Human–Computer Interaction Series

David Worrall

Sonification Design

From Data to Intelligible Soundfields

Springer

Human–Computer Interaction Series

Editors-in-Chief

The Human-Computer Interaction Series, launched in 2004, publishes books that advance the science and technology of developing systems which are effective and satisfying for people in a wide variety of contexts. Titles focus on theoretical perspectives (such as formal approaches drawn from a variety of behavioural sciences), practical approaches (such as techniques for effectively integrating user needs in system development), and social issues (such as the determinants of utility, usability and acceptability).

HCI is a multidisciplinary field and focuses on the human aspects in the development of computer technology. As technology becomes increasingly more pervasive the need to take a human-centred approach in the design and development of computer-based systems becomes ever more important.

Titles published within the Human–Computer Interaction Series are included in Thomson Reuters' Book Citation Index, The DBLP Computer Science Bibliography and The HCI Bibliography.

More information about this series at http://www.springer.com/series/6033

David Worrall

Sonification Design

From Data to Intelligible Soundfields

Springer

David Worrall
Department of Audio Arts and Acoustics
Columbia College Chicago
Chicago, IL, USA

ISSN 1571-5035 ISSN 2524-4477 (electronic)
Human–Computer Interaction Series
ISBN 978-3-030-01496-4 ISBN 978-3-030-01497-1 (eBook)
https://doi.org/10.1007/978-3-030-01497-1

This Springer imprint is published by the registered company Springer Nature Switzerland AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

For Bek
My Angkor, Wat

Preface and Acknowledgements

This book is about the contemporary design practice known as *data sonification,* which assists us to experience information by listening, much as we understand relationships between spatial features by viewing a graph of them. Data sonification begins with the observation that sounds can convey meanings in a multiplicity of circumstances: Exploring the structure of deep space, detecting a stock-market bubble, assisting injured patients to recover more quickly with less pain, monitoring the flow of traffic to help detect potential congestion, improving sporting performance, tracking storms, earthquakes and other changes in the environment … and the list could go on: hundreds of applications that rely on studies in acoustics and psychoacoustics, philosophy of perception, cognitive psychology, computer science, creative auditory design and music composition.

Data sonification grew out of an interest by composers in generating musical forms with the assistance of computers: *algorithmic compositions* based on both traditional musical languages, and the exploration of mathematical models of natural and abstract worlds. With ears searching for new means of expressing the contemporary world, they explored such fields as fractal geometry, neural networks, iterated function and reaction–diffusion systems, and flocking and herding. As computer processing speeds and storage capacity increased, and digital networks evolved, it became possible to use large datasets from real and real-time systems, not just abstract idealized models, and this led us into the currently emerging era of Big Data.

One of the motivations for this book was to understand some of the perceptual and conceptual correlates of intelligible data sonification so as to encapsulate the knowledge-bases that underpin them in software design. For example, the psychoacoustic, gestural and psycho-physiological substrates such as cognitive-load sensitivity and emotional valence, with low-level latent functions that can be compiled into higher level interactive modelling tools. Design is an inherently 'messy' and iterative activity that, while a process, may never be entirely procedural. So, the purpose here is not to trivialize the skills of an experienced designer, but to hierarchize the masking of many of the functional decisions made in designing, including such processes as equal-loudness contouring and modal convex pitch and time transforms. It is hoped that in doing so, novice designers might consider more adventurous possibilities and experienced designers will be enabled to implement complex procedures more flexibly: to test multiple different approaches to sonifying a dataset.

Part I: Theory

Chapter 1: The idea that sound can reliably convey information predates the modern era. The term data sonification has evolved along with its applications and usefulness in various disciplines. It can be broadly described as the creation and study of the aural representation of information, or the use of sound to convey non-linguistic information. As a field of contemporary enquiry and design practice, it is young, interdisciplinary and evolving; existing in parallel to the field of data visualization, which is concerned with the creation and study of the visual representation of information. Sonification and visualization techniques have many applications in 'humanizing' information, particularly when applied to large and complex sets of data. Drawing on ancient practices such as auditing, and the use of information messaging in music, this chapter provides an historical understanding of how sound and its representational deployment in communicating information has changed. In doing so, it aims to encourage critical awareness of some of the socio-cultural as well as technical assumptions often adopted in sonifying data, especially those that have been developed in the context of Western music of the past half-century or so. Whilst acknowledging the Eurocentricity of the enquiry, there is no suggestion that the ideas discussed do not have wider applicability.

Chapter 2: Encompassing ideas and techniques from music composition, perceptual psychology, computer science, acoustics, biology and philosophy, data sonification is a multi- even trans-disciplinary practice. This Chapter summarizes different ways sonification has been defined, the types and classifications of data that it attempts to represent with sound, and how these representations perform under the pressure of various real-world utilizations.

Chapter 3: One task of data sonification is to provide a means by which listeners can obtain new ideas about the nature of the source of derived data. In so doing they can increase their knowledge and comprehension of that source and thus improve the efficiency, accuracy and/or quality of their knowledge acquisition and any decision-making based on it. The purpose of this chapter is to develop an historical understanding of what information is as a concept, how information can be represented in various forms as something that can be communicated with non-verbal sonic structures between its source and its (human) receiver and thus retained as knowledge. Whilst a complete philosophical and psychological overview of these issues is outside the scope of the chapter, it is important, in the context of developing computational design strategies that enable such communication, to gain an understanding of some of the basic concepts involved. A quasi-historical epistemology of human perception and the types of information these epistemologies engender is followed by a discussion of the phenomenal nature of sounds and sonic structures their ability to convey information of various sorts.

Chapter 4: The previous chapter traced a path towards an understanding of the inadequacy of epistemological approaches to knowledge formation that do not account for the deeply embodied nature of perception, to succeed in solving any but highly constrained problems. The slower-than-expected rise in the effectiveness of artificial intelligence provided the impetus for a more critical examination of the ontological dimensions of perception and human knowledge, which necessitated the recognition of another form of truth which is not derived empirically but from meaningful action. This chapter enunciates this Pragmatist approach as it can be applied to sonification design: a generative activity in which bespoke design skills are supported by scientific research in biology, perception, cognitive science—in the field of conceptual metaphor theory in particular, and aesthetics. It proceeds to outline pertinent features of a broad design methodology based on the understandings developed that could yield to computational support.

Chapter 5: The need for better software tools for data sonification was highlighted in the 1997 *Sonification Report*, the first comprehensive status review of the field which included some general proposals for adapting sound synthesis software to the needs of sonification research. It outlined the reasons the demands on software by sonification research are greater than those afforded by music composition and sound synthesis software alone. As its *Sample Research Proposal* acknowledged, the development of a comprehensive *sonification shell* is not easy and the depth and breadth of knowledge, and skills required to effect such a project are easily underestimated. Although many of the tools developed to date have various degrees of flexibility and power for the integration of sound synthesis and data processing, a complete heterogeneous Data Sonification Design Framework (DSDF) for research and auditory display has not yet emerged. This chapter outlines the requirements for such a comprehensive framework, and proposes an integration of various existing independent components such as those for data acquisition, storage and analysis, together with a means to include new work on cognitive and perceptual mappings, and user interface and control, by encapsulating them, or control of them, as *Python libraries,* as well as a wrappers for new initiatives, which together, form the basis of *SoniPy*, a comprehensive toolkit for computational sonification designing.

Part II: Praxis

Chapter 6: Having established the design criteria for a comprehensive heterogeneous data sonification software framework in the previous chapter, this chapter introduces two pillars of such a framework, the *Python* and *Csound* programming languages, as integrated through a *Python–Csound* Application Programming Interface. The result is a mature, stable, flexible and comprehensive combination of tools suitable for real and non-realtime sonification, some of the features of which are illustrated in the examples of subsequent chapters.

Chapter 7: Despite intensive study, a comprehensive understanding of the structure of capital market trading data remains elusive. The one known application of audification to market price data reported in 1990 that it was difficult to interpret the results, probably because the market does not resonate according to acoustic laws. This chapter illustrates some techniques for transforming data so it *does* resonate; so audification may be used as a means of identifying autocorrelation in trading–and similar–datasets. Some experiments to test the veracity of this process are described in detail, along with the computer code used to produce them. Also reported are some experiments in which the data is sonified using a homomorphic modulation technique. The results obtained indicate that the technique may have a wider application to other similarly structured time-series datasets.

Chapter 8: The previous chapter explored the use of audification of a numerical series, each member representing the daily closing value of an entire stock market, to observe the cross-correlation (trending) within the market itself. This chapter employs parameter-mapping sonification to study the perceptual flow of all trades in individual stock groupings over a trading day by applying various filters and selection methodologies to a detailed 'tick' dataset. It outlines the use of a size/mass metaphorical model of market activity and the simultaneous use of two opposing conceptual paradigms without apparent conceptual contradiction or cognitive dissonance to demonstrate, given conducive conditions, the power of intention over perception and sensation, in auditory information seeking.

Chapter 9: The design of a real-time monitor for an organization's digital network can produce several significant design challenges, both from the technical and human operational perspectives. One challenge is how to capture network data with minimal impact on the network itself. Also, from an operational perspective, sounds need to perform *en suite* over long periods of time while producing only minimal listener fatigue. This chapter describes two related network data sonification projects which resulted in a set of audiovisual "concert" compositions (*Corpo Real*), an immersive installation, and a perceptual monitoring tool (*Netson*). This tool uses both sonification and visualization to present monitoring humans with features of data flow that allow them to experience selectable operational network characteristics. In doing so, it can be used to assist in the peripheral monitoring of a network for improved operational performance.

Code and audio examples for this book are available at https://github.com/david-worrall/springer/.

Acknowledgements

Gregory Kramer had a particular vision and commitment to establishing auditory display as a legitimate discipline. His organizing of the first conference in 1992, followed by the editing and publication of the extended proceedings, *Auditory*

display: Sonification, Audification, and Auditory Interfaces, produced a go-to reference for researchers in the field until the publication, in 2011, of *The Sonification Handbook*, with major contributions by many members of the community under the inciteful editorship of Thomas Hermann, Andy Hunt and John Neuhoff.

Over the past 10 years or so, various strands of the work in this book have appeared in papers for the International Conference for Auditory Display and I am grateful to many of the members of that diverse community for their annual collegiality, vigorous debate and general bonhomie. In 2009, my first attempt at a succinct overview of the field (Chap. 2) was published in *The Oxford Handbook of Computer Music and Digital Sound Culture*, edited by Roger Dean. Roger was brave enough, with Mitchell Whitelaw, to supervise my Ph.D. which eventually also formed the foundation for parts of Chaps. 3, 5, 7 and 8, for the latter of which, the Capital Markets Cooperative Research Centre in Sydney funded the experiments and provided the data. The securities trading data techniques discussed in Chap. 7 were first published in the Springer's 2010 *Lecture Notes in Computer Science* volume on *Auditory Display*. The research and development for the network sonifications reported in Chap. 9 was undertaken, in addition to other work with Norberto Degara, during 2013–16, whilst a Professorial Research Fellow, at Fraunhofer IIS, in Erlangen, Germany, at Frederik Nagel's International Audio Laboratories, under the leadership of the Institute Director, Albert Heuberger.

I am often struck by how even a small sense of the influences on a writer can provide meaningful insights into their work. In that spirit, I mention some. However, lest it resulted in too autobiographical an appearance, I omit details of years of music-making and restless inquisitiveness in the arts of mathematics and animal anatomy. I spent the mid-1980s as a member of the Composition Department in the Faculty of Music, and in the Computer Science Department at The University Melbourne, followed by 15 years of research and teaching at the School of Music and the Australian Centre for the Arts and Technology (ACAT) at the Australian National University in Canberra, and, more recently, in the Audio Arts and Acoustic Department at Columbia College Chicago. In order not to draw the wrath of any individual inadvertently missed from what would be a long list, most of the names accompany mine on conference papers, in journal articles and concert programs, so I defer to another time to name them all individually. I have been fortunate to work with a few people who have dedicated their talents to working behind the scenes in technical and assistive capacities, and without whom everything would have ground to a halt: Les Craythorn in Melbourne, Niven Stines and Julie Fraser in Canberra, and David Knuth and Maria Ratulowska in Chicago.

This work is grounded in an intellectual and artistic experimental tradition, from which I have been blessed with more than my fair share of excellent mentors. Out of respect for those traditions and in honor of them, I also invoke the spirits of those who have passed, thus: Richard Meale (and through him), Winifred Burston, Ferruccio Busoni, Frans Liszt, Carl Maria von Weber, Carl Philipp Emmanuel Bach and his father Johan Sebastian, John Bull, Carlos Gesualdo … That thread has been crisscrossed in my own life by various others, including Tristram Cary, Iannis Xenakis, Olivier Messiaen and Jean-Claude Risset. Richard was a mentor, friend

and as fierce critic as he was an experimentalist: in music, chess and cooking. Although we fought bitterly as he was overcome by the affliction of postmodernism in his latter years, he wrote some of the best music of his generation. He is deeply missed.

Thanks go to the editorial staff at Springer for their encouragement and long-suffering: tolerance way beyond reasonable expectations. I was not to know, at the time of discussing publishing with them in 2015, that this book would be written in ten residences on three continents. That it has appeared at all is a minor miracle, performed by my beautiful, gracious, strong and unbelievably perceptive wife, Rebekah, who, with Isaac and Catheryn have been my constant companions and family support throughout. As we have lived out our semi-nomadic existence, their day-long 'visits' to the local library, wherever we were, so "Dad could work on *the book*" has not been easy for them, and we're looking forward to a summer of bikes and music-making.

Oak Park, Illinois David Worrall
March 2019

> I like to listen. I have learned a great deal from listening carefully. Most people never listen.
>
> (Ernest Hemmingway, my Oak Park neighbor before I moved in.)

Contents

List of Figures

List of Tables

Table of Code Examples

Table of Audio Examples

Part I
Theory

Chapter 1
Data Sonification: A Prehistory

Yasyāmataṁ tasya mataṁ; mataṁ yasya na veda saḥ
["One who (thinks he) knows not, knows; one who (thinks he)
knows, knows not." (Muktananda 1972)].

Abstract The idea that sound can reliably convey information predates the modern era. The term *data sonification* has evolved along with its applications and usefulness in various disciplines. It can be broadly described as the creation and study of the aural representation of information, or the use of sound to convey non-linguistic information. As a field of contemporary enquiry and design practice, it is young, interdisciplinary and evolving; existing in parallel to the field of data visualization, which is concerned with the creation and study of the visual representation of information. Sonification and visualization techniques have many applications in "humanizing" information, particularly when applied to large and complex sets of data. Drawing on ancient practices such as auditing, and the use of information messaging in music, this chapter provides an historical understanding of how sound and its representational deployment in communicating information has changed. In doing so, it aims to encourage critical awareness of some of the socio-cultural as well as technical assumptions often adopted in sonifying data, especially those that have been developed in the context of Western music of the last half-century or so. Whilst acknowledging the Eurocentricity of the enquiry, there is no suggestion that the ideas discussed do not have wider applicability.

1.1 An Ancient and Modern Practice

We are at a time in the evolution of humans and their tools when the power of digital information processing and algorithmic decision-making is demonstrating an ability to radically change our lives: From genetic finger-printing, gene-splicing and pharmacology, to driverless vehicles, patterns in our consumption and how we amuse ourselves. Even now, so early in this new Dataist era, organizations with networked computational intelligence, already have access to more data about

D. Worrall, *Sonification Design*, Human–Computer
Interaction Series, https://doi.org/10.1007/978-3-030-01497-1_1

ourselves than we ourselves have access to, and are beginning to demonstrate a power to make better decisions for us than we make for ourselves. What, then, one might reasonably ask, is the use of exploring such ancient, intensely human-centered approaches to information-gathering and decision-making as listening? How old-fashioned; how quaint! This is an increasingly pertinent question and has been latently fundamental to why this book was written. The answers are not necessarily obvious as they lie at the heart of the difference between a conception of life merely in terms of information flow and data storage as might be imagined by the cognitivists (Worrall 2010), and one in which mind, body and (now technologically enhanced) consciousness play a fundamental role in active perception, knowledge acquisition, meaning-creation and decision-making.

In a contemporary media-saturated environment, sound plays a wider variety of different social and communicative roles today than it ever did in the past. Although computer-generated data sonification is a relatively recent formalized practice,[1] cultural antecedents can be identified in all periods of history. This chapter provides a focused description of how the deployment of sound and sonic structures to communicate information has changed over time. It the spirit of the adage "know from whence you came", doing so will assist us to critically examine some of the socio-cultural assumptions we have adopted in our own age.[2]

In the process of uncovering the accuracy or truth[3] of our assumptions about perception, it is not uncommon for commentators to be seduced into a type of sense war, in which hearing and listening are pitted against vision and seeing. Such casting can take many forms. What is to be gained by Marshal McLuhan's myopia, for example?

> The ear is hypersensitive. The eye is cool and detached. The ear turns man over to universal panic while the eye, extended by literacy and mechanical time, leaves some gaps and some islands free from the unremitting acoustic pressure and reverberation (McLuhan 1968, 168).

Thankfully, the discussion has moved on, at least in some circles.[4] Each have their place and so to be encumbered with such a burden is not useful, particularly when a culturally-driven focus on one sense assumes a decline of another. While this book

[1]The first international conference was held in 1992 (Kramer 1994a).

[2]As Jonathan Sterne indicates, while there is a vast array of literature on the history and philosophy of sound, it is without some kind of overarching, shared sensibility about what constitutes the history of sound, sound culture, or sound studies (Sterne 2003). It is not the purpose of this chapter to remedy that!

[3]Truth, as in the Heideggerian meaning of the term *alétheia*: a non-propositional unconcealment. This concept of 'truth' takes on the dynamic structure of uncovering that is disclosive rather than propositional or judgmental.

[4]A comprehensive survey of increasingly nuanced arguments is outside the confines of this work. In its notes and bibliography, *The Audible Past* Stern (2003) lists a significant amount of literature in English on the topic since the Enlightenment. Over a longer historical timeframe, Veit Erlmann's, *Reason and Resonance* traces historical changes in the understanding of the relationship between developing conceptions of sound and hearing physiology (2010).

is about the transmission of intelligibility through sonic structures, it does not contend that sound should replace vision in enquiry, any more than the assertion that music expression is best when unencumbered by the "strictures" of musical notation. This is not to deny that there are sensory differences; they all assist us differently, else it is unlikely evolution would have sustained their individual continuance. In fact, it can be instructive to identify them. For example, it is hearing, not vision, that affords omnidirectional coherent perceptual experiences. For all moving creatures, including our ancestors, both ancient and modern, in situations where sight is obscured, spatial auditory clarity plays a vital survival role in determining both from where the predator is approaching or to where the prey has escaped. On the other hand, we can suppose that, to a creature sleeping in a cave, being alerted early by the amplifying echoic resonance of the space that something of a certain mass was entering, was more important than details of its exact position. These evolutionary adaptations are sometimes very deep in our biology, such as the presence of defensive startle reflexes which are very resistant to habituation, and the fascinating orienting reflex (eliciting a "what is it?" reaction) (DeGangi 2017, 309–60; Sokolov et al. 2002) which have cultural resonances such as, for example, when the biological preferencing of the ears to lead the eyes[5] forms the basis of as a scene transition technique to enhance narrative continuity in film.

While sensory differences do exist, it is instructive to consider that the domination of one sense over another—especially when considering hearing and vision—is not a biological phenomenon, but a cultural one. As will be seen in the upcoming discussion of the role of sound in the Church's dominion over Europe, when the control of the whole community was through the sense of hearing, such control was more easily broken away from by individual 'voices' using another sense (vision) and, supported by development in visual technologies, which in-turn, through notation, supported the development of individual musical voices that were able to create complex sonic 'inputs' to the resonant cathedrals while maintaining 'signal' coherence which supported understanding.

1.2 Ancient Egyptian Use of Sound to Convey Information

There are many reasons why, in certain circumstances, sound might be the preferred representation and communication medium for information, including the known superiority of the hearing sense to discriminate particular kinds of structures. For example, it is easy for most of us to personally verify that a purely visual side-by-side comparison of two sets of written records requires high levels of concentration and that doing so is very prone to error, especially over extended periods of time. One the other hand, listening to vocalizations of such

[5]Resulting, presumably, from the superiority of the omni-directionality of hearing in visually-obscured environments.

representations is much easier. The presence of such auditing[6] can be inferred from records of Mesopotamian civilizations going as far back as 3500 BCE. To ensure that the Pharaoh was not being cheated, auditors compared the "soundness" of meticulous independently-scribed accounts of commodities such as grains moving in and out, or remaining, in warehouses (Boyd 1905). When granary masters, otherwise strictly isolated from each other, assembled before the Pharaoh and alternated in their intoning of such records, differences in their accounting records could be easily identified aurally. A faster and more secure method that eliminates any "copy-cat" syndrome in such alternation, is to have the scribes read the records simultaneously—a type of modulation differencing technique. Although we have no hard evidence that these techniques were practiced, such a suggestion does not seem unreasonable, and would represent possibly the earliest form of data sonification.

1.3 The Ancient Greek Understanding of Music

While sound has also played an important role in both theoretical and empirical inquiry for millennia, the ancient Greeks wrote extensively on the subject, notably in reference to music, the most complex and abstractly considered form of non-linguistic aural communication made by humans. It can be divided into three modes of enquiry that are of direct concern to sonification: numerical rationality, empirical experience and expressive power.

1.3.1 Numerical Rationality

At least as far back as Pythagoras (born $\sim$ 569 BCE), arithmetic was considered to be number in itself, geometry to be number in space, and harmony to be number in time. The concept of The Harmony of the Spheres, in which the Sun, Moon and planets emit their own unique "sounds" based on their orbital revolution,[7] played a unifying role in the development of the arts and sciences, and incorporated the metaphysical principle that mathematical relationships express qualities or "tones"

---

[6]Literally, *the hearing of accounts* from the Latin *auditus*.

[7]Known generally as the "music of the spheres". As Gaius Plinius Secundus observed, "… occasionally Pythagoras draws on the theory of music, and designates the distance between the Earth and the Moon as a whole tone, that between the Moon and Mercury as a semitone, …" the seven tones thus producing the so-called diapason, i.e. a universal harmony (Pliny [77AD] 1938). Ptolemy and Plato also wrote about this practice.

of energy ratios. Pythagoras' approach to music was as a numerically rational analysis of the way string lengths and sounds relate to each other physically, that is *acoustically*, and importantly, he classified the sensation of harmoniousness according to these ratios.

The application of number relations (i.e. ratios) and sound have been integral to the conceptualization and realization of Western music over all periods in radically different ways: From the development of richly fecund investigations in tuning and temperament to multiple voice polyphony and highly chromatic polytonalities; from serialized additive rhythms under group theory transformations to stochastic mappings controlled by Poisson's distribution.

1.3.2 Empirical Experience

Aristoxenus of Tarentum (c. 375–335 BCE) was a pupil of Aristotle. With the exception of his treatise on harmony, (Macran 1902) most of his writings have been lost. In contrast to Pythagoras, Aristoxenus' approach was more concerned with the structure of the listening *experience*, which he explained in terms of the various modes.[8] A pupil of his musician father and the later Pythagoreans who were keen on distilling their inherited scientific knowledge from its more mystical entrapments, Aristoxenus' own writings on music are somewhat empirical, perhaps influenced by Aristotle, with whom he also studied. He maintained, in contradistinction to the Pythagoreans, that the notes of the scale could be tuned by the ear rather than ratio measurement, and formulated a theory that the soul is related to the body as harmony to the parts of a musical instrument.[9] Here, for example, is his description of vocal pitch inflection such as glissandi:

> The continuous voice does not become stationary at the "boundaries" or at any definite place, and so the extremities of its progress are not apparent, but the fact that there are differences of pitch is apparent…; for in these cases we cannot tell at what pitch the voice begins, nor at what pitch it leaves off, but the fact that it becomes low from high and high from low is apparent to the ear. In its progress by intervals the opposite is the case. For here, when the pitch shifts, the voice, by change of position, stations itself on one pitch, then on another, and, as it frequently repeats this alternating process, it appears to the senses to become stationary, as happens in singing when we produce a variation of the mode by changing the pitch of the voice. And so, since it moves by intervals, the points at which it begins and where it leaves off are obviously apparent in the boundaries of the notes, but the intermediate points escape notice… (Vitruvius 1914: Chap. 4).

[8]Xenakis (1971, 183–189) has a more detailed explanation of Aristoxenus' modal thinking.

[9]The following quote from Vitruvius '*De architectura*' (Book V Chap. 4) contains a paraphrase of a extant fragment of a treatise on meter in writings on music that is attributable to Aristoxenus.

1.3.3 Expressive Power

Aristotle (384–322 BCE) was interested in the ability of sound (music and poetry), to *express* states of mind and evoke these states in the soul (mind) of the listener:

> …for when we hear [music] our very soul is altered; and he who is affected either with joy or grief by the imitation of any objects, is in very nearly the same situation as if he was affected by the objects themselves; … now it happens in the other senses there is no imitation of manners; … for these are merely representations of things, and the perceptions which they excite are in a manner common to all. Besides, statues and paintings are not properly imitations of manners, but rather signs and marks which show the body is affected by some passion… But in poetry and music there are imitations of manners; … those who hear them are differently affected, and are not in the same disposition of mind when one is performed as when another is; the one, for instance, occasions grief and contracts the soul, … others soften the mind, and as it were dissolve the heart: others fix it in a firm and settled state, … fills the soul with enthusiasm… The same holds true with respect to rhythm; some fix the disposition, others occasion a change in it; some act more violently, others more liberally (Aristotle 2018, Politics, VIII:V).

1.4 The Ear as the Central Organ of the Church

In pre-medieval Europe, the Roman Catholic Church eventually filled the political and spiritual vacuum caused by the collapse of the Roman Empire around 500 CE and went to extraordinary lengths to establish supreme papal power, resulting in the construction of massive reverberant cathedrals, which replaced the echoic caves of the past to become temples in which the resonant voice of an omnipresent God, through the Pope, and his priests and their choirs, was delivered to a largely passive, observing audience; silenced by the unintelligible Latin and the presence of the holy sacrament, for the safety of their immortal souls. Even though this church almost universally forbade the use of all instruments of "profane" music in worship, sometime during the tenth century large *Blockwerk* organs began to be permanently installed in churches and cathedrals.[10] By producing the lowest-pitched and most powerful musical sounds of the period, these organs resonated the buildings in which they were installed, and impressed on congregations a power found only in the church. In addition, these modern caves were engineered to produce an awe-inspiring, comforting, inclusive "community" feeling through spatial aural reverberant incoherence, and by which potentially questioning or dissenting individuals were subsumed into the anonymity of the Mass.[11]

When Johannes Gutenberg (1398–1468) invented his mechanical movable-type printing press in 1440, only six to seven percent of European adults were literate

[10]This somewhat convenient simplification of the role of music and musical instruments in the Middle Ages is more evenly discussed in Grout (1960).

[11]The pun is intended.

(Schlossberg 2011) and eighty percent of English adults couldn't even spell their names.[12] Thus, the closely connected (visual) activities of writing and reading were engaged in only by the educated, ruling classes while the overwhelmingly illiterate common people's only connection to writing was when it was read aloud to them. While they maintained their profane cultures by telling each other stories from within their rich oral/aural traditions, played musical instruments and danced, their only access to the word of God was through lectionary periscopes.[13] Gutenberg's invention flooded Europe with printed material. It made the Bible more available and encouraged an increase in literacy rates which eventually resulted in dissent from the Roman church's authority in the form of the sixteenth century's Protestant Reformation. The spread of literacy was accelerated by the emphasis on education in increasingly urbanizing societies; making way, in the seventeenth century, for an intellectual rebirth in the Europe in the form of the Renaissance which eventually lead, over an extended eighteenth century, to the Age of Enlightenment in which reason was advocated as a means of establishing an authoritative system of government, religion, ethics and aesthetics. This supported an intellectual and scientific impetus to obtain objective truths about the whole of reality, the spread of learning to the masses, and laid the material basis for modern knowledge-based economies.

Breaking the mental clutches that the church imposed by leveraging illiteracy and the use of sound to surround and sublimate independent voices, required men of vision, using empirical techniques, to "look" past the religious dogma of the church for objective theories of the natural world "out there". It also required the invention of observational instruments and techniques "favorable to the progress of the arts and sciences" (Wilson 1957, 227). For example, Galileo Galilei (1564–1642) built his own telescope in 1609, a year after it was patented in the Netherlands (Loker 2008, 15). Another such invention was the (re)discovery of the laws of linear perspective by the Filippo Brunelleschi of Florence (1387–1446) which enabled the depiction of visual depth on a planar surface. Lacking a theory of mathematical perspective, artists of the Middle Ages were more concerned with the static depiction of all-encompassing religious or spiritual metaphors rather than depicting the real, physical world, oriented towards an individual viewer. In addition to promoting and acknowledging the viewer, perspective provided a powerful technique to visually depict mechanical devices for the first time in a realistic manner, and assisted in the invention and dissemination of cosmographic scientific instruments for astronomy, surveying, navigation, map-making and time-telling.

In parallel to these advances in visual representation, the angelic voice-only sacred monophonic plainchants of the church were challenged by, and eventually subsumed into, the polyphonic complexities of the compositions of increasingly-individually-recognized composers who, in composing order on the

[12]This figure was calculated by combining literacy rates derived from Roser and Ortiz-Ospina (2018) and population demographics estimated by Urlanis (1941).

[13]Literally "cuttings" (from the Bible), from the Greek *perikopē* meaning a "section" or "cutting".

cathedral-induced echoic aural incoherence of multiple overlapping lines of the chants, explored, notated and extemporized new musical structures. These structures were not arbitrarily abstract however, but developed from careful attention to both the sounds and accents of the words being sung as well as the meanings in their texts.[14]

1.5 Music as Language and Rhetoric

The belief, that since music is a language and can be consciously treated as such, is not confined to post-medieval Europe. Many cultures have developed melodic and rhythmic modes and dance gestures through which associated concrete meanings and/or affective states are communicated to understanding audiences. For example, the ragas, talas and mudras (hand/finger gestures) of classical India are used to clearly communicate specific ideas, events, actions, or creatures. The ancient Greeks invented a whole system of how different musical elements affect the soul in different ways. The Renaissance of their ideals in Europe included the application of rhetorical devices, not only in speech, but in music and dance. A system of rhetorical devices, i.e. a representational vocabulary for making communication explicable and persuasive, was considered essential for organizing the syntax of (initially voice-only) compositions and, by addressing an audience's logical and emotional dimensions simultaneously, making the music semantically effective and able to communicate successfully. By the middle of the sixteenth century, such rhetorical devices had developed into often extravagant musical word paintings or *madrigalisms* as they were known.[15] Nevertheless, the study and use of rhetoric to make musical discourse more concretely meaningful continued through the seventeenth and well into the eighteenth century. Haydn, often described by his contemporaries as "a clever orator" and "the Shakespeare of music," is probably the last major European composer whose music was regularly discussed by his contemporaries in terms derived from the classical tradition of rhetoric (Beghin and Goldberg 2007; Saint-Dizier 2014). The artistry of his musical rhetoric is amply displayed in his oratorio, *The Creation* (Zbikowski 2019).

The purpose of this discussion is not to suggest that linguistic utterances are music, or vice versa. However while music and linguistic speech have different

[14]The isorhythmic motets of Guillaume de Machaut and John Dunstable, for example, in which each voice in a canon is in a different rhythm.

[15]Of special interest, is the extravagantly manneristic and harmonically experimental music of Carlo Gesualdo (1566–1613) which was also encouraged by the visionary experimental composer and music theorist Nicola Vicentino (1511–1575/6). For a fuller discussion, see Brown (1976). Also of interest is the composition of Eye Music, in which the use of black and white musical notations was used to suggest darkness and light; sadness and joy (Einstein 1949). See Sect. 1.7.

cultural functions, rhyme and rhythmic speech have also been important in musical devices in many cultures and the two forms of expression are often blended. The adulatory praise poems of Africa in which professional bards, who may be both praise singers to a chief and court historians of their tribe, chant a series of epithets in an attempt to capture the essence of an object, event or person (Weinberg 1978). In the tradition Chinese form *shuochang*, performances commonly intermix speaking and singing, accompanied by percussion and sometimes plucked or bowed string instruments. Such praise songs and "story-singing" are the precursors of Rap[16] music, a contemporary form of vocal delivery that incorporates rhyme, rhythmic speech, and street vernacular; performed or chanted over a backbeat, musical accompaniment or *a cappella*.[17] Stylistically, Rap occupies a gray area between speech, prose, poetry, and singing (Edwards 2009).

1.6 The Rise of Abstract Instrumental Music

During the European Middle Ages, secular public musical activity consisted of heroic and lyrical minstrel songs and instrumental music to accompany dancing. During the Renaissance, regal courts became significant centers of economic and cultural activity outside of the Catholic Church, which lead to its eventual failure to suppress most forms of instrumental music as undesirably profane. So it was, that the Baroque style, particularly in music, painting and architecture, was encouraged by the post-Reformation Church as a means to stimulate religious fervor.

As monarchs and their courtiers required intellectual stimulation and entertainment, the composers serving in the courts became more experimental. Using a greater variety of musical instruments, and innovative harmony, they increasingly considered musical forms as modes of personal expression. The codification of staff notation, supported by a burgeoning print industry, encouraged the extended development and dissemination of more abstract musical structures: The simile-like representations of word-painting and other rhetorical devices gave sway to these more conceptually metaphoric and self-referential (symbolic) motifs, which became built into the very fabric of Western musical language.[18]

While initially following in the spirit of courtly dances, the decline on the reliance of textual subjects in favor of formal reflexivity also became a means of embodying aesthetic affects (Erlmann 2010, 94). As a consequence, there arose a

---

[16]The word has been in British English since the sixteenth-century, meaning "to lightly strike" and is now used to describe quick speech or repartee (ACOD 1992). Since the late 20th century the term *shuochang* is used in China to refer to rap music (Wikipedia 2019).

[17]Literally "in the manner of the chapel" that is without instruments, thus emphasizing the point made in Sect. 1.4 concerning the Church's banning of musical instruments in worship.

[18]Some such motifs have become well known. Two such are J.S. Bach's "B-A-C-B^{b}" (The note B^{b} is represented by the letter 'H' in German), and Beethoven's "Fate" motif from the opening of his *Fifth Symphony*.

belief that music was not a language in the sense that it no longer had properties of defining and referring to specific meanings about which interpretations and responses could be made.[19] The need for some understanding of what we might call the semiological codes of music was apparent however, for without them it would have been difficult to claim for music to be more than just beautiful arrangements rather than important expressions of human experience and expression in the humanist tradition. In discussing musical language, Durant (1984, 10) says:

> On the one hand, there is 'language' as an assumed range of properties and associated effects, determined … by psychological or acoustical resources. These supply continuity for the 'language' and an overall framework for activity. Music made outside the framework is simply part of another 'language'. What would be most interesting about this 'language' would be the precise nature of the acoustic and psychological resources, even if these—failing some way of determining correspondences between forms and effects—will not explain very much about actual, particular pieces of music. In the other emphasis, there is language as a set of properties and associated effects (still conditioned by psychological and acoustic resources), whose most significant features are those of social, regional and historical variation around those resources.

And later (op.cit.: 13–14), recalling Claude Debussy's "Works of art make rules but rules do not make works of art" (Paynter 1992, 590), Durant observes:

> …In this latter emphasis, musical works do not simply exist within, or use, the musical 'language': they make and remake it, in particular realizations and directions of potential.

Following the rise of public concerts and music to be listened to for its own sake, the status of instrumental music grew in the eighteenth century, along with the Cartesian idea that abstract reflexive forms were the pre-eminent *raison d'être* of musical expression, so that "critics increasingly came to think of music as having emancipated itself from mimetic representation" (Erlmann op.cit.:189).

1.7 "Organizing the Delirium"

During the nineteenth and early twentieth century, there was also a shift by composers from musically representing various gestural and aesthetic affects (as portrayed in the different formal dance forms, for example) to the representation of individuals' changing emotional states. The principal structural means of organizing compositions was functional tonal harmony, which relies on a set of codes or functions to create expectations in the listener of, given what has just occurred, what might occur next. The German Romantic interest in the personal expression of intense emotional states led composers to search for new expressive means such as

[19] The evolution of the late Middle English term *motive* from the Old French *motif* (adjective used as a noun) has its origin in the Latin *motivus*, from *movere* 'to move' (ACOD 1992).

rapidly shifting modulations and eventually the abandonment of functional harmony as means of organizing both moment-to-moment flow within a work as well as the form of the entire composition—and eventually of tonality itself.[20]

In addition to its social and cultural imperatives, there is a distinct trend for Western experimental composers to conceive of their music as complex-patterned time-ordered series of acoustic events that vary in pitch, loudness and timbre; that are absorbed and elicit emotions when listened to. This paradigm is embedded in scored compositions that are abstractly composed and transmitted to listeners by expert musicians by a variety of means, including concerts, recordings, broadcasts, and internet streaming. Following the examples of the Second Viennese School in the first part of the twentieth century,[21] many composers used serial procedures based on permutation group orderings of the chromatic pitch gamut.[22] This led to a short but intense period when each of the (now "parameterized") dimensions of a composition (pitch, duration, dynamics and timbre) were all serially organized. The idea of music as a variable-parameterized space has been adopted by sonifiers engaged in parameter-mapping sonification, will be discussed below, and the attempt to use logical operations on parameterized datasets to make them perceptually coherent is discussed in detail in Chap. 3.

The composer/architect Iannis Xenakis, whose oblique contribution to sonification is discussed later, is a major and distinct European voice of musical composition of the second half of the twentieth century. He was clear that the serialists had made an important contribution to the search for new methods of organizing music by introducing abstraction (Xenakis [1971] 1991, 193). However, he did not agree with their declared necessity for using just serial techniques to organize musical material, and was critical of linearity of serial thinking: the way pitch dominated musical structuring, even when its influence on the sound was only secondary; that under serialism as they defined it, duration was, in general, even less structurally organized than in traditional forms, and that the resulting disjunct linear polyphony destroys itself by its very complexity.

John Cage often produced scores somewhat automatically by making marks on a page using the *I Ching*, or spatially related mappings such as those in a star map for *Atlas Eclipticalis* (1961), or the imperfections in the paper itself in *Music for Piano* (1952–62). In these works, the (often graphic) scores function as non-specific stimuli to performers rather than as a set of precise codes or instructions made by

[20]The connection between Freud and Schoenberg and their attempts to "organize the delirium" (as composer Pierre Boulez was known to have described it) relies on an affinity between Freud's structural model of the mind and Schoenberg's method, whose rules are based in the principles of symbolic logic, even if applied unconsciously. See Carpenter (2018).

[21]Principally Arnold Schoenberg, Anton Webern and Alban Berg.

[22]There are far too many composers to mention in this context as the techniques were employed and taught by numerous composers across the Western world. Two publications were influential: *Die Reihe* (Eimert and Stockhausen [1957] 1968, 1968) in English and German, and *Serial Composition and Atonality* (Perle 1962).

the composer to performers. As such, the relationship between the score and the sonic results are seldom perceivable by the listener, nor were they meant to be (Cage and Knowles 1969). This type of representational mapping has its historical precedents in Eye Music (Einstein 1949, 3:1971) and while, considered broadly, it is a type of sonification, it is outside the gamut of the current work.

1.8 From Program Music to Programmed Music

Not all composers were caught up in the psychological angst of the German romantics and expressionists, or the following existential modernist's position that ordinary musical listening was archaic and needed to be replaced with "structural listening" (Eisler and Adorno [1947] 1994, 20). So, the need to concretely "ground" musical expression in the earthly domain persisted in compositional references to the real, often rustic, world. In fact, compositions depicting nature in one way or another were so common that the term *Pastoral* is used as a description of a type of composition, and there are a number of recognized national Pastoralist schools.

While most critics and aestheticians of the time wrote disparagingly about the depictions of birdcalls and battles, which they regarded as a debasement of the art, such things were popular with the public, as were—and are—nature "atmospheres." Well known examples include Antonio Vivaldi's *Four Seasons* (1725), the second movement of Ludwig van Beethoven *Piano Sonata*, Op.28 (1801), Felix Mendelssohn-Bartholdy's overture *Fingal's Cave* (1833), Frederic Chopin's *Raindrop Prelude* (1838), and Nikolai Rimsky-Korsakov's *The Flight of the Bumblebee* (1900). But even after Arnold Schoenberg had composed *Erwartung* (1909) and *Five Orchestral Pieces* (1909), Ottorino Respighi composed *Pines of Rome* (1917), Ralph Vaughan Williams, *A Pastoral Symphony* (1922), Arnold Bax, *November Woods* (1917) and Charles Ives *Three Places in New England* (1912–16). A compilation list would include many of the best know works in the Western canon.

Although less common, the literal transcription of real-world sounds into music has been even more challenging to the abstract expressionists than Pastoral music. The cuckoo calls in Beethoven's Sixth Symphony (1808) and Olivier Messiaen's *Catalogue d'oiseaux* (1958) [Catalog of Birds] and his use of a wind machine in *Des canyons aux étoiles* (1974) [From the Canyons to the Stars] are obvious examples but the trend extended throughout the twentieth century with the inclusion of non-traditional sounds such as the fire siren in Edgard Varese's *Ameriques* (1921), the futurist's noise art[23] and then, especially following the invention of the tape recorder in the 1940s, by *musique concrète*—the term itself being coined by Pierre Schaeffer[24] to contrast the music he was making at *Radiodiffusion Française* in Paris with the "pure" (i.e. abstract) music of the period, especially that being

[23]As described in the 1913 manifesto *The Art of Noises* (Russolo 1916).

[24]Schaeffer's work is discussed in more detail in Chaps. 3 and 4.

produced by the German expressionist composers including those at the WDR in Cologne.[25] Schaeffer and others established the GRMC[26] in Paris in 1951 which attracted many notable composers of the period. He went on to establish the GRM,[27] one of several theoretical and experimental groups unified by the study of audiovisual communication and mass media.

