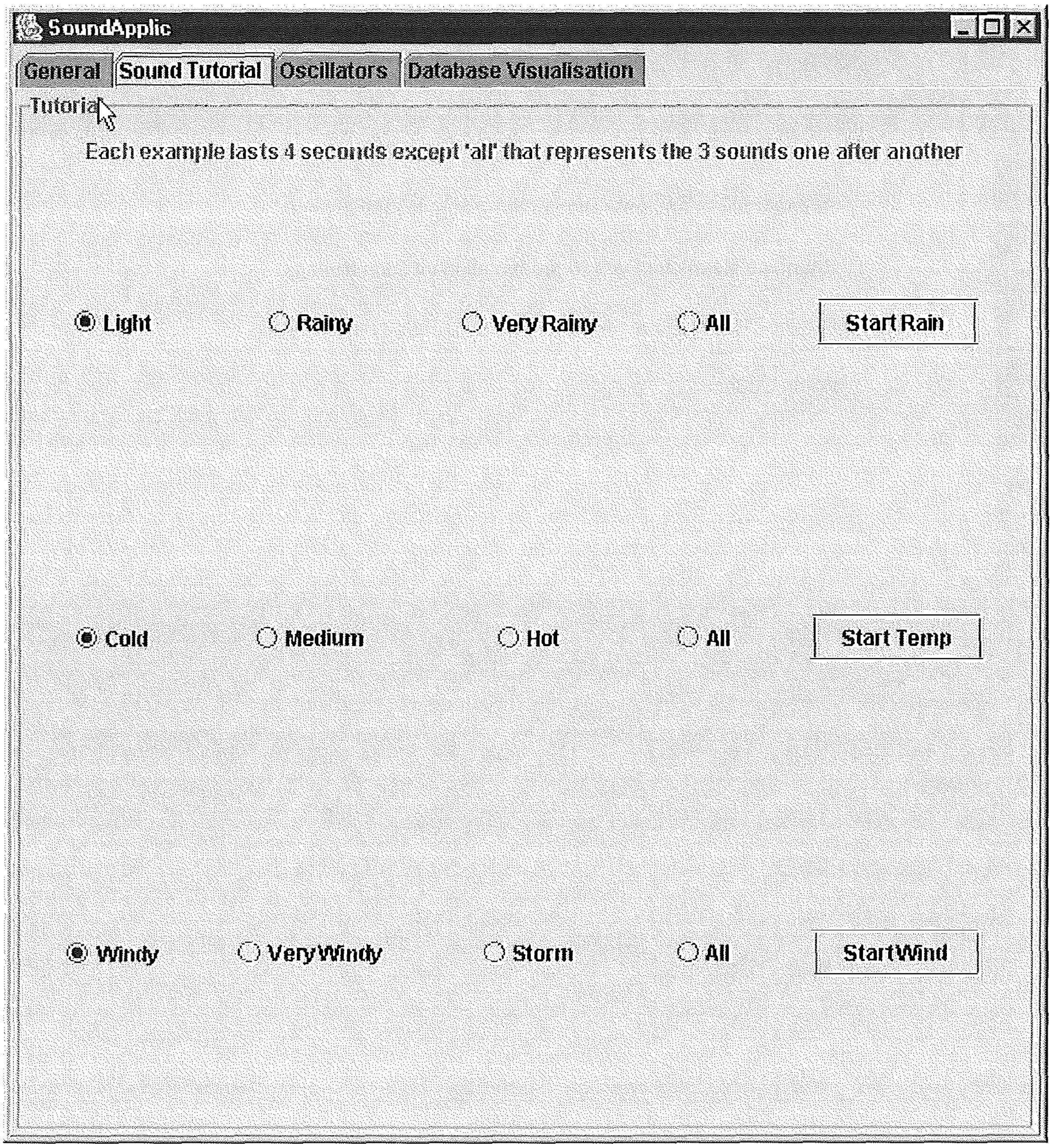


Figure 8.3: SoundApplic, Sound Tutorial

Oscillators panel

This main panel is divided in two parts. The first one presents the different oscillators that are the wind, the bird and the rain.
The second one shows the result of all these oscillators together. Each oscillator gives the possibility to accentuate the amplitude of each wave to give a better representation (by clicking on the v*2 button).

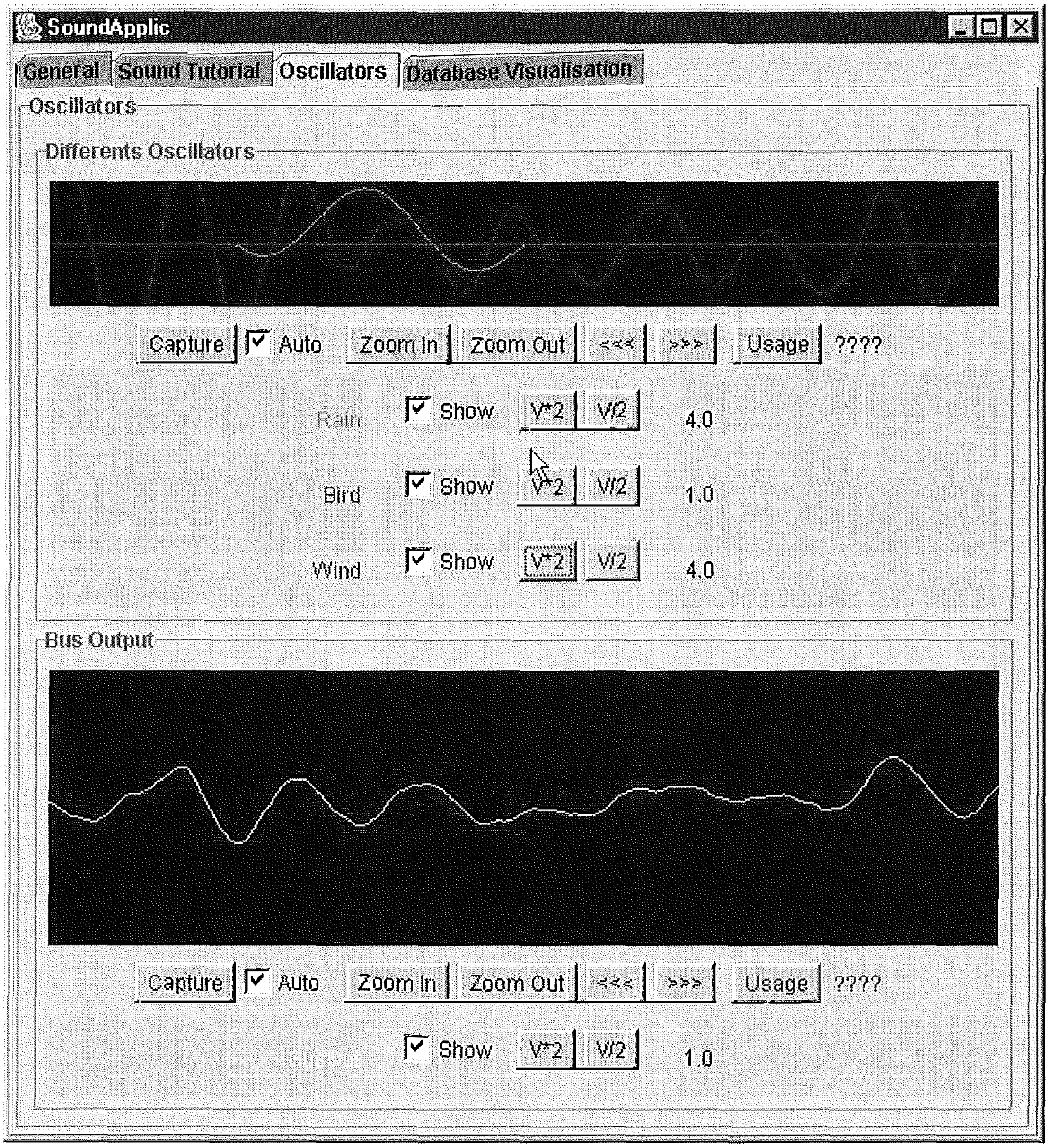
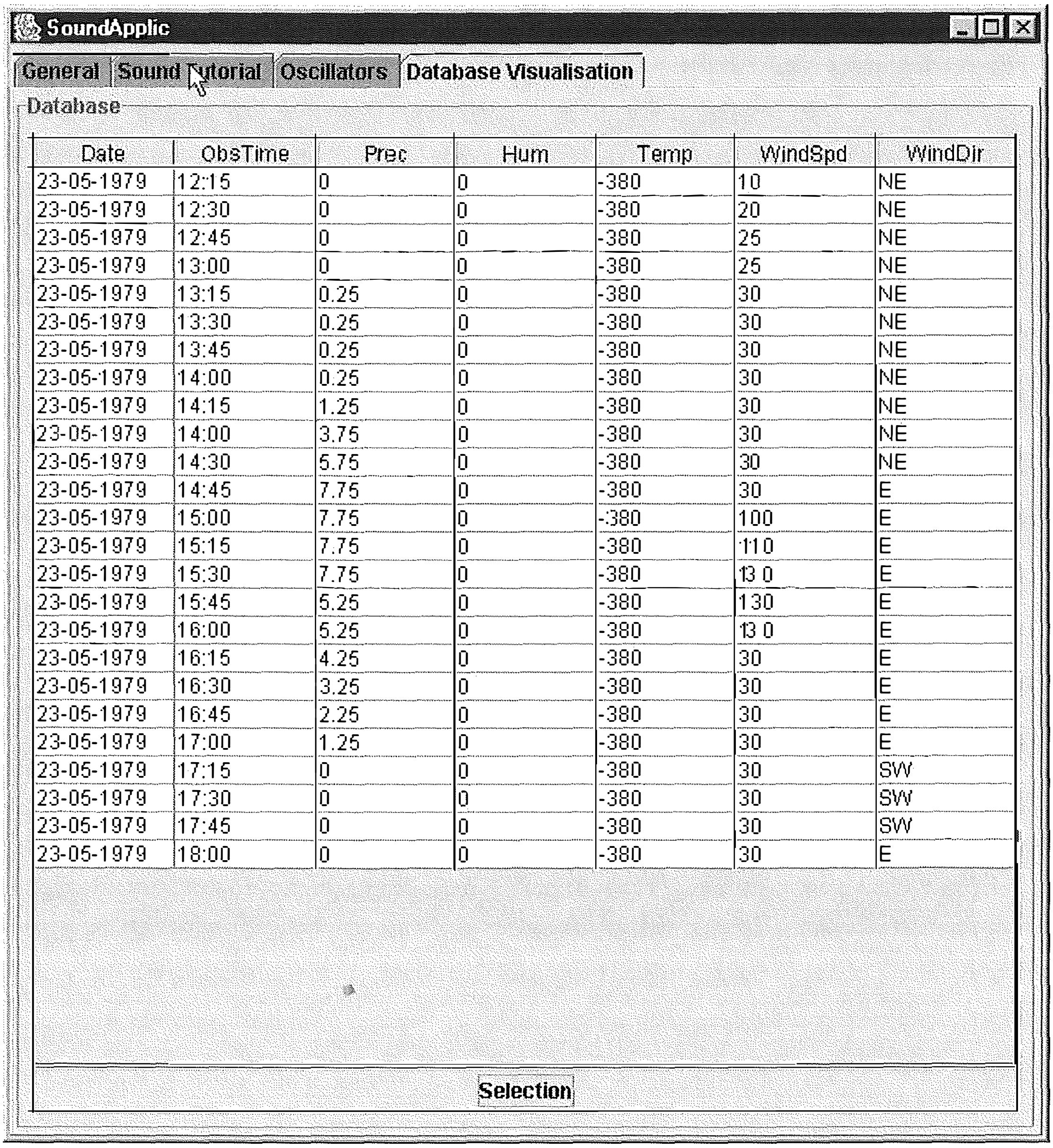