1.9 Algorithmic Composition and Data Sonification

Xenakis's focus on using mathematics in composition included the application of group theory, game theory, symbolic logic and stochastics to musical composition (Xenakis [1971] 1991) was seminal in establishing a compositional "style" known as algorithmic composition which sowed the seed for the idea of representing abstract data relations in sound for investigative purposes.[28] Xenakis was a participant at GRM but, following Schaeffer's refusal in 1963 to use mathematics and the computer in the studio there,[29] established EMAMu, later CEMAMu,[30] specifically in order to undertake research into the application of mathematical ideas to music composition.

The other notable early algorithmic music research of the period was by Lejaren Hiller and Leonard Isaacson at the University of Illinois at Urbana-Champaign. They used the university's ILLIAC I computer to develop a "rules-based" system to create *cantus firmi*, four-voice harmonies, various rhythmic and dynamic schemes,

[25]Studio für elektronische Musik des Westdeutschen Rundfunk (Studio for Electronic Music of the West German Radio). Following a brief period in Schaeffer's studio in 1952, Karlheinz Stockhausen joined Gottfried Michal Koenig and Herbert Eimert at the WDR. While the early electronic music he produced there employed purely electronically produced sounds (e.g. *Studie I* and *Studie II*), Stockhausen soon felt the need to work with a continuum between electronic and human vocal sounds (e.g. *Gesang der Jünglinge* in 1956) and by extension, the inclusion of materials other than sounds produced purely by electronic means (Stockhausen 1964). Electronic music from the Cologne studio thereby moved closer conceptually to the musique concrète from Paris.

[26]Groupe de Recherches de Musique Concrète, Club d 'Essai de la Radiodiffusion-Télévision Française at RTF.

[27]Groupe de Recherches Musicales.

[28]This musical style in also known as or generative- or procedural- or automated-composition. One assumes that today such compositions are composed, and perhaps realized, with the aid of computers, but this is not necessarily the case. Many of the calculations for Iannis Xenakis early works, such as *Metastaseis* (1953–54) and *Pithoprakta* (1955–56), were made using a hand-held calculator (personal communication). Using dice to randomly generate music from pre-composed options (Musikalisches Würfelspiel) was quite popular throughout Western Europe in the 18th century. (Wikipedia: Musikalisches Würfelspiel). Code can be written post mortem, however: Witness that for Wolfgang Mozart's German dances (Saunter 2018).

[29]Xenakis met Schaeffer in 1954 and composed five major pieces of *musique concrète* during the period (Gibson and Solomos 2013).

[30]Centre d'Etudes de Mathématique et Automatique Musicales (Center for Mathematical and Automatic Musical Studies).

and Markovian stochastic grammars which they realized in the *Illiac Suite* for string quartet in 1957.

Such computational proceduralism developed rapidly during the second half of the twentieth century in league with the development of cybernetics and cognitive science. The technical feasibility of being able to accurately repeat the synthesis of sounds by digital means enabled the birth of computer music, which was also heavily influenced by the "acoustic event" paradigm that became embedded in many of the compositional software tools used to create it. These tools have been widely adopted by researchers who use them in an attempt to obtain a better understanding or appreciation of relations in datasets of various sizes, dimensions and complexities—what is now called scientific sonification.

It is useful to distinguish data sonifications made for the purposes of facilitating the communication or interpretation of relational information in data, and data-driven music composition, ambient soundscapes and the like—the primary purpose of which is personal expression and other broader cultural considerations, whatever they may be. While scientific or pragmatic data sonifications and music compositions share a common reliance on the need to render structures and relations into sound, their purpose is often different, and so too the rationale for the evaluation of the sonic results. The current use of the term *sonification* to include such cultural concerns is somewhat unfortunate because it blurs purposeful distinctions. A desire to maintain these distinctions is not to suggest that there are not commonalities—the two activities can provide insights that are mutually beneficial. However, because the purposes of the activities are different, so too will be their epistemological imperatives and consequences, such as, for example, in the representational methodologies employed, in tool design, in user-interface and usability requirements and evaluation—all matters dealt with in subsequent chapters.

1.10 Purposeful Listening: Music and Sonification

There is no one-way, or reason, to listen to music. Even different musics have different contexts and thus require different ways of listening that may involve whole complexes of social dimensions that are simply not relevant to the perceptualization of data relations for pragmatic purposes. Furthermore, although music may be composed of syntactic structures, there is no universal musical requirement that these structures be made explicit, even aurally coherent. In fact, stylistic or even dramatic considerations may require the exact opposite—in the orchestration of spectral mixtures by melding of instrumental timbres, for example.

In contrast, clarity in information sonification is essential and rendering techniques that produce a kind of "sonic gumbo"[31] can be more successful.

[31] In gumbo cuisine, the ingredients are added—a little bit of this, a little of that—so that they do not meld together but remain sensorially distinct, yielding a 'rainbow' of flavors, aromas and textures rather than a uniform blend.

Consequently, sonifications in which the user-driven real-time exploration of datasets using dynamic scaling in multiple dimensions, perhaps with auditory beacons (Kramer 1994b), may not result in what is currently understood to be musically coherent sound streams. Even if listened to *as* music, information sonifications may provoke critical commentary about musical issues such as the appropriateness or formal incompleteness of the resulting sonic experience. Perhaps, as Paul Polansky suggested, the closest point of contact between such pragmatic data sonification and musical sonification is in compositions in which a composer intends to "manifest" mathematical or other formal processes (Polansky and Childs 2002). This "classical" algorithmic motivation is explicitly enunciated by Xenakis in his seminal book, *Formalized Music* ([1971] 1991), and many of the cultural concerns in his thesis defense (1985).

While numerous composers use mapping and other procedural techniques of one kind or another in their musical compositions, they are rarely interested in "featuring" the mapping explicitly. Nor do they use mapping in order to simplify the working process or to improve production efficiency, but so as to craft the emergence of musical forms. In order to gain a deeper insight into the way composers map conceptual gestures into musical gestures, Doornbusch (2002) surveyed a select few composers who employ the practice in algorithmic composition.

> I am not interested in projecting the properties of some mathematical model on to some audible phenomena in such a way that the model be recognized as the generator of some musical shape.

So, those interested in producing music of a certain complexity may shy away from simple mappings as they can be hard to integrate with other musical material of a substantial nature. On the other hand, as Larry Polansky explains:

> …the cognitive weight of complex mappings degenerates rapidly and nonlinearly such that beyond a certain point, everything is just 'complex'.

Even a suitably complex, structurally coherent mapping may not be musically sufficient if the composition relies on a (human) performer, as composer Richard Barrett (in ibid.) emphasizes:

> In a score one is always dealing with the relatively small number of parameters which can be recorded in notation, and which interact with an interpreter to produce a complex, 'living' result.

The importance of this embodied "living" aspect of music has often been forgotten, ignored, or even dismissed in many discussions of Western art music, including by some composers. While there are historical reasons and—consequences of doing so—such an approach to data sonification could have a major impact on the intelligibility of computer-rendered mapping-encoded artifacts.

1.11 Musical Notation as Representation

In Western art music, notation evolved, along with the notion of the *work*, from a performer's *aides-mémoire* to a tool of thought for defining works of increasingly abstract complexity (Goehr 1994). Notated scores came to be thought of as the encoded representation of sounds, even as a somewhat definitive objectification of a composer's thoughts. That we (at least in English) so frequently substitute the word 'note' for 'tone', and 'music' for 'score', exemplifies the strength of this conceptual elision. Indeed, in a number of intricately notated works of the twentieth century, it seems the performer is sometimes considered an unfortunate necessity. In others, notation functions as encapsulated stimuli by which the performers, as they attempt the impossible task of playing it "note-perfectly", enact a drama of physical and mental exertions.[32] Theodore Adorno noted a tendency to consider the bodily presence of the performer as a kind of contamination of musical experience, as a manifestation of a commodity fetishism where the "…immaculate performance … presents the work as already complete from the very first note. The performance sounds like its own phonograph record" (Adorno 1991).

Occidental art music today encompasses a wide range of motivations and listening practices, and reducing the intelligibility of such music to the conceptual level of scores and instruments enabled an unprecedented level of complexity. However, there is a growing recognition among music researchers, supported by a significant body of research in neuroscience,[33] that the conveyance of this complexity is reliant, at least to some extent, on embodied interpretation for effective communication (Goehr 1994). It was not until it was technically possible to construct musical compositions without the assistance of embodied interpreters that it was possible to meaningfully speculate on the extent to which a listener's perception of the structural characteristics of a piece of music are dependent on the sound-encoded gestures of performers, and not just the notated score. This has the unfortunate consequence that if sonifiers think of data as being the sonification equivalent of a parameterized musical score, and, following the path of least resistance, as most have been apt to do, use music composition software designed to produce abstract musical objects, their sonifications will lack the intelligence that is recognized as embodied in the (often micro-) gestures of musical performers (Worrall 2014).

This should act as a caution that, while adopting tools of one domain into another can be a very empowering, such adoption does not come value-free. If the intelligibility of much music is bound, not only to text, rhetoric, metaphor and formal devices such as phrase structure and the semantics of harmonic tension and resolution, but to the transmission through sound of the embodied foreknowledge of performers, in establishing the foundations for the practice of information

[32]This idea is integral to many of the works of a number of contemporary composers, including Mauricio Kagel, Luciano Berio, Brian Ferneyhough and Iannis Xenakis.

[33]Discussed in detail in Chaps. 3 and 4.

sonification, it is sensible to embrace approaches to forming sounds and their relationships that are supported, yet as unencumbered as possible, by the conceptual boundaries placed by tools for computer music-making as they currently exist. This, in turn, has the potential to enrich the practices for which the tools were originally meant.

Sounds and sonic structures demonstrate their weak proclivity to bind to casual inferences (Chion 2016) and strongly to metaphorical representations: From the rustle of dried leaves caused by escaping prey to the echoic resonances of a cathedral and a crowded street; from the phonetic structures of speech to word-painting in Renaissance madrigals; from musical rhetoric and Pastorals to instrumental expressions of affective states; from the turbulence of gas molecules to gravity waves and the folding and cracking of the earth's crust; from the movement of objects relative to each other to the flow of data through a network—sounds and sonic structures demonstrate homophonic[34] tendencies. With careful attention to details and their relationships to each other in context; to the isomorphic-heteromorphic bindings between sounds and their causes, sonification can render vibrant voices for unseeable things.

References

ACOD (1992) The Australian concise Oxford dictionary, 2nd edn. Oxford University Press, South Melbourne

Adorno T (1991) On the Fetish character in music and the regression of listening. In: Bernstein JM (ed) The culture industry. Routledge, London

Aristotle [c.BCE 335] (2018) Aristotle: Poetics, ethics politics and categories. Seltzer Books

Beghin T, Goldberg S (eds) (2007) Haydn and the performance of rhetoric. University of Chicago Press, Chicago, IL, USA

Boyd E (1905) History of auditing. In: Brown R (ed) A history of accounting and accountants. Edinburgh, T.L. & E.C. Jack

Brown H (1976) Music in the renaissance. In: Prentice hall history of music series. Prentice-Hall, Inc, Englewood Cliffs, New Jersey

Cage J, Knowles A (1969) Notations. New York: Something Else Press

Carpenter A (2018) Parallels between Schoenberg and Freud. In Music–Psychoanalysis–Musicology,edited by S. Wilson. New York: NY: Routledge.

Chion M (2016) Sound: an acoulogical treatise. Duke University Press, translated by Steintrager JA

DeGangi GA (2017) Treatment of attentional problems, 2nd edn

Doornbusch P (2002) Composers' views on mapping in algorithmic composition. Organised Sound 7(2):145–156

Durant A (1984) Conditions of music: language, discourse, society. Macmillan, London and Basingstoke

[34]Homophones are similar sounds arising from different causes (for example, the sound of audience clapping and a crackling fire place). Homonyms are a special case: different words with the same pronunciation but different spellings and meanings, for example, "weather", "whether" and "wether".

Edwards P (2009) How to rap. A Cappella Books, Los Angeles, CA
Eimert H, Stockhausen K [1957] (1968) Die Reihe. Universal Edition, London
Einstein, A. 1949. The Italian madrigal, vol 3. Princeton: The Italian
Eisler H, Adorno T [1947] (1994). Composing for the films. Athalone, London
Erlmann V (2010) Reason and resonance: a history of modern aurality. Zone Books, Brooklyn, NY
Gibson B, Solomos M (2013) Research on the first musique concréte: the case of Xenakis's first electroacoustic pieces. In: Proceedings of the electroacoustic music studies, Lisbon
Goehr L (1994) London: The imaginary museum of musical works. Oxford University Press, Oxford, UK
Grout D (1960) A history of Western music. W.W. Norton & Company
Kramer G (1994a) Auditory Display: sonification, audification, and auditory interfaces. Proceedings, Santa Fe Institute studies in the sciences of complexity, vol XVIII. Addison Wesley, Reading MA
Kramer G (1994b) Some organizing principles for representing data with sound. In: Kramer G (ed) Auditory Display: sonification, audification, and auditory interfaces. Proceedings, Santa Fe Institute studies in the sciences of complexity, vol XVIII. Addison Wesley, Reading, MA, pp 185–221
Loker A (2008) Profiles in colonial history. Solitude Press, Williamsburg, VA, USA
Macran HS (1902) The harmonics of aristoxenus. Clarendon Press, Oxford, UK
McLuhan M (1968) Understanding media: the extensions of man. McGraw-Hill Book Company, New York
Muktananda S (1972) Śri Guru Gitā. In: The nectar of chanting. SYDA foundation
Paynter J (1992) Companion to contemporary musical thought. Psychology Press Ltd., London
Perle G (1962) Serial composition and atonality: an introduction to the music of Schoenberg, Berg, and Webern. University of California Press, Berkeley
Pliny, The Elder [77AD] (1938) Natural history. Translated by Rackham H. Book II. Harvard University Press, Cambridge, Massachusetts
Polansky L, Childs E (2002) Manifestation and sonification. The science and art of sonification, Tufte's visualization, and the 'Slippery Slope' to algorithmic composition. An informal response to childs' short paper on Tufte and sonification, with additional commentary by Childs. http://eamusic.dartmouth.edu/
Roser M, Ortiz-Ospina E (2018) Literacy
Russolo L (1916) L'Arte dei rumori. In: Niuean pop cultural archive, Nov 1916. https://www.unknown.nu/futurism/noises.html
Saint-Dizier P (2014) Musical rhetoric: foundations and annotation schemes. Wiley, London
Saunter D (2018) Mozart dice game. http://donaldsauter.com/mozart-dice-game.htm
Schlossberg T (2011) Literacy rates. https://www.mcsweeneys.net/articles/literacy-rates
Sokolov EN, Spinks JA, Näätänen R, Lyytinen H (2002) The orienting response in information processing. Erlbaum, New Jersey, Mahwah, London
Sterne Jonathan (2003) The audible past: cultural origins of sound reproduction. Duke University Press, Durham, NC
Stockhausen K (1964) Texte zu eigenen Werken, zur Kunst Anderer. Aktuelles. Edited by Dieter Schnebel, vol 2. M. DuMont Schauberg, Köln. https://en.wikipedia.org/wiki/Studio_for_Electronic_Music_(WDR
Urlanis B (1941) Rost naselenila v Evrope : opyt ischislenila, vol 42379320. OGIZ-Gospolitizdat: Moskva. https://en.wikipedia.org/wiki/Medieval_demography
Vitruvius M (1914) De Architectura. Book V, Chapter IV
Wilson A (1957) Diterot: The Testing Years 1713–1759. https://archive.org/stream/diderotthetestin001232mbp/diderotthetestin001232mbp_djvu.txt
Weinberg P (1978) Ethnomusicology and its relationship to some aspects of music in Cetshwayo's time. Natalia 8:61–68
Wikipedia, contributors (2019) Shuochang. In: Wikipedia: the free encyclopedia. https://en.wikipedia.org/w/index.php?title=Shuochang&oldid=882921570

Worrall D (2010) Parameter mapping sonic articulation and the perceiving body. In: Proceedings of the 16th international conference on auditory display, 9–15 June, 2010, Washington, DC, USA

Worrall DR (2014). Can micro-gestural inflections be used to improve the soniculatory effectiveness of parameter mapping sonifications? Organised Sound 19(1):52–59

Xenakis I (1985) Arts/sciences: alloys. Pendragon Press, New York

Xenakis I [1971] (1991) Formalized music: thought and mathematics in music. The Pendragon Press, NY, New York

Zbikowski LM (2019) Music, metaphor, and creativity. In: Hidalgo-Downing L, Kraljevic-Mujic B (eds) Performing metaphor and creativity across modes and cultures. John Benjamins Publishing

Chapter 2
Sonification: An Overview

Abstract Encompassing ideas and techniques from music composition, perceptual psychology, computer science, acoustics, biology and philosophy, data sonification is a multi- even trans-disciplinary practice. This chapter summarizes different ways sonification has been defined, the types and classifications of data that it attempts to represent with sound, and how these representations perform under the pressure of various real-world utilizations.

A primary distinction can be made between so called *audifications*, which entail the direct amplification, filtering and/or time-warping of existing sounds, such as is accomplished with the esophageal stethoscope, and the use of sound to convey inherently silent abstractions such as *variables* and *data*. The non-chemical[1] definition of *sonification* has evolved over the last quarter-century as its use in auditory displays has developed. For the purpose of discussing multivariate data mappings, Bly (1994, 406) described sonification as *audio representation of multivariate data*. The sonification of univariate data is also possible, and thus Scaletti (1994, 224) proposed a more formal working definition for her investigation of auditory data representation, as

> a mapping of numerically represented relations in some domain under study to relations in an acoustic domain for the purposes of interpreting, understanding, or communicating relations in the domain under study.

In order to differentiate sonification from other uses of sound, Scaletti explicitly draws attention to two parts of her definition: a *technique* (mapping numerical data to sound) and an *intent* (to understand or communicate something about the world).

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

D. Worrall, *Sonification Design*, Human–Computer Interaction Series, https://doi.org/10.1007/978-3-030-01497-1_2

Barrass (1997, 29–30), reworking Scaletti's definition, *en route* to a definition of auditory information design ("the design of sounds to support an information processing activity"), emphasizes the idea of information (the content) over data (the medium):

> a mapping of information to perceptual relations in the acoustic domain to meet the information requirements of an information processing activity.

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

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

The first sentence of this definition appears to be the most succinct and widely used. Speech audio is specifically excluded, presumably so as to discriminate sonification techniques from speech-related practices, such as linguistics and text-to-speech software. While speech research is an extensive research field in and of itself, there is no reason why speech audio should *necessarily* be excluded from the definition. As Hermann argued (2002, 23), speech does have attributes which, if data-driven, could be useful for sonification purposes, and some research has suggested speech-audio displays could be used to convey non-verbal information upon which people can make useful decisions (Nesbitt and Barrass 2002). Further, speech that is temporally compressed until incomprehensible as speech[2] could significantly improve menu navigation in PDA devices (Walker et al. 2006).

While *the representation of data relations in sound relations* (Anderson et al. 2002) is likely the most succinct definition of sonification, it avoids the intent referenced in the 1999 definition: *for the purposes of facilitating communication or interpretation*. Nor does it quite capture distinctions between, for example, data sonification and data-driven music composition. While a purpose of music is the expression of musical knowledge and broader cultural considerations, whatever they may be, between composers, performers and listeners, the purpose of sonification, as the term was originally used, is to represent data in sound in such ways that structural characteristics of the data become apparent to a listener. This distinction is emphasized by the portmanteau expression *soniculation* (from *sonic articulation* Theusch et al. 2009) introduced by the current author (2010). In the early literature (Kramer 1994a, b), the term *sonification* is used as a shortened form of *data sonification* as a sub-category of *auditory display*, but the distinction, perhaps unfortunately, seems to have subsequently disappeared. All three terms are now used both nominally and verbally, as are the expressions *to sonify* and, though less common, *to audify*.

[2]Discussed in detail in Sect. 2.2.1.4.

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

> Data sonification is the acoustic representation of informational data for relational non-linguistic interpretation by listeners, in order that they might increase their knowledge of the source from which the data was acquired.

This, and other descriptions of data sonification, preserves some of the 'ambiguity' of the implications of the phrase *for relational interpretation by listeners*. For generality, it specifically does not intimate the fact that it is often the data *relations* that are sonified rather than the data itself. Relations are abstractions of, or from, the data and the sonification of them more explicitly implies sonifier intent. The expression *information sonification* expresses that distinction. The term *information* adds a complexity of its own; one that has changed over time, as discussed more fully in the next chapter. Throughout this book, when the distinction is not important to the argument, the term *sonification* is used without a qualifier.

2.1 Classifying Sonifications

Sonifications can be classified in a number of ways: distribution technology (medical, public arena, interactive) intended audience (anaesthetists, stock brokers, the visually impaired), data source (electrocardiograms, securities markets, in-ear navigation aids) data type (analog, digital, real-time, spatial, temporal) and so on. Twenty-five or so years since concerted research began, data sources and target applications for sonification have now become too numerous to review them in this context. They are also the easiest to locate using keyword searches on the world-wide-web.[3] For the current purpose, we partially follow Kramer (1994b, 21–29) in implying a description of the type of representation that is used to present information to the auditory system.

There are two broad representational distinctions at either end of an analogic–symbolic continuum. Analogy[4] is a high-level cognitive process of transferring

[3]For example, in June 2008, an internet search of the phrase "stockmarket sonification" revealed over 6,000 links (c.f. 50 million for "stockmarket"). A decade later, the numbers were 27,500 and 204 million, respectively.

[4]From Gk. analogia 'proportion', from ana– 'upon, according to' + logos 'ratio', also 'word, speech, reckoning' [OLED].

information from one particular subject (the analog or *source*) to another particular subject (the *target*) or of establishing a relation between *sources* (e.g., data) and *targets* (e.g., listeners). Similes and metaphors are both analogies. With similes, the transfer is explicit. ("The *source* 'is like' the *target*"). With metaphors, the transfer is not overt, often implicit. ("The *source* 'is' the *target*").[5] The purpose of analogic representation is to make the structure of the information better suited or simpler for the *target* than it was in its original form. Analogic representation is more connotative than denotative. A good example of analogic sonification is a Geiger counter that produces clicks in a loudspeaker at a rate proportional to the strength of radiation in its vicinity.

By contrast, a symbolic[6] representation is a categorical sign for what is being represented. It is more denotative than connotative and thus more abstracted from the source than an analogic representation. In symbolic representation, source data is aggregated and assigned to the elements of a schema (symbols) according to a set of rules that, as a result, implicitly makes the distinction between data and information. These symbols are then displayed to the *Target*, who processes them according to their knowledge of the schema used. By illustration, consider a stream of noise emanating from a source. On listening to the stream, a *Target* recognizes that some segments of it are symbols in one schema that they know, some sounds are symbols in another known schema and there are still other sounds for which they do not have a schema (other than the 'everything–I–do–not–have–a–schema–for' schema). Two important features of symbolic processes are (a) a variety of display schemata can be used to symbolically represent the same information, or different information by different discrete–element–aggregation of the data, and (b) the relationships between the symbols themselves do not reflect relationships between what is being represented. For example, there is no relationship between the ideas represented by the words *loss*, *boss* and *lass*, nor the number *1055*, even though the structures of the symbols are similar.

In data sonification, for which the *Source* is data and the *Targets* are listeners, the types of representations a sonifier might employ is dependent on the structure of the data, the kind of information needing to be extracted from it and how easily (efficiently and effectively) that information can be processed by the listeners' hearing systems: physiological, perceptual, cognitive and memoric. This description emphasizes the importance, for sonifiers, of developing a thorough working knowledge of the ways hearing systems organize acoustic waves into discrete events using *Targets'* ability to perform such fundamental perceptual and cognitive processes as auditory stream segregation and auditory stream integration (Bregman 1994). Ongoing research in the fields of psychophysics and experimental psychology is directed at the empirical investigation of these processes. As is to be

[5]For example, "Love is like a butterfly" (after Dolly Parton) is a simile. "The Lorenz Butterfly Effect" (after Edward Lorenz) and "The Big Bang" (after Fred Hoyle) are metaphors.

[6]From Gk. symbolon syn– 'together' + stem of ballein 'to throw', evolves from the 'throwing things together' to 'contrasting' to 'comparing' to 'token used in comparisons to determine if something is genuine'. Hence, the 'outward sign' of something [OLED].

expected, most of this work is directed at understanding the segregation and integration of simultaneous, or near-simultaneous, components constituting auditory streams, rather than the larger or more complex scenarios of real-world situations.

2.2 Data-Type Representations

The data-type representations described below, loosely based on de Campo's groupings (2007) are *discrete*, *continuous* and *interactive*. These types can be located at different points on the analogic-symbolic continuum discussed above. Discrete representations mostly function symbolically, continuous representations mostly analogically, and interactive representations, mostly analogically but in a discontinuous manner. In practice, a particular sonification may use one or more of these types, even simultaneously, to satisfy the needs of the specific application and the data/information representation model being employed.

2.2.1 Discrete Data Representations

Discrete data representations are representations in which every data point (datum) is sonified with an individual auditory event. Discrete data representations are strongly symbolic and can be used as signifiers when there is no single preferred ordering of the datum. User-defined subsets of the data can be formed at will and randomly iterated over, much as a visual scene can be scanned in whatever way the viewer chooses. This flexibility of access to the data assists the user to build up a mental representation of the data space and the position of each data point in it. Whether or not a representation *appears* discrete or continuous will often depend on the interplay between data scaling and perceptual thresholds.

For example, a sequence of discrete frequencies may appear as a continuous glissando if they are closely clustered, or the close examination a continuous representation may cause it to appear discrete when it is time-stretched. Such definitional conundrums emphasize that these explanations are more descriptive than strictly taxonomic. Another example is the way in which auditory beacons (Kramer 1994b) function within an otherwise continuous data representation adds further emphasis to this point. However, when individual isolated sounds are initiated as environmental events, such as in the context of Human Computer Interaction (HCI), the distinction is clearer and it is these we consider next.

2.2.1.1 Auditory Warnings: Alarms and Alerts

Many animals use sounds of various kinds to warn those in their vicinity to the presence of others. Alarm signals, known in animal communication sciences as anti-predator adaptations, can be species-specific, and are particularly effective when imminent danger is sensed. Boisterous alarms tend to provoke a *fright or flight* response, which, in ongoing monitoring situations, is perhaps not subtle enough, given the likelihood of false alarms[7]: the ensuing panic is more likely to lead to a silencing of the alarm, or when that isn't possible, adaptation by cognitive filtering.

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

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

2.2.1.2 Auditory Icons (Discrete Real-World Sounds)

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

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

[7]As exemplified by Aesop's fable of the boy who cried "wolf".

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

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

Real-world sounds easily convey simple messages that are cognitively processed without much learning if the sounds are easy to identify. Yet sound in an interface that "seems cute and clever at first may grow tiresome after a few exposures" (Gaver and Smith 1990). In evaluating *Mercator*, Mynatt also observed a number of limitations, including the difficulty users had in identifying cues, especially those of short duration; the need for high (spectral) quality of sound and the need to evaluate auditory icons in a complete set for maximum dissimilarity while maintaining ease of identification. Ballas discusses recognition delays caused by auditory icons being too similar (aural ambiguity) but which can be controlled for if the number of alternative causes can be identified (1996).

Both Gaver and Mynatt raise the question of whether or not the concept of auditory icons breaks down with virtual objects that do not have a clear counterpoint in the everyday world. Such objects are known phenomenologically as *mental* or *immanent* objects (Husserl 1927).[8] Mynatt concludes, at the same time as cautioning that the participant sample size was too small for formal statistical analysis, that while non-sighted users could successfully navigate the system, "the overall test results indicated that the goal of designing intuitive auditory icons was not satisfied in the prototype interface." There were suggestions, by Blattner et al. (1989), that the problems in memorizing a large number of such icons was due to there being no structural links between them. In addition, auditory icons can sound similar even though they arise from different physical events. Later research Sikora et al. (1995), Roberts and Sikora (1997), Bussemakers and de Haan (2000) showed that, when compared with earcons and visual icons, test participants found that auditory icons were the least pleasant and least appropriate for use in computer interfaces. This can be due to the cognitive dissonance arising from the obscuring of the source of sounds if an auditory icon is too similar or conflicts in some other way with the real-world sounds in the environment of the user.

[8]This issue is addressed more extensively in the next chapter.

2.2.1.3 Earcons (Sets of Discrete Synthetic Tones)

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

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

One of the most powerful features of earcons is that they can be combined to produce compound messages. So, symbolic sound elements (motifs) and their transformations can be used hierarchically to represent events or objects. A drawback, however, is that the deeper the structural hierarchy being represented, the longer a sequentially structured earcon takes to play, thus interfering with the speed at which a user can respond to or interact with the computer. Playing earcons more rapidly would risk recognition errors, though this could presumably be controlled for. Brewster et al. (1995) experimented with a different approach, namely to play the earcons at the same rate but present the information in parallel, taking the time of only a single earcon to do so. Their results indicated that parallel and serial earcons were recognized equally as well. This is an important finding because it indicated that cognitive processing of an auditory display for known informational content may provide greater immutability than psychoacoustic conflict might otherwise indicate.[9] In both cases, recognition rates improved with repeated usage, which is the case in most HCI applications. Interestingly, more *training* produced greater improvements for parallel earcons than serial ones–ninety-percent recognition rates being easily achieved. Furthermore, given their symbolic structures are similar to the semantic aspects of music, it is interesting to note that the hypothesis that musicians would perform better than non-musicians was not verified: there were no differences in performance between the two groups,

[9]This suggestion was tested in Experiment 5 in Chap. 8, Sect. 8.5 in which apparently conflicting metaphorical representations are combined without any apparent cognitive dissonance.

except that some of the non-musicians took slightly longer to learn the system. Results from several experimenters had confirmed that correct identification of earcons decreased markedly if they occurred concurrently. McGookin (2004) undertook a detailed investigation of this phenomenon and produced a set of design principles including the use of spatial separation and a 300 ms onset-time offset. His findings indicate that, even when using these guidelines, the accuracy of concurrent earcon identification still decreased from ninety-percent to thirty-percent when the number of earcons increased from one to four. However, these (2008) guidelines did improve the accuracy with which individual earcons were identified.

Polotti and Lemaitre evaluated the use of rhetorical earcons to map common operating-system functions in a graphical interface. They found that test subjects who were musicians performed better than non-musicians, and that rhetorical earcons led to better memorization performances for all participants (2013). Rhetorical figures made the sound-meaning association easier to remember for all participants suggesting that the representational character of this relationship in rhetorically guided design was a strong cue for the listeners. They surmised that melodically-structured rhetorical strategies were so powerful that they overcame the limitation of pitch memorization suggested by Brewster et al. (op. cit.).

Overall, the earcon studies reported here indicate that they can be an effective means of communicating hierarchical structures but that the number of them that can be usefully identified is quite limited when they are used concurrently. The human auditory system is so attuned to fine variations in information-rich sounds, that the devil is often in the details, or lack thereof.

2.2.1.4 Speech Noises

NoiseSpeech is made by digitally processing sounds so that they have some of the acoustic properties of speech (Dean 2005). It is made either by applying the formant structures of speech to noise or other sounds, or by distorting speech sounds such that they no longer form identifiable phoneme sequences. The 'hybrid' sounds that results from this process encapsulate some of the affective qualities of human speech, while removing semantic content. Empirical experimental evidence suggests that most listeners cluster the sonic characteristics of NoiseSpeech with speech rather than those of musical instrument or environmental sounds (Dean and Bailes 2006, 2009).

Less abstracted than NoiseSpeech, *spearcons* (speech earcons) are brief computer-generated text-to-speech (TTS) spoken phrases that have been time-compressed until they are not recognizable as speech (Walker et al. 2006). They were designed to enhance hierarchical menu navigation for mobile and screen-limited devices, in which they can be created automatically by converting menu item text (e.g., *Export File*) to synthetic speech via text-to-speech software. Keeping pitch invariant, this synthetic speech is then time-compressed, rendering it incomprehensible. Whilst each spearcon in an application is unique, phonetic similarity is maintained. For example, the initial phonemes are invariant in *Open*

File, *Open Location…*, and *Open Address Book* as are the endings of *Open File*, *Export File*. The relative lengths of the text strings are maintained by their spearcons and this assists the listener to and learn and identify the mappings. Later research (Walker et al. 2012) revealed that spearcons significantly improve navigation performance in advanced auditory menus.

Another more focused approach to assisting the aural navigation of alphabetically-ordered indexes (such as dictionary entries or lists of names) is the *spindex*. A spindex ('speech index') is a brief non-linguistic auditory cue based on the TTS pronunciation of the first letter of each item in the index (Jeon and Walker 2009). For example, the spindex cue for "south" would be the sound /s/, and the sound /ʃ/ ("sh") for shellfish.

Speech and non-speech sounds share such identifiable auditory characteristics as pitch, rhythm, articulation, and rugosity, as well as some larger Gestalten such as phrase and prosody, the ability to simultaneously listen to music and talk or read without confusion is well known and easily demonstrated. This is supported by cognition research that finds that the auditory cortices in the two hemispheres of the brain are relatively specialized enough to be able to exploit the temporal and spectral differences between speech and musical sounds (Zatorre et al. 2002). While these reasons emphasize the importance of not excluding speech-like sounds from their use in sonification, some caution is warranted so as to ensure that the privileging of speech and the particular cognitive and pre-cognitive processes in play for its processing as language are taken into account. For example, a speaker's native language influences all aspects of phonology, including discrimination difficulties in foreign phonetic contrasts (/r/-/l/for Japanese speakers, for example), and the insertion of illusory vowels between unrecognized consonant clusters (/ebzo/is perceived as/ebozo/) (Dehaene-Lambertz and Gliga 2004). Although the issue is not unique to speech, temporal order (auditory sequence perception/confusion) of patterns requires specific attention if speech is to be confidently used in sonification. It may be the case that an understanding of the specific ways spoken language is based on "system(s) of oppositions and differences within a certain distribution of sonorities" (Chion 2016, 45), can positively inform research into how to make more linearly independent (orthogonal[10]) parameter spaces.

[10]A simple description of orthogonality is that of vectors that are perpendicular, such as the *X* and *Y* axes of two-dimensional geometry. The etymological origins are from the Greek ὀρθός (orthos), meaning "straight", and γωνία (gonia), meaning "angle" (OLED). The term is used here in its more general vector space definition: Two vectors *x* and *y* in an inner product space V are orthogonal if their inner product is zero. Formally, a linear transformation $\mathrm{T}: \mathrm{V} \rightarrow \mathrm{V}$ is called an orthogonal linear transformation if it preserves the inner product. That is, for all pairs x and y in the inner product space V, $<\mathrm{Tx}, \mathrm{Ty}> \ = \ <\mathrm{x}, \ \mathrm{y}>$. Our concern here is to develop sonification spaces with an equally–spaced metric, that is, that preserve the psychophysical inner–product between a data vector and a linear transformation of it in that space. For a more formal definition of orthogonality, see Behnke et al. (1983, 1:273). There are subtle differences between orthogonality, dependence and correlation (Rogers et al. 1984) that need not particularly concern us here.

An interesting extension to the research on discrete data representation is the overlaying of text-to-speech-sounds derived from the name or functionally related words to the referent, with carefully composed earcons. Called *Lyricons* (lyrics +earcons), experimental results showed that the average accuracy rate for correctly identifying lyricons with their referents was almost double that of just their earcons. Furthermore, once users became familiar with lyricons, subjects could successfully use just the earcon component alone (Jeon 2013). A similar use of such mnemonic devises can be found in bird identification guides where the familiar call of a bird is represented by memorable phonetic or word sequences: *cheerup, cheerily, cheerily* for the American Robin and the *four-o-clock, chokk chokk* of the Australian Noisy Friarbird.

2.2.1.5 Morphocons

Morphocons (contraction of morphological earcons) are an interesting addition to the development (Parseihian and Katz 2012) of earcons. They are short audio units that aim at constructing a sound grammar based on the temporal evolution (morphology) of sound. Morphocons are designed to answer to end-user interface needs (Kraig 2010) in term of aesthetic and sonification customization by addressing issues of navigation performance, efficiency (learnability), and satisfaction (sense of comfort, pleasure, and well-being). Empirical research results indicate that both blind and sighted subjects were able to perceive temporal variation of acoustical parameters as an abstract form and to recognize several categories and subcategories of information on the basis of the morphological sound grammar.

As morphocons are described using simple morphological shapes, they can be applied to any type of sound. The creation of numerous sound palettes could thus be facilitated by processing a diverse collection of sounds with such processes as envelope modification and time- and pitch-stretching, in order to find adequate sound shapes within large sample databases.

2.2.1.6 Discrete Data Representations Compared

The postulation that, because of the connection to fundamental percepts, auditory icons should be easier to learn than earcons, was questioned by Lucas (2004) who found it took significantly less time to learn spoken messages than it did to learn either earcons or auditory icons, and unlike earcons and auditory icons, these spoken messages were consistently interpreted without error. Further, he found no significant differences in the amount of time needed to learn the meanings associated with earcons or auditory icons. This result is at odds with a simple interpretation of ecological theory Gibson (1966, 1979) and Ballas (1994) suggests that the activity of listening contains intermediate mental processes that take account of a listener's expectations and experience, and the context in which the auditing occurs. One could speculate that, because the listener is also immersed in a real auditory scene as well as the computer's virtual one, a differentiation between virtual and real elements requires greater cognitive processing than if the two scenes were not superimposed. Bussemakers and de Haan (2000) undertook a comparative study of earcons and auditory icons in a multimedia environment and found that having sound with a visual task does not always lead to faster reaction times. Although reaction times are slower in the experiment with earcons, it seems that users are able to extract information from these sounds and use it. Furthermore, users find real-life sounds annoying when they hear them frequently.

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

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

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

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

Seeking to extend Pierre Schaeffer's work on a descriptive vocabulary and classification grid for musical objects, researchers and composers of the Laboratory of Music and Informatics of Marseille developed the concept of Temporal Semiotic Units, which are defined as musical segments that convey meaning through their dynamic organization in time. The suggestion (Frémiot 1999) is that there are six classes of environmental sound possessing different morphological specificities. Contrary to the TSU, while the purpose of the morphocons is not to transmit a naturally perceptible meaning, they are based on a morphological description of acoustical parameters, such as envelope, rhythm, frequency, and sound length. So, while they do not allow, as with auditory icons, the establishment of an intuitive link between the sound and the associated represented information, they do, as with earcons, afford the construction of hierarchical sound grammars.

A weakness of making sonification decisions based the ontological classification of discrete data representations, such as earcons, icons, etc., is that an emphasis on obtusely veridical denotations (i.e. formally accurate classification) rather than often subtle, meaningful qualities, can have the effect of discouraging a user's active listening, and so significantly discourage their engagement with the auditory display altogether. Such subtlety is more likely to be achieved by the cohesive use of careful mixtures of real-world audio-graphic and syntactic constructions whose symbolic power comes not from their concreteness or abstractness but from their intuitively understood affect. The design challenge is how to do this without recourse to distractional splendiferous phantasmagorias and in ways that support the user's need to segregate their computed and physical environments with minimum cognitive load.[11]

There have been a lot of improvements in computer user-interface design since the initial auditory icon and earcon research was undertaken, not the least in the quality of sound synthesis, parametric audio playback control and overall system responsiveness. These improvements can support more complex additive synthesis routines and dynamic physical models which may well result in a re-evaluation of the effectiveness of these fundamental sonification structures.

---

[11]A cogent example of this can be found in the practice by automobile manufacturers of using auditory icons and earcons to promote a distinctive "image" of their model at the expense of functional utility or safety. While the "opening the throttle" sound on model SXY to make it throatier that model EXY is of no significant consequence, it is a different matter if, when attempting to enter an autobahn in a hired vehicle on a foggy day, an unseen vehicle approaching from the rear at 200 km per hour causes your vehicle's auditory display system to 'assist you' by issuing a sound you've never heard before and, have no way of being able to identify it in the time required. A standard set of discrete sounds, or better still a user-personalized or learned sound-set seems an eminently more satisfactory solution. Such is the somewhat fetishistic way that sound design is often employed, driverless vehicles are more likely to make the activity of human decision-making redundant before such a reform occurs.

2.2.2 Continuous Data Representations

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

Continuous data sonifications sometimes include discrete elements within an overall analogic mapping: symbolic representations such auditory icons employed to highlight particular informational features and *auditory beacons* (Kramer 1994b) to highlight features such as new maxima and minima, or absolute reference points in a sonification such as ticks, to indicate the regular passing of time. We consider a number of types of continuous data representation: *Direct Data Audification, Parametric Mapping Sonification*, *Homomorphic Modulation Sonification* and an interesting recent addition, *Wave Space Sonification.*

2.2.2.1 Direct Data Audification

Direct data audification is a technique for translating data directly into sound. Kramer (1994b, 186) used the unqualified *audification*, which he describes as "a direct translation of a data waveform to the audible domain for the purposes of monitoring and comprehension." Direct data audification may be applicable as a sonification technique for datasets that have an equally-spaced metric in at least one dimension. It is most easily applied to those that exhibit oscillatory time-series characteristics, though this is not a requirement. Because of the integrative capabilities of the ear, audification is useful as a technique for very large numerical datasets whose datum can be logically arranged as a time sequence of audio samples. These samples can be either stored in an audio file for delayed audition or streamed directly through the computer's audio hardware in real-time. On playback, any number of standard audio-signal processing techniques such as filtering, frequency shifting, sample interpolation, and time-and-amplitude compression can be applied, perhaps under user control, to enhance information detection. The inclusion of these techniques indicates that *direct* is used as a general descriptive, rather than taxonomic, classifier. So, while direct data audification allows for signal-processing techniques, the defining characteristic is that there are no sound-generating or acoustical models used to produce the find result.

Direct data audification has been shown to be effective in situations where the data is voluminous, such as that produced by monitoring physical systems such as in seismology. Speeth (1961) for example, audified seismic data for the 90%

successful differentiation of events caused by bomb blasts from those caused by earthquakes. Speeth's experimental results were remarkable because the task is apparently very difficult to achieve using visual plots of the data (Frysinger 2005). By time-compressing the signals to bring them into audio range, analysts could review twenty-fours-hours' worth of data in a few minutes. Hayward (1994), in describing the use of audification techniques on seismic data, found the technique very useful, but stressed that proper evaluation and comparisons with visual methods are needed. In summarizing a body of work on earthquake data, Dombois (2002) remarks that

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

In monitoring the operation of a complex machine, Pauletto and Hunt (2005) determined that key set of attributes (noise, repetitive elements, regular oscillations, discontinuities, and signal power) in helicopter flight data were discernable equally well using an audification as a visual spectrogram. Krishnan et al. (2001) undertook a comparative study of direct data audification and other sonification techniques to represent data related to the rubbing of knee-joint surfaces and did not find that audification was the best technique for displaying the difference between normal and abnormal signals.

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

In audification, the data *becomes* the signal and so is the technique most intimately connected to the source of the data being sonified. Despite this directness, it would be a mistake to assume that the technique is useful in uniquely identifying the source of the data *a priori*. However, while there is only a weak cognitive binding between a particular sound's origin as a result of specific physical activity and the sounds produced, as discussed in Chap. 1, there is an easy propensity to identify and remember quite intimate details of the characteristics of such activity once a realistic description or a strongly metaphorical representation is provided. For example, in informal research conducted by the author, when un-informed listeners are played examples of the audification of earthquake data and asked to identify the source of the sound, their responses frequently include quite detailed gestural profile of rain or hail on a metal roof, or the amplified scrunching of paper, but they have never identified an earthquake as the source of the data. However, having learned how the audio was produced, they were easily able to intelligently answer questions about the cracking and crunching of the earth's crust that had occurred in the earthquake itself.

2.2.2.2 Parametric Mapping Sonification

Parameter mapping is the most widely used sonification technique for representing high-dimensional data as sound. Parameter mapping sonifications are sometimes referred to as sonic scatter plots (Flowers et al. 1997; Flowers 2005) or nth-order parameter mappings (Scaletti 1994). Typically, data dimensions are mapped to sound parameters: either to physical (frequency, amplitude), psychophysical (pitch, loudness) or perceptually coherent complexes (timbre, rhythm). Analogic variations in the sound can result when mapping from a large data domain into a small perceptual range or when data is specifically mapped to acoustic modifiers such as frequency or amplitude modulators. Parametric mapping sonification is sometimes referred to as *multivariate data mapping*, in which multiple variables are mapped to a single sound. Scaletti describes one way of implementing it by "mapping of each component of a multidimensional data point to a coefficient of a polynomial and then using that polynomial as the transfer function for a sinusoidal input" (1994).

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

The main limitation of the technique is the lack linear independence or orthogonality in the psychophysical parameter space: loudness can affect pitch perception, for example. Though conceptually simple, in practice, parameter mapping requires a working knowledge of how the parameters interact with each other perceptually: At the very least, linear changes in one domain produce non-linear auditory effects, and the range of the variation can differ considerably with different parameters and synthesis techniques. These perceptual interactions, caused by coupled perceptual parameters, can obscure data relations and confuse the listener.

Flowers, an experienced multivariate data sonifier, observed that while "the claim that submitting the entire contents of 'dense and complex' datasets to sonification will lead to the 'emergence' of critical relationships continues to be made, I have yet to see it "work" (Flowers 2005). However, although a truly balanced multivariate auditory display may not be possible in practice Kramer (1994b) suggested that, given powerful enough tools, it may be possible to heuristically test mappings to within acceptable limits for any given application. Frysinger (2005) provided a useful overview of the history of the technique, and Flowers (2005) highlighted some of its pitfalls and possible future directions.

In audification, the relationship between data and audio is so direct that the data becomes, or is transformed, more or less directly into sound. However, in parameter-mapping sonification, analytically identifiable features in the data are used as inputs to mapping functions that *control* a set of sound synthesis parameters. The considerable flexibility that results in this less-direct approach comes at some cost, including, as discussed in the previous chapter, the privileging for cultural not psycho-perceptual reasons, of the qualities of the sounds themselves over the algebra of their temporal connectivity. So, the speculation is that orthogonality problems have two principal causes: the lack of necessary perceptual redundancies due to the restriction of mappings to individual psychophysical parameters, and the lack or too-simplistic application of sound event connectives such as micro-gestures, as discussed in more detail in Worrall (2014) and in Chap. 4. Nevertheless, parameter-mapping sonification remains the most common technique of data sonification in use today, and is thus the technique discussed in most detail throughout this book.

2.2.2.3 Homomorphic Modulation Sonification

A homomorphic mapping is one in which the changes in a dimension of the auditory space tracks changes in a variable in the dataset, with only as few mediating translations as are necessary for comprehension (Kramer 1994a, 26). This section describes a narrow interpretation of the term; here named *Homomorphic Modulation Sonification*. There is a subtle but important distinction between the mapping described here and the parametric mapping approach, described above, in which each datum is played as, or contributes to a separate tone with its own amplitude profile or 'envelope'. In the separate-tones case, the audio-amplitude profile of the resulting audible stream fluctuates from–and–to zero while with homomorphic modulation, a single continuous pulsed waveform results. In the case of frequency modulation, there is the opportunity for the amplitude formant to be held relatively constant. In research discussed in Chap. 7, this appears to assist the listener's auditory system to respond to it as a single modulating "tone" rather than a sequence of auditory objects individuated by rapid onset transients. This difference is illustrated in Fig. 2.1 and may result in lower perceptual loading, especially tone extended listening periods. Patterson's Warning-Design Guidelines study of navigation alerts mentioned earlier (Sect. 2.2.1.1), support this suggestion in recommending slower onsets/offsets times to avoid startling the auditor; to impress itself upon their consciousness without it dominating cognitive function.

Pre-attentive perceptual faculties are also applicable when the data is used to frequency– or amplitude–modulate a simple carrier signal. This is a relatively

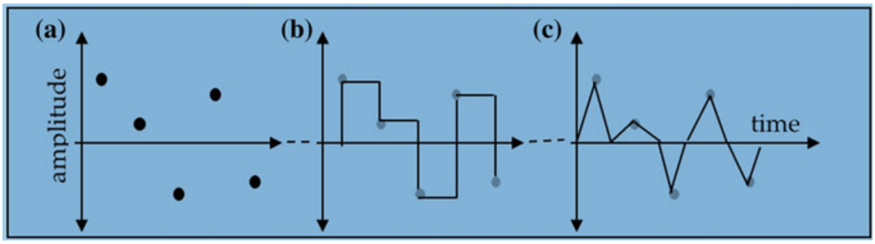


Fig. 2.1 The different amplitude profiles of **a** the sample the samples **b** being realized with amplitude modulation, and **c** individually enveloped events

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

This possible difference in perceptual loading is emphasized by the noticeable difference between the resultant effect of passing parameter mapped and homomorphic modulation sonification of signals through post-synthesis signal 'enhancing' modification such as reverberation. In the case of parametric mapping, the by-products appear to enhance the original, making it more embodied, while for the homomorphic mappings, they appear as distractional, noisy artifacts, as is discussed in the homomorphic modulation sonification study reported in Chap. 7.

The observation that (sonic) object continuity enhances auditory attention has received research attention in the wider context of object continuity in the attentional loading of perception (Best et al. 2008). The acoustic startle reflex (Ladd et al. 2000) is a gross version of the effect described here. This effect can be summarized as follows:

> The minimum stimulus required to elicit a response in normal rats is about 80–90 dB of 50 ms. duration, provided that the stimulus reaches full potential within about 12 ms. initial onset. In mammals, the amplitude of the startle response is proportional to the duration of the stimulus… The startle reflex has been observed in all mammals tested.

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

2.3 Interactive Data Representations

2.3.1 Sound Graphs

The term *auditory graph* is used in a variety of ways, often simply meaning the output of a multivariate data sonification. In order to provide a restricted meaning, the term *sound graph* is used here to refer to a sonic representation of a visual graph (Stockman et al. 2005a; Harrar and Stockman 2007). Other names by which it is known are *tone graph*, *auditory graph*, *tree-graph* and *auditory box plot*. Their function is to provide sonified interfaces to discrete datasets so that the relationships between the data points can be investigated interactively and asynchronously. The addition of auditory tick-marks, axes, beacons and labels to add context is not uncommon (Stockman et al. 2005b) as discussed earlier in this chapter.

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

2.3.2 Physical Models and Model-Based Sonifications

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

Model-based sound synthesis has proved effective in producing complex, ‘natural’ sounding acoustic events (see Pearson 1996 for an extensive bibliography). While the building of such models requires considerable skill, this would not be a barrier for their use in data sonification if templates were readily available. And, although they can be computationally expensive, such models tend to lend themselves to parallel computing. The application of these models to sonification have not yet been widely adopted, and the empirical testing of them against other sonification methods will, in time, determine the degree to which they become generally accepted techniques.

One reason for exploring such techniques is to try to improve the dimensional orthogonality of a sonification by integrating a dataset in, or through, a quasi-physical medium.[12] Attempts to produce acoustically–dimensioned linear psychophysical spaces, such as with equally-distanced timbres (Barrass 1997, 88–216), so as to differentiate multiple orthogonal dimensions, on which parameter-mapping techniques rely, have not been yet been very successful. The representational dimensions of physical systems, on the other hand, *appear* to be more promising as they are of a higher order than acoustic parameter-maps, involve the redundant use of acoustic parameters, and are structurally closer to those in ecological use in every day listening. In order for them to become more generally used, it will be necessary to make intuitive implementation tools more widely available.

There are two basic kinds of models available: those based on the known physical properties of acoustic resonators which are 'played' by data (*physical models*) and those in which the resonator is 'constructed' from data and interacted with ('played') in order to explore the structure of the data.

2.3.2.1 Physical Modeling

Digitally modeling the resonating components of a musical instrument (called *physical modeling*) is a relatively common practice in computer music. For such dynamic models, the temporal behavior of the components of the model are determined ahead of time by detailing their resonant properties and the way they are connected together to form a single resonator. The instrument is then excited with virtual bows, drumsticks, scrapers etc. and virtual contact-microphones placed at strategic places on the 'body' of the instrument to capture its resonance (Pearson 1996). Such models can be used in generating discrete sounds (for auditory icons and earcons, for example) but also as data-driven resonators capable of responding in aurally interesting and complex ways, consistent with a listener's intuitive understanding of the physical world (Farnell 2010).

2.3.2.2 Model-Based Sonification

Model-based sonification allows one to create virtual data instruments, which can then be investigated to reveal information about the dataset by the nature of the sounds produced in response a user's interaction with the model. In this sonification

[12]This matter is discussed more fully in Chap. 3.

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

Model-based sonification has been demonstrated to be effective in a number of situations, including in being able to intuitively perceive the complexity or dimensionality of a data set (Hermann and Ritter 2004), a problem that had been previously explored using parameter-mapping sonification (Bly 1994). Hermann (2011, 399–427), the author of the technique, has provided a comprehensive overview, examples, design and interaction guidelines and suggested interaction models.

2.3.2.3 Wave Space Sonification

A scalar field is a space in which each defined point has only a single magnitude. A scalar field associates a scalar to every sample point, and is often used to indicate some physical properties such as temperature, pressure or density. More complex multivariate systems produce several streams of data simultaneously for each (indexed) point in time. Examples of multivariate data streams include those from EEG, seismological sensors, and motion capture technologies.

In Wave Space Sonification, the scalar field is called a wave space and this space is scanned along a trajectory which is data-defined. The data are indexed according to a scalar variable; for example, time-indexed data, such as for an audio signal. A common visualization of such data is achieved by vertically stacking time-aligned plots of the functions of each variable of the system: f_1, $f_2 \ldots$ f_d. A time-ordered sequence of tuples of scalar values of all the functions in the system thus provides a trajectory in d-dimensional space:

$$([f_1(t_1), f_2(t_1), \ \ldots F_d(t_1)], \ \ [f_1(t_2), f_2(t_2), \ldots F_d(t_2)], \ \ [f_1(t_n), f_2(t_n), \ldots F_d\left(t_{n)}\right)])$$

Patterns resulting from different trajectories through the space represent different states of the system. For example, similar to the simpler techniques of Borgonovo and Haus for sound synthesis using two variable functions (1986), orbits in the space correspond to oscillations in the time-series. Various methods, such as embedding previous tuple sequences as delays in an evolving time series, can be employed to extend sequences of patterns which shift in time to reveal dynamic system states. Such techniques provide a means of studying data from evolving systems such as complex biological processes.

Other techniques are possible. For example, reminiscent of the technique developed for the scanned synthesis of musical sounds (Verplank et al. 2001), a morphing function can be applied to control exactly when a move along the trajectory takes place: increasing or decreasing the tuple sample rate—according to the rate of change of successive tuples. Other related sound synthesis techniques include those for Wave Train Synthesis (Mikelson 2000; Comajuncosas 2000).

The inventor of the Wave Space Sonification technique, Hermann (2018), describes several *instances* of it: Canonical, Sample-based, Data-driven Localized, and Granular. Overall, he classifies the technique as fitting conceptually between traditional audification and parameter-mapping sonification, with audification being a special case.

2.4 Sonification Research

2.4.1 Music Psychology

Although musical activity can raise useful suggestions when sonifying data relations, music is as an ancient, universal human activity with an astonishing breadth: culturally, structurally and expressively. Although music may be composed of syntactic-semantic structures, there is no universal musical language any more than music in an Ionian (minor) mode is implicitly sad. The 'rules' of functional harmony were developed iteratively through many generations of musicians across time, in the process of composing, not the other way around and, crucially, there is no absolute requirement that these structures be made explicit, nor even aurally coherent in a musical composition. A characteristic of all kinds of music is that they are continually being recontextualized, both technically and culturally, as they are influenced by, and influence, the fluctuating zeitgeist of the cultures in which they exist. Furthermore, there is no one way, or reason, to listen to music; different musics require different ways of listening in different contexts which may involve whole complexes of social dimensions that not be relevant to the perceptualization of data relations (Tuuri et al. 2007; Tuuri and Eerola 2012).[13] Attempts to impose such "musical-isms" on a sonification invariably appears false and more likely to increase blandness and confusion than assist under-standing or make it more palatable.

The advent of the Internet and the percolation of computational capacity into ordinary, everyday objects and activities, means that most of us live in increasingly datarized cultural environments. So, from a cultural, experimental composition perspective, data sonification is of significant interest and this introduction to data sonification would not be complete without acknowledging the use of the terms *sonification* and *data sonification* as a genre of music. Data Music is not the principal focus of this book and so its treatment here is for completeness and somewhat cursory. One way of classifying data music (also known as *data sonification music*)

[13]These modes of listening are explored in Chap. 4, Sect. 4.2.4.

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

At the representational end is data music that employs the techniques discussed in this book; techniques that are derived from a developing knowledge of psychoacoustics, and the cognitive sciences more generally; that are primarily used in pragmatic information interpretation for decision-making, system regulation, vision substitution etc. For composers who are interested in engaging with listeners as interpreters rather than the struggling receivers of obscure–or even blatant–messages, this research offers a stronger technical basis for their practice, not necessarily for overtly programmatic narratives but to better control the overall dramaturgy of a composition (Landy 2007, 36–38). For example, although the techniques for making adherent tone complexes have formed the basis of studies in orchestration for at least two centuries, in order to discover how to disassemble the ensemble, to compose coherent sounds that maintain their independent perceptual clarity and segmentation, what is needed is a kind of *unorchestration studies*, which results in what was referred to in Chap. 1 as *sonic gumbo*. Such studies would include psychoacoustic techniques for producing disaggregated complex sounds in situational environments (Bregman 1994, 458) and a developing understanding of auditory cognition.

2.4.2 Computational Tools

Computer-based composition tools have their origins in the desire to explore and express algorithmically-generated sonic structures at both micro-(sound synthesis) and macro-(gestural) compositional levels. As a cultural activity, making new music is not just about exploring new ways to put new noises together but about providing other people with opportunities to engage, both actively and passively, in that exploration. Once digital musical instruments became available, MIDI[14] was quickly adopted as a computer protocol suitable for capturing and transmitting data from human performance gesture and for coordination with other technologies, such as lighting controllers. However, not all interesting sources of data have MIDI interfaces or are related to human corporeality and proprioception. So, at a time in human history when there is an imperative for us to take responsibility for our

[14]Musical Instrument Digital Interface. See https://www.midi.org/.

relationships with both natural and man-made environments, the extension of sound synthesis software to aurally explore such datascapes opens the possibility of compositional access to, and thus cultural dialogue about, a broader spectrum of phenomena: interspecies, planetary and interstellar.

The ability for digital computers to generate, transform and present sound to users in a timely manner is fundamental to sonification research. The vast majority of the tools and techniques used for the computer synthesis of sound have been developed by composers and engineers engaged in the task of making new music, leading John Chowning to remark, “With the use of computers and digital devices, the processes of music composition and its production have become intertwined with the scientific and technical resources of society to greater extent than ever before” (Chowning in Roads 1996, ix). Because, as he goes on to say, “a loudspeaker controlled by a computer is the most general synthesis medium in existence”, these researchers were quick to begin exploring adaptations of, and eventually alternatives to, the static models of musical instrument tones as described in the literature of the time (Olsen 1967, 20–41; Olsen 2002) an experience which analog synthesis had already revealed to be deficient of the lively qualities of the sounds of acoustic instruments.

The original periodic synthesis techniques include additive, subtractive and modulation models (Mathews 1969), which were augmented by dynamic (non-periodic) techniques such as microphone-sampled waveforms, stochastic function and granular synthesis-generated timbres (Xenakis 1971, 242–254; Truax 1988; Roads 2004). As computing hardware became faster and object-oriented software tools became more prevalent, software models of physical resonance systems were developed, ranging from difference equations, mass-spring, modal, non-linear excitation, waveguide, and formant (after speech) synthesis to the highly individual Karplus-Strong techniques for plucked-string and drum-like sounds. Roads provides a brief introduction to these techniques (1996, 261–315), most of which are still under active development. Their usefulness to sonification is that they provide appropriate coupling mechanisms between data structures and the synthesis models being used to express them. As Hermann (2002) concludes, apart from the often–considerable time involved in setting up such systems, their major disadvantages are the expertise needed to set up the models, and the computing resources necessary to implement them. As experience grows and computation speeds increase, neither of these impediments is likely to prove permanent. Computing science and software engineering continues to play important roles in the development of both the theoretical models and efficient algorithmic processes within which these sound synthesis techniques can be implemented.

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

The formation and support of such embodied ‘in-formation’ is a difficult task and remains a long-term goal. It will be aided by continued active cross-fertilization between psychophysics, the relatively new field of cognitive science, and the older disciplines of the philosophy and psychology of perception.

2.5 Design

The systematic study of the design issues involved in using computers to *display* information to the human auditory system is a relatively young discipline. The first International Conference on Auditory Display (ICAD) was held in 1992 and Kramer’s extensive introduction to the *Proceedings* (1994b) provides an overview of interesting precedents, the then current state of the field, as well as some possible futures. Before the computing resources with which to process large amounts of data were readily available, sonification was restricted to relatively simple tasks, such as increasing the frequency of a metal detector’s audio oscillator in proportion its proximity to metals, and the timely production of attention-capturing beeps and blurbs. As the size and availability of datasets have grown and computational speed has afforded the increase in the complexity of communication tasks, so also have the expectations of sonification; requiring sonifiers to draw on skills and research knowledge from a widening variety of disciplines.

In such interdisciplinary and transdisciplinary enquiries, researchers from various fields come together to share and integrate their discipline-specific knowledge and in doing so reinvigorate themselves with new questions and new methods. For example, when introducing the concept of intelligent help using knowledge-based systems, (Pilkington 1992, vii) notes that different disciplines often adopt different forms of evidence. While psychologists prefer experiments, computer scientists prefer program and abstract algebras, composers prefer sound-organizing activities, and linguists prefer grammatical and phonetic rules, and this plethora of approaches can lead to communication failures when they try framing a theory that will satisfy each discipline’s criteria of proof. Moore (1990, 23) observes that in interdisciplinary research, the challenge is to do justice to several points of view simultaneously because, “An awareness of [a variety of] points of view is important not so much because ignorance of any one of them makes it impossible to do anything but because what may be done will eventually be limited by that lack of awareness.” As experience continually reaffirms, knowledge is not value free, so when these differing points of view come from disciplines with divergent aims, a lack of such awareness may not be obvious. The opportunities to more closely examine disciplinary assumptions that collaborative work provides to researchers can lead through apparent impasses and strengthen the foundations of domain knowledge.

In this context, it is worthwhile to speculate on the kinds of knowledge and experience a cultural activity such as music composition can bring to the data sonification process, and inversely, of what use are the results of experimental psychology to music composition. As numerous historical examples attest–from the ancient Greeks, who believed in a direct relation between the laws of nature and the harmony of musical sounds, through the numerical symbolism of the Middle Ages and beyond, musicians were sonifying numerical data long before computers and multivariate datasets occupied the Western mind.[15] While their approaches were perhaps more heuristic and teleological, surviving artifacts from past eras are more likely than not to have been either workable solutions to practical problems, or sufficiently novel to have induced a new way of thinking music. However, except of the very recent past, no sound recordings are extant, and the function of the musical score as something more than an aide-mémoire is only a relatively recent one.

So, although musicologists have described many of the un-notated performance practice and 'style' characteristics of the music from contemporaneous non-musical sources, such descriptions are rarely couched in the phenomenal language of modern science. From the contemporary design perspective, the appropriation of scripted musical gestures of the historical past, together with some a-historical harmony book's 'rules of composition' which omit reference to the tuning or temperament system employed, or the timbral distinctiveness of the instruments of the period, fails to grasp the cultural imperative: music is not science, and its aims are not always for clarity of line, or purpose. The sounds of music are social sounds and as society changes, so the ways its components interact are determined: as much by cultural forces as by psychoacoustics (Attali [1977] 1985, 46–86; Chion 2016; Erlmann 2010).

The parametric analysis of tone according to physical characteristics (frequency, loudness, etc.) and the availability of computational tools for additive synthesis from these parameters, belies the psychoacoustic evidence that, for all but the simplest sonifications, the results of design by simply assigning data dimensions to such physical parameters quickly leads to unsegregated, difficult-to-interpret clumping (Flowers 2005). Whether stream segregation is maintained during a temporal simultaneity of two or more separate spectra as components of textures or 'chords', or whether integration, occlusion or emergence results (Bregman 1994, 456–528) is dependent upon a combination of characteristics, including the system of tuning in use, the spectral density profiles (timbres) involved, the duration of the simultaneity, the previous and following directional profiles of the spectral contributors, their relative amplitudes, and so on. Furthermore, hearing is very sensitive to spectral congruence, as illustrated by the account-auditing example in the first chapter. Sequences of spectral events that are highly congruent, such as occurs when using General MIDI instruments, or auditory icons for example, quickly pall on the ear. The extra cognitive load needed to work within such soundscapes is not insignificant, and possibly contributes to the annoyance commonly reported by listeners when using such sonifications for extended periods of time. There are ways out of this impasse, as illustrated by the introduction to sound synthesis techniques in common use in computer music, such as physical modelling (including voice), granulation and stochastics, as well as careful attention to

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

second-order features such as reverberation and spatialization; these latter being employed in the environmental "embodying" of sound sources. The importance of embodiment in information sonification, is discussed in Chaps. 3 and 4.

Flowers (2005) comments that most of the problems that arise in data sonification stem from lack of adequate understanding about key properties of auditory perception and attention, and from inappropriate generalizations of existing data visualization practices. It is common to suggest that such generalizations arise more as a result of the overwhelming dominance of vision over audition in the perceptual psychology literature than from experimental evidence (Bregman 1994, 1–3). They can be overcome with more sustained applied research in auditory perception, including the rôle played by silence, and a deeper understanding of the relationship between linguistic speech and non-speech auditory processing (Slevc et al. 2008). For sonification design to advance, *sonification designing* needs to be practiced and critiqued as designs to be interpreted, not just listened to as *musique e d'ameublement.*[16] To that end, task and data analysis of information requirements (Barrass 1997, 35–45), together with generally available datasets and interesting mapping templates in software environments that combine sonic complexity with experimentation flexibility, will encourage the much-needed community-wide sharing of examples towards a catalogue of best practice.

How sonification designers attempt to achieve an effective communication solution with a sonification design task is affected by many things, including the imagined range of possible solutions for the design (the state space), which in turn, is affected by the tools and methodologies employed, and the skills applied in using them. Attempting to evaluate the extent to which the solution is optimal, and how effective it is in achieving a stated goal, remain open questions because the connection between the individual decisions that are made in the designing process and the way they interact in the solution space are non-linear: at best open to interpretation, and at worse a collection of individualized black-box heuristics. Such a description is rarely controversial, even by those calling for robust evaluation and scientific comparison of sonification methods, as it is understood that:

> In the context of data exploration, what can be potentially learnt from a sonification is unknown, or at least not defined properly, and therefore it is very difficult to specify an objective performance measure (Degara et al. 2013).

Design is a messy business and the relationship between decisions made in the process of 'tweaking' the contribution of individual parameters in the final result is rarely the sum of simple linear combinations. So, being able to evaluate the effectiveness of a sonification for a clearly defined purpose is not the same as being able to determine what aspects of the sonification are responsible for or contribute to that effectiveness. This is no different in form to being able to construct a neural-network to simulate an observable behavior without being able to understand the complexity of the individual contributions to the network itself.

---

[16]Literally "music of furniture", as so called by the composer Eric Satie.

2.6 Summary

This chapter began with an informal history of the evolution of the term *data sonification*. It emphasised the point that the development of the term has reflected a growing awareness in the data sonification research community of the importance of building the listener and their intentions into what became increasingly clear were better characterised as descriptions than succinct definitions of the practice of sonifying, and the sonifications that result. The bulk of the chapter is devoted to discussing various types of sonification according to their primary sonic identities, namely discrete representations: earcons, auditory icons, and various speech-related objects; continuous representations: direct data audification, parameter-mapping sonification and homomorphic modulation; interactive representations: sound graphs and model-based sonification and a new class of sonification techniques called wave space sonification. The chapter concluded with a brief introduction to some issues in three dimensions of the interdisciplinary practice of data sonification: music psychology, computational tools and auditory design.

Reaching back to the ancient Greeks, the next chapter provides an overview of Western civilization's attempt to understand the nature of perception and how human beings use it to construct a coherent understanding of their worlds. This endeavor was a fulcrum in developing the empirical method on which the last 400 years of scientific endeavor is based and continues to be a central theme of Western philosophy. It discusses the inadequacy of such approaches to creating intentional inexistent objects when they do not take into account the deeply embodied nature of perception, resulting in a more critical examination of epistemological and ontological dimensions; the cultural 'background radiation' against which new, more effective, sonification software can be designed.

References

Alten SR (2002) Audio in media, 6th edn. Wadsworth Thomson Learning, Belmont, CA

Anderson J, Sanderson P, Norris M (2002) The role of auditory attention and auditory perception in the design of real-time sonification of anaesthesia variables. In: Proceedings of HF2002 human factors conference, pp 25–27

Attali J (1985) Noise: the political economy of music. Translated by B. Massumi. University of Minnesota Press, Minneapolis, MN [1977]

Ballas JA (1994) Delivery of information through sound. In: Kramer G (ed) Auditory display: sonification, audification, and auditory interfaces. In: Proceedings, Santa Fe Institute studies in the sciences of complexity, vol XVIII. Addison Wesley, Reading MA, pp XVIII:79–94

Ballas JA (1996) Designing sound for an interactive experience. AES, Los Angeles, CA

Barrass S (1997) Auditory information design. PhD thesis, Australian National University. https://openresearch-repository.anu.edu.au/handle/1885/46072

Behnke H, Bachmann F, Fladt K, Süss W (eds) (1983) Fundamentals of mathematics, vol 1. The MIT Press, Cambridge, MA

Best V, Ozmeral EJ, Kopco N, Shinn-Cunningham BG (2008) Object continuity enhances selective auditory attention. Proc Natl Acad Sci USA 105(35):13174–13178

Blattner MM, Sumikawa D, Greenberg R (1989) Earcons and icons: their structure and common design principles. In: Human computer interaction, vol 4, no 1. Lawrence Erlbaum Associates, London, pp 11–44

Blattner MM, Papp AL, Glinert EP (1994) Sonic enhancement of two-dimensional graphics displays. In: Kramer G (ed) Auditory Display: sonification, audification, and auditory interfaces. Proceedings, Santa Fe Institute studies in the sciences of complexity, vol XVIII. Addison Wesley, Reading MA, pp XVIII:447–470

Bly S (1994) Multivariate data mappings. In: Kramer G (ed) Auditory Display: sonification, audification, and auditory interfaces. Proceedings, Santa Fe Institute studies in the sciences of complexity, vol XVIII. Addison Wesley, Reading MA, pp 405–416

Borgonovo A, Haus G (1986) Sound synthesis by means of two-variable functions: experimental criteria and results. Comput Music J 10(3):57–71

Bregman A (1994) Auditory scene analysis: the perceptual organization of sound. The MIT Press, Cambridge, MA

Brewster SA, Wright PC, Edwards ADN (1994) A detailed investigation into the effectiveness of earcons. In: Kramer G (ed) Auditory Display: sonification, audification, and auditory interfaces. Proceedings, Santa Fe Institute studies in the sciences of complexity, vol XVIII. Addison Wesley, Reading MA, pp 471–498

Brewster SA, Wright PC, Edwards ADN (1995) Parallel earcons: reducing the length of audio messages. International Journal of Human-Computer Studies 43(2):153–75. https://doi.org/10.1006/ijhc.1995.1039

Bussemakers MP, de Haan A (2000) When it sounds like a duck and it looks like a dog… auditory icons vs. earcons in multimedia environments. In: Proceedings of the tenth international conference on auditory display. Georgia Institute of Technology, Atlanta, USA, pp 184–189

Chion M (1994) Audio–vision: sound on screen. Edited and Translated by C. Gorbman. Columbia University Press, New York

Chion M (2016). Sound: an acoulogical treatise. Translated by J.A. Steintrager. Duke University Press

Comajuncosas JM (2000) Wave terrain synthesis with csound. In: Boulanger R (ed) The csound book: perspectives in software synthesis, sound design, signal processing, and programming. MIT Press, Cambridge, MA

de Campo A (2007) Toward a data sonification design map. In: Proceedings of the 13th international conference on auditory display, Montréal, Canada, pp 29–342

Dean RT (2005) NoiseSpeech, a noise of living bodies: towards Attali's 'Composition'. NMEDIAC: J New Media Cul 3. http://www.ibiblio.org/nmediac/winter2004

Dean RT, Bailes F (2006) NoiseSpeech. Perf Res 11(3):85–86

Dean RT, Bailes F (2009) When is noise speech? A survey in sonic ambiguity. Comput Music J **33**, 1–11

Degara N, Nagel F, Hermann T (2013) Sonex: an evaluation exchange framework for reproducible sonification. In: Proceedings of the international conference on auditory display (ICAD 2013). Łódź University of Technology, Łódź, Poland, pp 167–174

Dehaene-Lambertz G, Gliga T (2004) Common neural basis for phoneme processing in infants and adults. J Cogn Neurosci 16:1375–1387

Dombois F (2002) Underground sounds. An approach to earthquake prediction by auditory seismology. Geophys Res Abst 4:02–02741

Erlmann V (2010) Reason and resonance: a history of modern aurality. Zone Books, Brooklyn, NY

Farnell A (2010) Designing sound. MIT Press, Cambridge MA, USA

Flowers JH (2005) Thirteen years of reflection on auditory graphing: promises, pitfalls, and potential new directions. In: Proceedings of the first symposium on auditory graphs, Limerick, Ireland

Flowers JH, Turnage KD, Buhman DC (1997) Cross-modal equivalence of visual and auditory scatterplots for exploring bivariate data samples. Hum Factors 39:341–351

Frémiot M (1999) De l'objet musical aux unités sémiotiques temporelles. In: Ouïr, entendre, écouter, comprendre après Schaeffer. INA Buchet-Chastel, Paris, pp 227–246

Frysinger SP (2005) A brief history of auditory data representation to the 1980s. In: Proceedings of the first symposium on auditory graphs, Limerick, Ireland

Gaver WW (1994) Using and creating auditory icons. In: Kramer G (ed) Auditory display: sonification, audification, and auditory interfaces, vol XVIII. Santa Fe Institute Studies in the Sciences of Complexity: Addison Wesley, pp 417–446

Gaver WW, Smith RB (1990) Auditory icons in large-scale collaborative environments. In: Proceedings of human-computer interaction: interact 1990, Cambridge, UK, pp 735–740

Gibson JJ (1966) These senses considered as perceptual systems. Houghton Mifflin Company, Boston, MA

Gibson JJ (1979) The ecological approach to visual perception. Houghton Mifflin Company, Boston, MA

Harrar L, Stockman T (2007) Designing auditory graph overviews: an examination of discrete vs. continuous sound and the influence of presentation speed. In: Proceedings of the thirteenth international conference on auditory display, 26–29 June 2007, Montréal, Canada

Hayward C (1994) Listening to the earth sing. In: Kramer G (ed) Auditory Display: sonification, audification, and auditory interfaces. Proceedings, Santa Fe Institute studies in the sciences of complexity, vol XVIII. Addison Wesley, Reading MA, pp 369–404

Hermann T (2002) Sonification for exploratory data analysis. PhD thesis, Bielefeld University, Bielefeld, Germany

Hermann T (2011) Model-based sonification. In: Hermann T, Hunt T, Neuhoff JG (eds) The sonification handbook. Logos Publishing House, Berlin, Germany, pp 399–427

Hermann T (2018) Wave space sonification. In: Proceedings of the 24th international conference on auditory display (ICAD 2018). Michigan Technological University

Hermann T, Ritter H (1999) Listen to your data: model-based sonification for data analysis. In: Lasker GE (ed) Advances in Intelligent Computing and Multimedia Systems. The International Institute for Advanced Studies in System Research and Cybernetics, Baden-Baden, Germany, pp 189–194

Hermann T, Ritter H (2004) Neural gas sonification: growing adaptive interfaces for interacting with data. In: Proceedings of the information visualisation, eight international conference (IV 2004). IEEE Computer Society, Los Alamitos and Washington, DC, USA, pp 871–878

Husserl E (1927) Phenomenology. Article for encyclopaedia britannica. Translated by RE Palmer. Husserl's Shorter Works, vol 2, no 2, pp 77–90

Jeon M (2013) Lyricons (Lyrics+Earcons): designing a new auditory cue combining speech and sounds. In: Stephanidis C (ed) HCII 2013, CCIS 373, Posters, Part I. Springer, Berlin, Heidelberg, pp 342–346

Jeon M, Walker BN (2009) 'Spindex': accelerated initial speech sounds improve navigation performance in auditory menus. In: Proceedings of the human factors and ergonomics society. Annual Meeting; Human Factors and Ergonomics Society. Red Hook, New York and San Antonio, TX, USA

Kraig F (2010) The usability metric for user experience. J Interact Comput 22(5):323–327

Kramer G (1994a) Auditory Display: sonification, audification, and auditory interfaces. Proceedings, Santa Fe Institute studies in the sciences of complexity, vol XVIII. Addison Wesley, Reading MA

Kramer G (1994b) Some organizing principles for representing data with sound. In: Kramer G (ed) Auditory Display: sonification, audification, and auditory interfaces. Proceedings, Santa Fe Institute studies in the sciences of complexity, vol XVIII. Addison Wesley, Reading MA, pp 185–221

Kramer G, Walker B, Bonebright BT, Cook P, Flowers J, Miner PN, Neuhoff J (1999) Sonification report: status of the field and research agenda. Technical report. International Community for Auditory Display

Krishnan S, Rangayyan RM, Frank CB, Douglas BG (2001) Auditory display of knee-joint vibration signals. J Acoust Soc Am 110(6):3292–3304

Ladd CO, Plotsky PM, Davis M (2000) Startle response. In: Fink G (ed) Encyclopedia of stress, vol 3. Academic Press, San Diego, CA p 486

Landy L (2007) Understanding the art of sound organization. MIT Press, Cambridge, MA

Lucas PA (2004) An evaluation of the communicative ability of auditory icons and earcons. In: Kramer G (ed) Sante Fe institute studies in the sciences of complexity, proceedings, vol XVIII. Addison Wesley, pp 79–94

Mathews MV (1969) The technology of computer music. The MIT Press, Cambridge, MA

McCormick EJ, Sanders N (1984) Human factors in engineering and design, international student edition. McGraw-Hill

McGookin DK (2004) Understanding and improving the identification of concurrently presented earcons. PhD thesis, University of Glasgow, UK

Mikelson H (2000) Terrain mapping synthesis. In: Boulanger R (ed) The csound book: perspectives in software synthesis, sound design, signal processing, and programming. MIT Press, Cambridge, MA

Moore FR (1990) Elements of computer music. Prentice Hall

Mynatt ED (1994) Auditory presentation of graphical user interfaces. In: Kramer G (ed) Auditory Display: sonification, audification, and auditory interfaces. Proceedings, Santa Fe Institute studies in the sciences of complexity, vol XVIII. Addison Wesley, Reading MA, pp 533–556

Nesbitt K, Barrass S (2002) Evaluation of a multimodal sonification and visualisation of depth of market stock data. In: Proceedings of the eighth international conference on auditory display, 2–5 July 2002, Kyoto, Japan

Olsen HF (1967) Music, physics and engineering. Dover Publications

Olsen R (2002) High Frequency Data-an essential resource. In IFC Bulleting 13

Palladino DK, Walker BN (2007) Learning rates for auditory menus enhanced with spearcons versus earcons. In: Proceedings of the thirteenth international conference on auditory display, 26–29 June 2007, Montréal, Canada

Parseihian G, Katz Brian F G (2012) Morphocons: a new sonification concept based on morphological earcons. J Audio Eng Soc 60(6):409–418

Patterson RD (1982) Guidelines for auditory warning systems on civil aircraft. Civil Aviation Authority, London

Patterson RD (1989) Guidelines for the design of auditory warning sounds. In: Proceeding of the institute of acoustics, spring conference, vol 11, no 5, pp 17–24

Pauletto S, Hunt A (2005) A comparison of audio and visual analysis of complex time-series data sets. In: Proceedings of the eleventh meeting of the international conference on auditory display, 6–9 July 2005, Limerick, Ireland

Pearson M (1996) TAO: a physical modelling system and related issues. In: Organised sound, vol 1. Cambridge University Press, pp 43–50

Pilkington RM (1992) Intelligent help: communicating with knowledge-based systems. Oxford University Press, New York

Polotti P, Lemaitre G (2013) Rhetorical strategies for sound design and auditory display: a case study. Int J Des 7(2):67–83

Roads C (1996) The computer music tutorial. MIT Press, Cambridge, MA

Roads C (2004). Microsound. The MIT Press, Cambridge, MA

Roberts LA, Sikora CA (1997) Optimising feedback signals for multimedia devices: earcons vs. auditory icons vs. speech. In: Proceedings of the international ergonomics association congress, Tampere, Finland, pp 224–226

Rogers JL, Nicewander WA, Toothmaker L (1984) Linearly independent, orthogonal, and uncorrelated variables. Am Statistician 38(2):133–134

Scaletti C (1994) Sound synthesis algorithms for auditory data representation. In: Kramer G (ed) Auditory Display: sonification, audification, and auditory interfaces. Proceedings, Santa Fe Institute studies in the sciences of complexity, vol XVIII. Addison Wesley, Reading MA, pp 223–251

Sikora CA, Roberts LA, Murray L (1995) Musical vs. real world feedback signals. In: Proceedings of the ACM SIGCHI conference on human factors in computing systems, Association for Computing Machinery, New York, pp 220–221

Speeth SD (1961) Seismometer sounds. J Acoust Soc Am 33:909–916

Slevc LR, Patel AD, Rosenberg JC (2008) Language, Music, and Modularity: Evidence for SharedProcessing of Linguistic and Musical Syntax. In: Proceedings of the 10th International Conferenceon Music Perception & Cognition (ICMPC10). Sapporo, Japan

Stockman T, Hind G, Frauenberger C (2005a) Interactive sonification spreadsheets. In: Proceedings of the eleventh meeting of the international conference on auditory display, 6–9 July 2005, Limerick, Ireland

Stockman T, Nickerson LV, Hind G (2005b) Auditory graphs: a summary of current experience and towards a research agenda. In: Proceedings of the first symposium on auditory graphs, Limerick, Ireland

Truax B (1988) Real-time granular synthesis with a digital signal processing computer. Comput Music Jsthetics and ideals of data-sound 12(2):14–26

Tuuri K, Eerola T (2012) Formulating a revised taxonomy for modes of listening. J New Music Res 41(2):137–152

Tuuri K, Mustonen M, Pirhonrn A (2007) Same sound—different meanings: a novel scheme for modes of listening. In: Proceedings of audio mostly. Fraunhofer Institute for Digital Media Technology, Ilmenau, Germany, pp 13–18

Verplank B, Mathews M, Shaw R (2001) Scanned synthesis. J Acoust Soc Am 109(5):2400–2404

Vickers P (2005) Whither and wherefore the auditory graph? Abstractions & aesthetics in auditory and sonified graphs. In: Proceedings of the first symposium on auditory graphs, Limerick, Ireland

Vincente KJ (2002) Ecological interface design. Hum Fact 44(1):62–78

Walker BN, Nance A, Lindsay J (2006) Spearcons: speech-based earcons improve navigation performance in auditory menus. In: Proceedings of the twelfth international conference on auditory display, 20–23 June 2006, London, UK

Walker BN, Lindsay J, Nance A, Nakano Y, Palladino DK, Dingler T, Jeon M (2012) Spearcons (speech-based earcons) improve navigation performance in advanced auditory menus. Hum Factors 55(1):157–182

Warin C (2002) Sound as means for data representation for data mining. Master Degree Thesis, Facultés Universitaires Notre-Dame de La Paix, Namur

Watson MO, Sanderson P (2007) Designing for attention with sound: challenges and extensions to ecological interface design. Hum Fact 49(2):331–346

Whitelaw M (2004) Hearing pure data: aesthetics and ideals of data-sound. In: Unsorted: thoughts on the information arts: an A to Z. Sonic Acts/De Balie, Amsterdam, pp 45–54

Worrall DR (2004) Audification experiments using XAO returns. In: Auditory display as a tools for exploring emergent forms in exchange-trading data. Research Report to the Sonic Communication Group, University of Canberra. http://www.avatar.com.au/sonify/research/sonsem1/

Worrall D (2010) Parameter mapping sonic articulation and the perceiving body. In: Proceedings of the 16th international conference on auditory display, 9–15 June 2010, Washington, DC, USA

Worrall DR (2014) Can micro-gestural inflections be used to improve the soniculatory effectiveness of parameter mapping sonifications? Organised sound, vol 19, no 1. Cambridge University Press, pp 52–59

Xenakis I (1971) Formalized music: thought and mathematics in music. Indiana University Press, Indiana

Zatorre RJ, Belin P, Penhune VB (2002) Structure and function of auditory cortex: music and speech. Trends in Cog Sci, 6:37–46. Elsevier Science Ltd.