Figure 8.4: SoundApplic, Oscillators panel

Database visualization

This main panel shows the database associated with the file. You can select a part of this database by using ctrl and the arrows and just by pressing the selection button, you can launch the “sound” for the selected database.



Date	ObsTime	Prec	Hum	Temp	WindSpd	WindDir
23-05-1979	12:15	0	0	-380	10	NE
23-05-1979	12:30	0	0	-380	20	NE
23-05-1979	12:45	0	0	-380	25	NE
23-05-1979	13:00	0	0	-380	25	NE
23-05-1979	13:15	0.25	0	-380	30	NE
23-05-1979	13:30	0.25	0	-380	30	NE
23-05-1979	13:45	0.25	0	-380	30	NE
23-05-1979	14:00	0.25	0	-380	30	NE
23-05-1979	14:15	1.25	0	-380	30	NE
23-05-1979	14:00	3.75	0	-380	30	NE
23-05-1979	14:30	5.75	0	-380	30	NE
23-05-1979	14:45	7.75	0	-380	30	E
23-05-1979	15:00	7.75	0	-380	100	E
23-05-1979	15:15	7.75	0	-380	110	E
23-05-1979	15:30	7.75	0	-380	13 0	E
23-05-1979	15:45	5.25	0	-380	130	E
23-05-1979	16:00	5.25	0	-380	13 0	E
23-05-1979	16:15	4.25	0	-380	30	E
23-05-1979	16:30	3.25	0	-380	30	E
23-05-1979	16:45	2.25	0	-380	30	E
23-05-1979	17:00	1.25	0	-380	30	E
23-05-1979	17:15	0	0	-380	30	SW
23-05-1979	17:30	0	0	-380	30	SW
23-05-1979	17:45	0	0	-380	30	SW
23-05-1979	18:00	0	0	-380	30	E

Figure 8.5: SoundApplic, Database Visualization

8.8 Description in terms of principles

This section illustrates the application in terms of principles seen in the rule-based method. The five principles are {directness, type, level, organization, range}.

Directness

This principle is reached because the sounds for the precipitation and the wind can be understood almost immediately. They can be heard like the usual sound of precipitation and wind. The sound for the temperature is not very complex and can be understood with little training. A direct representation is proposed for each sound. We will add that to distinguish more easily the different sounds, it is recommend to listen to this application with some headphones.

Type

Difference: As seen before, the different parameters used for each kind of sound (precipitation, temperature and wind) can produced some different sounds. Each sound seems different from the previous one and can indicate a trend in the data.

Order: The order is easy to perceive for the precipitation and for the wind because it fluctuates like in the reality. For the temperature, it needs more attention to distinguish the order because, we need to be accustomed to the sound. But simply, the number of modulation (the number of times a modulated sound is played) is useful because 2, 5 and 12 modulations can give an order in the stream; it is like to count 2, 5 and 12.

Metric: If we look at the values for the rain, we can see a variation of an equal step for each parameter. But for a few parameters, it seems difficult to talk in terms of metric because of the following reason. With the goal to base on everyday sound the range for the different parameter has to be calibrated, outer this calibration sometimes the sounds does not look like the natural phoneme of rain, wind and temperature (bird). Furthermore, certain parameters could not be used in their entire range because they propose a sound that can fluctuate too much, we think at the modDepth. If the fluctuation is too big it can create a confusion between the categories. Also with the numbers of parameters in each category, it is difficult to talk in terms of metric. It should be easier if there were only one or two parameters. Whatever, it seems difficult to think in terms of metric when we hear the different sounds in each category. It is difficult to interpret a metric from one sound to another in each category but there are perceptually different that seems enough to distinguish all the sounds. To simplify, this principle works well for the following parameters: the amplitude, the frequency, the speed in a modulated sound but it has more difficulties with parameters like the depth in a modulated sound.

Zero: A natural zero is present for the precipitation and the wind because no sound is produced when there is no rain and no wind. We do not have to talk about zero for the bird because the temperature is divided in three categories that do not really require a zero.

Level
We can note the two following listening processes: a global level where all the sounds can be heard by the listener and can represent a global analysis and a global information of the data set, a local level where we can hear the details in each stream. We will note that the global analysis can represent some difficulties because of the ability to the listener to interpret all the sources. In the application, there are three sources that form a general information about the weather. It seems that three sounds or streams are already a maximum for the listener. The listener needs a full attention to listen to the sounds. According to that experiment, it seems that an information given by the way of sound can reach its maximum with only three different streams of information.

Organization
The different parameters used seem efficient to organize the sounds in the different parts of the spectrum. The Java synthesizer Jsyn and the synthesis parameters give the possibility to organize different sounds. We will add that the amplitude can be difficult to manipulate because it can overlap other sounds when it is at its maximum. The sounds can thus become a noise because of the variation of this parameter. So it can reduce the degree of freedom in the spectrum considered.

Range
This principle can be easily apply to the amplitude parameter even if we have to care about the noise it can provoke. In each case, the values are chosen to correspond to everyday sounds. As seen before, the entire range is not always used for other parameters because mainly of the idea of basing on everyday sounds.

8.9 Experimentation

An experimentation using the sounds found in the application has been proposed. The purpose of the experimentation is to determine how people can appreciate and understand the sounds that they hear. So, three sequences of sounds were proposed to the listener:

- A sequence from no rain to heavy rain
- A sequence from Cold to Warm
- A sequence from no Wind to Storm

It was simply asked to the listeners to say what they hear in each sequence and to describe the evolution in each sequence. With these two parts, the hope was that firstly we test the ability to understand the sounds proposed and secondly, we test the ability to hear some differences in the sounds. A difference that means for us, a change in a particular characteristic of the sound and by this way we can analyze the features that are easily perceptible. These features that are easily perceptible could be introduced in a new model like the TBP prototype and could be applied to map some type of data into sounds.

This test has been sent to a few persons via e-mail and the sounds has been recorded in three files that we could found in an Internet Web Site. By this way, the three sequences could be presented and evaluated by the visitor without interaction with the people who know the application. The files were available in the MP3 sound file format.

The experimentation has been tested with sixteen persons. This number seems enough because we can find a lot of similarities in the answers.

This is the analysis of the results. The results can be found in Annex B:

The analysis could be divided in two parts, as there are two distinct questions. The first question could be summarize like this: Could you say which phenomena you are listening in this sequence? The second question could be summarize like this: Could you describe the evolution of the sound in the sequence? According to these two questions, the answers that we expect could be the following for each file:

- File 1: I hear the rain or something like an impact on the floor or somewhere else, the speed and the number of impacts increases, it sounds more and more sharp.
- File 2: I hear the song of a bird, the sounds seem more and more long, quick, sharp and with more and more difference.
- File 3: I hear the wing blowing, it accentuates in the speed, the amplitude. It blows more and more fast and violent, there is more and more difference in the three sounds, it oscillates more.

With the results on the test, we will analysis the situation with the two questions proposed before.

The first question is: Could you say which phenomena you are listening in this sequence?

It seems that the first example (file 1; the rain) is very difficult to interpret. People have difficulties to say exactly what it is. Only two persons have answered the rain for this first file. Different answers have been proposed like an engine, some horses galloping, some impacts, a plane or a rocket, some water boiling, the whirring of a thunder. According to these answers, the stream for the precipitation should be improved because people are not able to discern the everyday sound that we wanted to base on. A better relation between the sound played and the everyday sound should be found to improve the recognition of this kind of event. The goal to base on everyday sounds is not reached for the first sequence.

The second example (file 2; the bird) gives a better answer because most of the people recognize the song of a bird, especially with the second and third sound played in the second sequence. It seems that the first sound in the sequence is a bit extreme to represent exactly the song of a bird but with the presence of the second and third sound, the user are able to recognize that it is well the song of a bird.

The third example (file 3; the wind) does not require a lot of concentration to recognize it because all the testers have found what it was. The wind is very easy to detect and they are able to discern it in a few seconds.

The following table resumes the answers for each file. The number represents the number of times the same answer has been proposed. For example: for the file 1, two persons have answered that it was some water boiling.

Remark: there are more than fifteen answers for the file 1 because the listeners have sometimes proposed two different answers.

	Water boiling	Horses	Rocket/ Plane/Reactor	Rock rolling	Popcorn	Thunder/ Storm/ Tornado	Impacts	Rain	Bird	Wind
File 1	2	2	3	1	1	5	1	2		
File 2									16	
File 3										16

Table 8.4: Answers for the first question

The "natural" relationships between the sounds played in the application and the everyday sounds seem efficient for the wind and the bird because the listeners can easily recognize it. These sounds are very quickly perceived and by this fact quickly learned. People do not need an important mental effort to remember the sounds and thus would not need an important mental effort to analyze data associated with the sounds. The sound of the rain is not directly perceived like some precipitation. It should be improved. The first time the user hears the sound of the rain, they cannot recognize it but with a little training they will be used to hear this kind of sound. Even if it does not really sound like some rain, the stream associated maybe used to show some trends in the data. We will see if it works in the second question.

The second question is: Could you describe the evolution of the sound in the sequence?

The second part of the analysis is interesting because it seems that most of the listeners were able to recognize the difference in the sounds played. By difference we mean that we are able to hear a difference in the speed, in the frequency (sharp sound, ...), in the amplitude and by this way distinguish different sounds. For the file 1, the listeners have particularly heard a more and more sharp, violent and quick sound. For the file 2, they have heard a more and more sharp and quick sound. For the file 3, the listeners have particularly heard an increase in the amplitude and a more and more violent sound.

The following table resumes the answers for the second question. For each file, the number of times a characteristic has been mentioned is counted. For each characteristic, it should be interpreted like an increase in this characteristic. For example: for the file 2, eight persons have heard a more and more sharp and quick sound.

Remark: This classification is not easy to do because of the general question asked. The answers can cause some confusions between two different characteristics.

	Sharp	Speed	Volume/ Amplitude	Number of sounds played/ Presence	Length	Nearness	Violence
File 1	4	10	1	4		2	4
File 2	9	13		3	3		
File 3	2	3	16			3	9

Table 8.5: Answers for the second question

Human listeners are sensitive to frequency changes, intensity changes but also there are able to discern that the sound accelerates. Maybe the feature that gives the less satisfaction is the one that manipulates the depth in the sounds modulated. This feature gives the feeling that the sounds have more difference. To give an example, for the wind a high depth will correspond to a non-linear wind, there will be more fluctuation in the wind. This feature stays interesting because it can intensify the difference between two sounds.

We have to note the following remark that was not really waited: Some listeners have interpreted an increase in the characteristics of the sounds like a nearness of the sound. First the sound seems far and comes nearer. This perception has been cited for the rain and the wind. It is difficult to say if this perception is due to only one parameter or all the parameters together. We will tend to this last solution because some different parameters were used in the flow for the rain and the wind with the same perception of nearness. To give an example, the amplitude was used with the wind but not with the rain.

People are able to discern the difference in each sequence. They are sensitive to a change in the different characteristics used. These characteristics could be introduced in a new model because they seem enough efficient to discern some sounds in a same type of sound. We mean that with these characteristics we were able to dissociate three different sounds in each sequence. If we add the perceptual zeros for the wind and the rain, it is four distinct sounds that can be played.

As these features seem interesting when they are manipulated, a model like a TBP model could be proposed. This is an idea of model:

- The vertical axle would be represented by the frequency because with this feature we can divide the entire range into a lot of steps. The frequency can create a lot of difference in the different sounds. According to the experiments made by Stephen Barrass, this characteristic suits well for this dimension (except that there is no natural zero see table 6.5). As seen in the application, the range could be calibrated to base on everyday sounds.
- The radial spokes would be represented by the amplitude (loudness).
- The Pedestral would be represented by a factor that accelerates the speed in the sound like a modulation rate. This feature could also be summarize by a factor like the tempo because with such a parameter you can speed up or slow down a sound. Another possibility could be the number of sounds in a modulated sound.
- A fourth dimension could be added like a modulation depth. This dimension will be there to accentuate the difference between the different sounds.

The radial spokes, the pedestral and the fourth dimension (if used) should be limited to three or four steps because listeners cannot identify correctly more than a few different intensities, changes in the speed,...

As seen in the application, a model based on these characteristics seems interesting and could be the base of a formal method as proposed here.

After having discussed about the sounds with a few testers, we have to add a few remarks.

- People have found it was difficult to describe the sounds because they do not know a lot of musical terms.

- According to the testers, it is easy to recognize that in each sequence they are three distinct "sounds".
- The testers have difficulties to recognize the sound of the rain but when they know what exactly it is or it should be they are more convinced. According to them, an easy way to improve the sound of the rain should be to hear more the impacts because they have the feeling that something is blowing and it is annoying for the whole sound.
- A few persons think it will be easier to have an idea on the file 1 after having heard the sequences 2 and 3.

8.10 Conclusion

This chapter has shown an application realized with Jsyn, a Java Synthesizer. The methods and specification learnt in the previous chapters have allowed to describe and to analyze this application.

The application and the experimentation have shown the importance of some parameters. These parameters could be introduced in a new model to map data into sounds.

To conclude in terms of features of the sound, we will say that frequency seems the more interesting characteristic because it can create the biggest difference and disposes of a large range. But this parameter has to be calibrated to respect the principle that we want to base on familiar sounds. The amplitude seems not so easy to manipulate. Firstly, because the range is not so important. Only three or four different sounds can be produced. Secondly, because the loudness can perturb the possibility to hear all the sounds played. Other parameters seem interesting and can be summarized like this: the speed at which a sound is played allows to make some differences and can be reinforced by the number of time the sound is played or by a bigger depth in the sound. We talk here in terms of a modulation of a sound.

We will notice that in this application, the manipulation of only one parameter cause some differences but not as big as we could expect. At each case, four or five parameters are used to create the different sounds. According to this remark, a prototype could be proposed including the useful dimensions that we have seen in this application. This prototype could propose obviously more than three dimensions and could be tested by some experiments made with the Java Synthesizer.

Chapter 9

Analysis of the Tada Method

9.1 Introduction

As seen, the different methods around the Tada method have a lot of advantages that we have to take care. Nevertheless, this section lists a few of disadvantages of these methods. The three methods are the case-based, the rule-based methods that support the TaDa method but also the TBP prototype. Mainly, we will look how it behaves in the practice. We will look separately at each method and will argue about the utility of the methods proposed. The three methods are the case-based method, the rule-based method and the TPB prototype.

9.2 Criticisms of the Tada Method and its tools

For each method, we list a few arguments to criticize the principles.

The case-based method that has been implemented in an Access database by Stephen Barrass gathers a lot of information about everyday listening. The specifications in its database as seen in this study are an important source of information. Nevertheless, the TaDa analysis is supported by an interface that forms an Access Database but it was not possible to test the opportunity and all the advantages of this tool.

The GeoViewer has shown the utility to retrieve a similar task and an appropriate mapping but it is the only example known.

It is difficult to generalize the effectiveness without testing the database that is why we will say that this database seems a very interesting tool but actually it is only marking our mind by its theory.

Furthermore and generally for the TaDa method, by applying the scenario description, the information requirements and the auditory characterization in the application, we have remarked that sometimes it is difficult to chose between the possibilities given in the different specifications. With more experience it should be easier to fill in the different specifications but as the GeoViewer is the only example that we know, it is the only example on which we could base.

Also, it takes a long time to write all the specifications and the GeoViewer has not shown the usefulness of all these specifications because the mapping proposed (on the timbre) was also proposed in the TBP prototype but this last one requires less time because it is just an application based on the information types.

The rule-based method offers another opportunity. The principles summarize a lot of ideas from the literature and proposed a systematic approach for mapping data into sounds. When we want to realize such a mapping, we think in terms of these principles. In despite of the advantages proposed by this method, it seems suffer of a few problems in practice.

As seen in the Java application, the principles of Order, Metric and Range were difficult to implement. The order for the bird song because it was difficult to find a usual sound for this kind of data. The song of a bird has no perceptual order that is why this principle was not an evidence for the implementation. The number of modulation can give an order. The Metric and Range were difficult to implement because the values were set to apply the general principle that we want to base on everyday sounds. An extreme value in the frequency for example proposes a sound that does not look like a usual sound. An extreme value for the amplitude can cause a sound pollution. It seems that this approach allows to think in terms of important principles but it is not the best solution in every case because another solution was adopted in the application SoundApplic.

Furthermore, we will notice that to be applied the principles have to be learnt and understood. They are not so easy to understand even if a sound with each principle has allowed to test the principles.

We will conclude that to apply these principles the user has to learn and understand it but also he has to find an appropriate mapping after thinking in terms of these principles. Obviously, it takes a long time and an important effort that the user maybe does not want to.

Also, we will notice that the principles never talk in terms of features of the sound. By this fact, it does not accelerate the realization in practice of these principles. This lack is solved in the TBP method.