Chapter 3
Knowledge and Information

Hamlet:	Do you see yonder cloud that's almost in shape of a camel?.
Polonius:	By th' mass, and 'tis like a camel indeed.
Hamlet:	Methinks it is like a weasel.
Polonius:	It is back'd like a weasel.
Hamlet:	Or like a whale.
Polonius:	Very like a whale.
Hamlet:	… They fool me to the top of my bent….

(Shakespeare: *Hamlet* Act III, Scene 2)

Abstract One task of data sonification is to provide a means by which listeners can obtain new ideas about the nature of the source of derived data. In so doing they can increase their knowledge and comprehension of that source and thus improve the efficiency, accuracy and/or quality of their knowledge acquisition and any decision-making based on it. The purpose of this chapter is to develop an historical understanding of what information is as a concept, how information can be represented in various forms as something that can be communicated with non-verbal sonic structures between its source and its (human) receiver and thus retained as knowledge. Whilst a complete philosophical and psychological overview of these issues is outside the scope of the chapter, it is important, in the context of developing computational design strategies that enable such communication, to gain an understanding of some of the basic concepts involved. A quasi-historical epistemology of human perception and the types of information these epistemologies engender is followed by a discussion of the phenomenal nature of sounds and sonic structures, and their ability to convey information of various sorts.

3.1 Introduction

In this chapter we attempt the somewhat convoluted task of overviewing the main conceptual and perceptual issues related to the thorny issues of perception and knowledge acquisition as they relate to data sonification. These fundamental ideas

D. Worrall, *Sonification Design*, Human–Computer Interaction Series, https://doi.org/10.1007/978-3-030-01497-1_3

have been in active discussion by some of the most and perceptive minds over a couple of thousands of years, yet many of them remain difficult and still somewhat controversial, because they relate to how humans, as special kinds of beings, perceive new ideas and integrate them into what they already know, and in doing so, modify their knowledge, and indeed themselves.

Unless made otherwise explicit, the term *listener* is used to mean a human being without gross auditory system abnormalities, who, in various ways, engages, perhaps interactively, in listening within an environment to sound that is generated or manipulated from data. While that engagement is primarily through listening, it is not exclusively so, because hearing is not an isolated sense. All our senses are integrated in our being: vestibulation, proprioception, hearing, sight, touch smell, and taste. Except under explicit alternate conditions, such as when users are blind, we assume that these sensory input mechanisms are essentially the same for all users. What is different is individual listener's knowledge of the source of the data being sonified: For each user, information conveyed by one means or another will have meaning according to their knowledge of the source of the data, not according to their special sensory processing or knowledge-acquisition skills.[1]

3.2 Knowledge

There are important distinctions to be made between the terms *information* and *knowledge*, despite them being frequently used to mean the same thing in common parlance.[2] In this ontology, knowledge is considered to be *a theoretical or practical understanding of a subject* Fromm (1947, 238); a coherent understanding of a way accepted facts about the subject relate to each other, whilst *information* is used to mean items of communicable knowledge (ideas). Accepted facts contribute to *truth,* that can be understood, after Fromm, to mean a functional approximation to reality; itself understood to be relative to an accepted body of knowledge. Thus, in this description, the term *knowledge* implies that its content is meaningful, 'acknowledged', and veridical (truthful), with Truth currently acknowledged to have multiple, and developing, ontologies. Formally,

> An individual S knows a proposition P only if S believes P and P is true. If P is false, S does not know P, even though S might know that they believe that they know P.

---

[1]An exception can be made for sonifications whose target users are musicians who have received often years of "ear training" as part of their education.

[2]Especially in English, where the Anglo-Saxon distinction between *witan* ('wit') and *cnawan* ('to know') barely survives. The distinction remains, in German, for example, with *wissen*, *kennen*, *erkennen* and in French, with *connaître* and *savoir*. See Footnote 20 which relates 'wit' to the Greek *eide*.

Epistemology, the theory of knowledge, especially with regard to its methods, validity, and scope is a central theme in Western philosophical discourse. In this context, we focus on two aspects of knowledge: The forms it is held in and the methods of acquiring it.

3.2.1 Types of Knowledge

The first epistemological distinction to be made is between the *logic* of knowledge that is concerned with logical relations, and the *psychology* of knowledge that deals with empirical discoveries. German philosopher Immanuel Kant (1724–1804) expressed the distinction like this:

> It is of the utmost importance to isolate the various modes of knowledge according as they differ in kind and in origin, and to secure that they be not confounded owing to the fact that usually, in our employment of them, they are combined. … It must be admitted, however, that the two elements of our knowledge—that which is in our power completely *a priori*, and that which is obtainable only *a posteriori* from experience—have never been very clearly distinguished, not even by professional thinkers and that they have therefore failed to bring about the delimitation of a special kind of knowledge, and thereby the true idea of the science which has preoccupied human reason so long and so greatly ([1787] 2007).[3]

Later called Psychologism, the philosopher of science Karl Popper (1902–1994) made the same point, more than a century-and-a-half later:

> The initial stage, the act of conceiving or inventing a theory, seems to me neither to call for logical analysis nor to be susceptible of it. The question how it happens that a new idea occurs to a man… may be of great interest to empirical psychology ; but it is irrelevant to the logical analysis of scientific knowledge (1959, 31).

Table 3.1 provides a summary of different types of knowledge that are acquired by individuals (not necessarily independently) when engaging in different activities for different purposes. For example, empirical science, as practiced, increasingly relies on authority in the form of peer review as well as observation, and intuition can play a large part in deciding which possible consequences to use inductively. From a deductive perspective, abduction is a logical fallacy, yet it has proved to be a successful method in artificially intelligent agents' interaction within their environments and is now considered by some to be at the root of human perception and cognition. Contemporary methods in scientific research, and particularly if they involve interdisciplinary practices, are usually an evolving mixture, and Reliabilism is seen to be a reasonable account of the methodology of many contemporary practices, including many that are described by their users as empirical.

[3]By *professional thinkers*, Kant is referring to Lock and Hume, neither of whom made the distinction.

Table 3.1 Different types of knowledge

Knowledge, (form of)	Description	Example
a priori	lit. from before—what a person can derive from the world without needing to experience the world directly.	The laws of arithmetic; F = ma.
a posteriori (empirical)	lit. from what comes later. K. based on observation: sensory experience followed by logic and reflection.	Wind instruments with cylindrical bores emphasize the odd-numbered partials more than instruments with conical bores.
Explicit	K. that is recorded and communicated between people through various media. Tends to be organized systematically.	Principles of psychoacoustics; K. stores in libraries and databases.
Genetic/Embodied	Produced by functions for, and activities in environments (acquired through inheritance and epigenetics).	Phylogenetic-proclivity for language learning; perfect pitch (Theusch et al. 2009).
Tacit	K. that can only be achieved through experience. Is probably impossible, to communicate it through any medium; Most closely resembles *a posteriori* K.	The ability (skill) to use a hammer, ride a bicycle or play a violin.
Propositional (Descriptive or Declarative)	K. of something, i.e. that can literally be expressed in propositions or declarative sentences; most closely resembles *a priori* and explicit K.	Instructions containing steps repeatable by a third person without knowing why the instructions are necessary.
Procedural (Non-propositional)	K. that is acquired ("hands-on") by doing. K. that can be used.	Methods that can be protected as intellectual property that can then, be sold, protected, leased, etc.

There is a further epistemological distinction between explicit and implicit knowledge; between knowing-that (2 + 2 = 4; Tuesday follows Monday etc.) and knowing-how (to play the violin; ride a bicycle etc.). The difference can found in the Greek distinction between *episteme* (theoretical truth) and *tekhne* (practical methods for effecting results). By the time the Latin Scholastics of the Middle Ages had rediscovered the Greek language and the philosophy of Aristotle, the term *science* had come to imply both, differentiated as the theoretical sciences (philosophy)[4] and the manual arts (practice).[5]

[4]In its oldest sense (c. 1300), *science* meant knowledge (of something) acquired by study, also a particular branch of knowledge, from O.Fr. *science*, from L. *scientia* knowledge, from *sciens* (gen. *scientis*), prp. of *scire* to know, probably originally to separate one thing from another… Main modern (restricted) sense of a body of regular or methodical observations or propositions … concerning any subject or speculation is attested from 1725; in 17c.–18c.; this concept commonly was called philosophy. (OLED).

[5]*ars*, from the Greek *artios* 'a complete skill as a result of learning or practice' (OLWD).

Explicit knowledge is also known as declarative, descriptive or propositional knowledge. Such knowledge is usually acquired reflectively by logical reasoning, mathematical proof and scientific methods, or by reference to historical or cultural practices. By the time Descartes formulated his theory of the separation of the mind or soul (*res cogitans*) and the 'extended' world outside it (*res extensa*), now known as Cartesian Dualism, the mind was clearly thought of as the seat of reason, God and the sacred, while the body was a fleshy machine and clearly inferior, or at least secondary. This attitude began to change slowly through a philosophical 'bottom up' exploration of the relationship between sensation and knowledge of the world, culminating in Kant's transcendental idealism as a way of resolving the idealist's dilemma of how true knowledge of the world was possible with the obvious success of empirical methods. We return to these ideas later in the chapter when considering information.

The polymath Michael Polanyi (1891–1976) thought that non-explicit know-how type knowledge is *tacit* in that it is embodied and cannot be fully described in words. Attempts to perform such analyses are often laborious, difficult and 'destructive' (1975). The fact that we know more than we can clearly articulate contributes to the conclusion that much knowledge is passed on tacitly by practical and/or genetic means. In such an operational approach to knowledge, truth is functional; an idea or hypothesis is true if and while it works.

3.2.2 Methods of Acquiring Knowledge

Knowledge acquisition, the process of gaining knowledge from information, can be understood as the integration of new information into that which is already coherently embodied. Considered in this way, the trans*form*ation of in*form*ation into knowledge, while there may be sonification techniques to enhance that process,[6] they lie outside scope of this book. The major methods of acquiring knowledge are summarized at the end of this Chapter in Appendix 3A and 3B. the latter being devoted to inferential methods. The *reasons* an individual, or community, or culture use one method, or mixture of methods, rather than another, in a particular circumstance is outside the scope of the current work. Also, whilst they are not an historical account of the *getting* of knowledge, the methods are presented in an order which reflects the occidental procession from a reliance on authority and revelation as the principle source, stimulated by occasional revelatory experiences, to individual discovery that is verifiable by others, whether through intuition or more formally, through inference.

We are all, in some way, committed to the basic truth of commonsense psychology and, hence, to the existence of the states to which its generalizations of the states to which its generalizations refer (Dretske 2000). Some, such as Fodor

[6]Such as those used to enhance learning by entraining the brain's beta frequencies.

(1987), also hold that commonsense psychology will be vindicated by cognitive science, given that propositional attitudes can be construed as computational relations to mental representations.

Churchland (1981) thinks that, as a theory of the mind, folk psychology has a long history of failure that can't be incorporated into the framework of modern scientific theories, including cognitive psychology. He argues that the states and representations folk psychology postulates simply don't exist; that folk psychology is comparable to alchemy and ought to suffer a comparable fate. On the other hand, Daniel Dennett seems prepared to admit that the generalizations of commonsense psychology are true and also indispensable, but denies that this is sufficient reason to believe in the entities to which they appear to refer. He supports this stance on that basis that there is nothing more to having a propositional attitude than to give an intentional explanation of a system's behavior by adopting an the intentional stance toward it. Assuming a system is rational,[7] if the strategy of assigning contentful states to it and predicting and explaining its behavior is successful, then the system is intentional and the generalized propositional attitudes we assign to it are true (Dennett 1987, 29).

In an age when many thinkers were still wrestling for intellectual[8] independence from the dominance of the church in matter of knowledge, understanding and wisdom, the eighteenth century philosopher Immanuel Kant considered understanding to arise rationally:

> Understanding may be regarded as a faculty which secures the unity of appearances by means of rules and reason as being the faculty that secures the unity of the rules of understanding under principles. Accordingly, reason never applies itself directly to experience or to any object, but to understanding, in order to give to the manifold of knowledge of the latter an *a priori* unity by means of concepts, a unity of which may be called the unity of reason, and which is quite different in kind from any unity that can be accomplished by the understanding ([1787] 1787, 303).

Knowledge acquisition, is not always explicit. However, it is an historically common 'spectator view' of knowledge that human experience is primarily a matter of contemplation. The American humanist philosopher Corliss Lamont suggests that this position is largely derived from an overemphasis on the role of vision. Further, he emphasizes the location of truth with respect to such knowledge:

> [W]orkability is the *test* of a truth, not the *source* of it. The truth of an idea does not lie *in* verification; we are able to prove it true *through* verification. An idea is true *if* it works, not because it works; for it already *was* true and corresponding to objective reality. New truths lie all about us waiting to be discovered by persons wielding scientific techniques; but the process of discovering does not *make* ideas true (1949, 243).

[7]Something is rational if behaves in accordance with the truths and falsehoods afforded by its environment.

[8]The noun usage of the term *intellectual* for persons arose at this time (ACOD).

On the assumption that our perceptual processes are in fact reliable in the way that we take them to be, Reliabalism. offers a seemingly straightforward account of how perceptual beliefs about physical objects and the physical world are justified.

It emphasizes the properties of the processes used to arrive at truths. In reliabilist approaches to knowledge acquisition, noticing a static relationship between a conjecture and a body of evidence, knowing that an hypothesis does not contradict the evidence, or even is in-accord with it, for example, is insufficient to warrant support for it from the evidence: additional account must be taken of how reliable the method that produce the hypothesis is known to be in producing truthful hypotheses. Reliabilist thinking underpins the greater acceptance of the diagnostic judgements of experts over laypersons, the preferential support for research programs with fecund histories and the scorn of *ad hoc* hypotheses.

The first, at the time unrecognized as such, formulation of a reliability account of knowledge was in the mathematician Frank Ramesy's writing on knowledge (1931). Ramsey said that a belief is knowledge if it's true, certain, and obtained by a reliable process. Several similarly subjunctive theories, such as Tracking Theory and Contexturalism, were developed in the latter part of that century, as discussed by Goldman, who notes that reliability theories of knowledge of varying stripes are still in active development and continue to appeal to many epistemologists for their robustness and flexibility. Permutations abound: Some theories Focus on modal reliability; on getting truth or avoiding error in possible worlds with specified relations to the actual one. They also focus on local reliability, that is, truth-acquisition or error avoidance in scenarios linked to the actual scenario in question (Goldman [2008] 2015).[9]

3.3 Information

Considering the current scientific, economic and social importance that is given to the concept of information, it is remarkable how unclear its meaning is in some contexts. Ronald Stamper, the founder of systems analysis studies at the London School of Economics, writes about his common experience of computers being used with great technical skill but with poor results. He argues that the overwhelming reason is that there is a lot of technical skill with information technology but little knowledge about the information it carries. While there is a general agreement about the nature of data, information, he remarks,

> …is a vague and elusive concept. … My greatest hope is that, one day, members of the information systems profession will stop talking about information as something 'obtained by processing data to produce meaning' (1996, 349).

---

[9]For an extended description, See http://plato.stanford.edu/entries/reliabilism/.

3.3.1 Information as Quantity of Improbabilities

Information Theory is an applied mathematical theory invented by Claude Shannon to assist in finding measures of the limits to compressing data and reliably communicating it within an existing telecommunications system (Shannon and Weaver 1949). In this theory, information is understood as an objective quantifiable entity, the information content of a message being a quantitative measure of its improbability. Information Theory

> …provides a measure for how much information is to be associated with a given state of affairs and, in turn, a measure for how much of this information is transmitted to, and thus available at, other points. The theory is purely quantitative. It deals with amounts of information - not, except indirectly and by implication, with the information that comes in those amounts. (Dretske 1981) cited in Hoffmeyer and Emmeche (1991).

In such quantification, the value of information in a statement (a message) reflects the statistical structure of the statement. The question of how information was created, or what should be meant by its significance, is not addressed through the theory. In human communication, most statements are only understandable at a semantic level. Information theory requires a finite set of possibilities, or it cannot assign precise probabilities to the possibility of any particular message. Thus Khinchin (1957, 90) more properly calls Information Theory a *theory of discrete information*.

Sound spectra can be considered as frequency encodings of information, with white noise, which contains no redundancy, at one extreme and a sine tone, which contains maximum redundancy, at the other (Moles 1966). One of the fundamental tasks that auditory systems are known to perform well is frequency spectra differentiation (Shepard 1999), and this capability can be used in data audification "to provide a measure of how much information is to be associated with a given state of affairs" by affording comparisons between the microsound qualities of spectra-encoded data and various recognizable encodings such as white, pink, blue, and Brownian noises. Experiments 1 and 2 in Chap. 7 articulate an experimental technique for the application of sonification to aurally identify the distinction between information-less sound (a uniform distribution) and one that exhibits correlational information.

In general, the concept of information as a quantitative measure of noise or redundancy is of little use in human-to-human sensible communication, since the statistical analysis of the probabilities to be ascribed to any definite statement is not only infeasible, but also theoretically not possible. While totally unforeseen events are demonstrably a part of life, their eventual appearance makes it impossible to ascribe distinct probabilities to any event.

3.3.2 Information: General, Scientific and Pragmatic

The term *information* is also commonly used synonymously for an instruction, an answer, a news message or announcement and, in general, people do feel that they are often *informed* through observation and conversation. The quantitative improbability definition of Information Theory does not account for this sense of information. In empirical science the concept *information* is as a statement (Roederer 2005). It is known as the semantic aspect of information, because it answers a pre-formulated question, such as "what is the fundamental frequency of this wave?" or it indicates the specific outcome of known alternatives, such as "Did the viola section enter one measure too early or too late?" or "Is the next card the queen of hearts?"

> It is the intended *effect* that ultimately identifies information.

Only *living* systems (and perhaps their cyber-primitives) have the capability of engaging in information-driven interactions with each other and the environment. This aspect of information is called *pragmatic* (Küppers 1990) and is often used by such systems to counterbalance the otherwise normal course of physical events. Some important characteristics of this information are

- *Information is the shape or pattern of something,* not the energy or forces used to give it that pattern. These forces and energy may be necessary to effect the pattern but they are separate from, and subservient to the purpose for which the information was formed.
- *Information always has a purpose*: to cause specific *change* somewhere or at sometime that would not otherwise occur, except, perhaps, by chance.
- *It is what information does that is important*, not how the form in which it resides is textured, or sounded. This is not to say that the form of information is not important for listener's perception of it, or ability to process it, especially when not heard in isolation.
- *The* form *in which information resides expresses* it*, but this expression is not the information.* This idea is expressed differently later on in this chapter; for example, in Husserl's 'bracketing' of the phenomenal world from the 'world–as–it–is' (Hoffmeyer and Emmeche 1991) and in Korzybski's "the map is not the territory (see Sect. 3.4, below). Music, for example, is not a score or a Compact Disk or audio streamed from a Cloud-computing resource, or the compactions and rarefactions of air pressure in a sound wave, or the neural activity of the brain. It is the intended *effect* that ultimately identifies information.
- *Altering the form in which the information resides may result in a different expression or may destroy the information.* For example, rearranging the notes of a melody destroys the melody, and the continuity of a glissando on a violin is not preserved when expressed by a clarinet.

- *Information does not exist in isolation.* Information always requires
  - *An origin*, source or sender where the original pattern is located or generated.
  - *A recipient*, where the intended change is supposed to occur.
  - *A transmission* from one to the other. For the purpose of a specific change to occur, a specific information–processing mechanism must exist and be activated.
- *Information can be stored* and, providing it has not been corrupted, retrieved, either in the form of the original pattern, or in some transformation of it.

3.3.3 Forms of Perceptual Information

Metaphysics is a process of thought, which has two subdivisions: epistemology and ontology. It is less directed to providing conclusions than it is to exploring a mental landscape; of organizing beliefs and experiences into holistic descriptions. Our aim here is less purely contemplative; it is to gain an understanding of the nature of certain types of mental states that can be induced to reside in the minds of listeners engaging with a sonification. We can then use our knowledge of psychophysics and cognitive load to choose how to induce them, or at the very least reduce impediments to their induction. The focus of the early part of this section is a primarily a summary of the historical epistemological approach to information and perception. However, because of the directed, or intentional and interactive nature of perception, such information is bound between the sounds of a sonification and the intentions and perhaps embodied non-intentions of listeners. Recent discoveries in neuroscience support the assertion that these phenomena are not always consciously directed or even explicit, especially with respect to perception and information-gathering, and so a phenomenological account of the ontological issues involved is outlined.

3.3.4 Platonic Ideals

The elusiveness of the term *information* is embedded in its historical epistemologies (Hoffmeyer and Emmeche 1991). In the Middle Ages it was used in connection with a person being informed or educated[10]–to become aware of the form (of something), give shape to, fashion (ACOD 1992). The concept of *form* extends back through Aristotle and Plato to at least Pythagoras who is responsible for

[10]From L. *informationem* (nom. *informatio*) "outline, concept, idea," noun of action from *informare* (OED); L *informere*—in (into) + forma (form).

adding it as a correlative concept to the Milesian's conception of 'matter'. Burnet (1904/1981, 35). This shift of information in the mental sphere moving to the physical sphere of substance and action, is paralleled in the shift of Form as a Platonic Ideal to something physical in which

> [f]orms were induced on, not derived from substance. They reflected human or Godly will. Thus, to bring something into form presupposed a person in whom the idea of the form must first have occurred. And this occurrence, the idea, was the root of information (Hoffmeyer and Emmeche 1991).

Burnet traces this thinking though Socrates' Doctrine of Forms, and Plato's Ideals which assert that individual objects are but imperfect examples of universals, divinely given and residing in man's soul (mind):

> There is a sharp distinction between the objects of thought and the objects of sense. Only the former can be said to be; the latter are only becoming. …We know what we mean by equal, but we have never seen equal sticks or stones. The sensible things we call equal are all 'striving' or 'tending' to be such as the equal, but they fall far short of it. Still, they are tending towards it, and that is why they are said to be becoming (Burnet 1904/1981, 126–27).

In contrast to Plato's position, which held that particular things are deduced from these *a priori* Ideals[11] by contemplation, Aristotle held that particular things contained universal essences, and such Forms (ideas) were induced from particular Category instantiations. He compares the act of perception to the pressing of a signet ring into wax:

> By a 'sense' is meant what has the power of receiving into itself the sensible forms of things without the matter. This must be conceived of as taking place in the way in which a piece of wax takes on the impress of a signet-ring without the iron or gold; we say that what produces the impression is a signet of bronze or gold, but its particular metallic constitution makes no difference: in a similar way the sense is affected by what is coloured or flavoured or sounding, but it is indifferent what in each case the substance is; what alone matters is what quality it has, i.e. in what ratio its constituents are combined (Aristotle 350AD, BkII, Ch12).

Upon the reappearance of Aristotle's works in the West, this position was adopted by medieval Scholastics such as Thomas Aquinas, for whom the representation of an external object in the mind and the object itself was the same object in two different forms of existence. This is because, put simply, they held the same form, or *conformed*. Medievalist Henrik Lagerlund explains:

> This 'conformality' account of mental representation is for Aquinas embedded in a much larger, causal theory of the reception of these forms into the mind or intellectual soul,

[11] Greek ιδεα. Plato seems to use the term interchangeably with *eidos*, which serves to designate any of those primary realities that have come to be known as the Forms. The term takes on a significant philosophical meaning from his writings onwards that the term seems to develop an elaborate life of its own (Taylor 1911). See also footnote 20.

> according to which forms are transmitted through the intervening medium between subject and object (the doctrine of the *species in medio* or 'species in the medium') and received in the external sense-organs and sense-faculty, which leads to the production of phantasms or sensible species and ultimately to the creation by the active intellect of a mental representation or intelligible species in the passive intellect (2008).

3.3.5 Materialist Ideals

3.3.5.1 Early Rational Approaches to Perception

In addition to the mental representations adopted from Aristotle, based on conformity and resemblance, the late Scholastics such as Ockham developed other notions based on causality and signification. Following the Scholastics, seventeenth century contemporaries Descartes, Spinoza and Leibniz continued to look for knowledge using rational thought, including the continued development of logic, mathematics and Euclidian geometry.

René Descartes' starting point was his own experience (*cogito ergo sum*). He thought that the mind reached out to external objects themselves, with the sense organs and nerves serving as literal mediating links between the objects and those brain events that afforded the perceptual awareness of them. He is clear however, that sensory awareness does not reach out to the physical things themselves; that in veridical sensation, the mind's ideals are not the immediate objects of awareness:

> …whence I should forthwith be disposed to conclude that the wax is known by the act of sight, and not by the intuition of the mind alone, were it not for the analogous instance of human beings passing on in the street below, as observed from a window. In this case I do not fail to say that I see the men themselves, just as I say that I see the wax; and yet what do I see from the window beyond hats and cloaks that might cover artificial machines, whose motions might be determined by springs? But I judge that there are human beings from these appearances, and thus I comprehend, by the faculty of judgment alone which is in the mind, what I believed I saw with my eyes ([1641] 1901).

3.3.5.2 Material or Perceptual Idealism

At the same time, and in contrast to a rational approach, try-it-and-see empirical[12] heuristics developed out of the methodologies of occidental physicians and alchemists, as exemplified by William Harvey (1578–1657) and Francis Bacon (1561–1626). Their increasing success led to attempts to apply such methods to

[12]Empirical. First use recorded in 1569, from L. *empiricus*, from Gk. *empeirikos* "experienced," from *empeiria* 'experience', from *empeiros* "skilled', from *en-* 'in' + *peira* 'trial, experiment'. Originally the name of a school of ancient physicians who based their practice on experience rather than theory [OLED].

understanding how we obtain knowledge itself. The Scottish physician John Locke (1632–1704) is credited as the founder of modern Empiricism, as he took the crucial step of denying all *a priori* knowledge (Gregory 1981, 338–39). Locke asked, if our understanding of the physical world was not descendant from Ideas but induced in the natural world, how did we acquire such understanding? He thought that a child's mind was like a blank tablet (*tabula rasa*) and all ideas came from experience, and then later, from reflection:

> Let us then suppose the mind to be, as we say, white paper, void of all characters, without any ideas: How comes it to be furnished? Whence comes it by that vast store which the busy and boundless fancy of man has painted on it with an almost endless variety? Whence has it all the *materials* of reason and knowledge? To this I answer, in one word, from *experience*. In that all our knowledge is founded; and from that it ultimately derives itself. Our observation employed either, about external sensible objects, or about the internal operations of our minds perceived and reflected on by ourselves, is that which supplies our understandings with all the *materials* of thinking. These two are the fountains of knowledge, from whence all the ideas we have, or can naturally have, do spring (Locke 1690a, 1: Book II, Ch.1).

Locke's concept of perception was simple: objects have characteristics, and perception reflects the world, much as a mirror reflects objects. The Irish bishop George Berkeley (1685–1753) went on to argue that if an empirical analysis of human knowledge was followed rigorously, it had to be admitted that all experience is nothing more than experience, that the qualities that the human mind registers are ultimately experienced in the mind itself, and there can be no conclusive inference from this as to whether or not some of those experiences represent or resemble anything outside it:

> There was an odour, that is, it was smelt; there was a sound, that is, it was heard; a colour or figure, and it was perceived by sight or touch. This is all that I can understand by these and the like expressions. For as to what is said of the absolute existence of unthinking things without any relation to their being perceived, that seems perfectly unintelligible. Their *esse* is *percipi*, nor is it possible they should have any existence out of the minds or thinking things which perceive them (Berkeley 1710/1957, Book II, Ch.3).

However, the similarity of individual's reported experiences of the world convinced Berkeley that these experiences indicated the presence of order that was not just subjectively determined by whimsical fantasy. He postulated that the order that inheres in the world depends on God (Tarnas 1991, 335–36).

Faced with things considered unknowable, both Locke and Berkeley relied on the same resolution as the majority of their predecessors—to a universal mind that transcends individual minds, which, as Gregory (1981, 346) puts it, *takes away the interest of this account*. Scottish philosopher David Hume (1711–1776) agreed with Berkeley's empirical arguments but saw no reason to accept the resolution of a universal mind. He reasoned that, because the intellect cannot deduce the veridicality of received sensations, or the connection between these sensations and the Truths in the intellect, the only truths of which the intellect is capable are tautological, that is, are not capable of producing new information. He goes on to argue that, not only can one not logically *deduce* the nature of things using reason alone, neither can one *infer* them from experience. The only thing that connects sequences of sensations together is the principle of causality, which relies on the experience of individual concrete events in temporal succession:

> When we look about us towards external objects, and consider the operation of causes, we are never able, in a single instance, to discover any power or necessary connexion; any quality, which binds the effect to the cause, and renders the one an infallible consequence of the other. We only find, that the one does actually, in fact, follow the other (Hume [1777] 1978, Book VII, Chap. I, Sect. 50).

Hume maintained that, whilst one might recognize the regularity of perceived events, the acceptance of the necessity for events to follow each other is not based on logical certainty, but on habit:

> When we say, therefore, that one object is connected with another, we mean only that they have acquired a connexion in our thought, and give rise to this inference, by which they become proofs of each other's existence: A conclusion which is somewhat extraordinary, but which seems founded on sufficient evidence (Hume [1777] 1978, Book VII, Chap. I, Sect. 59).

The mind draws an explanation of experience from itself and it cannot know what *causes* a sensation because it doesn't experience *cause* as a sensation. Further, he goes on to argue, the concreteness of these individual events cannot be logically asserted, so causality is made meaningless. If all human knowledge is based on empiricism, and induction cannot be logically verified, there can be no certain knowledge. So, Hume had begun by trying to eliminate the necessity for metaphysics in deductive rationalism to investigate man's reasoning by using the empirical experimental techniques employed so successfully by Galileo (1564–1642) and Newton (1642–1727), and ended up, like Berkeley, skeptically questioning whether objective certainty could ever be empirically achieved because inductive inference couldn't be logically justified Tarnas (1991, 340).

3.3.6 Transcendental Ideals and Phenomena

Although Immanuel Kant admired Hume's reasoning that, in his own terminology, causal judgments are synthetic, involving an act of the mind that connects the cause and the effect, he was also convinced that the (empirical) experimental methods of Newton had really revealed generalized knowledge, so he set about the task of trying to unify the rational and empirical methods. In other words, to show how (scientific) knowledge is possible, how both reason and experience contribute to that knowledge, and in doing so, refute the skepticism of Hume's claim that experience and reason are extremely limited in the kinds of knowledge they can provide.

In *Critique of pure reason* (hereafter CPR), Kant proposed that the 'world' that science explained was a world already ordered by the mind's own cognitive apparatus. He begins by agreeing with Locke and Hume that there are two kinds of mental propositions: those in the intellect based on sensations and those in the intellect alone, based on relations between concepts. For him, it was the (internal, conceptual) relations between intuited sensations that constitute forms of appearance:

> The capacity (receptivity) for receiving representations through the mode in which we are affected by objects, is entitled *sensibility*. Objects are *given* to us by means of sensibility, and it alone yields us *intuitions*; they are *thought* through the understanding, and from the understanding arise *concepts*. …
>
> The effect of an object upon the faculty of representation, so far as we are affected by it, is sensation. That intuition which is in relation to the object through sensation, is entitled empirical. The undetermined object of an empirical intuition is entitled appearance…
>
> That in the appearance which corresponds to sensation I term its *matter*; but that which so determines the manifold of appearance that it allows of being ordered in certain relations, I term the *form* of appearance. ([1787] 2007, 65).

Kant suggests that there are two 'pure' forms of sensible intuition that serve as principles of *a priori* knowledge: space and time:

> That in which alone the sensations can be posited and ordered in a certain form, cannot itself be sensation; and therefore, while the matter of all appearance is given to us *a posteriori* only, its form must lie ready for the sensations *a priori* in the mind, and so must allow of being considered apart from all sensation.
>
> I term all representations *pure* (in the transcendental sense) in which there is nothing that belongs to sensation. The pure form of sensible intuitions in general, in which all the manifold of intuition is intuited in certain relations, must be found in the mind *a priori*. This pure form of sensibility may also itself be called *pure intuition*. [For example,] if I take away from the representation of a body that which the understanding thinks in regard to it, substance, force, divisibility, etc., and likewise what belongs to sensation, impenetrability, hardness, colour, etc., something still remains over from this empirical intuition, namely, extension and figure. These belong to pure intuition, which, even without any actual object of the senses or of sensation, exists in the mind *a priori* as a mere form of sensibility ([1787] 2007, 65).

Kant considered space and time to be the two forms in which sensibility occur because they are obtained, not from sense data, but by the actions of the mind in coordinating the sensations it receives. In other words, we intuit the (real-world) objects of our senses by the detection of the manifold presentation of sensations (from hearing, touch etc.) of those objects by the two aspects of our minds that are not (empirical) sensations, namely space and time. Whatever is perceived, is perceived as having spatial and/or temporal relations.

In CPR, Kant sought to determine, despite Hume's conclusion otherwise, if it was possible to expand knowledge by proving the validity of propositions that are true without reference to experience. In the process, he accomplished what he called a 'Copernican revolution' in philosophy, now known as 'transcendental idealism':

Whereas philosophers had previously considered the mind to be a passive agent of knowledge of an objective world, he showed that the actual world cannot be known, but that the human understanding of reality is shaped by both our means of perceiving it (sensible intuitions) and pure (transcendental)[13] concepts which are available to us through our awareness of space and time. These transcendental concepts are derived from the translation of sensible Judgments of Quality, Quantity, Relation and Modality into transcendental Categories[14] (Kant 1787, 104–119).

Having described the process of form perception, Kant asserted that all necessity is grounded in a transcendental condition and therefore, so must consciousness, which he calls transcendental apperception,[15] *the synthesis of the manifold of all our intuitions*—the process whereby perceived qualities of an object are related to past experience, the inner self. Without transcendental apperception, the consciousness that is the foundation of the synthetic unity of experience, it would be impossible to think. Kant made a distinction between this transcendental apperception and empirical apperception which he described as the consciousness of the changing states of the inner self, or *[c]onsciousness of self according to the determinations of our state in inner perception is merely empirical, and always changing* and it is as a result of this mental plasticity, this openness of the inner self to change, that knowledge, *the final goal of the understanding in combining intuitions and concepts*, is possible (Kant 1787, 136).

Kant answers the question of how synthetic *a priori* propositions of mathematics are possible, with his Transcendental Aesthetic and the doctrine of the transcendental ideality of space and time. He answers the question of how synthetic *a priori* propositions of natural (empirical) science are possible in his Transcendental Analytic, where he demonstrates the essential role the categories play in establishing the possibility of knowledge from experience. In attempting to refute Hume's idealism, Kant concluded that we cannot be certain of the knowledge of the

[13]In Kant's philosophy, the adjective *transcendental* is applied to the condition of experience or anything related to it. Transcendental knowledge is possible, though transcendent knowledge is not (Runes 1942). Transcendental cognition is thus *a priori* knowledge. Kant's description of the expression 'transcendental knowledge' as *all knowledge which is occupied not so much with objects as with the mode of our knowledge of objects in so far as this mode of knowledge is to be possible "a priori"* (Kant 1787, 59), is the meaning used throughout the current work. This is not to be confused with the use of the term *transcendental* by the New Englanders Ralph Waldo Emerson, Henry David Thoreau, Margaret Fuller, Charles Ives etc.

[14]These categories are derived from Aristotle's Categories: Substance, Quantity, Quality, Relation, Place, Time, Position, State, Action, and Attention. *[O]ur primary purpose being originally identical with his, notwithstanding the great difference in the execution* (Kant 1787, 113).

[15]The term *apperception* was introduced by Leibniz to denote the introspective or reflective apprehension by the mind of its own inner states, in contrast to perception as the inner state of representation of outer things. In psychology, the term came to be used for the process by which an individual's new experience is assimilated into, and transformed by, the residuum of their past experiences to form a new whole (Runes 1942).

objects of our sense experiences 'as they really are' (*Ding an sich*, *noumena*[16]): that *a priori* knowledge of them can only be transcendental.

3.3.6.1 Kant's Refutation of Material Idealism

Kant identifies two type of material idealism: Dogmatic and Problematic. The Dogmatic Idealism of Berkeley asserts that things in space are merely imaginary entities. Whilst Kant acknowledged that this position is unavoidable if space is interpreted as a property of things in themselves, he showed through his Transcendental Aesthetic, that space is an *a priori* pure (non-empirical) intuition in us prior to any perception of objects and thus does not represent a property of objects ([1787] 2007, 65):

> In holding the empirical assertion that *I am,* the Problematic idealism of Descartes is pleading our inability to prove, through immediate experience, any experience other than our own.
>
> The required proof must, therefore, show that we have experience, and not merely imagination of outer things; and this it would seem cannot be achieved save by proof that even our inner experience … is possible only on the assumption of outer experience ([1787] 2007, 65).

Idealism assumes that the only immediate experience is inner experience from which we can only *infer* outer things, and then only uncertainly. Turning the idealists assertion on its head, Kant argues that, given that I am conscious of my own existence in time, and an awareness of time requires something permanent in my perception against which I can observe the progress of time, this permanence cannot be *in* me, that is, cannot be a property of me, since it is only through it that I can be conscious of my existence in time. So, it follows that

> perception of this permanent is possible only through a *thing* outside of me and not through the mere *representation* of a thing outside of me; and consequently the determination of my existence in time is possible only through the existence of actual things which I perceive outside of me. … [and *ipso facto*] the consciousness of my existence is at the same time an immediate consciousness of the existence of other things outside me ([1787] 2007, 65).

However, although *I am* immediately includes the existence of outer things, it does not follow that every intuition of outer things implies their existence, because a mental representation (thought) of them might have, for example, been imagined, dreamt or hallucinated. Kant thinks such (imagined etc.) representations are the reproduction of previous outer perceptions ([1787] 2007, 65).

In showing the way reason and empirical enquiry fitted together, Kant demonstrated the role of the human mind in *constructing* reality and knowledge and his

[16]Many, most famously Schopenhauer, in his 1818 *Criticism of the Kantian Philosophy* (Schopenhauer 1844), consider Kant to have inappropriately appropriated the Greek word *noumena*, 'that which is thought', and used it to mean 'things-in-themselves'.

insight affected all subsequent epistemological enquiry. In the course of addressing Hume's skepticism, that the inherent nature of things (the so-called *thing–in–itself*) is always beyond the reach of our empirical capacity to know, Kant had divided the world between the *phenomenal* world of empirically unknowable things (the world measured by Newtonian physics, for example) and the ever-unreachable *noumenal* world of *things-in-themselves* which we know exists if we exist. A phenomenon, then, is an object of empirical knowledge that is conditioned by space, time and the categories.

3.3.7 Brentano's Mental Phenomena and Intentional Inexistence

The general view of post-deists was that consciousness was composed of simple or basic elements–experiential events–that combine to form complex experiences. Both Locke and Hume considered that such simple events were parts of experiences and must, given the seeming unrestricted possibilities for their recombination, have a certain independence vis-à-vis other parts with which they might be realized:

> When the Understanding is…stored with…simple Ideas, it has the Power to repeat, compare, and unite them even to an almost infinite Variety, and so can make at pleasure new complex Ideas (Locke 1690b, 2:Bk II, Chap. 2. Sect. 2).

> [A]ll simple ideas may be separated by the imagination, and may be united again in what form it pleases (Hume [1777] 1978, Bk 1, Part I, Sect. IV).

The neo-scholastic philosopher and psychologist Franz Brentano (1838–1917) was widely influential as a teacher and mentor and his work is often considered precursory to Phenomenology,[17] a philosophy principally developed from his descriptive psychology by his student Edmund Husserl who used the term to mean

> the reflective study of the essence of consciousness as experienced from the first-person point of view (Smith 2007 cited in Wikipedia).