The TBP method is articulated around three dimensions that are the timbre, the brightness and the pitch. The method gives an interesting point of view because it is easy to apply in practice because a few tools exist where you can manipulate these three aspects of the sound, see Csound for example [Cso].

Nevertheless, as seen, the three dimensions offer a general method that can be applied in practice but it is difficult to say if this method should be applied in every case. Other mappings could be proposed and could be more appropriate in a few cases, that is why we can argue about the generality of the method.

The choice of the second dimension, the brightness, could be argue because this aspect of the sound corresponds to an implementation (the brightness property has been realized with Csound) that improves the nature of the sound. Typically, it could be interpreted like a better gain in the sound played that is why we could ask us if the choice of the amplitude was not a better choice. Indeed, the amplitude compared to the brightness is a feature know by everybody and a feature on which we know the reaction when we manipulate it.

Furthermore, the brightness was the only property tested for the second dimension, the radial spokes.

Also, the method is articulated around only three dimensions. The Java Application has shown that more than three dimensions could be used. So some other methods based on more than three dimensions could be considered.

In the end, we will say that the use of sound seems diminish the possibility of difference in the data compared to a method based on the visualization. For example, the pitch was tested with only twenty categories that represents twenty steps for this dimension, compared to one hundred steps for the vertical axle in the Information-Perception Space. Only the timbre seems efficient to represent the height step of its dimension.

9.3 Conclusion

All the methods propose a lot of advantages and disadvantages. One of the main aspects is the possibility to apply the methods in practice. By this argument, the TBP method offers an opportunity compared to the Earbenders that exists for this study only in papers and the rule-based method that needs intellectual efforts to apply it in practice. The TBP method could be applied in most of the examples even if a best mapping could exist and could be based on more than three dimensions. Nevertheless, the TBP method offer an opportunity to map data into sounds without understanding difficult principles and is important in term of time. The only thing to know to apply the TBP prototype is the information type that does not require a long time to understand it.

Further research should be based on this method to discover new pattern with new features of the sound. The experiment part seems important to compare the different propositions on each dimension.

Chapter 10

Conclusion

The objective of the dissertation was to discover in the literature some principles, methods, current models which sound can be used in supporting explorative data mining and analysis. In this recent research, we have found several techniques that could help to translate data into sounds. Some formal methods in developing sound support for data mining has been analyzed. Also, this project was connected with the implementation of a model proposed. The implementation has included a prototype where some key elements of the model were presented and where the features of sound used and the mapping for each data type were described.

Different methods have shown that the main features of the sound (pitch, timbre, amplitude,...) can be exploited to develop some interesting models. These models can be extended until the use of sounds in a 3D world. The advantages of methods based on sound are not negligible despite some drawbacks like the fact of using both visualization and sound can reduce the performance. The plurality of sound features supposes that a lot of different models based on it could be implemented, many possibilities are offered.

The previous approaches have allowed to discover several techniques or principles that are based on organization of auditory elements, metaphorical meaning of the sound, psychoacoustic discrimination of the sounds, cultural and aesthetic implications of the sound, transportability and optimization of the sounds between different devices. Even if these methods are not really some formal methods, they can be considered like a starting point for a future process. They are a gathering of ideas that turn around the sound and its possibilities.

The TaDa method has shown the importance of the scenario description and the requirements analysis. By this way, a case-based synthesis can be released and is based on a database that includes everyday listening. The requirements are used to look-up examples from this database called the Earbenders database. Furthermore, a matching on the fields of the requirements can be applied to find some audio characterizations that have been used in a similar example. Also the audio characterization is a specification for the process design and for a further use. The rule based method is used to look-up a principled mapping of information relations to acoustic relations. It has been thought because of the lack of systematic approach for designing the data-to-sound mapping. From the literature, a taxonomy of mapping has been generalized and these principles can capture and formalize knowledge. These principles are organized around five fields: reading, type, level, organization and range.

Although the rule-based method shows some helpful principles and guidelines. It is difficult to apply in practice without keeping in mind the principles. They have to be learnt and understood. That is why an alternative representation of the principles has been presented. The approach used is similar to the HSS model (Hue, Saturation and Skill) that is commonly used in visualization techniques. The Information-Perception Space makes the Hearsay principles more direct and easy to apply in auditory design practice. According to the information type of the data, the auditory relations are described in terms of timbre, pitch and brightness. These three dimensions form the Information-Sound Space.

A Java application has been developed to illustrate the TaDa method. The application has been described using the different parts of the Tada method: scenario description, requirements for the task analysis, information analysis, data analysis and audio characterization. Also, the application has allowed to illustrate the Hearsay principles and to understand in a more formal way the sounds produced. An experimentation has been realized to test the recognition of the sounds played and to analyze the sensibility to changes in the characteristics of the sounds.

The application and the experimentation have shown the importance of some features present in the sounds produced. We can list here like important characteristic: the frequency, the amplitude, the speed of sound, the depth of sound and the number of times the sound is produced. As seen in the theory, some features can be added like the timbre. The timbre has not been tested in this study because we did not use different instruments to play the sounds. These features that could be the source of a new prototype like the TBP prototype.

Chapter 9 argues around the different methods proposed in the Tada analysis that are the case-based method, the rule-based method and the TBP prototype. All the methods suffer of a few disadvantages that have been listed in this chapter. These disadvantages appear when we try to apply all the methods in a more practical way.

In conclusion, as seen, the Timbre Pitch Brightness Information-Sound Space but also the application implemented show that useful applications can be released based on the properties of sound. Sound proposed a new perspective to treat and analyze data. According to the experiments, a lot of mapping techniques could be released and the mappings shown here in the study represent only a few one and are maybe not the most efficient. According to the number of properties of sound, there are a lot of other possibilities of mapping techniques even if some particular properties seem more suitable than others.

Bibliography

[Anr99] Koen Anrijs, The use of sound in 3d representations of symbolic objects. Master's Thesis, Facultés Universitaires Notre-Dame de la Paix, Namur, 1999

[Alt93] James Alty, Multimedia: We have the Technology but do we have a Methodology?. LUTCHI Research Centre, Educational Media 1993, Available from http://www.dis.port.ac.uk/~callear/CBT82230.htm.

[Bar97a] Stephen Barrass, "Previous approach" in Auditory Information Design. A thesis submitted for the degree of Doctor of philosophy of the Australian National University, pp. 7-27, July 1997

[Bar97b] Stephen Barrass, "Tada: task and data analysis of information requirements" in Auditory Information Design. A thesis submitted for the degree of Doctor of philosophy of the Australian National University, pp. 35-45, July 1997

[Bar97c] Stephen Barrass, "Earbenders: case-based design from stories about listening" in Auditory Information Design. A thesis submitted for the degree of Doctor of philosophy of the Australian National University, pp. 47-67, July 1997

[Bar97d] Stephen Barrass, "Hearsay: principles for auditory information design" in Auditory Information Design. A thesis submitted for the degree of Doctor of philosophy of the Australian National University, pp. 69-88, July 1997

[Bar97e] Stephen Barrass, "Information-Sound Space: a cognitive artifact for auditory display" in Auditory Information Design. A thesis submitted for the degree of Doctor of philosophy of the Australian National University, pp. 89-126, July 1997

[Bar97f] Stephen Barrass, Tada! Demonstration of Auditory Information Design. Available from http://www.icad.org/websiteV2.0/Conferences/ICAD96/proc96/barrass.htm, 1996

[Bar97g] Stephen Barrass, Auditory Information Design. Available from http://viswiz.gmd.de/~barrass/thesis/, 1997

[Barh] Stephen Barrass, Some Golden rule for designing auditory displays.

Available from http://iem.kug.ac.at/ritsch/Vorlesungen/manuals/boulanger/site/2%20cd1/cd%20chapters/barrass/

[Bar95i] Stephen Barrass, Personify: a Toolkit for Perceptually Meaningful Sonification. Proceedings of Australian Computer Music Conference ACMA'95, Australian Computer Music Association, Inc., 1995

[Bar96j] Barrass S., Earbenders: Using Stories About Listening to Design Auditory Displays. Proceedings of the First Asia-Pacific Conference on Human Computer Interaction APCHI'96, Information Technology Institute, Singapore, 1996.

[Bar94k] Barrass S. A Naturally Ordered Geometric Model of Sound Inspired by Colour Theory. Proceedings of Synaesthetica `94, Australian Centre for the Arts and Technology, 1994

[BWE94] Stephen A. Brewster, Peter C. Wright and Alistair D.N. Edwards, The Design and Evaluation of an Auditory Enhanced ScrollBar. Available at http://www.dcs.gla.ac.uk/~stephen/papers/CHI94.PDF, 1994.