It became clear, largely through the work of Brentano and his students (principally Carl Stumpf (1848–1936) and Edmund Husserl (1859–1938)), that such a strong form of independence among parts of experiences is unnecessary; in fact, various

[17]Not to be conflated with Phenomenalism, the doctrine that all human knowledge is confined to the appearances presented to the senses. See Sect. 3.3.1 4. Martin Heidegger (a former pupil of Husserl) believed that Husserl's approach overlooked basic structural features of both the subject and object of experience. He developed a more ontological approach to phenomenology that influenced the development of existentialism. Hegel also used the term to describe a dialectical phenomenology that begins with an exploration of that which presents itself to us in conscious experience (phenomena) as a means to finally grasp the ontological and metaphysical Spirit that is behind them.

kinds of dependency relations are possible. Brentano published the first volume of his large-scale work on the foundations of psychology, *Psychology from an Empirical Standpoint*, in 1874 ([Brentano 1874] 1995). He emphasized that all our scientific knowledge should be based on direct experience. However, for him, doing empirical psychology meant describing what one directly experiences in inner perception from a first-person point of view, in contradistinction to contemporary empirical science, which attempts to take a third-person approach. Brentano's aim was to provide a scientific basis for psychology which, in his *Psychology*, he defined as *the science of mental phenomena* ([1874] 1995, 18), and which he distinguishes from physical phenomena because they are the exclusive objects of inner perception, they always appear as unities, and they are always intentionally directed towards objects (Huemer 2008). He later renamed his approach as *descriptive,* to distinguish this first-person analyses of subjective processes from the third-person *genetic* approach, in which causal or genetic laws are developed to explain what phenomenology merely describes.

Brentano was essentially an idealist, in that he maintained that external, sensory perception could not tell us anything about the de facto existence of the perceived world, though we can be absolutely sure of our internal perception. When I hear a police siren, I cannot be completely sure that there is a police siren in the real world, but I am absolutely certain that I do have an internal perception, that I hear. External sensory perception can only yield *hypotheses* about the perceived world, not the truthfulness of such perception nor the existence of the apparent external origin of such perception.

In considering the separability of experiential parts of consciousness, Brentano distinguished between *merely distinctional* and *actually separable* parts, the latter of which are either *one-sided* or *mutually* (two sided) separable ([1874] 1995, 15). An example of one-sided separable parts is in evidence in the way one can hear (notice) a sound without listening to the sound itself, but can't listen to a sound without noticing it. Multi-sensorial experiences have frequently mutually–separable parts. For example, the visual and aural experiential parts of a violin performance can be experienced separately, as can the aroma and sight of a banana.

Tracing the idea back through the Scholastics, Descartes and the Greeks, Brentano introduced the notion of the directed intention of mental objects into contemporary philosophy:

> Every mental phenomenon is characterized by what the Scholastics of the Middle Ages called the intentional (or mental) inexistence of an object, and what we might call, though not wholly unambiguously, reference to a content, direction toward an object (which is not to be understood here as meaning a thing), or immanent objectivity ([1874] 1995, 88).

Rather than developing a full and systematic description of intentionality, Brentano's goal was to outline the criteria for distinguishing mental and physical phenomena and he used the terms *mental or intentional inexistence* to refer to what today is known as a characteristic of consciousness: the mind's capacity to *refe*r or be *directed* to objects that existing solely in the mind. Although Brentano's formulation of intentional objects does not address their ontological status, this was addressed, at least to some extent, by his students, principally Alexius Meinong (1853–1920) and Edmund Husserl. Meinong's concern was with the intentional *relation* between the mental act and an object. He maintained that such a relation existed even when the object external to the mental act towards which it is directed doesn't exist, such as Pinocchio, Orpheus, Unicorns and the Fountain of Youth. Earlier, Hume had considered the concept of non-existent objects contradictory, Kant and Gottlob Frege (1848–1925) considered it logically ill-formed, while later, Betrand Russell (1872–1970) adopts the idea (Reicher 2006).

These examples seem more like Platonic Ideals than non-existent objects. In data sonification, a dataset can be considered as a particular ('accidental') object, so it seems reasonable to assume that an exploration of the rôle of non-existent objects in data sonification is only of relevance if can be reasonably concluded that sounds, or the other objects of data sonification software, cannot be ontologically accounted for in other ways.

According to Brentano's theory, mental acts cannot have duration. This brings up the question of how we can perceive temporally extended objects (*events*) like melodies. Brentano accounted for these cases by arguing that an object towards which we are directed does not immediately vanish from consciousness once the mental act is over. It rather remains present in altered form, modified from *present* to *past*. Every mental phenomenon triggers an 'original association' or *proteraesthesis*, as he calls it later, a kind of memory which is not a full-fledged act of remembering, but rather a part of the act that keeps lively what was experienced a moment ago. When I listen to a melody, for example, I first hear the first tone. In the next moment I hear the second tone, but am still directed towards the first one, which is modified as past, though. Then I hear the third tone, now the second tone is modified as past, the first is pushed back even further into the past. In this way Brentano can explain how we can perceive temporally extended objects and events.

3.3.8 Husserl's Transcendental Phenomenology

The notion of intentionality also played a central role in Husserlian phenomenology. Applying his method of phenomenological reduction, however, Husserl addresses the problem of directedness by introducing the notion of *noema*, which plays a role similar to Frege's notion of *sense*.

Edmund Husserl was the first to apply the term *Phänomenologie* (phenomenology) to a whole philosophy and his usage of the term has largely determined the sense of it in the twentieth century.[18] As the term underwent

[18]In contradistinction to Hegel's 1807 use of the word in his *Phänomenologie des Geistes* (Phenomenology of the spirit), in which is expressed a radically different concept.

development since his early use of it, the present comments are derived from Husserl's own mature introduction (1927)[19] in which he outlines the two principal uses of the term: Firstly, in application to "a new kind of descriptive method which made a breakthrough in philosophy at the turn of the twentieth century, and secondly, a new psychological discipline parallel to the first in method and content: the *a priori* pure or "phenomenological" psychology, which raises the reformational claim to being the basic methodological foundation on which alone a scientifically rigorous empirical psychology can be established."

Husserl's aim was that the empirical approach, which he calls *an objective science of nature* and which, significantly, he considered as a branch or anthropology or zoology in that it was decidedly *egoical* (anthropocentric) and physicalist, would lead to a better understanding of philosophical phenomenology, which he calls *pure psychology*, after Brentano's *descriptive psychology*. While the term has fallen into disuse in favor of *philosophy*, it does exemplify the importance that early scholars involved in the burgeoning science of psychology placed on dynamic interplay between empirical investigation and its philosophical underpinnings (Husserl 1927, Chap. 5). It is the contention here, as exemplified by the very existence of this chapter in the current work, that this relationship remains important, especially when creating, as we do with data sonification software, phenomena (objects, events) that need not conform to any resonating material Kantian *thing–as–it–is*.

Building on Brentano's outline, the notion of intentionality played a central role in Husserl's development of phenomenology. He understood intentionality as a fundamental attribute of subjective processes, maintaining that phenomenology must describe these processes, not only with respect to their immanent cognitive components, but also with respect to their intended (external) objects:

> In unreflective holding of some object or other in consciousness, we are turned or directed to-wards it: our *intentio* goes out towards it. The phenomenological reversal of our gaze shows that this being directed [*Gerichtetsein*] is really an immanent essential feature of the respective experiences involved; they are "intentional" experiences (Husserl 1927, Chap. 2).

Husserl describes how "a multiple and yet synthetically unified intentionality" arises from the phenomenological reflection on the synthesis of appearances of an object as orientation to in changes. In considering his example of a die, he mentions orientations of left-and-right, near-and-far, front-and-back. These orientations form an intentional structure comprised of the perceptions of the *actually-seen*, *undetermined* (such as the back side) and *unseeable* which has to be inferred because they are obscured. This 'directness of appearance' formed the basis of Gestalt psychology, yet Husserl was thinking of more than the direct appearances:

> The intentional structure of any process of perception has its fixed essential type [*seine feste Wesenstypik*], which must necessarily be realized in all its extraordinary complexity just in order for a physical body simply to be perceived as such. If this same thing is intuited in

[19]A summary of the historical development of Husserl's connotation of the term is available in Runes (1942).

> other modes for example, in the modes of recollection, fantasy or pictorial representation—to some extent the whole intentional content of the perception comes back, but all aspects peculiarly transformed to correspond to that mode (op cit).

He was also not restricting himself to visual phenomena:

> This applies similarly for every other category of psychic process … these constitute themselves, with fixed essential forms corresponding to each process, in a flowing intentionality (op cit.).

but was aiming to

> investigate systematically the elementary intentionalities, and from out of these [unfold] the typical forms of intentional processes, their possible variants, their syntheses to new forms, their structural composition, and from this advance towards a descriptive knowledge of the totality of mental process, towards a comprehensive type of a life of the psyche (op cit.).

It follows, then, that a phenomenal sound object is a *multiple and yet synthetically unified intentionality* whose appearance is revealed through time, and consists of *actually-being-heard*, *undetermined* (previously heard) or *unhearable*–a missing fundamental, for example.

Husserl's phenomenology was a starting point and major influence on Pierre Schaeffer's treatise on musical objects as objects in their "flux of modes of appearing and the manner of their 'synthesis'" (Schaeffer 1966). Husserl's concern was to provide a theory of the purely phenomenal, psychical, multiple 'appearances' of objects; each appearance being a unit or component of meaning accruing to the phenomenal object as each of these appearances occurs. In order to do this he applied an attitude of 'self-restraint' (*epoché*) in which he 'bracketed-off' consciousness from the 'world-as-it-is' in order to develop what he later called an *eidetic* science of essential forms[20] that is, one that involves no assertion of actual material existence; pure, in the sense that pure logic is pure[21]. In constructing such phenomenological psychology, it could be

> exclusively directed toward the invariant essential forms. For instance, the phenomenology of perception of bodies will not be (simply) a report on the factually occurring perceptions or those to be expected; rather it will be the presentation of invariant structural systems without which perception of a body and a synthetically concordant multiplicity of perceptions of one and the same body as such would be unthinkable. (op cit. Chap. 4).

While Kant's proof of the existence of things outside of him, even though their nature was uncertain; of the necessity for 'things as they are' as a chink in the

---

[20]Husserl calls these forms *Eide*, singular *edios*. Greek origin: "By *eidos* I mean the essence of each thing and its primary substance" (Aristotle, *Metaphysics*). The verb is *eido* 'to see' appears in the Latin verb *video*. The term is related to Sanskrit's 'veda', a cognitive activity like 'knowing', and the Old English *wit*, 'to know'.

[21]Husserl described such eidetic phenomenon as *noetic*, that is, *noemata* (singular *noema*) conceived in the stream of consciousness (*noesis*) entirely by reason (*nous*).

Cartesian mind/body separation, had an impact on Husserl, his phenomenally pure psychology is still firmly bound to the mind. In his four criteria for the phenomenally pure psychology, there is no recognition of the body at all:

(1) The description of the peculiarities universally belonging to the essence of intentional mental process, which includes the most general law of synthesis: every connection of consciousness with consciousness gives rise to a consciousness.
(2) The exploration of single forms of intentional mental process which in essential necessity generally must or can present themselves in the mind; in unity with this, also the exploration of the syntheses they are members of for a typology of their essences: both those that are discrete and those continuous with others, both the finitely closed and those continuing into open infinity.
(3) The showing and eidetic description [*Wesensdeskription*] of the total structure [*Gesamtgestalt*] of mental life as such; in other words, a description of the essential character [*Wesensart*] of a universal "stream of consciousness."
(4) The term "I" designates a new direction for investigation (still in abstraction from the social sense of this word) in reference to the essence-forms of "habituality"; in other words, the "I" as subject of lasting beliefs or thought-tendencies–*persuasions*–(convictions about being, value-convictions, volitional decisions, and so on), as the personal subject of habits, of trained knowing, of certain character qualities (Husserl [1929] 1999).

To some, Husserl struggled, despite his protestations to the contrary, to separate his *eide* ('essential forms')[22] from Platonic Ideals (Smith 2007; Martin 2007).

3.3.9 Gestalt Psychology and Group Theory

Gestalt psychology arose from experimental findings of perceptual *Gestalten*, principally in the empirical experimentation laboratories established by another of Brentano's pupils, Carl Stumpf (1848–1936). This occurred at a time when what we now think of as experimental psychology was separating itself from philosophy into its own discipline. Max Wertheimer (1880–1943), Kurt Koffka (1886–1941, and Wolfgang Kohler (1887–1967) were instrumental in the founding of Gestalt psychology (Hergenhahn 1992). They considered perceptual *Gestalten* as phenomena arising naturally and directly from the physical nature of sensation.

Although initially Gestalt theory began as a theory of perception, Wertheimer clearly recognized the fundamental epistemological issue, and Kohler and Koffka presented Gestalt theory explicitly on epistemological dualist grounds. Curiously, as Harvard cognitive-neuroscientist Steven Lehar indicates, the Gestaltists did not reference Kant as the originator of the idea of *Gestalten*, possibly because of their

---

[22]Husserl calls these forms *Eide*, singular *edios*. Greek origin: "By *eidos* I mean the essence of each thing and its primary substance" (Aristotle, *Metaphysics*. The verb is *eido* 'to see' appears in the Latin verb *video*. The term is related to Sanskrit's *veda*, a cognitive activity like *knowing*, and the Old English *wit*, 'to know'.

confusion over his Idealist position (Lehar 2000). Lehar also suggests that one of the most controversial and pivotal aspects of Gestalt theory is Wertheimer's principle of isomorphism, which was elaborated by Kohler in 1924 as the hypothesis that every perceptual experience is "not only blindly coupled to its corresponding physiological processes, but is akin to it in essential structural properties" (Kohler, quoted in Lehar (2000)) which, he suggests "is a direct consequence of the indirect realist foundations of Gestalt theory, whereby phenomenal experience is a direct manifestation of neurophysiological processes in our physical brain, and therefore it cannot help but be similar in structure, since they are identical in ontology."

The French phenomenological philosopher Maurice Merleau-Ponty (1908–1961)[23] described the structure of perception as sensation and conscious awareness, between which phenomenology uncovered a 'figure-ground' invariant. As he put it: "To be conscious = to have a figure on a ground–one cannot go back any further" (Merleau-Ponty 1964a, 191).

In the second quartile of the nineteenth century, the mathematical notion of the group was being discussed, though its full importance as an organizing and clarifying principle was yet to be realized. According to the neo-Kantian philosopher Ernst Cassirer (1874–1945), the first attempt to apply speculations concerning the *concept of group* to psychological problems of perception was made by the physician and physicist Hermann von Helmholtz in 1868 (Cassirer 1944). Amongst his many enterprises, Helmholtz (1821–1894) was interested in tracing Kant's philosophical theories in fields such as physiology and the newly emerging empirical psychology. He recognized that at the root of perception is the concept of the *constancy* of perceptual objects under changing sensory conditions; of relations that remain unchanged or invariant under transformation. He endorsed Kant's thesis of space as a "transcendental form of intuition" but for him this was the beginning of, not a solution to, the problem of finding the most general form of invariance, that is, in which systems of points in a multi-dimensional manifold may be displaced relative to one-another without changing their forms.

Henri Poincaré (1854–1912) recognized the concept of the group as a fundamental *a priori* one, derived from an intuition that is imposed on us "not as a form of our sensibility, but as a form of our understanding" (Poincaré 1913, 79), and as such, one that precedes and underlies all experience. This led him to conclude that the axioms of geometry and empirical statements derived from observation and measurement cannot be compared because there is an irreducible difference between them; they belong to entirely different objects. It is a characteristic of perception that it can never abandon the here–and–now (*hic et nunc*), since its task is to apprehend it as precisely and completely as possible. Helmholtz is similarly clear that laws cannot be responsible for the causes of natural phenomena; that

[23]Merleau-Ponty was strongly influenced by Husserl and Heidegger, and closely, albeit disagreeably, associated with Sartre and de Beauvoir. He was the only major phenomenologist of the first half of the Twentieth Century to engage extensively with the sciences, and because of it, his writings have become influential with the recent project of naturalizing phenomenology, as discussed in Sect. 3.3.7.

explanations are the urge of our intellect (our intent, or what he calls “judgment”) to bring our perceptions under its control. Further, that though the perceptual world does possess a structure, objective stimuli are not simply copied in perception, but are transformed in a certain direction.

Cassier understood, following Ewald Hering’s (1834–1918) inquiries into the sense of light, that perceptual content is characterized by the reduction of dissimilarity in objective stimulus rather than construction of similarity. Thus perception integrates the impressions of stimulii rather than being bound to their flux. His question is whether it is a mere accident that the concepts of invariance and transformation belonging to group theory, recognized as a being fundamental to mathematics, appears in the exposition of psychological facts, even if the connection from such extrapolation is a “mediate” one (1944, 12). He concludes that the relationship between the formation of invariants in perception and in geometry are analogous; that the differences may be characterized by an expression which Plato used to define the opposition of perception to thought in which all perception is confined to the “more or less”; that the realization of perceptual constancy is never ideally complete, but always remains within certain limits, and beyond these limits, there is no further “transformation.”

3.3.10 Pragmatic Perception and the Meaning of Truth

Like Husserl before him, Ernst Mach (1838–1916)[24] was interested in the phenomenon of melody. Mach understood sensations to be not simply raw experiences but the interaction of experience with a pre-formed cognitive structure. For instance, when we hear a known melody, we recognize it no matter what its transposition or even mode: It can be hummed, whistled, or plucked from a harp. Furthermore, even if one or more notes are incorrect, we still recognize it. In *Analysis of Sensations* ([1886] 1906) Mach asks, “What constitutes a melody?” It seems empirically incorrect, he argued, to say, as we must, in recognition of the above, that a melody exists in our ability to recognize it and not, it having been formed as an idealization by experience of one or more examples of it, in sounds. This idealization captures not the actual sounds, but the relationships of the sounds to one another (Mach [1886] 1906). For Mach, this process is at the basis of all perception. Experience requires an *a priori*, but that *a priori* is itself formed by experience.

John Dewey (1859–1952)[25] also thought that immediately felt or sensed experiences constitute by far the larger part of total human experience, but which are on a different level from the knowledge experience. Sensory qualities, he contended, “are

[24]Mach lends his name to the speed of sound and was considered by Einstein to be his forerunner on the theory of relativity (Einstein 1954, 26).

[25]Humanist philosopher, educator and founder, with Peirce and the psychologist William James, of the Pragmatism school of philosophy.

not objects of cognition in themselves, … [but] acquire cognitive function when they are employed in specific situations as signs of something beyond themselves" (1938, 147). This Pragmatist philosophy was extensively developed by the psychologist William James (1842–1910). Pragmatists asserted that our world primarily reveals itself to us, not as something that merely is, reduced to the physical/abstract/scientific properties of things and processes, but as something to utilize and navigate through. "Theories thus became instruments, not answers to enigmas". By extension, the abstract noun *truth* is not a static relation between our ideas and reality, but refers to ideas which successfully guide our behavior towards goals.

> The truth of an idea is not a stagnant property inherent in it. Truth happens to an idea. It becomes true, is made true by events. Its verity is in fact an event, a process, the process namely of verifying itself, its verification. Its validity is the process of its validation. James (1909, vi)

This definition of truth was met with almost instant and complete rejection, in part perhaps because of his non-technical enunciation of it is his public lectures (James [1907] 1995). The pragmatic method, which we discuss in more detail in Chap. 4, is an

> attitude of looking away from first things, principles, 'categories', supposed necessities and of looking towards last things, fruits, consequences, facts ([1907] 1995, 22).

3.3.11 The Immediate Perception of Objects Through Sensation

Most of the epistemological discussions that followed the idealists accepted a subjectivist or internalist explanation of perceptual experience; namely, that while the veridicality of perceived objects cannot be ascertained on the basis of sense data alone, the experience of *internal* representations of objects was such as to justify inferring the existence of the corresponding *external* objects. What follows provides and overview of the principal non-naïve,[26] non-skeptical[27] explanations for how immediate perception of physical objects and the physical world generally can be warranted on the basis of sensory experience: *perceptual subjectivism*, consisting of *phenomenalism*, *representionalism* (otherwise known as *indirect realism*) and

---

[26]The naïve position, known as *naïve realism* is that physical objects are directly experienced; rejected by a large proportion of the philosophers as discussed in the previous section.

[27]The *skeptical* position is that sensible experience provides no evidence of external substances. Arising in the fourteenth century, it was used by those for whom the only certitudes are those of immediate experience and those of principles known *ex terminis* (by definition), together with conclusions immediately dependent on them. Most sceptics usually accepted a degree of probabilism, namely, that probability is the only guide to belief and action. Despite this lack of direct influence, the skeptical arguments of fourteenth century thinkers bear marked resemblances to those employed by Berkeley and Hume discussed earlier (Runes 1942).

direct realism, and less subjective *external* or *process reliabilism. Perceptual subjectivism* has a longer history and is widely accepted as the most justifiable (James 1995), while *direct realism* grew out of Gestalt psychology. The development of *reliabilism* began in the twentieth century and continues today.

Historically, a distinction has been made between sensations and perceptions: sensations are basic experiences elicited by simple stimuli, while perceptions are experiences elicited by complex, often meaningful stimuli. Being more complex, perceptions are often considered to be the result of the integration of simpler sensations. They may also involve other processes such as memory, thus ensuring the possibility of them being influenced by a knowledge of past experiences, and independent of the methods by which that knowledge is acquired, as discussed later.

The relationship between personal knowledge and perception has been extensively researched. For example, a faint tone is easier to hear if a listener knows what pitch to expect (Green 1961) cited in (Sekuler and Blake 1985, 423), and comprehension is enhanced for native speakers of a language when they understand that someone is speaking with a strong foreign accent. Other examples include the increasing ease, on perseverance, in reading someone's scrawling handwriting and the ease, having acquired the skill, of reading music notation or riding a bicycle. On the other hand, existing knowledge can impair or mislead accurate perception through mistaken or inferential expectations and ambiguities.

The term *immediate perception* is use to denote the content of those propositions that arise in the mind directly from external sensation, as distinct from those, for example, that arise as a result of memoric reflection; both types of which are (Kantian) phenomena. Also, the expression *public* is taken to mean external to the perceiver and *physical* or *material object* as a signifier to mean some *thing* with a non-virtual public existence in matter or energy, such as a table, a cloud or a violin tone.

3.3.12 The Sense-Datum Theory of Immediate Perception

In the literature, the term *sense-datum* or sometimes *sensum*[28] is used to denote an immediate un-analyzable private object of sensation; a non-physical entity that actually possesses the various sensory qualities that a person experiences. The adverbial or sense-datum theory argues that the direct or immediate object of an experience is an entity produced at the end of a causal process and is thus distinct from any physical object, if any, that initiated the process. Examples of the veridicality of this distinction include seeing the light from a star that no longer exists, viewing a straight stick that is immersed in water and so looks bent, feeling an itch in a previously amputated limb and hearing the voice of a deceased friend.

---

[28]Plural forms are *sense-data* and *sensa.*

Epistemologist Laurence BonJour (1943-) suggests such immediate experiences can be classified according to the following perceptual qualities (2007):

- **Relativity**. What is perceived has different qualities under different perceptual conditions, even though the relevant physical object does not change.
- **Illusion**. Qualities are experienced that the relevant object clearly does not possess.
- **Hallucination**. Qualities are experienced in a situation in which there is no physical object of the relevant sort present in the sensory field.

While it argues for the existence of sense-data, sense-datum theory doesn't account for the existent nature of it (its ontology), or the *relation* between it and the experiencing mind. The sense-datum is an object immediately present in experience and has the phenomenal qualities that it appears to have. The natural thing to say is that sense-datum somehow influences the state of mind of an individual in a way that reflects the sense-datum's specific character. It is this resulting state of mind that the adverbial theory describes.

In contrast to the sense-datum theory, the adverbial theory:

> … has no need for such objects and the problems that they bring with them (such as whether they are physical or mental or somehow neither). Instead, it is suggested, merely the occurrence of a mental act or mental state with its own intrinsic character is enough to account for the character of immediate experience.
>
> … I sense in a certain manner or *am appeared to* in a certain way, and it is that specific manner of sensing or way of being appeared to that accounts for the specific content of my immediate experience (BonJour 2007) (my italics).

The essential feature of adverbial theory is that there need not be an object or entity of any sort in the material world. They are objects of awareness, modes, or states of the mind and do not exist independent of it. To say that an idea is *an object of my awareness* is just a grammatically convenient way of saying that the idea is *that of which I am aware*.

In comparing the sense-datum and adverbial account of the contents of experience, BonJour thinks that the adverbial account is most likely to be more correct, because:

> if sense-data somehow affect the mind in a way that reflects their character, then the resulting adverbially characterizable states of mind are really all that matter, making the sense-data themselves superfluous; and if they do not affect the mind in such a way, then their apprehension by the mind is difficult or impossible to make sense of. … any characterization of sensory experience that can be given in sense-datum terms can equally well be adopted by an adverbial theorist, simply by construing a comprehensive sense-datum description of one's sensory experience as characterizing the specific manner in which one is adverbially "appeared to" (BonJour and Sosa 2003, 78–79, n3).

So, the sense-datum and adverbial theories are just different ways of expressing the same idea, *sense-datum* being the nominal or objective way of expressing the subjective content of sensing, or *being appeared to*. The adverbial form is almost always more unwieldy in English, so for the sake of simplicity I will use the simpler *sense-datum* to imply both, unless a distinction is called for.

3.3.13 Representationalism (Indirect Realism)

Indirect, or representative realism is the hypothesis that there is a justification for believing that our immediately experienced sense-data, when taken with further beliefs that we arrive at on the basis of these sense-data, constitute a representation or depiction of realm of material objects independent of our sensing of them.

One striking contrast between the representative realist's explanatory hypothesis and the others considered here, is that under the representative realist view there is a clear intuitive sense in which the qualities of the objects that explain our immediate experience are reflected in the character of that experience itself in such a way that these, albeit indirect, experiences can be said to be of the qualities of the objects.

3.3.14 Phenomenalism[29]

Phenomenalism is a theory of a perceptual subjectivity that maintains that the characteristics and relations of sense-data is *all* that constitutes the content of propositions about immediate (unreflective) perceptual experiences of public material objects. The theory, more correctly labeled *ontological phenomenalism*, grew as a radical form of empiricism with roots in Berkeley and Hume's subjective idealism, as discussed earlier, and developed by Ernest Mach in the nineteenth century, to be later refined by Bertrand Russell and the logical positivists (Tarnas 1991, 383), and not confused with Kant's *epistemological phenomenalism* which doesn't deny the existence of objects not experientially knowable (*noumena*). Ontological phenomenalism maintains that to believe that public material objects exist is to believe that various sorts of sense-data have been, and/or would be experienced under certain specifiable physical conditions, such as those that would intuitively permit the public objects to be perceived; the (relatively) *permanent possibilities of sensation* as John Stuart Mill expressed it (1865, 225–32). Phenomenalists offer no reasons for *why* such sense-data is *permanently possible*, and on the assumption that it would be immediately available to the perceiver if they happened to be there when a public material object was present, they do admit sense-data that is confined to a specific time and place.

Because the claim of phenomenalism is that the content of propositions about perceptual experiences of public material objects is given *entirely* in terms of

[29]Phemonenalism is to be differentiated from Phenomenology, a philosophical movement initiated by Edward Husserl, as discussed earlier.

sense-data, the sense-data, and of those environmental factors that are aspects of the order of the immediate experience, must be premised by other sense-data. For example, being able to assert that there is a violin in the room from actual or obtainable sense-data from a violin in the room is premised on sense-data of the room, as well as the violin because the room does not exist to the perceiver external of sense data. The need for sense-data to be specifiable only in terms of other sense-data, as the phenomenalist position seems to imply, leads to an infinite regress, so it fails as a theory of knowledge. Further, an offer of a specified route to the location of the violin (sense-data), would require a guarantee that such a sensory route exists to that location, difficult enough in present ('now') time, but unfathomable about objects and events in the (distant) past. BonJour reports Roderick Chisholm's generalized argument that

> there is in fact *no* conditional proposition in sense-datum terms, however long and complicated the set of conditions in the "if" part may be, that is *ever* even *part* of the content of a material-object proposition. This is shown, he claims, by the fact that for any such sense-datum proposition, it is *always* possible to describe conditions of observation (including conditions having to do with the state of the observer) under which the sense-datum proposition would be false, but the material-object proposition might still be true (BonJour 2007) (his italics).

A deeper problem is that our sense-data are obviously not random, but it is not clear *how* they are ordered without any reference to public material objects. Lastly, it seems that, as the above discussion implies, Phenomenalists, by only being able to infer knowledge of public objects by their own immediate experience of them, put themselves in the untenable solipsist condition that they cannot know of the existence of other minds or mental states. In addition, Kant's refutation of material idealism showed this position to be false.

3.3.15 Direct Realism and Ecological Psychology

The application of physicalism to psychology was the logical basis for the method known as Behaviorism and was used to support a theory of perception known as Direct Realism in which representations or ideas are not thought of as being themselves the immediate objects of awareness, but instead as directly constituting the act or state of awareness itself. Hence, at least in the case of veridical perception, the immediate object of awareness is the external thing itself, and not a representation of that thing.

> Cognitive states [i.e., representations], are not cognitive relations with objects, nor are they themselves peculiar objects supposed to mediate the occurrence of cognitive relations. They are simply the perceiver's awareness of possible objects.
>
> The immediate object of awareness is always the ordinary object and not some special object, and that therefore, for example, 'Intuitions… are the immediate awarenesses of… ordinary objects', rather than themselves objects of awareness (Aquila 1983, xi).

The most important immediately preceding realist philosopher was the Scot Thomas Reid (1710–1792) who wrote with extreme common sense in trying to refute his countryman David Hume's skepticism (Gregory 1981, 349–51). The psychologist James Gibson (1904–1979) developed Direct Realism into a theory of perception founded on the understanding that the senses had developed in an environment that thus latently afforded actions, whether or not these affordances were recognized (Gibson 1966, 1979). In a thorough analysis of the metaphysical roots of Gibson' psychology, Lombardo notes that

> …over a 50-year period he came to challenge both mind-matter dualism in his ecological theory and the epistemology of indirect perception in his direct realist philosophy of perception. Gibson's theory of ecological reciprocity avoids both the absolute philosophical dichotomies inherent in Platonic thinking, and the one-sided treatments of reality in monistic philosophies. … In examining the growth of Gibson's ecological psychology, one can find numerous roots. Aristotle, as a single theorist, probably anticipates Gibson more than any one, and in modern times, functional psychology, process ontology, Gestalt wholism, and the evolutionary-ecological view of nature have all influenced Gibson's thinking. (1987, 4).

Gibson is one of the few thinkers to explicitly defend a naïve view of perception, however he was forced to make seemingly implausible assumptions about the perceptual process in its defense, including a denial of the general materialist view that the sensory organs transmit sensory information into the brain, where neurophysiological processes compute a perceptual representation of the external world. Instead, Gibson suggested that perception occurs somehow out in the world itself, rather than in the physical brain. Exactly how this occurs, or what this actually means however, he could never explain very satisfactorily.

While direct realism may be a useful model for some perceptual mechanisms, it does not seem to be a plausible model of the complete experience of material objects, and so cannot be the basis of a justification for believing that our immediately experienced sense-data constitute a representation or depiction of an independent realm of material objects. That being said, JJ Gibson's postulation that the mind directly perceives environmental stimulii without additional cognitive processing, was influential support for perceptual phenomenologists such as Merleau-Ponty, and the emerging field of embodied cognition.

3.3.16 Information as Relations Through Signs

Fallibilism is the term the philosopher and scientist and founder of pragmatic semiotics Charles Peirce (1839–1914) described the unreliability of such empirical methods to provide meanings truthfully stronger than probabilistic propositions outside of the mind:

> …the doctrine that our knowledge is never absolute but always swims, as it were, in a continuum of uncertainty and of indeterminacy. Now the doctrine of continuity is that *all things* so swim in continua (Peirce 1868).

One consequence of the exploration of idealism was the understanding that, although sensations are experienced immediately, sense perceptions are not ideas and they do not give an instant knowledge of things; they are simply non-materialistic natural events that are neither true nor false in–and–of– themselves. Sense perception only becomes informatively meaningful when it stands for, becomes a sign of, something more than, or other than, itself, in some respect or other for somebody, such as when the perception of a high pitched repeated tone comes to signify *bird* to a human in one context and *alarm-clock* in another. Peirce labelled such hypothetical inferences as *abductions*: "[T]he content of consciousness, the whole phenomenal manifestation of mind, is a sign resulting from inference" (Peirce 1868).[30] Peirce developed a *pragramatical*[31] approach to the study of semiotic frameworks–the relationships between signs and their impacts on those who use them. Following Kant's *Categories,* Peirce developed a system of three existential Ceno-Pythagorean[32] categories:

1. *Firstness*. Reference to a ground (a ground is a pure abstraction of a quality). Essentially monadic; Informally: Quality of feeling.
2. *Secondness*. Reference to a correlate (by its relate). Essentially dyadic; Informally: Dependence.
3. *Thirdness*. Reference to an *interpretant*. (An *interpretant* is the product of an interpretive process, or the content of an interpretation.) Essentially triadic; Informally: Representation.

According to Peirce, a sign is something that stands for something else in some manner or other for somebody. Thus, the sign relation is triadic, involving:

1. A *causal relation* between a sign-user and something (an Object) that stands for something else (a sign, or in Peirce's terminology, a *Representamen*);
2. A *semantical relation* between the sign and the something else it stands for; and
3. A *pragmatical relation* between the sign and the thing it stands for, and the user (an Interpretant)–that is, the sign in the mind that is the result of an encounter with a sign.

Peirce distinguished three different ways in which things might stand for other things in some respect or other;

1. *Iconic relations*, in which things resemble other things; auditory icons and spearcons, for example;
2. *Indexical relations*, of causes and effects; earcons, for example, and
3. *Symbolic relations*, in which things are associated in essentially arbitrary ways; sounds and the letters of the alphabet, for example.

[30]A fuller explanation of abductive inference is outlined in the Forms of Knowledge section, earlier in the Chapter and in Appendix 2.

[31]A term Peirce invented to distinguish it from the more widely used term *pragmatic*.

[32]The prefix *ceno-* is from the Greek word *kainos*, which means "new" or "recent" and Peirce calls them 'Pythagorean" because they have to do with number.

While iconic and indexical relations exist in nature whether or not anyone notices them, they can only function as signs when the relationships by which they can be associated is noticed by someone as standing for some associated properties of other things. For example, words such as "wind" and "pipe" do not resemble the forms of moving air or a cylindrical tube for which they stand, so the words are not icons; tree rings are an effect of the aging of tree but can only act as an indication of (an indexical sign) of this aging once the relationship has been observed.

Peirce's semiotics is dependent on his deep understanding of phenomenal experience. It is useful to recognize that he thought that all processes of consciousness, including perceptual consciousness, involve or are sign processes, that is, are relational. According to Innis (1994), Peirce's central idea is that "the content of consciousness, the entire phenomenal manifestation of mind, is a sign resulting from inference…" (Peirce 1868, 53). Peirce emphasizes that he thought "Thirdness, [that is law-governedness] pours in upon us through every avenue of sense…*There can be no perceptual object without a unifying factor that distinguishes it from the 'play of impressions'*" (Innis 1994, 13) [my italics].

This is an important understanding with respect to the perception of immanent abstract objects, such as the mental impressions resulting from the sonification of multivariate datasets that have no direct perceptual correlates in the physical world. If it is necessary for such objects to have such a unifying factor, it bears some consideration how sonification can offer any possibilities in that regard.

3.3.17 Information in Networks and Connections

Cognitive science arose as an interdisciplinary research endeavor in the early 1960s around the idea that the brain could be identified with hardware, and on which cognition was computed. Because of its physical approach, Direct Realism has been very influential in the development of the application of neural networks and other connectionist models—in pursuit of environment-sensing robots capable of automatic navigation, for example. This has proved remarkably successful up to a point. In its broadest interpretation, Cognitivism considers that the mind is rational, autonomous and independent of the body; ideas strongly valued in Western culture, as evidenced by the long-running fascination with talking dolls, robots and other automata (Wood 2002). However Cognitivism's identification of the individual as central to thought is currently being challenged by the growing body of evidence from artificial intelligence research that embodiment is important to cognition, perhaps more important than vision (Varela et al. 1991; Hutchins 1995); that different sense modalities provide access to different types of information, and even that artificially intelligent agents require ontologically mediated schemata, as seminal artificial intelligence researcher Derek Partridge clearly infers. (Partridge 1991, 171–227).

The connectionist models favored by the cognitivists was in parallel with the music composition modelling experiments of those computing composers who saw the connectionist paradigm as offering "a new and unified viewpoint from which to investigate the subtleties of musical experience. Connectionist systems employ 'brain-style' computation, capitalizing on the emergent power of a large collection of individually simple interconnected processors operating and cooperating in a parallel distributed fashion" (Todd and Loy 1991, ix).

Though it is still a very active field, Cognitivism, no longer enjoys the limelight it once did. Computer scientist and editor of the *Journal of Consciousness Studies*, Joseph Goguen provides a succinct, if somewhat unbalanced, critique of Cognitivism and its relationship to the emerging field of consciousness studies. As the approach continues to fail to deliver results beyond the elementary, there is a growing awareness that it will not turn out to be the hoped–for panacea. Goguen thinks the next step, developing a computational model of conscious awareness, will not be achievable without a more sophisticated model of mind (Goguen 2002). David Chalmers, an influential philosophers of consciousness agrees:

> The easy problems are those of finding neural mechanisms and explaining cognitive functions: the ability to discriminate and categorize environmental stimuli, the capacity to verbally report mental states, the difference between waking and sleeping. The hard problem is that of experience: why does all this processing give rise to an experienced inner life at all? While progress is being made on the easy problems, the hard problem remains perplexing (1995).

3.4 An Attempt at Integration

Of the various attempts to integrate the three main threads: Idealism, Realism and Information Theory, probably the most widely known is the work of the British anthropologist and semiotician Gregory Bateson (1904–1980), who coined the definition of information as "the difference which makes a difference." Bateson adopted Alfred Korzybski's concept that "the map is not the territory":

> But what is the territory? Operationally, somebody went out with a retina or a measuring stick and made representations which were then put on paper. What is on the paper map is a representation of what was in the retinal representation of the man who made the map; and as you push the question back, what you find is an infinite regress, an infinite series of maps. [T]he process of representation will filter [the territory] out so that the mental world is only maps of maps, ad infinitum. All 'phenomena' are literally 'appearances' (Bateson 1972).[33]

Bateson considered that the great dualist dichotomy of epistemology has shifted under the impact of cybernetics and information theory, and that, with the discovery of cybernetics, systems and information theories was a formal base "enabling us to

[33]The original source of the quotation is from Bateson's Nineteenth Annual Korzybski Memorial Lecture entitled *Form, Substance and Difference*, delivered January 9, 1970, under the auspices of the Institute of General Semantics. and printed in the *General Semantics Bulletin*, No. 37, 1970. It was republished in (Bateson 1972).

think about mind and enabling us to think about all these problems in a way which was totally heterodox from about 1850 through to World War II" (op.cit.). In discussing the origins of "the map is different from the territory", he emphasized that the idea came out of ancient Greece:

> It all starts, I suppose, with the Pythagoreans versus their predecessors, and the argument took the shape of "Do you ask what it's made of—earth, fire, water, etc.?" Or do you ask, "What is its *pattern*?" Pythagoras stood for inquiry into pattern rather than inquiry into *substance*. That controversy has gone through the ages, and the Pythagorean half of it has, until recently, been on the whole the submerged half. The Gnostics follow the Pythagoreans, and the alchemists follow the Gnostics, and so on. The argument reached a sort of climax at the end of the eighteenth century when a Pythagorean evolutionary theory was built and then discarded—a theory which involved Mind. (op.cit).

He does not mean to imply by this that he is supports traditional Cartesian dualism, but a new approach in which the

> individual mind is immanent but not only in the body. It is immanent also in pathways and messages outside the body; and there is a larger Mind of which the individual mind is only a subsystem. This larger Mind is comparable to God and is perhaps what some people mean by "God", but it is still immanent in the total interconnected social system and planetary ecology (op. cit.).

Bateson's overt aim is one of raising his listener's awareness of ecological issues, so perhaps his primary concern is best described as 'functional epistemology'. Suggesting that the resolution lies in the God Mind is a familiar way of collapsing the unresolved mystery back into itself and further emphasizes that epistemology is a process of thought, less directed to providing conclusions than it of exploring a mental landscape, of organizing beliefs and experiences into holistic descriptions.

3.5 Perception and the Neural Correlates of Consciousness

Whilst a considerable amount is now known about the structure and functions of individual neurons, the fundamentals of how macro effects emerge from populations of neurons are still largely obscure, despite considerable effort over the last decades. As the field develops, there is a growing realisation that the phenomena associated with *consciousness*, *nonconsciousness* and *cognition* are too diverse to continue to be meaningfully subsumed under the same ill-defined terms (Churchland 2005). For example, given the verifiable presence of nonconscious antecedents to an intention (Libet 1985), it is unclear how formed our decisions are when we become aware and think of ourselves as mentally "creating" them.

The search for the neural correlates of consciousness has been aided by the ease of Functional Magnetic Resonance Imaging (fMRI) of cortical activity. However, it is suggested by Churchland and others (Damasio 1995; Llinas 2001) that the ready availability of such technologies has contributed to a cortical "chauvinism" that tends to concentrate on conscious perception at the neglect of the role they have in servicing behavior. Specifically that, in service of keeping the body alive, the nervous systems of animals, as *movers*, function to support planning, deciding and executing these plans *in movement*. Importantly, much of the brain's input is consequent upon the dynamical feedback loop between observed phenomena and an organism's own movements, exploratory and otherwise. This loop extracts vastly more information about the causal properties of the external world in a given time interval, leading to greater predictive prowess, that is, skills regarding the causal structure of the world, than could a purely passive system.

Time is an essential component of causal knowledge, and predicting durations, interception intervals, velocities, and speeds of various body movements is critical to an animal's survival. Efference copy (being aware that a movement is one's own and not the world's) is also thought to be critical, as perhaps is the nonconscious "analysis" and memory of the movement of other movers, such as in predator–prey/pursue–evade relationships, for example. In contradistinction to the conventional wisdom that "the sensory pathways are purely sensory", according to the Guillery and Sherman hypothesis, messages to the thalamus and cortex also carry information about ongoing instructions to the organism's motor structures (2001). Consequently, as a developing organism begins to interact with the world, sensory signals also "carry" gestural predictions: as an animal learns the consequences of a particular movement, it learns about what in the world will probably happen next, and hence what it might do after that.