[BWE95] Stephen A. Brewster, Peter C. Wright and Alistair D.N. Edwards, Experimentally Derived Guidelines for the Creation of Earcons. Available from http://www.dcs.gla.ac.uk/~stephen/papers/HCI95.pdf, 1995

[Cso] Csound, Available from http://sunsite.univie.ac.at/pub/sound/csound/newest/

[DS01] Christophe Demoulin, Olivier Schöller, Sonification of Time-Dependent Data. Master's Thesis, Facultés Universitaires Notre-Dame de la Paix, Namur, 2000-2001

[FJ01] Doug Fulton and David A. Jaffe, Basic Sound Concepts. Available from http://www.musickit.org/MusicKitConcepts/basicsoundconcepts.html, 1999-2001

[FWG01] Jhon H. Flowers, Laura E. Whitwer, Douglas Grafel & Cheryl A. Kotan, Sonification of Daily Weather Records: Issues for Perception, Attention and Memory in Design Choices. Proceedings of the 2001 International Conference on Auditory Displays, Available at www.acoustics.hut.fi/icad2001/proceedings/ papers/flowers.pdf, 2001

[Gav] William W. Gaver, The SonicFinder, An Interface That Uses Auditory Icons. University of California, San Diego and Apple Computer, Inc,

[Glu00] Mike Gluck, The use of sound for data exploration. Available from http://www.asis.org/Bulletin/June-00/gluck.html, june/july 2000

[FPS] Usama M. Fayyad, Gregory Piateysky-Shapiro, Padhraic Smyth, From Data Mining to Knowledge Discovery: An Overview

[Had01] Tom Hadson, Sound waves and music. Available from http://www.glenbrook.k12.il.us/gbssci/phys/Class/sound/u1111d.html to http://www.glenbrook.k12.il.us/gbssci/phys/Class/sound/u1115d.html, 1996- 2001

[Her99] Thomas Herman, Data Exploration by Sonification. Available from http://www.techfak.uni-Bielefeld.DE/ags/ni/projects/datamining/datason/datason_e.html, 7-14-99

[Hur91] David Huron, Auditory Scene Analysis: The Perceptual Organization of Sound by Albert S. Bregman. Available from http://www.music-cog.ohio-state.edu/Huron/Publications/huron.Bregman.review.html, 1991

[Jav99a] The Java 3D API Specification. Available from http://java.sun.com/products/java-media/3D/forDevelopers/j3dguide/j3dTOC.doc.html, juin 99

[Jav99b] Getting started with the java 3d api, Sun Microsystems, 1999

[Jav00] Java 2 sdk, standard edition documentation, Sun Microsystems, 2000

[Jsy97] Jsyn, Audio Software Synthesis API and Plugins for Java. Available from http://www.softsynth.com/jsyn/, 1997

[KWT+97] Gregory Kramer, Bruce Walker, Terri Bonebright, Perry Cook, John Flowers, Nadine Miner, John Neuhoff, From the sonification report : Status of the Field and Research Agenda. Available from http://www.icad.org.websiteV2.0/References/nsf.html, 1997

[Mad90] Tara Maja Madhyastha, A Portable System for Data Sonification. Thesis submitted in partial fulfillment of the requirements for the degree of Master of Science in Computer Science in the Graduate College of the University of Illinois at Urbana-Champaign, 1992. Available at http://www-pablo.cs.uiuc.edu/Project/VR/DataSonification/Sonification.htm

[Noi00] M. Noirhomme-Fraiture, Le son dans les interfaces IHM: Application à la représentation de données multivariées complexes. 2ème journée multimédia, Namur 2000, Presse Universitaire de Namur

[Por] Porsonify, available at http://nscp.upenn.edu/parallel/www/oldchanges/1995-08.html

[Sto00] M.Storey, Graphic Design and Information Representation for the Web. Available at http://www.csc.uvic.ca/~csc485c/course_notes/graphic_design.pdf., 2000

[Sha97] Shappel, Human Factors in Aviation/Aerospace industry, Week6:Aviation Displays and Cockpit Automation. Available from www.ec.erau.edu/cce/faculty/MAS604_6.doc, 1997

[Tka96] Alejandro Tkaczevski, Auditory Interface Problems and Solutions for Commercial Multimedia Products. Available from http://www.icad.org/websiteV2.0/Conferences/ICAD96/proc96/tka5.htm, 1996

[TS] Cindy Tonneson and Joe Steinmetz, 3D Sound synthesis. Available from www.hitl.washington.edu/scivw/EVE/I.B.1.3DSoundSynthesis.html

[Wol] Joe Wolfe, What is sound spectrum. Available from http://www.phys.unsw.edu.au/~jw/sound.spectrum.html

[Wil] Sheila Williams, Perceptual Principles in Sound Grouping. Available from http://www.create.ucsb.edu/ken/A/aus/williams.html

Appendix A

Basic Concepts

A.1. Introduction

The purpose of this chapter is to recall the reader the main aspects of sound. First of all, we introduce the definition of sound and we identify the features of the sound representation. These features are important when we think in terms of methods and principles as seen in the different chapters. Secondly, we briefly introduce the definition of the sonification and we list the motivations and advantages of the methods based on sound.

A.2. What is sound?

Sound is a physical phenomenon produced by the vibration of matter [FJ01]. The matter can be almost anything: a violin string or a block of wood, for example. As the matter vibrates, pressure variations are created in the air surrounding it. This alternation of high and low pressure is propagated through the air in a wave-like motion.
When the wave reaches our ears, we hear a sound.

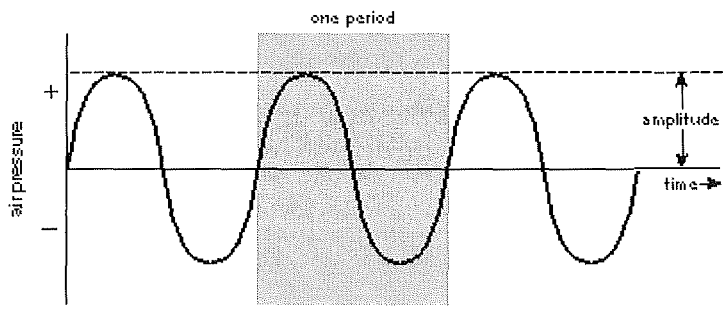


Figure A.1: oscillation of a pressure wave over time.

The pattern of the pressure oscillation over time is called a waveform. The waveform above repeats the same shape at regular intervals; the gray area shows one complete shape. This portion of the waveform is called a **period**. Then, a periodic waveform is a waveform with a clearly defined period occurring at regular intervals. Sound waveforms are never as perfect as shown in Figure 2-1 however sounds that display a recognizable periodicity tend to be more musical than those that are non-periodic. Here are some sources of periodic and non-periodic sounds:

Periodic

- Musical instruments other than unpitched percussion
- Vowel sounds
- Bird songs
- Whistling wind

Non-periodic

- Unpitched percussion instruments
- Consonants, such as "t," "f," and "s"
- Coughs and sneezes
- Rushing water

A sound spectrum is a representation of a sound - usually a short sample of a sound - in terms of the amount of vibration at each individual frequency. It is usually presented as a graph of either power or pressure as a function of frequency. The power or pressure is measured in decibels and the frequency is measured in vibrations per second (or Hertz, abbreviation Hz) or thousands of vibrations per second (kiloHertz, abbreviation kHz) [Wol].

Amplitude
The distance from the center of the wave to its crest is the wave's **amplitude**. Amplitude is a property subjectively heard as **loudness [FJ01]**.

Frequency and pitch
The frequency is the number of times the pressure rises and falls or oscillates in a second is measured in hertz (Hz). A frequency of 100 Hz means 100 oscillations per second. Abbreviation: 1 kHz (kilohertz) equals 1000 Hz.

The frequency range of normal human hearing extends from around 20 Hz up to about 20 kHz. This range is divided into three bands: high frequencies (5-20 kHz), mid frequencies (200 Hz-5Khz) and low frequencies (20-200Hz). Any sound with a frequency below the audible range of hearing (less than 20 Hz) is known as an **infrasound** and any sound with a frequency above the audible range of hearing (more than 20 000 Hz) is known as an **ultrasound.**

As the frequency is simply the reciprocal of the period, a sound wave with a high frequency will correspond to a pressure time plot with a small period and conversely[Had01].

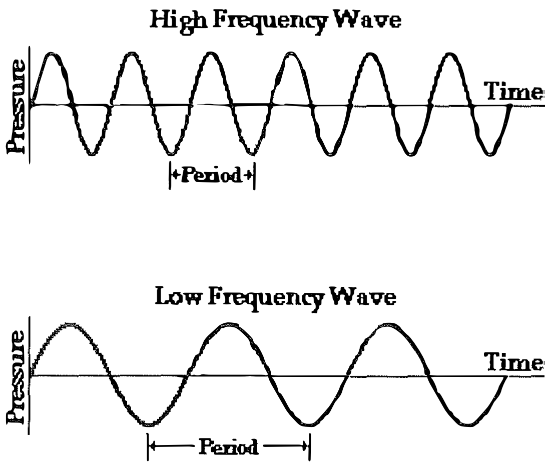


Figure A.2: High and low frequency waves

The sensations of these frequencies are commonly referred to as the **pitch** of a sound. A high pitch sound corresponds to a high frequency and a low pitch sound corresponds to a low frequency. Many people, especially those who have been musically trained, are capable of detecting a difference in frequency between two separate sounds that is as little as 2 Hz. Most people are capable of detecting the presence of a complex wave pattern resulting from the interference and superposition of the two sound waves played simultaneously with a frequency difference of greater than 7 Hz. **Consonant** sound waves are a sound that produce a particularly pleasant sensation when heard. Such sound waves form the basis of **intervals** in music. For example, any two sounds whose frequencies make a 2:1 ratio are separated by an **octave** and result in a particularly pleasing sensation when heard. Similarly two sounds with a frequency ratio of 5:4 are separated by an interval of a **third** and also sound good when played together. A fourth is represented by a 4:3 ratio and a fifth by a 3:2 ratio.

Intensity and decibel scale

The amount of energy that is transported past a given area of the medium per unit of time is known as the **intensity** of the sound wave expressed in Watts/meter^2. [Had01]

$$\text{Intensity} = \frac{\text{Energy}}{\text{Time} * \text{Area}} \quad \text{or} \quad \text{Intensity} = \frac{\text{Power}}{\text{Area}}$$

The amount of energy that is transferred to the medium is dependent upon the amplitude of vibrations. As a sound wave carries its energy through a two-dimensional or three-dimensional medium, the intensity of the sound wave decreases with increasing distance from the source. The mathematical relationship between intensity and distance is sometimes referred to as an **inverse square relationship**.