Damasio's studies of efference copying of one's own thoughts and empathy with others provide even more evidence for this thesis that perception, learning and memory are not just cerebral processes but are embodily integrated into an organism as, what Polanyi called, tacit knowledge (Damasio 1999; Polanyi 1975).

3.5.1 Mirror Neurons

Kohler et al.'s finding, not only that that certain neurons in the ventral premotor area will fire when a macaque monkey performs a single, highly specific action with its hand: pulling, pushing, tugging, grasping, picking up and putting a peanut in the mouth etc., but that that "mirror neurons" will also fire when the monkey in question observes another monkey (or even the experimenter) performing the same action, offers some neurological basis for a theory of cultural inheritance, "mind reading" empathy, imitation learning, and even the evolution of language (Kohler et al. 2002). As Churchland observes, by shifting perspective from "visuocentricity" to "motor–sensory-centricity," the singular importance of temporality takes center stage in an hypothesis that "time management," for want of a better term, is the key to the complexity of tasks of thalamic nuclei, and very probably also to a range of conscious phenomena (op.cit. 2005).

More recent studies have demonstrated that a mirror neuron system devoted to hand, mouth and foot actions, is also present in humans. Buccino, Solodkin, and Small review this literature, and that of the experimental evidence on the role of the mirror neuron system in action understanding, imitation learning of novel complex actions, and internal rehearsal (motor imagery) of actions (Buccino et al. 2006). The finding that actions may also be recognized from their typical sound, when presented acoustically, has important implications for embodied soniculation research. Besides visual properties, it was found that about 15% of mirror neurons, called *audio-visual mirror neurons*, also respond to the specific sound of actions performed by other individuals even if only heard (Kohler op. cit. 2002). It has been argued that these neurons code the action content, which may be triggered either visually or acoustically. Phillips-Silver and Trainor demonstrated an early cross–modal interaction between body movement and auditory encoding of musical rhythm in infants (Phillips-Silver and Trainor 2007). They found that it is primarily the way adults move their bodies to music, not visual observation, that critically influences their perception of a rhythmic structure. Their results suggest that while the mere visual observation of a conspecific's goal-directed movement (e.g., reaching for an object or hand–to–mouth action) is sufficient to elicit a neuronal representation of the action, this does not transfer to the domain of metrical disambiguation (Wilson and Knoblich 2005). So it appears that, either this type of rhythmical body movement is not an example of the kind of object-directed action that activates the mirror neuron system, or the information provided by the mirror neurons is not strong enough to influence the later-recalled auditory metrical representation of a rhythmic pattern.

In an experimental study of gestures, subjects of various ages were able, with a high degree of accuracy, on only hearing different individual human's walking and running on various kinds of surfaces, to determine their sex (Bresin 2003). A consequential inference is that differences in ambulatory action, presumably resulting from relatively small differences in skeletal anatomy, is tacitly 'available' to listeners. Also consequent to these findings is the need for better models of multi-modal sensory input, particularly with respect to the integrative functions of vestibulation and proprioception, which some empirical evidence suggests are available to listeners though aural means alone (Varela, Thompson, and Rosch 1991).

3.5.2 Critical Discussion

While knowledge of the structure and functions of individual and clusters of neurons is increasing, there are billions of them, each with tens-of-thousands of connections so there is no certainty, even when the overall functioning of the neural system is significantly better understood that it currently is, that such an understanding will be able to adequately account for the ability to synthesize perceptual objects. In fact, if the *rate* at which pulses are transmitted turns out to be the minimum unit in an account of the relevant activity of the nervous system (von Neumann 1963) and the diameter of an axon, which might be a function of the recency of a signal passing down it, plays a crucial role in processing information by acting as a filter (Rosenblith 1966), there may be no reason to believe that information processing at neurological-level can ever be formally described (Dreyfus 1992).

The mentalist approach has failed to find any means by which mental representations can be reliably accumulated for conscious reflection, at least not without a good deal of training and effort. This somewhat explains some of the difficulties reported in parameter-mapping sonification research that the vast majority of ordinary listeners, for whom a low conceptual loading is necessary for continued engagement, are precluded from making ‘sense’ of them, just as the attempt to use computers to develop an ‘artificial intelligence’ based on computational theories of mind that rely on a classical reductionist approaches such as “mind-is-to-software as brain-is-to-hardware” failed to be able to understand even the simplest stories—principally because of the unrepresentability of the background knowledge and the specific forms of human “information processing” which are based on their way of being in the world. This suggests that, when compared with the ease with which everyday sounds are identified; the ease with which a myriad of melodies are learned, remembered and identified; that the mentalist approach is inadequate at best. More likely, that it is just wrong.

The relatively recent availability of tools to abstract sound from its origin in the physical action of objects, and the development, alongside that of “good-old-fashioned-AI” (GOFAI) (Haugeland 1985) of seminal software for computer music (Mathews 1969), has blurred the functional distinction between sound and music, much as often occurs between data and information, and information and meaning. Sound recording enabled Schaeffer, building on the philosophical foundation of Husserlian (that is mentalist, *époché*) phenomenology, to propose a musical analysis based on *reduced listening*, that is, listening to sounds “analytically” for their own sake, as sound objects, removed from their real or supposed sources and meanings (Schaeffer 1966). It is of particular interest, in the light of the previous discussion of the role of time and causality in perception, that while Schaeffer does discuss tempo and temporality, he makes almost no reference to pulse and rhythm.

3.6 The Perceiving Body

Husserl’s pupil Martin Heidegger (1889–1976) was critical of the subject/object split that pervades the Western tradition and that is in evidence in the root structure of Husserl and Brentano’s concept of intentionality; that is, that all consciousness is consciousness *of* something, and (the idealist notion that) there are no objects without some consciousness beholding or being involved with them. Heidegger encompassed terms such as “subject”, “object”, “consciousness” and “world” into the concept of a mode of "being-in-the-world" as distinct from an essentially Positivist "knowing" of objects in the universe that is required for navigating the environment–measurement, size, weight, shape, cause and effect etc. His Being-in-the-world is characterized as "ready-to-hand":

> … the kind of dealing which is closest to us, not a bare perceptual cognition, but rather that kind of concern which manipulates things and puts them to use; and this has its own kind of 'knowledge.' (Heidegger [1927] 1962, 95).

In other words, participatory, first-hand experience: familiarity, tacit know-how, skill, expertise, affordance, adaptability etc. Heidegger argues that our scientific theorizing of the world is secondary and derivative and he exposes an ontology that is far broader than the dualistic Cartesian framework. He stresses the primacy of the readiness-to-hand, with its own kind of knowing, or relating to the world, in terms of what matters to us; what is significant and meaningful, and to which we attend. It follows, from Heidegger's perspective, that human action is embodied; that human knowing is enactive, and participatory.

The Hungarian scientist and philosopher, Michael Polanyi (Introduced in Sect. 3.2.1) proposes a type of participative realism in which personal knowledge plays a vital and inescapable role in all scientific research, indeed, in all human knowing:

> Let us therefore do something quite radical … let us incorporate into our conception of scientific knowledge the part which we ourselves necessarily contribute in shaping such knowledge. (Polanyi 1975, 28–29).

By stressing the tacit nature of participatory knowing, Polanyi claimed that "we know more than we can tell". In this way he emphasized knowledge that is implicit to tasks, situations and attitudes. He used the term *tacit knowledge* to refer to those things we can do without being able to explain how, that is, in the absence of explicit rules or calculative procedures. The "indwelling" nature of tacit knowledge is important in the development of the skill of reflexivity, such as is needed in the sifting through and interpreting of qualitative data.

Heavily influenced by both Husserl and Heidegger, Merleau-Ponty produced a much more developed understanding of the body and its role in non-conceptual perception (Merleau-Ponty [1945] 1962, [1964b] 1969). As perhaps the only major phenomenologist of the first half of the twentieth century to engage extensively with the sciences, he was able to systematically demonstrate the inability of the mentalist and empiricist explanations to adequately account for observed phenomena. In doing so, he produced a theory of perception in which the body and the world are entwined; in which perception occurs through the "intentional tissue" of the "body schema" (*schéma corporel*); much as epigenetic alterations occur in a phenotype by the osmotic transduction of molecules through semipermeable membranes.

Samuel Todes (1927–1994) builds on Merleau-Ponty's work by beginning to work out a detailed phenomenological account of how our embodied, non-conceptual perceptual and coping skills open a world to us. He then works out twelve perceptual categories that correspond to Kant's conceptual categories, and suggests how the non-conceptual coping categories can be transformed into conceptual ones (Todes 2001).

3.7 Summary

The relationship between our sensing of a variegated world and the way we interpret it has been a major theme in Western philosophy since its recorded beginnings. For Plato, all experiences of the world were understood by the extent to which they conformed to pre-existent Ideas that were 'received' at birth and ever-present in the mind of the perceiver. Aristotle's dissatisfaction with this 'unworldly' view led him to speculate about, and importantly experiment with, the structure and function of the world as he experienced it. On regaining Grecian scholarship from the Near East following the European Dark Ages, the Scholastics, began to question the received wisdom of the church, which many identified as a form of misplaced Platonism. For the sake of brevity, the insights of this period were not discussed in this chapter, important though some of them may turn out to be for an understanding the relationships between emotion and awareness that are currently engaging neuroscientists (Damasio 1995, 2003).

The uncertainty of obtaining reliable knowledge of the world through direct sensory experience that the idealists had exposed, undermined the Cartesian confidence in human reasoning about the perceived world as the foundation for truth and sent philosophers in various directions, including a search for new ways to underpin traditional rhetorically-based logic with mathematical foundations. At the same time as Hume demonstrated the unreliability of the senses, and on the basis of which he thus postulated the unknowableness of the world, empirical methods were being used to increase knowledge *of* that world, and obtain power over its natural forces. Kant's resolution of this apparent conflict, by demonstrating the special relationship between space, time and perception, which he called *Transcendental Idealism*, was a lens through which later philosophical investigations were obtained.

According to Kant's understanding, what, exactly, is meant by *information* is embedded in relationships between the sensation, perception and apperception of phenomena; what he called *appearances*[34]: things as they are for humans, as opposed to things as they are 'in–or–of–themselves'. From this perspective, information can be simply characterized as phenomena, or thoughts about

[34]*Erscheinungen* in German.

phenomena, in the mind of some person. Brentano and Husserl developed their pure (contemplative) psychology of perception, known as phenomenology, by bracketing off the ‘world–as–we–know–it’, Kant’s ‘things–as–they–really–are’ (*Ding an sich*, *noumena*), to try to determine whether or not the properties of phenomena were capable of being formally (that is logically) organized; firstly in the mind of the perceiver and secondly as sharable with others—a characteristic Husserl called *intersubjectivity*. Husserl’s aim was to develop an *eidetic* science of essential, invariant phenomenal forms that involves no assertion of actual material existence, but he struggled to keep them conceptually separate from Plato’s Ideals. At the same time, Peirce thought that there can be no perceptual objects without a unifying factor that distinguishes them from the ‘play of impressions’.

The success of methodologically rigorous empirical approaches in the natural sciences wedged the study of the psyche away from its purely first-person introspective philosophical roots. The most significant early discovery that resulted was that of perceptual *Gestalten*, which their discoverers considered as phenomena arising naturally and directly from the physical nature of sensation. Though the Gestaltists had found an important invariant, they had difficulty in extending the idea past the *noumenal* world. However, it is a characteristic of phenomenal forms (such as melodies) that their properties can remain unchanged when the objective stimuli upon which they rest undergo certain modifications (such as transposition). This phenomenon of identity is related to a much more general problem found in abstract mathematics; of invariances with respect to variations of the primitive elements out of which a form is constructed. The particular kind of identity that is attributable to apparently heterogeneous figures, because they can be transformed into one another using operations that define a group, exists in the domain of perception and permits us to grasp perceptual “structures”. The mathematical concept of transformability corresponds to the concept of transposability in perception. So, by accepting “form” as a primitive concept, Gestalt psychology made an attempt to free psychological theory from contingency on the mere mosaic of perceptions.

Without considering cultural differences,[35] not all group-theoretic transformations of perceptual objects are equally cognized, nor are the same transformations as easily perceivable in different sense modalities. For example, symmetry group transformations of pitch and temporal structures, such as transposition, inversion and retrogradation, occur frequently in music though they seem not to be all equally evident to the casual observer: under non-extreme pitch transposition and tempo

[35]Which needs to be done cautiously. The structural characteristics of music from a wide variety of cultures seem generally comprehensible, suggesting that cultural differences are more likely to be of degree rather than kind. Assumptions need to be treated skeptically, however. Werker and Vouloumanos (2001) document, many studies that indicate linguistic speech plasticity in infants is subject to cultural influences and in a classic study, (Segall et al. 1966) showed that the Müller-Lyer illusion is culturally determined, not neurophysiological.

acceleration, a melodic structure remains strongly invariant; pitch contour inversion and rhythmic retrogradation are common occurrences but not as strongly invariant, while rhythmic inversion seems not to be perceptually invariant, or even generally defined.

Gibson reacted to the findings of perceptual gestalts by questioning the existence or importance of the phenomenal. Instead, he based his Direct Realism theory of vision on an evolutionary approach whereby organisms and their environments develop reciprocities over genetic and diurnal time; their senses develop in an environment that thus latently afforded actions, whether or not these affordances were recognized. While Direct Realism may be a useful model for some perceptual mechanisms, it does not seem to be a plausible justification for believing that immediately experienced "sense-data" constitute a representation or depiction of an independent realm of real material objects.

The connectionist/cognitivist approach has been to build sensing, then responding, then perceiving, then environmentally aware machines, in the hope that, given enough neural connections, zombies would appear that could function in the world around them with intent.[36] These models have failed to bridge the divide between the easy and the hard problems of consciousness, and so researchers are beginning to explore other models of mind and the rôle emotions plays in intent.

If it is the *intended* effect that ultimately identifies information, as Küppers suggests, and that intent is *by* the listener *of* a sonification, presumably on the basis of what is meaningful to them, understanding what the principles are by which sonifications of abstract multivariate datasets can be made to reliably afford such intent, seems to require more, or something other, than what idealist and rationalist philosophies and their abstract codifications in the cognitive sciences have so far revealed about the processes involved. Pragmatic methods are inductive and appear, supported by current and developing models of human perception and cognition as embodied, goal-oriented processes, to be more fecund with possibility than all other approaches discussed. Thus, the next chapter explores in some detail how such a methodology can be applied to the task of using sonification to find relevant information in datasets. This then lays the groundwork for Chap. 5, which is a discussion of the requirements of a computational framework necessary to apply that methodology.

[36]According to Chalmers, "A zombie is physically identical to a normal human being, but completely lacks conscious experience. Zombies look and behave like the conscious beings that we know and love, but all is dark inside. There is nothing it is like to be a zombie." http://consc.net/zombies.html.

Appendix 3A General Methods of Acquiring Knowledge

General methods of acquiring knowledge

Authority: Leaders, considered knowledgeable and wise, decide what is true for everyone, sometimes during periods of special inspiration, insight or perhaps revelation. The surrender to authority is common in most arenas of human endeavour. It acts to stabilize and encapsulate a corpus of traditional knowledge against which new ideas can be tested.

Revelation: Often used by seers and prophets who employ magic and divination techniques. What is usually assumed is that the practitioners have ability to access knowledge through inspired communion with supernatural beings. Such access usually requires mindful techniques such as faith (belief), and is sometimes induced by body renunciation techniques.

Intuition: The direct, immediate and certain apprehension of truths without the intervention of conscious reasoning or related sensory perceptions. The difference between intuition and a revelation is the person doing the intuiting does not assume that the source of the knowledge is external to their own brain.

Heuristics, Folklore and Commonsense (Also known as Intentional Realism)[37]: The generalizations we apply in everyday life in predicting and explaining each other's behavior, often collectively referred to as folk psychology. They are both remarkably successful and indispensable. A person's 'personal knowledge' i.e. what they believe, doubt, desire, fear, etc. is a highly reliable indicator of what they will do, and we have no other way of making sense of each other's behavior than by ascribing such states and applying the relevant generalizations.

It recognizes that we all, in some way, are committed to the basic truth of common-sense psychology and, hence, to the existence of the states its generalizations refer to (Dretske 2000). Some, such as Fodor, also hold that common-sense psychology will be vindicated by cognitive science, given that propositional attitudes can be construed as computatinal relations to mental representations (Fodor 1987).

Churchland (1981) thinks that, as a theory of the mind, folk psychology has a long history of failure that can't be incorporated into the framework of modern scientific theories, including cognitive psychology. He argues that the states and representations folk psychology postulates simply don't exist; that it is comparable to alchemy and ought to suffer a comparable fate. On the other hand, Dennett (1987) seems prepared to admit that the generalizations of commonsense psychology are true and also indispensable, but denies that this is sufficient reason to believe in the entities to which they appear to refer. He supports this stance on that basis that there is nothing more to having a propositional attitude than to give an intentional explanation of a system's behavior by adopting an the intentional stance

[37]For an extended description, See http://plato.stanford.edu/entries/reliabilism/.

toward it. Assuming a system is rational,[38] if the strategy of assigning contentful states to it and predicting and explaining its behavior is successful, then the system is intentional and the generalized propositional attitudes we assign to it are true (Dennett 1987: 29).

Appendix 3B Inference Methods of Acquiring Knowledge

Inference covers a number of forms of reasoning in which conclusions are drawn or judgments made on the basis of circumstantial evidence and prior conclusions rather than purely on the basis of direct observation or knowledge arrived at by direct observation. The conclusion may be correct, incorrect, partially correct, correct to within a certain degree of accuracy, or correct in certain situations. Five distinct inferential methods are recognized: deduction, induction, Bayesian inference (which is really a form of induction), abduction and reliability. Inferential methods are the principal methods of science, although not all types are undisputedly considered applicable to all fields of inquiry.

Inference methods of acquiring knowledge

Deductive: Assuming a system is rational, if the strategy of assigning contentful states to it and predicting and explaining its behavior is successful, then the system is intentional and the generalised propositional attitudes we assign to it are true (Dennett 1987: 29). If a then b. [b is a consequence of the assumption of antecedent a].

Inductive: Assuming a system is rational, if the strategy of assigning contentful states to it and predicting and explaining its behavior is successful, then the system is intentional and the generalised propositional attitudes we assign to it are true (Dennett 1987: 29). There are three types of induction:

(a) Simple: All observed a are b, therefore all a are b. (enumerative induction).

(b) Proportonal: P(g), a percentage of known g's in group G, have attribute A. Individual i is another member of G, therefore there is a P(i), corresponding to P(g), that i has attribute A.

(c) Anaogic: a is similar to b. a has attribute X, therefore b has attribute X.

Bayesian: Assuming a system is rational, if the strategy of assigning contentful states to it and predicting and explaining its behavior is successful, then the system is intentional and the generalised propositional attitudes we assign to it are true (Dennett 1987: 29).

If $P(b|a) = x$ (if the probability of b given a is x), then $P(a|b) = x \cdot P(b)/P(a)$.

(then the probability of a given b equals x multiplied by the probability of b divided by the probability of a).

[38]Something is rational if behaves in accordance with the truths and falsehoods afforded by its environment.

Abductive: Assuming a system is rational, if the strategy of assigning contentful states to it and predicting and explaining its behavior is successful, then the system is intentional and the generalised propositional attitudes we assign to it are true (Dennett 1987: 29).

If D is a collection of data (facts, observations, *a priori* assumptions), and H is a collection of possible hypotheses (H1, H2, H3, … Hn) for explaining D, then the Hn that explains more D in the best, most elegantly is probably true

Reliability S knows P if and only if S truly believes P, and S's belief that P was produced by a reliable belief-forming process: Reliableness is a method for acquiring knowledge based on various belief-forming processes. It justifies the belief in the veridicity of perceptual sensations if the resulting perception is known to lead to a suitably high proportion of true beliefs. It is not a requirement of users of the method, or anyone else, to know that the process is reliable or have any sort of knowledge of its reliability–all that is required is that it is, in fact reliable. Thus, no appeal to sensory experience is required, effectively short-circuiting the issue that divides Representationalism and Phenomenalism. Reliabilism thus rejects the issue on which all three of the more traditional theories attempt to respond to: the issue of how sensory experience provides a reason for thinking that perceptual beliefs are true. On the assumption that our perceptual processes are in fact reliable in the way that we take them to be, it offers a seemingly straightforward and account of how perceptual beliefs about physical objects and the physical world are justified.

Reliabilism emphasizes the properties of the processes used to arrive at truths. In reliabilist approaches to knowledge acquisition, noticing a static relationship between a conjecture and a body of evidence, knowing that an hypothesis does not contradict the evidence, or even is in-accord with it, for example, is insufficient to warrant support for it from the evidence; additional account must be taken of how reliable the method that produce the hypothesis is known to be in producing truthful hypotheses. Reliabilist thinking underpins the greater acceptance of the diagnostic judgements of experts over laypersons, the preferential support for research programs with fecund histories and the scorn of *ad hoc* hypotheses.

The first, then unrecognised as such, formulation of a reliability account of knowledge was in the mathematician Frank Ramesy's writing on knowledge (Ramsey 1931). Several similarly subjunctive theories, such as tracking theory and contexturalism, were developed in the latter part of that century, as discussed by Goldman, who notes. Reliability theories of knowledge of varying stripes are still in active development and continue to appeal to many epistemologists for their robustness and flexibility. Permutations abound. [Some theories] focus on modal reliability; on getting truth or avoiding error in possible worlds with specified relations to the actual one. They also focus on local reliability, that is, truth-acquisition or error avoidance in scenarios linked to the actual scenario in question Goldman (2008).

References

ACOD (1992) The Australian Concise Oxford Dictionary, 2nd edn. Oxford University Press, South Melbourne

Aquila RE (1983) Representational mind. Indiana University Press, Bloomington

Aristotle. (350BC) (2008) De Anima (On the Soul). Translated by J.A. Smith. http://psychclassics.yorku.ca/Aristotle/De-anima/de-anima2.htm

Bateson G (1972) Steps to an ecology of mind, vol 8. Ballantine Books. Accessed, New York

Berkeley G (1710) (1957) A treatise concerning the principles of human knowledge, Part I. Edited by Turbayne CM. Forgotten Books, London, United Kingdom

BonJour L (2007) Epistemological problems of perception. In: The stanford encyclopedia of philosophy. Metaphysics Research Lab, Stanford University. http://plato.stanford.edu/archives/sum2007/entries/perception-episprob/

BonJour L, Sosa E (2003) Epistemic Justification: Internalism vs. Externalism, Foundations vs. Virtues. Blackwell Publishing, UK

Brentano F (1874) (1995) Psychology from an empirical standpoint. In: descriptive psychology. Routledge, London

Bresin R, Dahl S (2003) Experiments on gestures: walking, running, and hitting. In: D, Fontana Rocchesso, F

Buccino G, Solodkin A, Small SL (2006) Functions of the mirror neuron system: implications for Neurorehabilitation. Cognitive Behav Neurol 19(1)

Burnet J (1904) Greek philosophy. From Thales to Plato. MacMillan Press, London

Cassirer E (1944) The concept of group and the history of perception. Philos Phenomenol Res 5:1

Chalmers DJ (1995) Facing up to the problem of consciousness. J Conscious Stud 2(3):200–219

Churchland PM (1981) Eliminative materialism and the propositional attitudes. J Philos 78:67–90

Churchland PS (2005) A Neurophilosophical slant on consciousness research. In: Casagrande VA, Guillery R, Sherman S (eds) Cortical function: a view from the thalamus. Vol 149. Progress in Brain Research. Elsevier, Amsterdam

Damasio AR (1995) Descartes' error: Emotion, reason, and the human brain. Penguin, NY

Damasio AR (1999) The Feeling of what happens. Harcourt Brace, NY, New York

Damasio AR (2003) Looking for Spinoza: Joy, sorrow, and the feeling brain. Harcourt, Orlando, Fl

Dennett D (1987) The intentional stance. MIT Press, Cambridge, MA

Descartes R (1641) (1901) Meditations. Translated by Veitch J. http://www.wright.edu/cola/descartes/meditation2.html

Dewey J (1938) Logic: the theory of inquiry. Holt, New York

Dretske FI (1981) Knowledge and the flow of information. Blackwell, Oxford

Dretske FI (2000) Perception, knowledge and belief. Cambridge University Press, Cambridge

Dreyfus H (1992) What computers still can't do. MIT Press, Cambridge, MA

Einstein A (1954) Ideas and opinions. Books, New York, Bonanza

Fodor JA (1987) Psychosemantics. The MIT Press, Cambridge, Mass

Fromm E (1947) Man for himself. An inquiry into the psychology of ethics. Rinehart, New York

Gibson JJ (1966) These senses considered as perceptual systems. Houghton Mifflin Company, Boston, MA

Gibson JJ (1979) The ecological approach to visual perception. Houghton Mifflin Company, Boston, MA

Goguen JA (2002) Consciousness and the decline of cognitivism. In: Accessed. The UCSD distributed collective practices symposium

Goldman A (2008) 2015 Reliabilst epistemology. In: The stanford encyclopedia of philosophy. Metaphysics Research Lab, Stanford University. https://plato.stanford.edu/

Green BF Jr (1961) Using computers to study human percpetion. Educ Psychol Measuremnt 21 (1):227–233

Gregory RL (1981) Mind in science. Cambridge University Press, UK

Haugeland J (1985) Artificial intelligence: the very idea. The MIT Press, Cambridge, MA

Heidegger M (1927) (1962) Being and time. Translated by Macquarie J. Blackwell, Oxford. Blackwell, Oxford, UK

Hergenhahn BR (1992) An introduction to the history of psychology. Wadsworth Publishing Company, Belmont, CA

Hoffmeyer J, Emmeche C (1991) Code-duality and the semiotics of nature. In: Anderson M, Merrell F (eds) On semiotic modeling. Mouton de Gruyter, Berlin and New York, pp 117–166

Huemer W (2008) Franz Brentano. In: Zalta EN (ed) The stanford encyclopedia of philosophy, Fall 2008. Metaphysics Research Lab, Stanford University. http://plato.stanford.edu/archives/fall2008/entries/brentano/

Hume D (1777) (1978) An enquiry concerning human understanding. Edited by Selby-Bigge LA. The Posthumous Edition. Project Gutenberg. http://www.gutenberg.org/cache/epub/9662/pg9662.txt

Husserl E (1927) Phenomenology. Article for Encyclopaedia Britannica. Translated by Palmer RE. Husserl's Shorter Works Pages 2(2):77–90

Husserl E (1929) (1999) Cartesian meditations: an introduction to phenomenology. Translated by Cairns D. Vol 2. Kluwer Academic, Netherlands

Hutchins E (1995) Cognition in the wild. The MIT Press, Cambridge, MA

Innis RE (1994) Consciousness and the play of signs. University of Illinois Press, Bloomington and Indianapolis, Illinois

James W (1909) The meaning of truth. David Mckay Company Inc, NY

James W (1907) (1995) Pragmatism: a new name for some old ways of thinking. Dover Publications, Inc

Kant I (1787) (1787) Critique of pure reason. Translated by Meiklejohn JMD. 2nd ed. http://www.gutenberg.org/dirs/etext03/cprrn10.txt

Kant I (1787) (2007) Critique of pure reason. Translated by Smith NK. 2nd ed. Macmillan Palgrave, London

Khinchin AI (1957) Mathematical foundations of information theory. Dover, New York

Kohler EC, Keysers MA, Fogassi UL, Gallese V, Rizzolatti G (2002) Hearing sounds, understanding actions: action representation in mirror neurons. Science 297(5582):846–848

Küppers BO (1990) Information and the origin of life. MIT Press, Cambridge

Lagerlund, H. 2008. "Mental Representation in Medieval Philosophy." In The Stanford Encyclopedia of Philosophy, edited by E.N. Zalta. Metaphysics Research Lab, Stanford University

Lamont C (1949) The philosophy of humanism, 8th edn. Humanist Press, Amherst, New York

Lehar S (2000) The function of conscious experience: an analogical paradigm of perception and behavior. http://cns-alumni.bu.edu/~slehar/webstuff/consc1/consc1.html

Libet B (1985) Unconscious cerebral initiative and the role of conscious will in voluntary action. Behav Brain Sci 8:529–566

Llinas RR (2001) I of the vortex: from neurons to self. MIT Press, Cambridge, MA, USA

Locke J (1690a) An essay concerning humane understanding. Book 1. 2nd ed. Vol 1. 2 vols. Project Gutenberg. http://www.gutenberg.org/ebooks/10615

Locke J (1690b) An essay concerning humane understanding. Book 2. Vol 2. 2 vols. Project Gutenberg. http://www.gutenberg.org/ebooks/10616

Lombardo TJ (1987) The reciprocity of perceiver and environment. The evolution of James J. Gibson's ecological psychology. Lawrence Erlbaum, Hillsdale, NJ

Mach E (1886) (1906) Analysis of sensations. Edited by Waterlow S. Translated by Williams CM

Martin (2007) Review of D.W. Smith's 2006 Edition of Husserl. http://ndpr.nd.edu/review.cfm?id=10923

Mathews MV (1969) The technology of computer music. The MIT Press, Cambridge: Cambridge, MA
Merleau-Ponty M (1945) (1962) The phenomenology of perception. Translated by Smith C. Routledge & Kegan Paul, Oxford, UK
Merleau-Ponty M (1964a) The primacy of perception and other essays on phenomenological psychology, the philosophy of art, history and politics. Edited by Edie JM. Translated by Cobb W. Studies in phenomenology and existential philosophy. Northwestern University Press, Evanston
Merleau-Ponty M (1964b) 1969. The visible and the invisible. Edited by J.M. Edie. Translated by A. Lingis. Studies in phenomenology and existential philosophy. Northwestern University Press, Chicago, IL, USA
Mill JS (1865) An examination of Sir Wlliam Hamilton's philosophy. Longman Green and Co., Elibron Classics, London
Moles A (1966) Information theory and esthetic perception. Translated by Cohen JE. Ubana University of Illinois Press, and London
Neumann J von (1963) Probablistic logics and the synthesis of reliable organisms from unreliable components. In: AH Taub (ed) Collected Works, 5:372. Pergamon Press, NY, New York
Partridge D (1991) A new guide to artificial intelligence. Intellect Books, Bristol UK
Peirce CS (1868) Some consequences of four incapacities. J Specul Philos, 2http://www.peirce.org/writings/p27.html
Phillips-Silver J, Trainor LJ (2007) Hearing what the body feels: auditory encoding of rhythmic movement. Cognition (Amsterdam: Elsevier) 105:533–546
Poincaré H (1913) The foundations of science. Translated by Halsted GB. The Science Press, New York. http://www.archive.org/details/scienceandhypoth00poinuoft
Polanyi M (1975) The tacit dimension. Doubleday, Garden City, NJ
Popper KR (1959) The logic of scientific discovery. Revised Edition. Hutchenson, London
Ramsey FP (1931) Knowledge. In: Braithwaite RB (ed) The foundations of mathematics and other essays. Harcourt Brace, New York
Reicher M (2006) Nonexistent objects. In: The stanford encyclopedia of philosophy. metaphysics research lab, Stanford University. http://plato.stanford.edu/archives/fall2008/entries/nonexistent-objects/
Roederer JG (2005) Information and its role in nature. Springer, Heidelberg
Rosenblith WA (1966) On cybernetics and the human brain. The American Scholar, 247
Runes DD (1942) Dictionary of philosophy. Philosophical Library, New York
Schaeffer P (1966) Traité des objets musicaux: Essai interdisciplines. Editions du Seuil, Paris
Schopenhauer A (1844) The world as will and representation. Translated by Kemp J, Haldane RB. Kegan Paul, Trench, Trubner & Co, London
Segall M, Herskovits MJ, Ferguson Cabrera DS (1966) The influence of culture on visual perception. The Bobbs-Merrill Company, New York
Sekuler R, Blake R (1985) Perception. A. Knopf, Inc, New York: Alfred
Shannon C, Weaver W (1949) The mathematical theory of communication. The University of Illinois Press, Urbana: Ill
Shepard R (1999) Pitch perception and measurement. In: CookMusic PR (ed) In. The MIT Press, Cambridge, MA, pp 149–165
Sherman SM, Guillery RW (2001) Exploring the thalamus. Academic Press, San Diego, CA
Smith DW (2007) Husserl. Routledge, London-New York
Stamper R (1996) Signs, information, norms and systems. In: Holmqvist (ed) Signs of work: semiosis and information processing in organisations. Walter de Gruyter, New York
Tarnas R (1991) The passion of the Western mind. Random House, London
Taylor AE (1911) The words Eido. In Varia Socratica, First Series. James Parker, Oxford
Theusch E, Basu A, Gitschier J (2009) Genome-wide study of families with absolute pitch Reveals linkage to 8q24.21 and locus heterogeneity. Am J Hum Genet 85(1):112–119
Todd PM, Loy G (1991) Music and connectionism. MIT Press, Cambridge, MA
Todes S (2001) Body and world. The MIT Press, Cambridge, MA

Varela F, Thompson ET, Rosch E (1991) The embodied mind. The MIT Press, Cambridge, MA
Werker J, Vouloumanos A (2001) Speech and language processing in infancy: a neurocognitive approach. In: Luciana M (ed) Handbook of developmental cognitive neuroscience. The MIT Press, Nelson, CA
Wilson M, Knoblich G (2005) The case for motor involvement in perceiving conspecifics. Psychol Bull 131(3):460–473
Wood G (2002) Living dolls: a magical history of the quest for mechanical life. Faber and Faber, London

Chapter 4
Intelligible Sonifications

The domain of the unknown surrounds us, like an ocean surrounds an island. We can increase the area of the island, but we never take away much from the sea.

J. B. Peterson (1999)

Abstract The previous chapter traced a path towards an understanding of the inadequacy of epistemological approaches to knowledge formation that do not account for the deeply-embodied nature of perception, to succeed in solving any but highly constrained problems. The slower-than-expected rise in the effectiveness of artificial intelligence provided the impetus for a more critical examination of the ontological dimensions of perception and human knowledge, which necessitated the re-cognition of another form of truth which is not derived empirically but from meaningful action. This chapter enunciates this Pragmatist approach as it can be applied to sonification design: a generative activity in which bespoke design skills are supported by scientific research in biology, perception, cognitive science–in the field of conceptual metaphor theory in particular, and aesthetics. It proceeds to outline pertinent features of a broad design methodology based on the understandings developed that could yield to computational support.

Reaching back to the ancient Greeks, Chap. 3 provided an overview of Western civilization's attempt to understand the nature of perception and how human beings use it to construct coherent understandings of the world as it appears to them. This evolving understanding was a fulcrum in the development of the empirical method, on which the last four- to five-hundred years of scientific endeavor is based, and which continues to be a central *modus operandi* of our age. The empirical method played a central role in the establishment of psychology as a separate disciple from philosophy, and has led to the establishment of new fields, including those of neuroscience and artificial intelligence.

There is considerable research to show that a user's engagement with a stimulus can influence their perception and cognition of it, as well as the information and meanings they derive from it when doing so (Merleau-Ponty [1945] 1962; Gibson 1966; Lombardo 1987; Dretske 2000; Lehar 2000; Kohler et al. 2002). Such

D. Worrall, *Sonification Design*, Human–Computer Interaction Series, https://doi.org/10.1007/978-3-030-01497-1_4

engagement is interactional and intentional, or attentional, on multiple levels, including when mediated through technological interfaces. Because the goal of auditory displays is to convey meanings to listeners through sound, knowledge of the perceptual and cognitive processes that afford their engagement is crucial to the development of intelligible, sustainable sonifications. This is implicit in the Sonification Report's definition of sonification as... *the transformation of data relations into perceived relations in an acoustic signal for the purposes of facilitating communication or interpretation* (Kramer et al. 1999).[1] This goal is equally important, whether undertaken by an expert designer of bespoke sonic data representations, or computational systems that model classes of design solutions which a non-specialist might use.

The quest to understand the nature of perception, and how human beings use it to construct a coherent world is a defining quality of Western thinking. Through it, we have come to appreciate that the central questions reside in an understanding of consciousness, especially the consciousness concerned with the phenomenal perception of inexistent objects–of which information in data is an incumbent example– and our *intention* towards them.[2] This leads us to the central theme of this chapter, which is to determine, bearing in mind John Flowers' salutary observation that while *the claim that submitting the entire contents of* '*dense and complex*' *datasets to sonification will lead to the* '*emergence*' *of critical relationships continues to be made, I have yet to see it* '*work*' (2005), how sonification designers, by considering both the conceptual and practical issues and a productive software sonification framework, can support a listener to "position" themselves with respect to a dataset, so they can maximally extract information from it as efficiently and "truthfully" as possible.

4.1 Two Forms of Truth: Material and Functional

4.1.1 Material: Rational and Empirical Truth

Most Enlightenment scholars considered rational thought to be independent of the body. In fact, one of the overt purposes of rationality was to dispense with such incidentalities as the body's motivations, emotions, and other subjective demands. Until the last decades of the 20th century, the dominant paradigm was that, when we emerge into the natural world as infants, we begin to construct a mental model of it, and operate in it in ways that this mental model suggests we should. We then compare what actually occurs in the world to what our model predicted and, if

[1]The inclusion of the terms *perceived relations* and *communication or interpretation* in this definition highlights the need for an understanding of perceptual and cognitive processes in the development of effective auditory displays, although interestingly, the working group for this report initially disagreed over including perception in the definition of sonification.

[2]See Chap. 3, especially Sects. 3.3.7 and 3.3.8.

necessary, modify the mental model. From this physical perspective, the worlds[3] in which we live can be conceived of as collections of objective phenomena. Immanuel Kant thus conceptualized the material world as composed of events and objects as *things-in-themselves*. That is, the truth of their existence and their relationships with other objects are independent of observation: A *thing-in-itself*[4] thus being an object devoid of all values, but having what Heidegger was to later call *presence-at-hand*[5] ([1927] 1962, 132).

In as much as they were concerned with understanding the objective world, the developing empirical sciences inherited this conceptualization. According to classical materialist thinking, scientific understanding was the sole means of gaining objective knowledge of phenomena and of the things that undelay them as the "order of nature."[6] So, in order to understand the world, abstractions of it should be objective representations, independent of an observer's subjectivity. It follows then, that there is no scientifically rational justification for attaching any value system to the results of scientific enquiry. In fact, the scientific method positions researchers as abstract agents, attempting to remove all such attributable values (biases) from their *a priori* description of objects and activities, until only descriptions conceived to be value-free remain. Having so removed such values in the process of an enquiry, it is thus not logically possible to scientifically extract any value propositions from the results. This is not to infer that scientific results cannot be informative in making value decisions, but that the decision to use such results is dependent on values (such as the desire to avoid catastrophe) that are causally independent of the results themselves. For the empirical rationalists, then, truth is just the system of propositions which have an unconditional claim to be recognized as valid; a label for all those judgements we are obliged to accept about the world that was abstracted from the more chaotic real world prior to, or in the process of, formulating the rationalization.

4.1.2 Functional: Truth as a Value Proposition

Pragmatists assert an additional kind of truth.[7] They consider that something is true to the extent that it can be validated, corroborated and verified by any means, and

---

[3]A distinction is expressed by using the word "world" for a human centered domain, whilst "earth" and "planet" are used for the physical environment and "universe" for the sky and stars etc. A world is a place in an environment (see Sect. 4.2), not just an abstract space.

[4]German: *Ding an sich*, which Kant called noumena as distinct from a thing as it is knowable by the senses through its phenomenal attributes. See Chap. 3 for a more detailed discussion.

[5]German: *Vorhandenheit*.

[6]This classical scientific attitude was revealed to be artificial in the twentieth-century by quantum mechanics. For some of the early considerations of the consequences of these discoveries for brains, minds and consciousness see William James (1992, 227).

[7]Chapter 3 contains a perception-oriented introduction to Pragmatism.

false if it cannot James ([1907] 1995, 88; 77). From a pragmatic perspective, the world primarily reveals itself to us, not as something that merely is, reduced to the physical/abstract/scientific properties of things and events, but as something to utilize and navigate through. Truth has a value proposition attached to it, as in *true to form*; *straight and true*; *to thine own self be true*. That is, on-target or in agreement with an intent or desire. Such truths emerge from facts (data), which themselves are not true or false, they just are.[8]

> The truth of an idea is not a stagnant property inherent in it. The truth *happens* to an idea. It *becomes* true, is *made* true by events. Its verity *is* in fact an event, a process: the process, namely, of its verifying itself, its veri*fication.* Its validity is the process of its valid*ation* (W. James [1907] 1995, 79) (his italics).

Given that there can be numerous potential solutions to any problem, the Pragmatist's solution to how to decide if an ideational solution is accurate or not, is to construct an interpretative framework that contains a model based on a mapping strategy that has a target or goal. Such models of interpretation are not random or even abstract, but are subject to significant constraints because they have to perform in an embodied listener's world. If 'running' this formulated model doesn't produce the intended outcome, either the ideational solution is not accurate (true) or the model is ideationally or psycho-perceptually deficient (false). This is the truth of the existence in objects of something more than just their *thingness*, but their *equipment* quality, something Heidegger calls *readiness-to-hand*[9] ([1927] 1962, 98).

Propositional truth is the truth of Keats's *Ode*:

> Beauty is truth, truth beauty—that is all
>
> Ye know on earth, and all ye need to know.[10]

In Keats' tradition, truth never signifies correctness of intellectual statements about things, or the meaning of material truth, as now understood by science. It denotes the wisdom by which people live, as John Dewey's comment on the *Ode*, emphasizes:

> The critical words are "on earth"—that is amid a scene in which "irritable reaching after fact and reason" confuses and distorts instead of bringing us to the light. Ultimately there are but two philosophies. One of them accepts life and experience in all its uncertainty, mystery, doubt, and half-knowledge and turns that experience upon itself to deepen and intensify its own qualities…([1934] 1952, 34).