Humans are equipped with very sensitive ears capable of detecting sound waves of extremely low intensity. The faintest sound that the typical human ear can detect has an intensity of $1*10^{-12}$ W/m^2 (the **threshold of hearing**). This intensity corresponds to a pressure wave in which a compression of the particles of the medium increases the air pressure by a mere 0.3 billionth of an atmosphere. A sound with an intensity of $1*10^{-12}$ W/m^2 corresponds to a sound that will displace particles of air by a mere one-billionth of a centimeter. The most intense sound that the ear can safely detect without suffering any physical damage is more than one billion times more intense than the threshold of hearing.

The scale for measuring intensity is the **decibel scale** and is based on multiples of 10 (logarithmic scale).. The threshold of hearing is assigned a sound level of 0 decibels (0 dB). A sound that is 10 times more intense ($1*10^{-11}$ W/m^2) is assigned a sound level of 10 dB. A sound that is 10*10 or 100 times more intense ($1*10^{-10}$ W/m^2) is assigned a sound level of 20 dB, etc. While the intensity of a sound is a very objective quantity that can be measured with sensitive instrumentation, the **loudness** of a sound is more of a subjective response that will vary with a number of factors. The same sound will not be perceived to have the same loudness to all individuals. Age is one factor that affects the human ear's response to a sound. Obviously, for your grandparents, the same intensity sound would not be perceived to have the same loudness to them as it would to you. Furthermore, two sounds with the same intensity but different frequencies will not be perceived to have the same loudness. Because of the human ear's tendency to amplify sounds having frequencies in the range from 1000 Hz to 5000 Hz, sounds with these intensities seem louder to the human ear. Despite the distinction between intensity and loudness, it is safe to state that the more intense sounds will be perceived to be the loudest sounds.

Speed of sound

The speed of a wave is defined as the distance that a point on a wave travels per unit of time (**speed = distance/time** expressed in units of meters/second (m/s)). [Had01]

Like any wave, a sound wave has a speed that is mathematically related to the frequency and the wavelength of the wave. The mathematical relationship between speed, frequency and wavelength is given be the following equation:

Speed = Wavelength * Frequency

Even though wave speed is calculated using the frequency and the wavelength, the wave speed is **not** dependent upon these quantities. An alteration in wavelength does not affect wave speed. Rather, an alteration in wavelength affects the frequency in an inverse manner. A doubling of the wavelength results in a halving of the frequency.

Interference and Beats

Wave interference is the phenomenon that occurs when two waves meet while travelling along the same medium. The interference of waves causes the medium to take on a shape that results from the net effect of the two individual waves upon the particles of the medium. If two crests having the same shape meet up with one another while travelling in opposite directions along a medium, the medium will take on the shape of a crest with twice the amplitude of the two interfering crests. This type of interference is known as **constructive interference**. If a crest and a trough having the same shape meet up with one another while travelling in opposite directions along a medium, the two pulses will cancel each other's effect upon the displacement of the medium and the medium will assume the equilibrium position. This type of interference is known as **destructive interference**. The diagrams below show two waves - one is blue and the other is red - interfering in such a way to produce a resultant shape in green. [Had01]

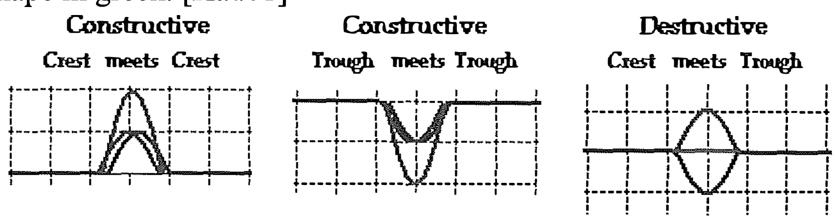


Figure A.3: Constructive and destructive interference

Beats are the periodic and repeating fluctuations heard in the intensity of a sound when two sound waves of very similar frequencies interfere with one another. The diagram below illustrates the wave interference pattern resulting from two waves (drawn in red and blue) with very similar frequencies. A beat pattern is characterized by a wave whose amplitude is changing at a regular rate. Observe that the beat pattern (drawn in green) repeatedly oscillates from zero amplitude to a large amplitude, back to zero amplitude throughout the pattern. Points of constructive interference (C.I.) and destructive interference (D.I.) are labeled on the diagram. When constructive interference occurs, a loud sound is heard; this corresponds to a peak on the beat pattern. When destructive interference occurs, no sound is heard; this corresponds to a point of no displacement on the beat pattern. Since there is a clear relationship between the amplitude and the loudness, this beat pattern would be consistent with a wave that varies in volume at a regular rate. [Had01]

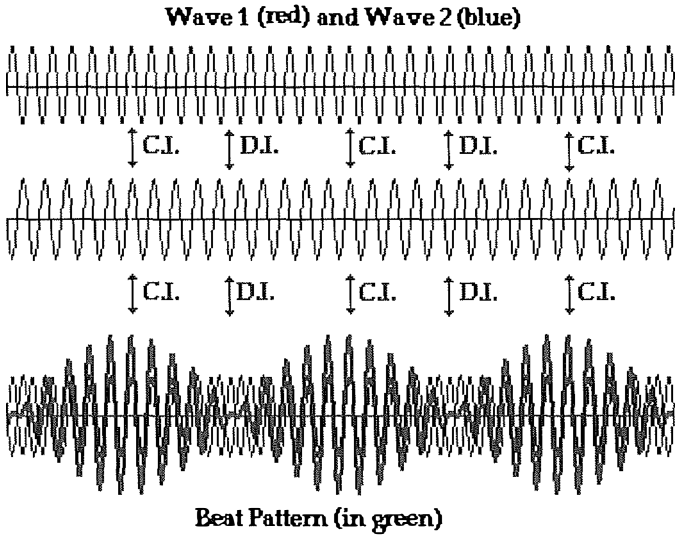


Figure A.4: Beat pattern

The **beat frequency** refers to the rate at which the volume is heard to be oscillating from high to low volume. For example, if two complete cycles of high and low volumes are heard every second, the beat frequency is 2 Hz. The beat frequency is always equal to the difference in frequency of the two notes that interfere to produce the beats. So if two sound waves with frequencies of 256 Hz and 254 Hz are played simultaneously, a beat frequency of 2 Hz will be detected. The human ear is capable of detecting beats with frequencies of 7 Hz and below. [Had01]

Boundary Behavior
As a sound wave travels through a medium, it will often reach the end of the medium and encounter an obstacle or perhaps another medium through which it could travel. When one medium ends, another medium begins; the interface of the two media is referred to the **boundary**. The behavior of a wave (or pulse) upon reaching the end of a medium is referred to as **boundary behavior**. [Had01]
Possible boundary behaviors:

- Reflection (the bouncing off of the boundary)
- Diffraction (the bending around the obstacle without crossing over the boundary)
- Transmission (the crossing of the boundary into the new material or obstacle)
- Refraction (occurs along with transmission and is characterized by the subsequent change in speed and direction).

Reflection of sound waves off of surfaces can lead to one of two phenomenons - an **echo** or a **reverberation**. A reverberation often occurs in a small room with height, width, and length dimensions of approximately 17 meters or less. Why the magical 17 meters? The effect of a particular sound wave upon the brain endures for more than a tiny fraction of a second; the human brain keeps a sound in memory for up to 0.1 seconds. If a reflected sound wave reaches the ear within 0.1 seconds of the initial sound, then it seems to the person that the sound is *prolonged.* The reception of multiple reflections off of walls and ceilings within 0.1 seconds of each other causes reverberations - the prolonging of a sound. Since sound waves travel at about 340 m/s at room temperature, it will take approximately 0.1s for a sound to travel the length of a 17 meters room and back, thus causing a reverberation (t = v/d = (340 m/s)/(34 m) =

0.1 s). This is why reverberation is common in rooms with dimensions of approximately 17 meters or less.

Reflection of sound waves also leads to **echoes**. Echoes are different than reverberations. Echoes occur when a reflected sound wave reaches the ear more than 0.1 seconds after the original sound wave was heard. If the elapsed time between the arrival of the two sound waves is more than 0.1 seconds, then the sensation of the first sound will have *died out*. In this case, the arrival of the second sound wave will be perceived as a second sound rather than the prolonging of the first sound. There will be an echo instead of a reverberation.

Natural Frequency

Nearly all objects, when hit or struck or plucked or strummed or somehow disturbed, will vibrate. If you drop a meter stick or pencil on the floor, it will begin to vibrate. If you pluck a guitar string, it will begin to vibrate. When each of these objects vibrate, they tend to vibrate at a particular frequency or a set of frequencies. The frequency or frequencies at which an object tends to vibrate with when hit is known as the **natural frequency** of the object. If the amplitude of the vibrations is large enough and if natural frequency is within the human frequency range, then the object will produce sound waves that are audible. [Had01]

All objects have a natural frequency or set of frequencies at which they vibrate. The quality or **timbre** of the sound produced by a vibrating object is dependent upon the natural frequencies of the sound waves produced by the objects. Some objects tend to vibrate at a single frequency and they are often said to produce a pure tone. A flute tends to vibrate at a single frequency, producing a very pure tone. Other objects vibrate and produce more complex waves with a set of frequencies that have a whole number mathematical relationship between them; these are said to produce a rich sound (for example: a tuba). Other objects will vibrate at a set of multiple frequencies that have no simple mathematical relationship between them. These objects are not musical at all and the sounds that they create are best described as noise. When a meter stick or pencil is dropped on the floor, it vibrates with a number of frequencies, producing a complex sound wave that is noisy.