[8]Continuing the theme of Footnote 3, the analogy is that empirical truth is to pragmatic truth as space is to place.

[9]German: *Zuhandenheit*.

[10]*Ode of a Grecian Urn*. (1820).

4.1.3 Truth and Sensory Perception

For biological creatures that actively function in their worlds, there currently appears to be no methodology for answering questions with respect to materially-objective truth and functionally-valid truth simultaneously. While it is not feasible to remove one in favor of the other, according to Husserl's central concept of *intentionality*, which directs or relates consciousness to its objects, there is always an inseparable correlation between the "intentions" of "constituting" consciousness—the source or origin of our experience—and the objective "constituted" meaning structures they intend, and in which they are united.[11] It is, after all, from this constituted world that our theoretical, objective world arises (Mabaquiao 2005).

So, from the Pragmatist perspective, our worlds are not just collections of abstract objects. Living beings are of a different order to inert matter; they are active, and behave with meanings and motivations. Those with complex nervous systems may also be conscious, thinking, imagining, beings: *Dasein* in Heideggerian terminology.[12] The relations that develop concomitant to our awareness of our proximity to objects, situations and to other beings, develop influential meanings in the process of us orienting ourselves in the world. For example, even though the sense of "closeness" is mutable, it is no less real between individuals, than are metrical measurements: The closeness of the bond between a mother and her child, is ignored at peril to both of them; trust between people is developed by them speaking truthfully to each other, which is not the same as them uttering empirical truths … or vacuous truisms.[13] Events and objects manifest themselves to us as things because of what they afford us; because of our capacity to use them. Maurice Merleau-Ponty asserted that the process of getting a *maximal grip* on things in the world, mirrors processes that are built into the fundamental structures of our perceptual systems:

> My body is geared into the world when my perception presents me with a spectacle as varied and as clearly articulated as possible, and when my motor intentions, as they unfold, receive the responses they expect from the world (Merleau-Ponty [1945] 1962, 250).

The problems of understanding exactly *what* is in a space, or how events unfold, are not just physical and conceptual problems, they are deeper biological ones, which

---

[11] See Chap. 3, especially Sects. 3.3.7 and 3.3.8. Husserl's conception of intentionality differs from his teacher's Brentano's, through which it came from the Scholastics. This difference can be traced in the thought of William James (Mullane 1973, 48).

[12] *Dasein* in vernacular German means existence. However, Heidegger, following Nietzsche and others sought to use it as a term for the primal nature of Being (*Sein*); always engaged, not just subjectively or objectively, but coherently as Being-in-the-world. This favoring of active, practical engagement with one's environment is thus ontologically opposed the Cartesian concept of abstract agency. Collins and Selina (1998, 48–61)

[13] Such as "Here is ear to hear", which is not at all vacuous, except for the empty-headed.

we share with all living creatures: They are integral to, and embedded in, the very nature of perception itself. The spaces we encounter are simply too complex to be perceived *in toto*. There are an infinite number of things in any space, just as there are, for all intents and purposes, an infinite number of relations in a dataset. Our nervous systems, including a large part of our brains, deals with the integration of the sensory world as an inseparable part of our bodies. Only a very narrow range of the electromagnetic spectrum is un-instrumentally available to us, and so our visual perception has evolved to notice things in front of us, of a certain size, or that are “handy,” that take precedence. And, being, mechanically constrained, our ears are a part of a body that can only binaurally transduce omnidirectional acoustic energy within limited frequency and amplitude ranges within limited proximities. So, what of the universe is open to us perceptually is dependent on our bodies, which “filter out” much of the currently unusable chaos in favor of potentially relevant events and objects. Such filtering is thus as necessary as it is inevitable; it is a determining characteristic of the nature of the perceptual process itself.

4.2 Cognition: Physical and Psychophysical

Cognition can be defined as “the mental action or process of acquiring knowledge and understanding through thought, experience, and the senses” [OED]. Studies in the cognitive sciences encompass such topics as attention, sensation, perception, mental action, information, memory, judgement, intelligence, emotion, affect, thinking, reasoning, comprehension, and knowledge. Following the re-emergence[14] of the central role of the body in contemporary Western intellectual thought, our understanding of the meanings of, and interrelationship between, many of these terms has deepened, and in some case radically shifted.

Without diminishing the undeniable power of logical reasoning, as clearly demonstrated by the success of the scientific method, such reasoning is, biologically speaking, a rarity, and, for approximately all biological creatures, not at all necessary or sufficient for survival. The paradox of cognitivist’s way of thinking, uncovered by artificial intelligence (AI) research in thesecond half of the twentieth century, is that a purely computational approach rarely works with anything but the simplest problems. Today, most expert AI systems are based on knowledge and pattern-recognition, not computational reasoning (Brooks 2002). The world is overwhelmingly chaotic and complex, and logical reasoning takes time, which increases exponentially in proportion to the number of variables involved. So, it is understandable that nature’s solution has been to have creatures reach survival-oriented decisions based on various forms of memory which require a

[14]The reference it to the Middle Ages, and prior, during which the mind and body were seen to interlink in ways that have remarkable modern resonances (Saunders and Fernyhough 2016).

minimum of reasoning capability.[15] Further, thinking is not separated from action: In order to stop the possibility of the body acting out thoughts while dreaming during paradoxical or rapid-eye-movement (REM) sleep, it is necessary for the musculature of the body to be relaxed; so as to induce paralysis known as the somatic or vegetative phenomena.[16]

Considered in this way, the function of thinking can be understood, not so much the objective abstract representation of the world, as the conceptualization and practice of the proper way of being *in* it; motivated to successfully manifest those actions that a living thing has to manifest in order to continue to exist. Abstract thinking does occur–in the brain's prefrontal cortex which is devoted to planning, and this ability to abstract, affords the separation of thinking from action, which is necessary for the ability to think about things *not* to do. Voluntary thought is a way of representing possible scenarios in the hypothetical worlds, which is necessary to imagine, plan and explore the survival possibilities or advantages of potential activities in the world without having to implement them in actions in the world. Such thinking is very much a broader and more embodied activity, than mere computational prowess. An understanding of the limited role logical deduction plays in our ability to negotiate and survive in the world, indicates the absurdity of thinking that it alone might be the basis of, or sufficient for, successful sonification designs.

4.2.1 The Neurophysiology of Perception and Memory

Neuropsychological research is able to assist us in the quest to understand cognitive processes. It can now demonstrate, for example, that the materialist's proposition that reason and emotion are separate, is not supported by the biological evidence: Instead, that they are integrated and mutually interdependent (Damasio 1995, 1999). While a detailed contemporary description that differentiates the functions, interconnections and integrations of the various sensory systems is beyond the present context, the purpose here is to provide a broad description of the neurophysiology of perceptual sensation, and the role of emotion in what we perceive and remember. This will assist in developing an understanding of how, and under what

[15] Even skillful chess, once the holy grail of AI, seems to rely more on memory that computational power. De Groot found that, after only five seconds exposure, grandmasters were able to reproduce real-game board configurations with a high degree of accuracy while less-skilled players were far less accurate (1966). Further, Chase and Simon found that there were no differences between expert players and others if the original configuration of the pieces was random, rather than from real games (1973).

[16] This paralysis is triggered by action of structures in a dorsal area of the pons under the influence of noradrenalin. If a cat's cerebral cortex is removed, this paralysis is not induced and when dreaming, the animal moves around (Jouvet and Jouvet 1963, 1967; Bjursten et al. 1976; Louie and Wilson 2001).

conditions, the memories of processes and events of a sonification are stored and reused. Our interest is principally in gaining an understanding of the perception and memory of auditory processes and informational phenomena; specifically, to use the language of Brentano, of *intentional inexistent phenomena*[17] such as the informational content of datasets.

The left and right hemispheres of the brain's outer layer or cerebral cortex have distinct fissures or crevices which divide each hemisphere into four lobes or regions: frontal, parietal, occipital and temporal, as illustrated in Fig. 4.1a. In combination, these regions may be regarded as two primary functional systems, known as the *motor cortex* and the *sensory cortex*, as illustrated in Fig. 4.1b, c. The *motor cortex* is located towards the back of the prefrontal lobe of the neocortex. It is composed of, front-to-crown, the prefrontal, premotor and motor lobes, which are located forward of the central sulcus, the fissure between the frontal and parietal lobes, as illustrated in Fig. 4.1b. The motor cortex is primarily concerned with the planning, control and execution of voluntary movements. In humans, this region is larger and more complex than in any other species, which partly accounts for our intelligence and behavioral adaptability, as demonstrated in our capacity to translate intentions into complex action plans, and then organize and regulate their execution.

The *sensory cortex* consists of the parietal lobes, located beneath the parietal bones in the top-back region of the neocortex, the occipital lobes beneath the occipital bone over the cerebella, located at the lower back of the head above the neck, and the temporal lobes, beneath the temporal bones, roughly located on the left and right sides, in region around and behind the ears, as illustrated in Fig. 4.1a, b. The sensory cortex is responsible for the construction of the sensory system; principally the visual cortex on the occipital lobes, shaded red in Fig. 4.1a and yellow in Fig. 4.1b, hearing in the auditory cortex on the temporal lobes (shaded green in Fig. 4.1a, b and touch, and their integration into a unified conscious experience. Other primary senses, not included in this discussion, include those for taste, smell and touch.

The *limbic system*, also known as the *paleomammalian cortex* is a set of structures located deep in the midbrain, forward of the third ventricle, above the brainstem, as partially illustrated in Fig. 4.1d, e, f.[18] The limbic system supports a variety of functions, including the primordial emotional processing of input from the sensory system, and long-term memory creation. Its major components are summarized in Table 4.1.

Animals have evolved, physically and psychologically, to survive in environments that consist of what is recognized and meets their expectations, and what is not recognized that could be either dangerous or advantageous–or both.[19] There are

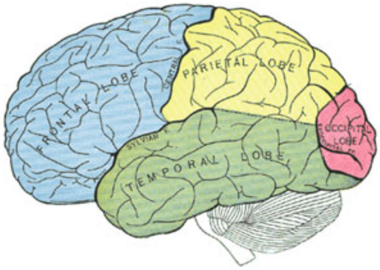
---

[17]See Chap. 3, Sect. 3.3.7.

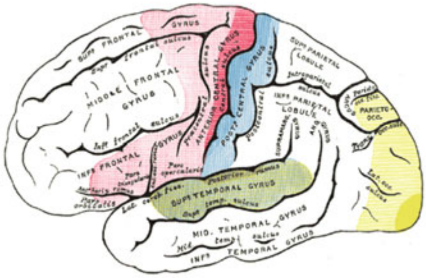
[18]From Latin limbus meaning "edge" or "border" [in this discussion, between the cerebral cortex and the brain core], related to limb, limbo, limber, limen, and perhaps limp [OED].

[19]This discussion of the limbic system is informed by Jorden Peterson's insightful Maps of Meaning (1999, 48–61), pp. 48–61 supplemented details from many Wikipedia articles on various brain structures. Other sources are referenced in the normal manner.

The sensory-motor cortex of the brain



(a) Locations of the four lobes of the cerebral cortex. The striated region shown beneath the temporal lobe is the cerebella. Gray, Fig.728

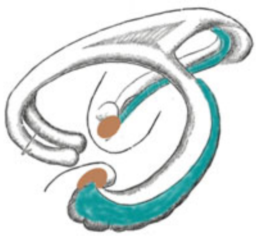


(b) The motor system (in reds) towards the back of the prefrontal lobe of the neocortex. Gray, Fig. 576.

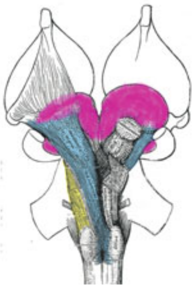
(c) The motor tract, back view. From Grey, Fig. 764.

Parts of the limbic system

(d) A part exposure of the left hemisphere hippocampus. From Grey, Fig. 739.



(e) The 'C-shaped bundle of nerve fibres, the fornix,showing the amygdalae (in brown) and hippocampii. Adapted from Gray, Fig. 747.



(f) The thalami (in cerese) are arched over by the fornix (e). From Gray, Fig. 690. 90° rotat.

Fig. 4.1 Illustrations of some major components of the human brain's motor, sensory and limbic systems involved in learning and memory. Illustrations, some retouched, from Gray's *Anatomy of the human body* (1918)

Table 4.1 The limbic system: Primary structural components and functions

Structure	Function
Hippocampus	Two seahorse-shaped structure that converts short-term to long term (episodic) memory. Figure 4.1d, e.
Amygdala	Two almond-shaped masses of neurons, one on either side of the thalamus at the lower, frontal end of the hippocampus. Regulates fear, aggression, and interest, attention and pleasure. Figure 4.1e.
Thalamus	Relaying sensory signals, including motor signals to the cerebral cortex. Regulation of consciousness, sleep, and alertness. Figure 4.1f.
Hypothalamus	Homeostatic regulation of temperature, appetite, thirst, pain, pleasure, sexual satisfaction, anger, aggression, autonomic nervous system.
Olfactory bulb	Olfaction (odor detection). The extremely old primary olfactory cortices, also beneath the temporal lobes, are not shown in Fig. 4.1.

an infinite number of potential events we might encounter, so it is impossible to monitor our chaotic worlds for all such potentialities. To be too sensitive to the flux of this chaos is to risk fear, and anxiety without adequate reward. At the other extreme, to be too insensitive, is to suffer from inertia, with similar outcomes.[20] It appears that the unmoderated limbic state is experienced as that primal of all emotions—fear—which is generated by the lack of suppression of amygdalic function. If sufficiently moderated, a state of curiosity, inquisitiveness, or hope to novel situations can emerge.

A major function of the hippocampus is to monitor particular locations in the sensory system for indications or signals of what Jordan Peterson calls *motivational significance* or *affective importance* (1999, 51): a complex decision-making process in which patterns in the signals are compared with learned expectations of the significance of similarly-structured signal patterns held in memory. If these signals are deemed irrelevant, they are ignored, even suppressed by the limbic system so as to reduce the need for attentional awareness to only those things that potentially require it (Peterson 2017).[21]

Memory thus plays a vital role in determining the relevance of sensations. A major function of the hippocampi is to monitor the sensory system for the presence of known (memorized) signal patterns, recognize and store new patterns or modify existing patterns and categories of patterns, as appropriate. This also play an important role in the consolidation of information from short-term memory (in the pre-frontal cortex) to long-term memory, and in spatial memory that enables

[20]This convex (outward-oriented) need can in some way be considered as a partner to the concave (inward-oriented) need for homeostatic regulation, as performed by the hypothalamus.

[21]This is a biological version of the Pragmatist's solution for determining if an ideational solution is accurate or not by constructing an interpretative framework that contains a models based on a mapping strategies that has a targets or goals. See Sect. 4.1.2.

Table 4.2 Regions of the brain used to store memories of various kinds

Memory type	Data type	Location
Explicit (Declarative)	Events and facts "know-what"	Medial temporal lobe and diencephalon (greater thalamus region)
Implicit (non-Declarative) i.e. has no conscious component	Skills and habits: "know-how"	Striatum
	Priming	Neocortex
	Conditioning-simple/ Pavlovian Emotional responses Skeletal/Muscular	Amygdala (no cortex) Cerebellum
	Non-Associative Learning (Habituation, Sensitization)	Reflex pathways

navigation (Petersen and Sporns 2015; Gray 1995; Swanson 2000). The clinical evidence is that, while the hippocampal region of the brain, illustrated in Fig. 4.1d, is important in the formation of new memories, they are not stored there, but, according to current understanding, in other cortical regions, as summarized in Table 4.2, after (Byrne 1997). People with lesioned or removed hippocampi show no impairment in recalling prior memories, but little or no ability to remember events after the intervention, even to the extent that new skills may still subsequently learned, but without conscious recognition that they have been acquired.

Through observing children's development, Jean Piaget (1896–1980) developed a theory about how people form and transform representations of the world. He concluded that there was a standard process: Infants produce a low-resolution representation that functions for them at the time. As they develop and acquire more experiences, they refine that representation by assimilating micro-alterations to their model, or transform the internal structure of their representation completely by accommodating a newly experienced action within the world. Piaget's genetic epistemology is a study of the origins of knowledge, the goal of which is to link the validity of knowledge to a model of its construction. The way he considered the development of factual ideas, at least in part, was that a person accumulates a set of ideas that they consider to be facts, and then those ideas would be superseded by a different theory within which that original theory would be nested (Piaget [1936] 1963). This theory is consistent with research on the use of schemas for the permanent store of knowledge in long-term memory discussed later.

The amygdala combines many different sensory inputs. It involved in primitive emotional learning and the retention of both fearful and pleasurable memories. Much of what is known about the amygdala and its role in emotional learning and memory comes from fear conditioning in which the cortex is not needed (Byrne

1997). Behavioral learning (conditioning) occurs in one of two ways: *classical conditioning* (respondent conditioning), or *operant conditioning* (instrumental conditioning). In classical conditioning, learning is an involuntary association formed between two (or more) different stimuli. Ivan Pavlov (1849–1936) famously demonstrated such conditioning by showing that if a bell was rung a short time after presenting food to a salivating dog, it would be conditioned (involuntarily learn) to salivate upon hearing the bell, even without the presence of the food. In *operant learning*, changes in behavior (learning) happens as a result of a positive (reinforcing) or negative (punishing) experience occurring after a response to a stimulus. Reinforcement itself, can be positive or negative. A positive reinforcement *adds* a reward, to encourage a repeat of the behavior, and negative reinforcement removes something undesirable. For example, receiving a smile in response to a greeting is a positive reinforcement; removing a headache as a reward for taking aspirin is a negative reinforcement, and being imprisoned is punishment, irrespective of whether the imprisonment is justified or not. Fear and pleasure can be experienced not just in response to learned experiences, but in response to the context in which the behavior is learned. The primary responsibility for contextual conditioning is the hippocampal structures. Stanley Kubrick's film, *A Clockwork Orange* contains multiple examples of all these conditioning, including with and related to music.

Although other components of the brain participate, the amygdalic and hippocampal systems are critically involved in production of brain waveforms associated with the *orienting response* or *orienting reflex,* which is part of an organism-wide immediate reaction, not only to a change in its environment when that change is not sudden enough to elicit an unconscious, largely defensive *startle reflex*, but when there is a discrepancy between the incoming signal and vestiges of similar past signals.[22] Both novelty and significance are implicated in the generation of an orienting response which is thus believed to play an integral role in constructing affective states of varying valence (preference), arousal, and motivational intensity. Specifically, the emotional significance of a stimulus, defined by its level of pleasantness, or otherwise, can affect the intensity of the attentional focus of attention on a subject (Sokolov et al. 2002; Ohman 1979). Peterson considers the orienting response to be fundamental, not only to perceptual consciousness but, given that it is engaged in the process of learning the consequences and thus motivations of the behavior that induced the response, behavioral directives:

> The orienting reflex is the general instinctual reaction to the strange category of *all occurrences which have not yet been categorized* – is response to the unexpected, novel or unknown *per se*, and not to any discriminated aspect of experience, any specifically definable situation or thing. The orienting reflex is at the core of the process that generates (conditional) knowledge of sensory phenomena *and* motivational relevance or valence. Such knowledge is *most fundamentally* how to behave, and what to expect as a

[22]The acoustic startle reflex is thought to be caused by an auditory stimulus greater than 80 decibels with as little as 10 ms. latency. Reactions to it include invoking muscular contractions that protect the cochlear itself (Musiek 2003).

> consequence, in a particular situation, defined by culturally-modified external environmental circumstance and equally-modified internal motivational state (1999, 52) (his italics).

He discusses the complexity of the morphing of the response into more complex learning over extended periods of time and how that may be used to adjust interpretive schema for transformation, potentially positioning it as an initial impetus in very complex learning process (2017). The role that the orienting response plays in perceptual consciousness is thus vital, not only in giving rise to perceptual processes and objects, but to being (*Dasein*) itself.

4.2.2 The Temporal Domain: Short- and Long-Term Memory

Different kinds of learning and retention and remembering impose different physical and cognitive demands on the learner. Short-term memories, like a sequence of unfamiliar sounds, or, in conversation, a stranger's name, may sometimes last for only a few moments or perhaps several minutes at best, and long-term memories, such as your own birthdate or the melody of a folksong, can last effortlessly for, weeks, years or even a lifetime. The duration of the memory is dependent on how transient the underlying biochemical changes are. Long-term memories involve changes in protein synthesis and gene regulation and in many cases, neuronal modifications and the growth of new synapses, whereas short-term memories do not (Byrne 1997, s7.4). Although the capacity of short-term memory can be mediated with segmentation or "chunking" and mnemonic devices, it is volatile, extremely limited in capacity (Miller 1956), requires considerable attentional awareness, and a high "refresh" rate–usually through repetition, or with pharmacological assistance. This is defined in the learning literature as imposing a *high cognitive load* (Sweller et al. 2011).

Like the acquisition of most physical skills, learning to play a musical instrument involves engaging in a range of interactive perceptual, cognitive, and proprioceptive activities in order to move note, phrases and gestures from short-term explicit memory to long-term memory using implicit "know-how" skills that were acquired over perhaps thousands of hours of practice. The initial learning phase involves a high cognitive load which diminishes as the mental representations in the short-term memory of the performer are replaced with embodied *skillful coping* (Ungureanu and Rotaru 2014). These processes are also enacted each time a performer begins to learn an unfamiliar composition (Lehmann and Ericsson 1997). They postulate that once something has been explored and a mental representation has been built that's functional, that functional representation then stands (in) for the thing itself. The thing it itself can thus be perceptually ignored, or at least perceived at a lower resolution, thus diminishing the cognitive load. This process applies equally well to abstractions such as musical structures as to physical gestures: Once we have cognitively learned to recognize various chord types (major, minor, augmented etc.) we no longer need to employ short-term, high cognitive load processes to identify them when they are presented to us. And also, at an even more abstracted schematic level: phrase-structure, rhythmic agogics and so forth.

It will be apparent from this discussion, that the bodily location of a declarative memory is not *just* determined by the type of information (numerical, spatial, verbal, gestural, abstract etc.) but by what the memorizer already knows, again schematically speaking, from their past experience and the emotional conditions under which such knowledge was acquired. Consider, for example, the memory of the timbres of various musical instruments. It would be rare for someone to confuse the timbre of a fife for that of a drum: Blowing and hitting are so elemental that the differences in spectral profiles are immediately apparent to anyone with minimal apparent cognitive effort. More knowledgeable listeners (those with a greater variety of memory schemas developed through appropriate amount of prior implicit learning) will similarly be easily able to distinguish between the timbre of a violin and a viola, between a B-flat and an E-flat clarinet, between a sarod and a sitar, a Stradivari violin and a Guarneri, a Fender Stratocaster strung for Jimi Hendrix or for Buddy Holly.

4.2.3 Implicit Aural Cognition: Gesture

Our understanding of the biological substrate of mind, and thus mentality, has been extended through the discovery of mirror neurons in the brain's various sensory cortexes, including the premotor cortex which projects directly to the spinal cord, and the amygdalae clusters located deep within the temporal lobes that perform a primary role in the processing of memory, decision-making and emotional responses, as discussed.[23] There is also a growing body of research in neuroscience that the aural conveyance of complexities such as most music, is reliant, at least to some extent, on an embodied interpretation (via performers) for effective communication with embodied listeners.[24] It was not until it was technically possible to produce musical compositions of any complexity that could be heard without the assistance of human performers, that it was possible to meaningfully speculate on the extent to which a listener's perception of the structural or emotional characteristics of a piece of music are dependent on, not just the notated abstract pitch-time structures, but the sound-encoded gestures of performers bodily realizing them.

> For many centuries, people learned to listen to sounds that had a strict relation to the bodies that produced them. Suddenly, all this listening experience accumulated during the long process of musical evolution was transformed by the appearance of electronic and recorded sounds. When one listens to artificially generated sounds he or she cannot be aware of the

[23] This research and its implications are discussed in more detail in Chapt. 3: Sects. 3.5 and 3.6.
[24] A summary of that work can be found in (Worrall 2010).

> same type of concrete and mechanic relations provided by traditional acoustic instruments since these artificial sounds are generated by processes that are invisible to our perception. These new sounds are extremely rich, but at the same time they are ambiguous for they do not maintain any definite connection with bodies or gestures (Iazzetta 2000).

Much research to do with physical gestures in the performance of music has concentrated on understanding the role of extra-notational aspects of music, particularly on emotional expression and affect. Godøy's and other's research also suggests that there are gestural components in the mental recoding of musical sounds (Godøy and Leman 2010; Godøy et al. 2012). In extending Schaeffer's idea of the sonorous object to the *gestural-sonorous object*, he found considerable evidence to support the hypothesis that when we listen or even just imagine music, we trace features of what we hear with our fingers, hands, arms and so on.

> This means that from continuous listening and continuous sound-tracing, we actually recode musical sound into multimodal gestural-sonorous images based on biomechanical constraints (what we imagine our bodies can do), hence into images that also have visual (kinematic) and motor (effort, proprioceptive, etc.) components (Godøy 2006, 149).

The association of body movement with music appears to be universal and independent of levels of musical training, and in gestural 'sound tracing' studies there seemed to be a significant agreement in the spontaneous drawings that people with different levels of musical training made to musical excerpts (Godøy 2010). Musical instrument designers have taken up the call for a more embodied computer music; for better interactive tools for performing with or on computers, and such gestural controllers have found applications in sonification, such as by providing means to interact with data-derived resonator models in model-based sonifications (Hermann and Ritter 2005). Despite their being a very large corpus of research on gesture, this corpus needs to be carefully examined and adapted in order to be used in sonification domain, particularly for non-interactive sonification. Micro-gestural inflections are regarded as a basis for musical expressiveness (Kima et al. 2010) which some studies reveal are aurally 'available' to listeners, even if only subconsciously (Fyk 1997). This appears to be the case not just for agogics (notes *inégales* and other micro-timings) but also for *elksis* (pitch interval flexibility when tones are intentionally micro-sharpened or micro-flattened, depending on their direction in a line.

Concentrating on the sound synthesis of responsively complex resonators, by, for example, constructing sophisticated physical models in software, is only part of the equation. While most research on the use of human gestures in computer music and sonification production have concentrated on interactive control interfaces that employ gross body gestures such as arm waving, professional performers know that these gross actions are used in order to control much finer gestures: elbow elevation for finger legato on the piano; wrist extensions for idiophones and for the fine continuous control of bow strokes, for example.[25] Humans have been interacting

[25]This is not to discount performative gesticulations, such as for dramatic visual effect.

with sound-generation devices through a variety of interface technologies for a long time. These devices transmit performer-controlled energy through resonators and attentive listeners are often able to infer the detailed and nuanced actions of performer through the sound alone. Research on the development of parametric sound controllers provides evidence that, with multi-parametric interfaces, people performed better on more complex tasks than with single parameter controllers, perhaps because they allowed people to think gesturally (Hunt and Wanderley 2002). What is needed is similarly complex models of activation that integrate subtle gestures that evolve the shaping of sonic objects and processes in accord with their function in the expression of informational structures in a data sonification, and research towards this goal is still in its infancy.

4.2.4 Modes of Listening

Listening modes are distinctive approaches or intentions that can be brought to bear on a listening experience. Since they explicate an understanding of how meanings can be conveyed in effective design, taxonomies of listening modes have been considered as useful tools in the field of sound design (Gaver 1989). A review led by Kai Tuuri (2007) reveals several other approaches to such a taxonomy and a recent update suggested that there are nine modes, each of which constitutes a different meaning-creation, as summarized in Table 4.3, based on Tuuri and Eerola (2012). This list is not the only classification possible. David Huron, for example, listed twenty-one styles and strategies for listening to music: Distracted, Tangential, Metaphysical, Signal, Sing-along, Programmatic, Allusive, Reminiscent, Identity, Retentive, Fault, Feature, Innovation, Memory scan, Directed, Distance, Ecstatic, Emotional, Kinesthetic and Performance listening (Huron 2002).

Whilst these taxonomies detail ways human listeners can attend to sounds, listening is not a sequence of discrete modes and the continuum of experience dictates that some of these modes can be employed simultaneously for stimuli to which different kinds of meanings are attached. Furthermore, because habituation de-privileges some sounds in some circumstances, (for example, after an initial awareness, we cease to be aware of the sound of air-conditioning until it ceases), two considerations that are vital to the success of a sonification, especially those listened to over long periods of time are (a) the careful control of the attentional aspects of the sounds chosen to contribute to the sonic environment: to what extent they can be attentionally isolated when necessary, yet pose a minimal cognitive load elsewise, and (b) the auspicious–perhaps sparse–use of sounds that have high cognitive loads only when absolutely necessary, and then for limited periods of time.

Table 4.3 Descriptions of nine listening modes

Name	Description	Example
Reflexive	A quickly evoked, innate action–sound reaction affordance based on an automated (or 'hard-wired') schema due to the evolutionary adaptation to our ecology	Stepping off a roadway when hearing the blast of a car horn.
Kinaesthetic	A gestural sense of motor-movement, based on processes that manifest innate or early developed schemata concerning bodily movements, coordination and postures. Musical listening contains multiple levels of such imitations	It appears likely that related to the experience of body movement. e.g. swaying, foot tapping, air guitar "performance".
Connotative	Active projections of action-relevant values as resonances of conscious or learned schemata based on natural and/or cultural constraints	Attention the pitch of the "woosh" that represents a file's size when it is trashed. (Auditory Icons).
Causal	The intention to recognize (features of) the source of sounds	Listening to the snap of twig to determine how dry it is.
Empathetic	Perceiving and signifying affective states that could signal someone's emotions and intentions inferred empathetically from body gestures	Listening for the emotion being expressed when someone slams a door.
Functional	Context-oriented listening focused on the purpose of sounds	Is the ambulance approaching or receding?.
Semantic	The intention to recognize sounds as signs that stand for something due to socio-culturally shaped and learned codes	Earcons. The pre-message alert signal from an emergency broadcast system.
Reduced	The intention to divorce the phenomena of sounds from their everyday contextual meanings so as to attend to the qualities of sounds themselves (after Schaeffer)	Listening to the attack time or decay profile of a sound envelope.
Critical	Reflective self-monitoring in order to verify the appropriateness or authenticity of responses given the context	To questions such as "Am I too loud?" or "Is it appropriate to whisper here"?.

4.2.5 Attention and Perceptual Gestalt Principles

The historical background and context in which Gestalt theory arose is outlined in Chap. 3. While there is no definitive list, and principles such as the figure/ground grouping is not usually referred to as a Gestalt principle (Wertheimer 1938, 1958; Metzger 2006), some of those most commonly discussed are summarized in Table 4.4. The classical assumption is that *Gestalt principles* are integral to the

Table 4.4 Perceptual *Gestalten* in audition with succinct examples

<table>
<tr><th>Principle</th><th>Description</th><th colspan="2">Examples and principles</th></tr>
<tr><td rowspan="7">Figure-Ground articulation</td><td rowspan="7">Separation of the aural field into foreground and background perceived as articulated into two components. Figure has an object-like appearance, the background having less perceptual saliency. Often stratified in depth and influenced by attention mechanisms such as perceptual focus and information reduction due to repetition.</td><td>Foreground</td><td>Background</td></tr>
<tr><td>Melody</td><td>Accompaniment</td></tr>
<tr><td>Louder</td><td>Softer</td></tr>
<tr><td>Brighter tones</td><td>Duller-tones (LP filtered)</td></tr>
<tr><td>Direct signals</td><td>Echoes</td></tr>
<tr><td>More intricate</td><td>Less intricate</td></tr>
<tr><td colspan="2">Background noise punctuated by a brief loud sound.</td></tr>
<tr><td>Proximity</td><td>Elements that are near to each other (relative to other elements) are perceptually grouped. Some percepts may be only partially present.</td><td colspan="2">Sequences of tones group into pitch or time sub-group with near (pitch/time) neighbors, for e.g.
******
=> ** ** ** => *** ***
Sounds emanating from the same or close by are considered to be related.</td></tr>
<tr><td>Similarity</td><td>Elements that appear similar are grouped. Increasing the distance between element groupings strengthens the salience of the organization.</td><td colspan="2">Instruments with similar sounds are groups. Eg. All clarinet sounds are identified as clarinets. Near frequencies are perceived as being at the same pitch. Ditto for slight variations in onset times. e.g. **oo**o**oo sub-wholes to groups of *'s and o's.</td></tr>
<tr><td>Continuity</td><td>There is a perceptual bias to perceiving continuous forms rather than disconnected segments. Oriented units or groups tend to be integrated into perceptual wholes if they are aligned with each other.</td><td colspan="2">Lengthy continuous tones interrupted by an abrupt event appear to continue during the abrupt event if it appears after the abrupt events ends.
The balance between Continuity and Proximity in the formation of salient sub-wholes may be shifted by varying similarity, which can be accomplished, for example, by orchestrating independent melodic lines differently.</td></tr>
<tr><td>Closure</td><td>The perceptual system attempts to close disparate figures into single elements. Elements tend to be grouped together if they are parts of a closed figure. Elements can be camouflaged or shift the wholes into which they are integrated, based on symmetry simplification.</td><td colspan="2">Separate harmonic frequencies perceptually form into a tone with timbre; simultaneous tones into chord complexes; tone sequences into melodies. Camouflage produces hidden objects without occlusion. Instead, they are perceptually subdivided and repartitioned.</td></tr>
</table>

(continued)

Table 4.4 (continued)

Principle	Description	Examples and principles
Symmetry	Time between temporally distributed tones are perceived with as simple rhythmic relations as possible.	Duple and triple temporally distributed tones (lines/ melodies) are perceived with as simple rhythmic relations as possible.
Past experience	Elements tend to be grouped if they were often together in the past experience of the observer.	Leading tone => tonic expectation; cadential formulae.
Convexity principle	Convex rather than concave patterns will tend to be perceived as figures.	Scales, chords, harmonic reductions and octave equivalencies, if represented in a tone space. (Honingh and Bod 2011)
Common fate	Elements tend to be perceptually grouped together if they move together.	Objects that move together are perceived as grouped or related. (Parallel lines in a harmony reduce to a single line + timbre).

perceptual system. However, some research has indicated that the assignment to figure and ground, for example, is often ambiguous depending on the environmental exposure of particular listeners to various stimuli (Peterson and Skow-Grant 2003). When more than one (*Gestalten*) favor the same grouping in a percept (perceptual scene), that grouping will tend to be strengthened (Palmer 1999). When they disagree, however, one principle will usually one dominate, or the organization of the percept is unclear. The theoretical problem of how to predict which principle will dominate and in which circumstances is being addressed,[26] but to-date remains unresolved: Whether, as generally favored by the Gestaltists, the principles are fundamental properties of the perceptual system, or are heuristically derived from a perceiving being's experience with the properties of things in the world (Rock 1975). The two-truths dichotomy has resonances of the mid/body problem, harkens back to that over Platonic Ideals and continues to engage perceptual psychologists and realist philosophers. Objects, (including sonic objects) in the world are usually located 'in front of' or enhanced against some background (figure-ground articulation), have an overall texture different from the texture of the background (similarity), consist of parts which are near each other (proximity), move as a whole (common fate), and have closed contours (closure) which are continuous (continuity).

Whether it is attention that creates forms, or that existing forms, organized in accord with *Gestalten* draw attention, organizing sonic objects in an auditory field with due attention to the perceptual effects of their interaction with respect to the principles of proximity, continuity and closure etc. has the potential to enhance the intelligibility of data sonifications.

[26]See, for example, Kubovy and Valkenber (2001).

4.3 Analogies, Conceptual Metaphors and Blending

In the elaborately elegant and approximately effortless means by which we perceive, conceive, store and recall our experiences, as broadly outlined earlier in this chapter, short-term working memory plays and important role in refining and elaborating our knowledge of the world and in conscious reflections of it, but it is not the primary or essential means by which we unconsciously orient ourselves from moment-to-moment, for every moment of our lives. Over the millions of years of our evolution, Nature has found more efficient means, which it has successfully iterated across species and across time that rely on long-term memories, which are not collections of snapshots of the objective world but dynamic schemata, encoded with values and meanings: propositional or emotionally valenced truths. This understanding varies considerably from the earlier value-free approaches of cognitive science and, as George Lakoff and Mark Johnson have argued, the embodied nature of meaning-making, that is, the process by which a person assigns meaning to their experience, makes scientific knowledge and research possible. Further, an approach to scientific research that recognizes that the embodied nature of cognition, firmly rooted in metaphorical thinking, is more empirically rigorous than the disembodied alternative (Lakoff and Johnson 1999). There is now strong empirical evidence that the same neural circuitry used to run our bodies physically, including the perceptual *Gestalten* discussed earlier, also structures our reasoning processes about all events and actions, both physical and abstract (Lakoff 2012).

4.3.1 Imagination and Making Sense of the Unknown

Imagination is the faculty for creating mental images and constructs of a distinctive nature.[27] To begin to cope with what is not understood, the human imagination searches for sensations, impressions, memories and feelings, with which to "fill" itself, supplemented and supported by the biological constructs discussed earlier in this chapter. This appears to be a structure similar to that which predisposes human beings to language: When infants begin to speak, they babble the set of all physiologically possible phonemes, rapidly reducing these utterances to a subset of only those phonemes in the linguistic speech they hear around them, even to the point of eventually not being able to distinguish the difference between many of the "discarded" phonemes from those they have retained. This process of *neural*

[27]Imagination is related to but distinct from imager, aural imagination, for example.

commitment to the native language is so fundamental, it appears to be built into the biology (Kuhl et al. 2008). More generally, perhaps as a pattern of adaptive behavior, humans manifest the *potential* to possess all possible alternatives, but retain only those forms and related ideations to which they're exposed in their culture.

There are numerous ways of perceiving and conceiving events and objects. We need a means by which to identify known objects and anticipated occurrences as quickly and coherently as possible, while peripherally identifying unusual (low-redundancy) structures, so as recognize them as something new and potentially interesting, or to dismiss them as mirages or misperceptions. This process occurs in listening to a sonification as much as it does to listening to linguistic speech, looking at a visual image, or exploring a physical environment. When events and objects appear to us indistinctly, our imaginations engage with their appearances, drawing them to us, so to speak, in an attempt to get a *maximal grip* on them; adjusting our relation to them, by physically and/or attentionally moving towards them in what Merleau-Ponty, following Brentano and Husserl, called an *intentional arc*.[28] As experiences and memories in the world accumulate in the neurobiological substrates of the imagination, the ideas and knowledge associated with them become available in interpretive frameworks, or *schema*, for imaginary thought in various forms: symbols, maps and metaphorical models, as discussed below.

Just because there are many potential perceptual objects or events (information in a dataset) at any juncture, does not mean that any one of them is as propositionally valid as any other. The strength of a propositional truth is dependent upon myrmecologist Edward Wilson as *consilience* (Wilson 1999) and so can be tested for using a Multitrait-Multimethod Matrix (MTMM) approach for assessing the construct validity of a set of measures (Campbell and Fiske 1959). This process might include a variety of scaling functions and metaphorical or other analogicalmappings to determine value structures and perceptual directions that enable listeners to get a *maximal grip* on the sonified information.

4.3.2 Analogies and Metaphors

Analogies compare things in order to understand an *explicit* relationship between them. An *analogy* is an explicit mapping between elements of a source domain (sometimes referred to as the base domain) and elements of a target domain; of establishing a structural alignment between two represented situations and then projecting inferences: one-to-one correspondence between the mapped elements in

[28]Merleau-Ponty describes the *intentional arc* as the way our successful coping continually enriches the way things in the world reveal themselves to us ([1945] 1962, 137).

the source and target (Bowdle et al. 2001). In data sonification, explicit mappings between, say the height variable in a dataset and pitch, and between thickness and loudness can be considered analogical mappings between data parameters in source domain and psychoacoustic parameters in the target domain.

A *metaphor* is a rhetorical device used to directly, or *implicitly*, represent one thing in terms of something else; a figure of linguistic speech in which one thing is compared to another by being spoken of as though it were that thing. There are two principle kinds of metaphor: literary and cognitive. A *literary metaphor* is a poetic or rhetorical turn of phrase that adds character and color to otherwise pedestrian language. Example: A SEA OF TROUBLES.[29] *Conceptual metaphor*, sometimes called *cognitive metaphor*, is a generic term used in cognitive science to denote the expression of one domain of action or knowledge in terms of another. Cognitive metaphors are thus projections of conceptual structures that occur in cognition as a way of comprehending certain abstract *experiences* in terms of other more concrete experiences. For example, a metaphorical sonification mapping may represent the temperature of a gas with a cloud of sonic grains, whose composition and entropy consists of a complexity of parametric determinates that are individually not relevant.

4.3.3 Literary Metaphors as Classical Rhetorical Devices

Classical rhetoric was considered an essential discipline in the training of educated Europeans for more than twenty centuries, until its eventual demise under literary criticism of it as just the study of flowery language, rather than Aristotle's description as "the faculty of observing in any given case the available means of persuasion" (Kahn 1985). De Oliveria traces its history in European education and observes that the eventual expulsion of the traditional study from curricula coincided with the advent of scientific positivism, under the sustained influence of Auguste Comte in the early nineteenth century (de Oliveira 2014).[30] Metaphor was one of a number of rhetorical devices designed to stimulate readers to engage in a process of practical judgement, "a pleasurable activity, exercise, or praxis, which educates us in the very act of reading at the same time that it moves us to the

[29]The expression "pedestrian language" is itself a flourish, a "flowery" way of saying "ordinary language, as heard when walking on the street." By convention, metaphors are capitalized in the cognitive science literature.