Examples of frequency for a flute, a tuba and a dropped pencil [Had01]:

Flute	Tuba	Dropped Pencil
200 Hz	200 Hz	197 Hz
	400 Hz	211 Hz
	600 Hz	217 Hz
	800 Hz	219 Hz
	1000 Hz	287 Hz
		311 Hz
		329 Hz

Forced Vibration - Resonance

Musical instruments and other objects are set into vibration at their natural frequency when a person hits, strikes, strums, plucks or somehow disturbs the object. A person or thing puts energy into the instrument by direct contact with it. This input of energy disturbs the particles and forces the object into vibrational motion - at its natural frequency. The tendency of one object to force another *adjoining* or *interconnected* object into vibrational motion is referred to as a **forced vibration**.

Resonance occurs when two interconnected objects share the same vibrational frequency. When one of the objects is vibrating, it forces the second object into vibrational motion. The result is a large vibration, and if a sound wave within the audible range of human hearing is produced, a loud sound is heard. [Had01]

Standing Wave Patterns

As mentioned earlier, all objects have a frequency or set of frequencies with which they naturally vibrate when struck, plucked, strummed or somehow disturbed. Each of the natural frequencies at which an object vibrates is associated with a standing wave pattern. When an object is forced into resonance vibrations at one of its natural frequencies, it vibrates in a manner such that a standing wave is formed within the object. In that unit, a **standing wave pattern** was described as a vibrational pattern created within a medium when the vibrational frequency of a source causes reflected waves from one end of the medium to interfere with incident waves from the source in such a manner that specific points along the medium appear to be standing still. Such patterns are only created within the medium at specific frequencies of vibration; these frequencies are known as harmonic frequencies or **harmonics**. At any frequency other than a harmonic frequency, the interference of reflected and incident waves results in a resulting disturbance of the medium that is irregular and non-repeating.

So the natural frequencies of an object are merely the harmonic frequencies at which standing wave patterns are established within the object. These standing wave patterns represent the lowest energy **vibrational modes** of the object. While there are countless ways by which an object can vibrate (each associated with a specific frequency), objects favor only a few specific modes or patterns of vibrating. The favored modes (patterns) of vibration are those that result in the highest amplitude vibrations with the least input of energy. Objects favor these natural modes of vibration because they are representative of the patterns that require the least amount of energy. Objects are most easily forced into resonance vibrations when disturbed at frequencies associated with these natural frequencies. [Had01]

Fundamental Frequency and Harmonics

Previously, it was mentioned that when an object is forced into resonance vibrations at one of its natural frequencies, it vibrates in a manner such that a standing wave pattern is formed within the object. Each natural frequency which an object or instrument produces has its own characteristic vibrational mode or standing wave pattern. These patterns are only created within the object or instrument at specific frequencies of vibration (the **harmonics**). At any frequency other than a harmonic frequency, the resulting disturbance of the medium is irregular and non-repeating. For musical instruments and other objects that vibrate in regular and periodic fashion, the harmonic frequencies are related to each other by simple whole number ratios. This is part of the reason why such instruments sound musical rather than noisy.

First, consider a guitar string vibrating at its natural frequency or harmonic frequency. Because the ends of the string are attached and fixed in place to the guitar's structure (the bridge at one end and the frets at the other), the ends of the string are unable to move. Subsequently, these ends become nodes - points of no displacement. In between these two nodes at the end of the string, there must be at least one anti-node. The most fundamental harmonic for a guitar string is the harmonic associated with a standing wave having only one anti-node positioned between the two nodes on the end of the string. This would be the harmonic with the longest wavelength and the lowest frequency. The lowest frequency produced by any particular instrument is known as the **fundamental frequency**. The fundamental frequency is alternatively called the **first harmonic** of the instrument. [Had01]

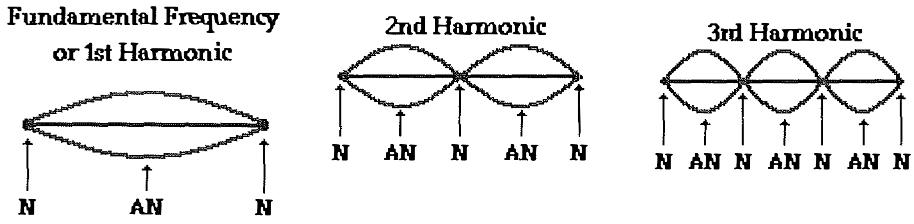


Figure A.5: 1st, 2nd and 3rd Harmonics

Each harmonic results in an additional node and antinode, and an additional half of a wave within the string. If the number of waves in a string is known, then an equation relating the wavelength of the standing wave pattern to the length of the string can be algebraically derived.

The mixture of harmonics determines the **timbre** or quality of sound that is heard. If there is only a single harmonic sounding out in the mixture (in which case, it would not be a mixture), then the sound is rather pure-sounding. On the other hand, if there are a variety of frequencies sounding out in the mixture, then the timbre of the sound is rather rich in quality.

Cues that aid in human sound localization

According to the recent metaphor in 3D space used in data mining, we have to add here some properties of sound in 3D space. These properties could obviously be mixed with the properties seen before. The distance, the direction and the source of the sound seem to be important and could be attached to a special meaning in a 3D world.

In order to gain a clear understanding of spatial sound, it is important to distinguish monaural, stereo, and binaural sound from 3D sound. A **monaural** sound recording is a recording of a sound with one microphone. No sense of sound positioning is present in monaural sound.

Stereo sound is recorded with two microphones several feet apart separated by empty space. Listeners of stereo sound often perceive the sound sources.

Binaural recordings sound more realistic as they are recorded in a manner that more closely resembles the human acoustic system. Binaural recordings are made with the recording microphones embedded in a dummy head, and yield sounds that sound external to the listener's head. Binaural recordings sound closer to what humans hear in the real world as the dummy head filters sound in a manner similar to the human head.

Humans use auditory localization cues to help locate the position in space of a sound source. There are eight sources of localization cues: interaural time difference, head shadow, pinna response, shoulder echo, head motion, early echo response, reverberation, and vision [TS]. The first four cues are considered static and the others dynamic. Dynamic cues involve movement of the subject's body, affecting how sound enters and reacts with the ear.

Interaural time difference describes the time delay between sounds arriving at the left and right ears. This is a primary localization cue for interpreting the lateral position of a sound source. The interaural time delay of sound sources that are directly in front or behind a subject are approximately zero, while sound sources to the far left or right are around 0.63 ms. The frequency of and the linear distance to a sound source factors into the interaural time delay as well.

Head shadow is a term describing a sound having to go through or around the head in order to reach an ear. The head can account for a significant attenuation (reduced amplitude) of overall intensity as well as provides a filtering effect. The filtering effects of head shadowing cause one to have perception problems with linear distance and direction of a sound source.

Pinna response describes the effect that the external ear, or pinna, has on sound. Higher frequencies are filtered by the pinna in such a way as to affect the perceived lateral position, or azimuth, and elevation of a sound source. The response of the pinna "filter" is highly dependent on the overall direction of the sound source.

Shoulder echo

Frequencies in the range of 1-3kHz are reflected from the upper torso of the human body. In general, the reflection produces echoes that the ears perceive as a time delay that is partially dependent on the elevation of the sound source. The reflectivity of the sound is dependent on the frequency; some sources do not reflect as strongly as others. The shoulder echo effect is not a primary auditory cue; others have greater significance in sound localization.

Head Motion

The movements of the head in determining a location of a sound source is a key factor in human hearing and quite natural. Head movement occurs more often as the frequency of a sound source increases. This is because higher frequencies tend to not bend around objects as much and is harder to localize.

Early echo response/reverberation

Sounds in the real world are the combination of the original sound source plus their reflections from surfaces in the world (floors, walls, tables, etc.). Early echo response occurs in the first 50-100ms of a sound life. The combination of early echo response and the dense reverberation that follows seems to affect the judgment of sound distance and direction. Research in this area is still emerging and will hopefully shed some light that will allow more accurate sound synthesis.

Vision helps us quickly locate the physical location of a sound and confirm the direction that we perceive.

All of these cues in some way contribute to the ability to spatially locate a sound in 3D space. 3D sound synthesis needs to deal with these cues in order to provide accurate sound immersion. The difficulty in doing this is great; researchers do not fully understand exactly how the brain interprets the signals it gets from the ear, nor do they understand all of the characteristics that cause sound to be perceived in 3D space. As research continues, we will hopefully gain a better understanding of the human ear and how to emulate it.

A.3 Sonification

We have to introduce what sonification is because it could be defined as an auditory representation of data and is thus importing according to the topic of this study.

Sonification is the transformation of data relations into perceived relations in an acoustic signal for the purposes of facilitating communication or interpretation.

A.4 Motivations of methods based on sound

This is some recent developments from [KWT+97] [Glu00] [Her99] [Noi00] that motivates the use of sound.

- The need to comprehend an abundance of data
- Increasingly powerful and available media techniques concurrent with the flood of data have been the emergence of powerful audio technologies, which are available across a wide range of computer platforms. Scientific visualization was in a similar situation a decade ago with the rapid development of computer graphics technology. The field of sonification is now in a position to leverage the new computer audio technology to solve many existing problems of scientific display. The wide availability of audio technology (e.g., in multimedia computers) makes auditory data representation a viable option for large numbers of users.
- Scientific visualization techniques are often insufficient for comprehending certain features in the data.
  Example: Voyager 2. During the Voyager 2-space mission there was a problem with the spacecraft as it began its traversal of the rings of Saturn. The controllers were unable to pinpoint the problem using visual displays, which just showed a lot of noise. When the data was played through a music synthesizer, a "machine gun" sound was heard during a critical period, leading to the discovery that the problem was caused by high-speed collisions with electromagnetically charged micrometeoroids.
  Although scientific visualization techniques may not yet be exhausted, some believe that we are approaching the limits of users' abilities to interpret and comprehend visual information. Audio's natural integrative properties are increasingly being proven suitable for presenting high-dimensional data without creating information overload for users. Furthermore, environments in which large numbers of changing variables and/or temporally complex information must be monitored simultaneously are well suited for auditory displays.
- Auditory perception is particularly sensitive to temporal characteristics, or changes in sounds over time. Human hearing is well designed to discriminate between periodic and a-periodic events and can detect small changes in the frequency of continuous signals. This points to a distinct advantage of auditory over visual displays. Fast-changing or transient data that might be blurred or completely missed by visual displays may be easily detectable in even a primitive, but well-designed auditory display.
- Unlike visual perception, perception of sound does not require the listener to be oriented in a particular direction. Auditory displays can therefore be used in situations where the eyes are already busy with another task.
  But, the combination of visual and auditory information may prove less effective in some circumstances than information presented in one sensory modality.

- Sound can be used to substitute the visualization and is a promising area for visually impaired users.
- Educational applications are also promising. Studies show that most people can understand trends, clustering, correlations, and other simple statistical features of a data set just as well by listening to it as they could by reading a graph. There are indications that using sonification to present information to students in primary and secondary schools can provide a more engaging learning experience. Rhythm and music are used as a mnemonic device for teaching young students concepts such as the alphabet and the number of days in each month. Similarly, it may be possible to harness the underlying components of this learning dynamic to assist students in grasping more sophisticated concepts such as common curves in calculus or distributions in statistics. Representing concepts and data through sound provides a means of capitalizing on strengths of individual learning styles, some of, which may be more compatible with auditory representations than more traditional verbal and graphical representations. Adult education, training and generalized information presentation may likewise benefit.
- Sounds are easily used as an alerting tool to quickly gain an analyst's attention to a strange detail or emergency condition. Tools of this sort have been called earcons, serving the purpose of icons to indicate important events. Sound can also be used to locate an object much as we hear a train coming and going. One minor but far from insignificant use of sonification is for those with visual impairments.
- The use of sound in graphic interface presents a lot of advantages. This is a few advantages. Sound can be used in parallel with the graphics. It disposes of a large dimensionality because of all its features. It can be heard in every position. The ears are more sensitive than the changes in the vision and can be used to detect some trends and relations in data sets. A message heard is more held back in memory than a message seen on the screen.

To give an easy summary of the promise of sound, we will list here a few objectives:

- Parallel listening (ability to monitor and process multiple auditory data sets).
- Rapid detection (especially in high-stress environments).
- Affective response (ease of learning and high engagement qualities).
- Auditory gestalt formation (discerning relationships or trends in data streams).

A.5 Conclusion

Sound properties offer interesting ways of communicating and interacting with the computer. Different models could be implemented using these features. Sound become an important support for data analysis and can extend the methods based on visual displays and even replace them. We can think in terms of discovering patterns, information, trends,... in data represented as sound by using characteristics like frequency, amplitude, speed, ...

Appendix B

Answers

The following pages show the answers of the two questions stated in chapter 7. The two questions were : Could you say which phenomena you are listening in this sequence? Could you describe the evolution of the sound in the sequence?

Nom: Cathy Javaux
File 1. On dirait un peu le feu de réacteurs. Il s'intensifie tout en restant continu.
File 2. On dirait le chant d'un oiseau. Le 1er son est fort. Le son augmente et devient plus rapide.
File 3.: On dirait du vent; au départ le son vient de loin on a l'impression qu'il vient de l'extérieur. Il y a 3 intensités différentes, elles sont de plus en plus forte. La troisième est plus aiguë.

Nom: Sébastien Clarinval
File 1. le premier son me fait penser à une tornade que l'on aurait enregistré avec un micro. On dirait que le son vient de loin au départ puis qu'il se rapproche de plus en plus et devient de plus fort ou violent. plus ca se rapproche plus on dirait une tornade.
File 2. le second me fait penser a une flûte mal réglée au départ puis à un oiseau par après. Le cris au départ sonne relativement faux puis il devient de plus en plus répétitif et de plus en plus strident.
File 3. le troisième me fait penser à du vent. On dirait que le vent souffle de plus en plus fort et devient de plus en plus proche. le volume du vent augmente par à coup et par dégradé continu.

Nom: Maxime Tihon
File 1. un avion ou une fusée qui décolle, son de combustion plus frottement de l'air de plus en plus aigu
File 2. oiseau qui siffle, son de plus en plus aigu et accéléré
File 3. vent qui souffle, diverses variations du souffle de plus en plus fort, violent et aigu

Nom: Moreau Florence
File 1. j'entends des impacts, petits au début puis la fréquence augmente
File 2. Il y a 3 sons, on est passé du grave à l'aigu, ca va du plus lent au plus rapide, du plus court au plus long, le 1er son parait artificiel, le 2eme et 3eme semblent être des chants d'oiseaux, les sons sont de plus en plus différents.
File 3. C'est un bruit de vent, le volume augmente 2 fois, il y a 3 sons au total, on a l'impression que c'est de plus en plus rapide et plus violent.

Nom: Dury Jeannine
File 1. ??? moteur, chevaux qui galopent, ca va de calme à violent, plus en plus rapide
File 2. C'est un oiseau, ca va de grave à aigu et de lent à plus rapide
File 3. C'est du vent, ca va de plus en plus fort et c'est de plus en plus violent

Nom: Warin Christian
File 1. ???, de plus en plus fort et plus rapide
File 2. C'est un oiseau, de plus en plus aigu, de plus en plus vite
File 3. C'est le vent, de plus en plus fort et plus violent

Nom: Solarski Johan
File 1. pas facile, des chevaux au galop ou de l'eau qui bout, le son s'accélère
File 2. le chant d'un oiseau, le son évolue en allant de + en + aigu et de + en + vite
File 3. Le son me fait penser à du vent qui souffle. Le son évolue en allant de + en + fort

Nom: Karl Barbier
File 1. tempête ou orage, le son commence de loin puis se rapproche, le bruit est grave et sourd, il va de + en + vite
File 2. C'est un oiseau, le son passe de grave à aigu et est de + en + rapide
File 3. C'est le vent, on a l'impression que ça part d'assez loin et que ça se rapproche de plus en plus, ça souffle de plus en plus vite et de plus en plus fort

Nom: Frédéric Bartiaux
File 1. ronronnement du tonnerre, ça passe de grave à aigu, le son change en 2 temps, il y a 3 phases et ça accélère
File 2. sifflement d'oiseaux, la durée augmente
File 3. vent, à nouveau 3 phases, c'est de plus en plus intense en volume et violent

Nom: François Gérard
File 1. genre de grondements qui deviennent de plus en plus aigus toutes les 9 secondes
File 2. des oiseaux, vitesse et nombre de bruits augmentent à chaque fois
File 3. vent dont la puissance augmente chaque fois

Nom: Jan Minet
File 1. le premier me fait penser a un bruit de réacteur d'avion, le bruit me semble continu et toujours de la même intensité, ca semble accélérer.
File 2. le début du son me fait penser à un genre de sifflet tandis que la fin me ferait plutôt penser à un oiseau qui chante. Le son change constamment de volume , la vitesse change...
File 3. Le troisième me fait penser au soufflement du vent pendant une tempête. En première partie le son est faible et continu ensuite il devient plus fort d'un seul coup. Apres le son est encore une fois augmenté d'un seul coup.

Nom: Cynthia Solarski
File 1. Ca me fait penser à une grosse pierre qu'on fait rouler à terre, il y a de plus en plus de présence, c'est de plus en plus rapide, ça passe de grave à plus aigu
File 2. Le 1er son me fait penser à l'interférence d'un gsm lorsque tu le mets contre un téléphone fixe puis ce sont des oiseaux, c'est aigu, ça passe de lent à rapide
File 3. C'est le vent, ça souffle de plus en plus fort, + de volume et + rapide

Nom: Rémy Rondeux
File 1. Ca me fait penser à de la pluie, ça tombe de + en +, de + en + de bruit et c'est + violent
File 2. Ce sont des oiseaux, c'est de + en + vite et + long
File 3. C'est le vent, c'est de + en + fort en volume, il y a 3 temps, + en + violent

Nom: Joseph Robert
File 1. Ca me fait penser à du popcorn dans une casserole ou quelque chose qui frétille, ça frétille de + en +, il y a plus d'impacts et de vitesse
File 2. Ce sont des oiseaux, ca va de + en + aigu et + rapide
File 3. C'est le vent, c'est de + en + fort et violent, à la fin on dirait une tornade

Nom: Amori Astire
File 1. On dirait de l'eau qui bout, ca va de + en + fort
File 2. Des oiseaux, c'est de + en + aigu et de + en + insistant
File 3. C'est le vent, c'est de + en + violent et de + en + fort

Nom: Thierry Lambilotte
File 1. Le son évoque celui d'un grésillement, voire d'une pluie sur un toit ou une fenêtre. Le son augmente de manière progressive en trois étapes. La changement entre les phases est audible, mais pas brusque.
File 2. Le son est composé de plusieurs sifflements. Le premier monte dans les aigus et ensuite redescend dans les basses. Le second est constitué de trois sons aigus semblables puis d'un "sifflement d'oiseau". Le troisième est composé de sifflements plus aigus et rapides. Les trois sons se terminent par un fondu.
File 3. Le son est celui du vent. Il y a trois phases pendant lesquelles l'intensité du son augmente. Les changements entre les phases sont fort marquées, sans transition.