[30]Positivism's principle tenet was that information derived from sensory experience that is interpreted through reason and logic (that is empirical evidence), is the exclusive source of all ("positively") certain knowledge. Comte's most controversial claim was that, not only does the natural world operate according to this tenet, but so also does society (Macionis 2012). Positivism is part of a more general ancient dispute between philosophy and poetry, laid out by Plato and enunciated in different forms over the intervening period, including as a quarrel between the sciences and the humanities.

application of prudence in human affairs" (Kahn 1985, 40).[31] The widespread use of literary metaphor exemplifies the widely held understanding that linguistic speech and meaning are distinct: Saying what we mean is not the same as meaning what we say. While being distinct, they are related, because 'saying' and 'meaning' inform each other, and according each their due can only occur if we consider them mutually inter-dependent.[32] The use of rhetorical metaphors to develop musical forms–though such devices as word painting–has played a significant role in the development of classical musics, as discussed in Chap. 1 Sect. 1.5, which enabled composers to express even contradictory emotional or mental states simultaneously.

4.3.4 Metaphors and Propositional Thinking

Encouraged by advances in understanding the ubiquity of metaphor in thinking made by George Lakoff and others in the 1980s, researchers began to study how cognitive linguistics worked in the brain, that is, a neural theory of language (Regier 1996) followed by a considerably developed neural theory of metaphors (Lakoff and Johnson 1999). In their work on the theory of *conceptual metaphors* George Lakoff and Mark Johnson postulated that metaphors are so common in the language of our everyday lives that mappings between conceptual domains correspond to neural mappings in the brain (1999, 569–83). Empirical studies–discussed in (Lakoff 2012)–have shed some light on the neural underpinnings of these cognitive faculties. They show that embodied schemata, conceptual metaphors and conceptual blending recruit neural networks in the human brain and sensorimotor system which are associated with bodily perception and action to perform a sensorimotor mimesis of the patterns of neural activity associated with gesture, perception and proprioception within the nervous system.

In the earlier discussion of functional truth (Sect. 4.1.3) we came to understand that there is simply too many events and things in the world for us to perceive them all, so events and objects manifest themselves as things that we can use. This is also the case when thinking. There are an extremely large number of possible relationships between things and actions and the process of connecting thoughts as they arise is more than just a process of stringing together a sequence of separate independent ideas and impressions, as John Dewey so eloquently expressed:

---

[31]As traditionally, 'prudence' here means "caution"; not to be confused with 'prudish', which today has negative connotations.

[32]This is known in natural language scholarship as *under-determinacy*, which occurs not just in the metaphorical A-B relations but within a metaphor itself. For example, the expression "it is all downhill from here" could mean two opposite things depending upon context. In general, it is a feature of metaphors that they exhibit noticeable *context-sensitivity*. That is, they are *as plain as the nose on one's face,* or *stick out like a sore thumb.*

> Thinking goes on in trains of ideas, but the ideas form a train only because they are much more than what an analytic psychology calls ideas. They are phases, emotionally and practically distinguished, of a developing underlying quality; they are its moving variations, … subtle shadings of a pervading and developing hue.…The experience, like that of watching a storm reach its height and gradually subside, is one of continuous movement of subject-matters. Like the ocean in the storm, there are a series of waves; suggestions reaching out and being broken in a clash, or being carried onwards by a cooperative wave. If a conclusion is reached, it is that of a movement of anticipation and cumulation, one that finally comes to completion. A "conclusion" is no separate and independent thing; it is the consummation of a movement ([1934] 1952, 39).

Cogent focused thinking is difficult to sustain; it requires all the assistance it can muster, including the power of our language structures and articulate speech. This is a compelling reason why the freedom to speak and mis-speak is necessary to clarifying thinking. Metaphorical concepts are thus a way of restricting the characteristics of events and objects about which we are attempting to communicate, to only that subset of all possible relationships that are relevant to constructing a coherent, in context, thought sequence. A single example should suffice: The metaphor MUSIC IS THE FOOD OF LOVE contains many ideas and relationships within it: that (one of the properties of) music is that it sustains; that (one of the things) love needs in order to be sustained is nourishment; that, like food, the quality of love is dependent on the characteristics of the nourishment, and so on.

4.3.5 Conceptual Metaphors as Frameworks for Thinking

The theory of generative grammar holds that words are syntactically combined in a special purpose language 'module' in the brain through a set of 'hard-wired' inferential, truth-preserving rules. Chomsky maintained that the principles of universal grammar ensure concepts and thoughts consist of manipulating sentence-like discursive mental representations in accord with these rules (Chomsky 1972). This abstract, formalist approach to understanding the structure of language inherits Ferdinand de Saussure's proposition that the relationship between the meaning and the form of a sign is arbitrary (Burbank and Steiner 1997, 18) is very much in alignment with the materialist position outlined earlier, and is mirrored in the active development of aurally complex abstract serial musical techniques over the same period. Both these developments influenced the distinctly connectionist approach to computer music beginning in the 1960s (Todd and Loy 1991), as discussed in Chaps. 1 and 3. In data sonification, an emphasis on, or the predominant use of, earcons and parametric sonification is, in general, thus somewhat in alignment with positivist thinking.

In the late 1970s and '80s a new theory of language emerged that focused on how concepts are organized *through* language. It postulated that words do not just play a functional role in syntactic structures, but are meaningfully combined in strategies for conceiving concrete situations. This cognitive linguistic approach asserts that thought is a matter of manipulating unconscious mental imagery in which concretely pictured physical objects and processes take the place of more abstract objects and situations. In doing so, we redeploy familiar, easy patterns of thinking about one familiar sort of thing to direct sensorimotor representation in a second, more elusive sort of thing. In this way, verbal and non-verbal signs such as gestures take on meanings which depend on the occasion of their use (Croft and Cruse 2004).

To the cognitive linguist, the essence of metaphors is not just their (analogical) representational function, but, derived from their thoroughly embodied nature, the forceful directness with which they express one kind of thing in terms of another. By externalizing tacit knowledge, metaphors and analogies have a prominent part to play in the process of turning ideas are into practical reality (Narayanan 1997). Lakoff and Johnson also argued that our use of metaphorical concepts is not restricted to language, but pervades the perceptions, thoughts and actions that govern our most mundane connections with, and reasonings about, the world (2003, 5). The term *cognitive metaphor* or *conceptual metaphor* (hereafter) refers to the understanding of idea and relationships between ideas in one conceptual domain, or *frame*, in terms of idea and relationships in another. The structure of this relationship is known as a cross-domain mapping *or framework.* In such a framework, the mapping of an organized system of concepts from an abstract, conceptual, domain frame to another more concrete, target domain frame is called a *schema* (plural *schemata*).

4.3.6 Mental Spaces and Conceptual Blending

A *mental space* is a theoretical construct consisting of an arrangement of discrete concepts or images; a partial locational and temporal mental aggregation that represents some recurring, and perhaps familiar types of situations. It is constructed as we think and talk for the purpose of understanding and for action. It has been hypothesized that, at the neural level, mental spaces are sets of activated neuronal groupings and that the connections between elements correspond to activated neural assemblies and that linking between elements corresponds to neurobiological bindings, called coactivation-bindings (Petersen and Sporns 2015). From this viewpoint, mental spaces operate in working memory but are built up partly by activating structures available from long-term memory. Mental spaces represent ways things can be *conceived* to be. Gilles Fauconnier used the mapping of cognitive spaces as an epistemological organizer of mental space. He accomplishes this by establishing four mental spaces: the source domain, the target domain, the blended space and the generic space. The generic space is created from two input spaces–a source domain space and a target domain space, and becomes the foundation for drawing inferences and elaborating ideas in the form of metaphorical relationships between elemental structures in that space (1985).

For example, in considering the metaphor THE SURGEON IS A BUTCHER (Fauconnier and Turner 1996, 144), the source domain (knowledge of surgeons and their practices and tools) and the target domain (knowledge about butchers and their practices and tools) are connected to a space that contains generic information about the roles of surgeons and butchers. This generic space is more abstract and schematic than the two input spaces. When the source domain and the target domain are brought together in to a blended space, they create a new space in which the surgeon's performance in an operating theatre on a living human is imagined to resemble a butcher's performance in a butchery on a carcass. The contrast between generic space and the blended space is emphasized by understanding that if the source and target domains were reversed (THE BUTCHER IS A SURGEON), the generic space would essentially be the same but the blended space would be completely different. Conceptual blending is a basic mental operation that leads to new meaning, global insight, and conceptual compressions useful for memory and manipulation of otherwise diffuse ranges of meaning. It plays a fundamental role in the construction of meaning in everyday life. Fauconnier and Turner suggest that the capacity for complex conceptual blending is the crucial capacity needed for thought and language.

4.3.7 Metaphors for Sonification: INFORMATION IS SOUND

Following the advances made in conceptual blending in linguistics, research into the application of music theory and analysis, spearheaded by Lawrence Zbikowski, has uncovered some interesting challenges related to the different grammars of spoken languages and music. In particular that, because spoken language and music have different cultural functions, they are supported by different forms of reference. For example, unlike non-linguistic sounds, linguistic utterances enable a speaker to direct the attention of another person to specific objects or events within a shared frame of reference. At the same time, musical structures make it possible to represent dynamic phenomena, ranging from inner psychological processes of individuals and groups, to guiding the steps of a dance and trajectory of bodies through space. Zbikowski points to forms such as *shuochang* (and by extension *Sprechstimme* and rap)[33] as providing fertile territory for future exploration (Zbikowski 2012). Another dissimilarity between linguistic speech and sound perception, which is built into the logic of musical theories, is the absence of a logical complement or negation operator; difference without negation: one does not, except perhaps through locally-restricted conceptual inferencing, hear a not-major chord or orchestrate with a 'not-clarinet'.

[33]Discussed in more detail in Chap. 1, Sect. 1.5.

Table 4.5 Space–time metaphors for moving objects and observers

Space (source domain)		Time (target domain)	
Moving objects	Moving observer	Moving objects	Moving observer
Objects	Path locations	Time values	
Object motion	Observer motion	The passage of time	
	Distance moved		Time passed: amount
Location of observer		The present	
Space in front of the observer		The future	
Space behind the observer		The past	

There is some empirical support for the assertion, following Conceptual Metaphor Theory, that we think about abstract domains, like time, in terms of relatively more concrete domains, like space but not *vice versa* (Kranjec et al. 2019), and this is corroborated by evidence that even in the technical language of the time-based arts time is frequently conceived and referred to metaphorically using the language of space. Although embodied cognition has been addressed by a variety of researchers in electroacoustic music, research into the comprehensive application of conceptual metaphors and blending to embodied techniques for data sonification, is in its infancy. In discussing his Temporo-spatial Motion Framework, Stephen Roddy has laid out a foundation. He discusses, and empirically tests, the effectiveness of several conceptual metaphors, including: MOVING TIME, MOVING OBSERVER, MOVING SONIFICATION, SONIFICATION LANDSCAPE, SONIFICATION AS A MOVING FORCE and TIME-SERIES DATA AS PHYSICAL MOTION (2015; Roddy and Bridges 2019). By way of illustration, Table 4.5, summarized from Roddy (ibid.) contrasts two of these: MOVING TIME and MOVING OBSERVER for representing time in terms of space as exemplified in the expression "Now that yesterday is behind us, we can move forward into a brighter future."[34]

Several examples of metaphorical mapping are discussed in Chaps. 7, 8 and 9, including an interesting blend of psychoacoustically contradictory, but conceptually coherent metaphors in Sect. 8.6.

4.4 Towards a Design Methodology

In our increasingly datarized lives, as the amount of data increases and alternative approaches to perceptualizing them are contemplated, sonification is becoming understood as a viable means of doing so, and excellently, especially, but not

[34]"Brighter" here is also clearly metaphorical; a term and also used to represent sounds that have broader spectral distribution of energy as experienced, due to dorsal-ventral (front-back) asymmetry of our hearing: Sounds in front of us appear brighter than those behind.

exclusively, with multivariate time-series data. A common method employed by a naïve sonifyer is to somewhat unsystematically find some sound-making software, which can, hopefully, be easily configured to read their dataset and magically make pleasant, informative sound of it. Some are easily mistaken, others are overwhelmed and make an early decision, to 'outsource' the problem to an experienced sound designer who, in a seemingly *ad hoc* manner, heuristically, produces a bespoke sonification that addresses as many of the implicit, explicit, and intuited criteria as possible. This solves a short-term problem but is unsatisfactory for non-specialists who prefer the empowerment of engaging in the quest for intelligible sonification design solutions themselves.

One of the motivations for this current work is to try understand some of the perceptual and conceptual correlates of intelligible data sonification and encapsulate the knowledge-bases that underpin them in software design. For example, the psychoacoustic, gestural, and psycho-physiological substrates such as cognitive-load sensitivity and emotional valence, with low-level latent functions that can be compiled into higher-level interactive modelling tools. Design is an inherently 'messy' and iterative activity that, while a process, may never be entirely procedural. So, the purpose of such software is not to trivialize the skills of experienced designer, but to hierarchize the masking of many of the functional decisions that need to be made in designing process; assimilations such as equal-loudness contouring[35] and modal convex pitch and time transforms. It is a hope that in doing so, novice designers might consider more adventurous possibilities, and experienced designers will be enabled to implement complex procedures more flexibly to test multiple approaches to sonifying a dataset in order to strengthen the comparative analyses of such sonifications for their effectiveness as design solutions.

4.4.1 The Auditory Environment of a Sonification

Edmund Husserl characterized an environment as an *Umwelt* ("lifeworld" or "surrounding world"); a world of entities that are meaningful to us personally because they exert "motivating" forces on us. Earlier in the chapter, we observed that object perception does not occur in abstracted isolation but in places of enfolded meanings; in *environments* where perceivers and object interactions happen. In realistic sonic environments there are almost always multiple sound-producing sources and sound-transforming reactions and their resultants.[36]

[35] As defined by the International Standards Organization ISO 226.

[36] The use of 'environment' rather than 'auditory scene' is deliberate. Because, in common parlance, a scene is in front of us; something we look at; to which we are paying attention, it seems unnecessarily obtuse or inappropriate to apply the term to the synthesis of an omnidirectional soundfield. The term "soundfield" tends to be used as a substitute for auditory space, i.e. the physically dimensioned space in which sounds or sound-streams can be placed, using appropriate

So it is, that sounds are not perceived in abstraction. To a listener in an environment, sound carries a great deal of information that enables them to orient towards events, and to move in accordance with them. For example, when I hear the sound of a siren as a police emergency, I recognize it as a warning that a speeding vehicle to which it is attached is approaching and which should not be needlessly impeded in its passage along the roadway. Thus, the sound has value for me by encouraging me to locate myself appropriately. Others also comprehend the sound of the siren in the same way, and so, in a social context, it acquires an intersubjective use-value; appreciated as serving a specific public purpose. In designing such a siren sound, while account is clearly to be taken of the psychophysical properties of the sound itself (loudness, timbre, glissando, pitch etc.), it does not become a siren to respond to, nor can it be evaluated as such, until *placed* in its environmental contexts. Such considerations are also pertinent in the design of sonifications for aquatic or outdoor use, such as a training aid for Olympic rowers (Barrass, Schaffert, and Barrass 2010).

Although the acoustic input to the human auditory system is a one-dimensional temporally varying air pressure waveform, we can perceive auditory scenes involving multiple sources of human speech, vocal and instrumental music, animal sounds and other natural and industrial noises, occasionally all occurring at the same time, each with their own structures and recognized transformations (Kubovy and Van Valkenburg 2001). In his ground-breaking analysis, Albert Bregman (1994) provides a comprehensive description of analytic and synthetic listening in terms of auditory stream integration and segmentation. While any systematic sonification meta-design schema needs to accommodate the knowledge summarized in this work, there is yet to be written such a generalized exposition appropriate for many sonification tasks. In particular, significant work is needed to be able to *synthesize* virtual sonic environments that exhibit perceptual cohesion, while maintaining aurally differentiable virtual sound sources. It remains a task of *soniculation* research to develop robust models of listener's perceptual organization that can be reliably reverse-engineered to produce affordances that solicit listeners to behave in ways that assist them to get enough 'grip' on these sound structures for them to be perceived as cohesive, evolving auditory objects in relation to each other. The literature is growing but still relatively sparse. Significant work includes that by Zwicker and Fastl (2014), Neuhoff and Heller (2005), and by Ferguson, with others, towards the use of data-to-sound mappings that are aligned with the listener's perception of the auditory connotations of the data in a controlled computational way (2006, 2017) (Cabrera et al. 2007).

("soundfield" technologies) such as loudspeaker arrays. The term 'soundscape' has historical connotations which associate it with acoustic ecology. According to The International Organization for Standardization, there is a diversity of opinions about its definition and aims and the use of the term has become idiosyncratic and ambiguous (ISO 2014). So, given the distinction made in the text between 'space' and 'place', 'environment' seems the least undesirable term in this context.

4.4.1.1 An Example: Sonorealism

A number of models for virtual sonification environments are possible, including every-day listening, hyper-real, abstractly musical, reduced-listening and abstract process-oriented procedures. These models are not mutually exclusive–it being possible to conceive of hybrids of these and of others. Rather than provide a theoretical model for each of them, we concentrate on describing one type which we call *Sonorealism*: a style of sound design that uses natural sounds, which we extend to physical models of such sounding objects, that are transformed using Foley-like (virtual) processes. Sonorealism is understood to encompass a range of techniques from the recording of acoustic sounds (including musical instruments) and their context-sensitive transformation on ‘triggering’, to physical models (Farnell 2010). An important feature of the style as it relates to sonification, is the way it uses Gestalt and other physical affordances to characterize the behavior of sound objects. Not only does a violin have a sound box (or body) that amplifies and modifies the interaction of moving strings with the body of the resonator to which they are attached, but the sound of that resonating system occurs in another ‘space’ with its own acoustic characteristics. A third space is engaged if this coupling of string-to-resonator-to room is recorded and reproduced (played back) in another acoustic environment. Each of these components of the heard result constitute a resonance region of an environment, and the human auditory system is frequently capable of ‘hearing out’ these differences: Discrete resonators placed in different locations of a space resonate that space according to the acoustic properties of that space. When resonators with different spectral characteristics are activated nearby each other, the space can be heard to have a character, that is, to be recognizable as a *place* having different soundfields, each with identifiable regions.

4.4.1.2 Auditory Events: Auditory Icons

Hearing is qualitatively different from vision in that we see objects and infer events, but hear events and infer objects. *Auditory Icons* are a popular way of displaying discrete information *through* sound.[37] They depend on a listener’s ability to effortlessly recognize the qualities of different types of sounds, and to learn and remember the meaning of specific sounds in context. Relying on a listener’s real-life experiences, simple transformations of certain types of sounds can be easily, often intuitively, understood to correspond to different informational values or states. For example, if a ‘scrunching’ paper sound is used to represent the deletion of a computer file, it is almost effortless to infer that a relatively short, high-frequency ‘scrunch’ might represent a small file, and the deletion of proportionally larger files by lower-frequency, durationally longer ‘scrunches’.

[37]Different uses of sound for the purpose of data sonification are introduced in Chap. 2.

William Gaver described the perception of events from the sounds they make as *everyday listening* (1988, 3). A remarkable feature of this perception is the propensity listeners have to identify and interpolate quite intimate details of the characteristics of such activity once a realistic description or a strongly metaphorical representation is provided. This occurs even when there is a weak cognitive binding between a sound and the particular physical object producing it (Ballas 1996). However, once the source of a particular sound is identified as that of a dish being hit with a spoon, for example, the sound transformations directly covey data from which the listener can infer information, such as the type of material (wood, metal etc.), the size and perhaps shape of the dish, the force of the impact, and whether the dish is empty or not. The embodied knowledge on which this is based is integrated into our biology and the long-term memory of our everyday experiences of the world. If the listener has been forewarned, or has a good aural memory, they might also be able to identify the mass and material composition of the actuator's contact (the spoon) and even some sense of whether the event occurred in the outdoors, or the size or echoic properties of the room in which the sound was made. If the strike was repeated, with a change in the content of the pan, most listeners would be able to recognize whether the sound was the same or even be able to verbalize in what way it was different: If there was a change in pitch, whether it was higher or lower etc. All this knowledge is easily available to an everyday listener. This incredibly detailed, easily accessible knowledge that required minimal effort and no specialist training, accounts, in part, for the delight most people gain from the sonic nuances of expertly played musical instruments: sound-making devices that were constructed so as to be acoustically responsive to minutely varying actuations by performers trained in the art.

4.4.1.3 Symbolic Events: Earcons

Earcons are a common way of displaying discrete information *with* sound. They are small motivic structures generated by controlling synthesized tones, usually through basic psychophysical parameters–pitch, loudness, duration and timbre–into structured, non-verbal "message" combinations. One of the most powerful features of earcons is that they can be used hierarchically to represent events or objects and combined to produce compound messages. In addition, a musically-trained listener might be able to identify symbolic manipulations of the sound elements such as the interval size of a change in pitch. An advantage of earcons is that these symbolic manipulations lend themselves to structural modelling and thus are extensible into classical parameter-based data sonification.

In music, there is also a web of possible structural relationships between symbolic and ecological (i.e. 'natural') sound transformations–roughly pitch-time structures and timbral and textural transitions. For example, in Baroque contrapuntal music, symbolic notation takes precedence to the extent that many compositions can be played on a variety of instruments without detriment to the intelligibility of the symbolic logic of the music. The abstract structures of functional harmony[38] on which that music is based, all-but completely dominate the music of the Western world–classical, jazz, folk and pop. Most people's engagement with music is through song, which is not at all abstract, and the formulaic abstractions of functional harmony are used to provide the song with emotional impetus and direction. The power of this functional pitch-time hierarchy has driven the evolution of counterpoint and harmonic thinking to unparalleled degrees over hundreds of years; so much so, that functional harmony is commonly referred to as *the* language of (Western) music (Cooke 1959), and its ubiquity plays a large role in the acceptance and (false) expectation of it as *the* formal means to organize musical expression.

When a musical form principally encodes information in the structural relations between sonic events (motifs), as it is with music based on functional pitch-time hierarchies, the timbral quality of the carrier can be much reduced without significantly degrading the formal information. Such is the case even when immutable sounds triggered by untouched-by-human-hands translations of musical notation to MIDI[39] so quickly pall on the ear. This may be acceptable if the musical structures are strong or simple, and listening is functional or semantic (see Table 4.1), but is unlikely to be acceptable for sonifications that convey information not in accord with the logics of functional harmony, especially if users need to listen over extended periods of time. A scan of the sonification research literature over the last forty years, reveals a steady stream of reports of such dissatisfaction among research project participants, including that, after repeated or continual exposure to such materials, some people simply stop listening. This issue was a "hot topic" at the ICAD 2004 conference in Sydney, which included both a major community sonification design project (Barrass, Whitelaw, and Bailes 2006), and an informally reported general increase in understanding by conference participants that experienced sound designers brought a set of skills to research projects which led to clear improvements in the listenability of sonification designs without compromising the integrity of the research.

There are some musical styles that rely on the dynamic timbral qualities of the sounds involved to convey a vital dimension of the music. In Japanese classical shakuhachi music, for example, there is a fluid evolution in which pitch functions as a component of evolving timbral modulations. The idiosyncratic noise compositions of the Futurists were quickly followed by fixed-media tape compositions using the recordings of noises. Even in Western classical music, there was a noticeable evolution over several centuries from the former towards the latter: The orchestration of keyboard-composed works for the instrumental resources available for an occasion was superseded by writing for the orchestra directly. In Claude Debussy's 1912 ballet score *Jeux*, for example, the symbolic structure of the music and the orchestration seem so closely linked that they appear to have been conceived in

[38] Tonal harmony based on major and minor keys in which chords function in various schema to produce quasi-predictable flow sequences such as cadences (interruptions and closures).

[39] Musical Interface Digital Instrument. See http://www.midi.org/

tandem. In 1911, Arnold Schoenberg's *Harmonielehre*, outlined the principles of *klangfarbernmelodie* (sound-color melody) which had a seminal influence on many composers of middle half of the twentieth century (Schoenberg 1978, 421).

Some sonifiers have utilized the language of functional harmony in their work–often, perhaps in order to create an easy 'pleasant' connection to the aurally familiar. Music-theoretic concepts and how they relate to psychophysical percepts is a field of research in its own right, and the simplistic application of it at the phrase level is fraught with difficulties. Timbre, for example, is a problematic musical concept to psychoacoustically scale, as is consonance and dissonance, especially if no account is taken of the tuning, temperament, tessitura or spectral composition of the resonator being used (Sethares 2005), to say nothing of the physical influence of critical bandwidth (Plomp and Levelt 1965) and other context sensitivities (Cazden 1980). Such an approach rarely works, which is not to suggest that the elements of Western tonal music cannot be usefully employed, but that to do so requires experienced ears–at least until computational design paradigms are developed that can accommodate the complexities of meanings and associations these elements bring to a sonification. Besides, there are plenty of less familiar opportunities, including alternatives to twelve-note equal temperament, harmonic and inharmonic spectra, drones, additive time structures and metric modulations.

4.4.2 Data

4.4.2.1 Analog and Digital

While the two terms *data* and *information* are used somewhat interchangeably in common parlance, there is a sense that *data* is somehow more elemental, and that information is somehow extracted from it.[40] Considering it like this, data could simply be defined as a set of values or variables which can be either qualitative or quantitative. Unlike quantitative properties, which are discrete, and can be expressed by assigning a numerical value to the them, qualitative properties are continuous cannot be expressed numerically. Analog data, then, is qualitative; data that is represented or is analogous to a direct or recorded observation made in a physical way, such as the surface grooves on a vinyl record. And digital[41] data are sets of individual, discrete discontinuous symbols which represent information of some kind. To convert an analog signal to digital form it has to be sampled, usually at evenly spaced time intervals. In the reverse process, a digital-to-analog converter

[40]The term *data* (singular datum) comes from Latin *dare*, meaning 'something given'. The term *information* comes from Latin *information (n-)*, from the verb *informare,* meaning to shape, fashion or describe into a form. Thus 'information' contains the sense of informing, of fitting (as in a mold) into some recognizable "thing". The relationship between data and information is discussed in more detail in Chap. 3.

[41]From *digitus*, the Latin word for finger.

transforms a finite precision number into a physical quantity, such as voltage or pressure. The most important concepts and practical techniques necessary for converting data between analog and digital formats are to be found in sampling theory.[42]

Unless otherwise indicated, the term *data* in this book is used to mean digital, symbolic data such as numbers (in various forms: integers, floating-point (real), complex, as well as numerals, alphabetic and other textual characters. Important features of data for sonification are the size, range, variability, dimensionality, format, location (source) and accessibility of it, ranging from small finite data-sets that are easily maintained on a portable computer, often in character-separated (CSD) format, to source-propriety formatted realtime internet 'feeds'.

4.4.2.2 Data Sets: Small, Big and Thick

Small: When the ability to count and measure is limited, we *ac*count for only the most important things. Until recently in human history, collecting and analyzing data "by hand" was expensive and time-consuming, so, for reasons of practicality, new questions and hypotheses had to be well defined before the data collection process was embarked upon. Then, following some initial tests and analysis, it was not unusual to find inadequacies such as sampling errors, biases, which necessitated new data to be collected and analyzed. The expense and difficulties involved in iterating over this collection-and-analysis sequence has been an encouragement for researchers to study in detail what they are trying to determine before embarking on the process of data collection. So, it is that, until recently, data collection and analysis methodologies have been focused on (a) the precision and careful curation of the data in databases that are designed so as to retrieve information both as accurately and as quickly as possible and (b) the use of statistical methods that were developed to extract the most information possible from the smallest amount of data. If data collection or generation is a task to be performed, it is imperative to understand that sampling precision improves most dramatically with quantifiable randomness not increased sample size. The implication is that it is more important to avoid assumptions in choosing the sampling space, than to collect more samples. In addition to elementary statistical analysis, such techniques include, interpolation, compression, curve-fitting, principal component analysis for dimension reduction, and factor analysis for identifying interrelationships between variables.

Big: The ready availability of tools and the digitization of many social interactions now generates more data than can be processed using techniques for small datasets. The term *big data* is usually reserved for datasets whose size is beyond the ability of typical database software tools to capture, store, manage, and analyze.

[42]While recommending texts to anonymous readers is a thankless task, *The Computer Music Tutorial* (Roads et al. 1996) contains and excellent outline of most of the standard digital sound and music analysis and synthesis concepts and techniques. It's 1200-odd pages is well written for readers with average mathematical skills.

Table 4.6 Summary of some qualitative aspects of big data sets

Monitoring	A lot of big data streams from sensors or tracking (multiple) devices that monitor continuously and indiscriminately without regard to environmental influences, and network transmission losses and blockages.
Filtering	Collection control structures are frequently unavailable, leading to inaccuracies, or incompleteness in unpredictable ways. This makes subset comparisons and analyses more difficult to validate.
Integrity	The sheer volume of the data can mask incompleteness in multiple dimensions, making the location of predictably-structured search paths unimpeded by 'dead ends' more difficult.
Context sensitivity	Data collected by third parties for new tasks unrelated to the original reason for collecting it may require adaptions which present interpretive challenges.
Merging	When datasets from different sources are combined, problems can arise related to data field format integrity and inconsistences. The lack of precise field definitions can easily mask misaligned objectives.

One of the core technical challenges of big data is that, because it is often obtained from instruments not designed to produce data amenable to scientific analysis, its validity can be overstated (Redman 2014). Some of qualitative characteristics of big data sets which can be challenging are summarized in Table 4.6.

In working with big data, the most important technical shift has been from the study of *causality* to the statistically simpler *correlation*. Patterns and correlations in data can offer valuable insights: That two things regularly occur in sequence in close temporal proximity, can be valuable information, even though it is not known precisely why it occurred. Much individually-targeted internet marketing, for example, is based on the study of such patterns of behavior (Mayer-Schönberger and Cukier 2013).

Thick: The term *thick data* comes from a combination of *big data* and *thick description*, which is an anthropological research methodology that documents human behavior in its context. The original reference for the term *thick data* is attributed to the philosopher Gilbert Ryle (1900–1976), who argued that the thick description of what someone can be said to be doing, will normally be more informative than the 'thin' description. When the opening and closing of an eye is (in context) a wink, there is a lot more than a single bit of data being communicated. Thick data is generated using qualitative, ethnographic research methods that reveal people's meaningful actions, emotions and models of their worlds. It relies on human learning, in contrast to the machine learning that is usually necessary to work with big data. Thick data is considered to be better for exploring unknown territory, that is, learning patterns of behavior that are not already known, or not inferable from big data analysis (Boyd and Crawford 2012). The development of methodologies for 'blending' or integrating big and thick data requires understanding intentional and strategic processes–the blending of insights. At the time of writing, techniques for achieving this are in active development (Bornakke and Due 2018).

4.5 Sonification Designing: The First Steps

Because data to be sonified doesn't, in itself, provide a solution to what to do with it, there is a need for overt methodologies to ensure that perceptualization aims, directions, and motivations are explicit. Because focused vision makes a privileged contribution to spatial processing, and there are space-time asymmetries in our languages and cultures, and perhaps more deeply in our biology, the use of sound to support non-linguistic communication does not come with the same tacit knowledge that vision does. Hearing is as equally sophisticated as vision, but fundamentally different: More acutely three-dimensional, capable of finer time and frequency resolution, and always aware–even in sleep, to itemize some obvious differences. This section outlines some of the challenges facing the designer in the process of sonifying datasets. In is not comprehensive–that would be a different kind of book–but exposes enough of the issues to enable us to begin to specify the types of software requirements necessary to improve the currently available sonification tools.

4.5.1 *Tuning Your Listening: Ear-Cleaning and Ear-Tuning*

It is not an accident that proficient sound designers often have a background in music composition. Having spent thousands of hours, sometimes from an early age, in detailed listening and composing aural structures, they bring skills learned in the process to sonification tasks. Effective design begins with learning to listen to unfamiliar material analytically, using reduced listening techniques (see Table 4.3). This is not at all the same as "listening" passionately to one's favorite music, but affectively.

A period of *ear-cleaning* and *ear-tuning* can invigorate and accelerate the learning and development of the listening skills needed to produce intelligible sonification designs. An effective exercise to accelerate the learning of the type of listening required that can be undertaken, simultaneously with the data integrity tasks discussed below is outlined in (Table 4.7):

Table 4.7 An ear-cleaning and ear-tuning exercise

Ear-cleaning	Stop listening to all music, other than that used in the ear-tuning, for the duration of the exercise (one week, perhaps).
Ear-tuning	(a) Listen to a recording of (a section of) a piece of instrumental music based on unfamiliar structural principles.
	(b) Repeat (a) until you can anticipate upcoming events.
	(c) While listening, draw a simple one-page free-form graphical mapping of what you hear, filling in more details on each listening. Dividing the page into equal time divisions can aid the process.
	(d) At times when you would normally listen to music, substitute active listening to the sounds of your environment.

Because of their timbral clarity, and brevity, Anton Webern's music is excellent for this exercise: A movement or two from his *Symphony* (Opus 21), *Concerto* (Opus 24), or *Five pieces for orchestra* (Opus 10).

4.5.2 Data: Cleaning and Statistical Analysis

A dataset is the foundation of a well-functioning sonification design. Integrity checking and data-cleaning are essential for anyone embarking on the task. With large datasets, these processes have to be undertaken algorithmically, using pattern-matching techniques together with robust exception handling.[43] Extensive testing is especially important with realtime data or for time-critical applications, and, considering the issues identified in Table 4.6, powerful tools designed to assist with the task are a distinct asset.[44]

Coming to an early understanding of some of a datasets' quirks and foibles through elementary statistical analysis, can assist in developing a confidence that will pay dividends in later stages of the design process. Because data cleaning, especially of large data-sets, is often iterative, it may be astute to build such statistical analysis into the cleaning and/or exception-handling processes themselves. In the case of realtime data, it is also a good practice to invoke separate capture-to-file routines for more leisurely, repeatable, explorations. This technique is discussed in more detail in the network data sonification examples in Chap. 9.

To reiterate, data is the foundation of a well-functioning sonification design. It may seem strange in this context to say this, but experience teaches that the appropriate attitude to data preparation is at least attentional *care*, or perhaps *taming*, in the spirit of the Little Prince's fox: "One only understands the things that one tames" (de Saint-Exupéry 1943).

4.5.3 Scoping: Defining the Purpose

Many of the early stages of a data sonification task are project-managed with generic skills and practices, and so are not discussed in any detail here. Understanding the exact purpose of a sonification is the first step towards making it. A wise move is to have the purpose or use of the sonification in writing in the form of a scoping document and, if others (called *the client* hereafter) are involved in defining this scope, to have them agree to such in a written statement. When, due to the complexity of the project, exact outcomes are uncertain, it is better to make the

---

[43] *Python* has excellent pattern-matching tools. The `re` module provides regular expression matching operations similar to those found in *Perl*. Both patterns and strings to be searched can be Unicode strings as well as 8-bit strings.

[44] See, for example, the Anaconda Python distribution of `pandas`.

distinction between those outcomes that are essential, and others. At this relatively early stage in the development of the field, most clients have only a limited experience of data sonification, if any, and, especially if the engagement for the task is a result of excellent proselytizing on the part of the designers, ignorance, fueled by unfettered imagination, can quickly morph to unrealistic expectations. If unchecked, such enthusiasm has the potential to lead to expectation creep, which becomes difficult to manage the longer it is tolerated.

Two features of scoping that later cause the most distress if they are not clearly itemized, are the level of interactivity, and the sophistication of the user-interface required. These two items alone can be all consuming and need to be managed carefully, for, as Jef Raskin, the founder, in 1979, of the Macintosh project at Apple, has been reported to have said, *As far as the customer is concerned, the interface is the product*.

4.5.4 Design Criteria: Representation–Figurative or Conceptual?

Data Sonification functions as an external, independent or abstracted representation of data in order to derive information from it. The aim of using it as a perceptualization process is to reveal information in a dataset to an attending listener at least as efficiently as some other form of presentation, such as linguistic speech, numerical tables, visual graphs, or animated visualizations. While representation is necessary, is not always sufficient; it needs to be more than simply symbolic. For a sonification to represent information meaningfully, the information must be part of the experience of the representation. That is, that the sonification should be able to be *experienced* in terms of what it *represents*. The experience-representation binding between a dataset and a sonification can vary significantly depending on its purpose. In an artistic sonification, that binding might be loose, even in the form of a cultural commentary on the environmental impact of the source of the data; perhaps even just invoked in the minds of listeners through a program-note.[45] A much tighter binding between experience and representation is necessary in the use of sonification for scientific purposes, for which is it necessary, following Hermann (2008) that:

1. The sound reflects *objective* properties or relations in the input data.
2. The sonification is *reproducible*: given the same data and identical interactions, or triggers, the resulting sound has to be perceptually identical.
3. The transformation is *systematic*. This means that there is a precise definition provided of how the data (and optional interactions) cause the sound to change.
4. The system can *intentionally* be used both with different data, and in repetition with the same data.

[45]This issue is discussed in Chap. 2 (2.5.1).

4.6 Aesthetic Considerations

While much has been written about aesthetics, the term is rarely defined, which is problematic because there is not a single accepted definition, as the etymology reveals:

> aesthetic (n.) 1798, from German … ultimately from Greek *aisthetikos* "of or for perception by the senses, perceptive," of things, "perceptible" (by the senses or by the mind), to feel "… "to perceive." … Popularized in English by translations of Kant and used originally in the classically correct sense "science which treats of the conditions of sensuous perception" [OED]. Kant had tried to reclaim the word after Alexander Baumgarten had taken it in German to mean "criticism of taste" (1750s), but Baumgarten's sense attained popularity in English c. 1830s (despite scholarly resistance) and freed the word from philosophy (Harper 2001).

As can be inferred from the etymology of the term, differences in interpretation abound. Rather than a comprehensive review, the intent here is more directly positional, which amounts to contrasting Baumgarten's aesthetic of the tasteful, and Kant's concept of aesthetics (which harkens back to Aristotelean thought), and updating his account for some contemporary developments. Completely absent from this discussion are the contributions of 19th century thinkers, including Arthur Schopenhauer's aesthetics of The Sublime.

Enlightenment theories of aesthetics isolated the aesthetic appreciation of objects or events by placing them in a realm of their own, disconnected from other modes of experiencing. Like many Enlightenment thinkers, Kant held that because we experience the structure of the world through reason, aesthetic judgements are 'reasonable' judgements of taste. According to Kant ([1790] 2007), such judgements have four distinguishing characteristics (or *moments* as he calls them):

1. They are *disinterested*: we do not find something beautiful because it pleasurable, but we find it pleasurable because we judge it to be beautiful.
2. They are *universal*: we expect reasoning others to agree with us.
3. They are *purposeful*, either of utility–for (external) things that are designed to accomplish something, or if not of utility, of 'perfection'–in (internal) appreciation of how something is meant to be; also thought of as purposeful purposeless.
4. They are *necessary*, that is, according to a commonly-held, or *a priori* subjective principle of what is beautiful. This subjective principle corresponds to what was thought of as the purposiveness of nature.

Society is much more culturally diverse today than it was for Kant, and intellectual reason is no longer considered the only, or most important, valid basis for aesthetic judgement. Despite this, unfortunately Baumgarten's assertion that aesthetic qualities inhere in objects of attention due to the balance and proportionality of their form, remains the meaning of the term 'aesthetic' in common usage today. However, Kant's dualist position, which considers aesthetic experience as the subjective appreciation of beautiful objects, does not withstand critique, unless his concept of *universal* is expanded to include styles and cultural influences of which he would not have been aware.

Today, the term *aesthetic* is used to identify the quality of unifying sensory experiences; of an event, an object, or an environment. Life proceeds not merely *in* an environment, but because of it. A person's destiny is bound up with their interchanges with their environments, not just in external manifestations, but in private ways (Savage et al. 2012). Different listeners may have similar or wildly different emotional experiences with the same artefacts. In as much as the quality of the sensory experience evokes emotions in the listener, those emotions are components of the listener's own aesthetic, bound by their cultural, time, place and perhaps other environmental factors. So, it is that the virginal galliards of John Bull[46] were perceived as wildly passionate in their day, and boring, sewing-machine clatterings by some aficionados of punk music.

In parallel with theories and research in embodiment in the second half of the 20th century, a theory of Aesthetic Engagement has developed which recognizes that a person's participation in the appreciative process contains more creative perceptual involvement than the subjective appreciation of beautiful processes and objects than are recognized in Kant's theory, and certainly more than considered by Baumgarten (Berleant 1991). Aesthetic Engagement rejects the separation between the perceiver and the object of perception in favor of a contextual field in which the reciprocity and continuity of appreciative experience is a perceptually active, direct, and intimate. In this theory, the multimodal activity of sensory perception is primary: in all sensible experiences, including the temporal, kinesthetic and somatic. Aesthetic Engagement offers a model of environmental appreciation and everyday events and objects, and considers the central concern to be, not difference between art and non-art, but between the aesthetic and the non-aesthetic.

Mind coordinates sensory activity into meaningful experiences, and sense covers a wide range of sentient ideas–from physical and emotional shock: sensor, sensory, sensitivity, to the meaning of things present in immediate experience: sensible, senseless, sensitive, sentimental, sensuous, sensational. That aesthetic experiences can have emotional effects or affects, suggests that an aesthetic experience is not the same as an emotional experience although they may have common fates.[47] Emotions are not *in* the sound work, but evoked in the listener; nor are emotions 'added-in' to the perception by the listener, they are a part of it. We have emotional experiences. Our emotions permeate through our perceptions: we perceive emotionally.

---

[46]John Bull (1562-3–12 March 1628) was renowned for his technical experiments and the solution of unusual musical problems—enharmonic modulations, for example, and asymmetrical rhythmic patterns. Some of his music is in *Parthenia* (1612), the first published book of English virginal music.

[47]A piece of music can be formulaically transformed to reliably elicit one of several emotion (Friberg et al. 2000; Kirke and Miranda 2009).That type of transformation may be applied to a sonification to enhance cohesiveness.

The idea that an aesthetic can be added *post hoc* to an otherwise unpalatable object of experience, like sugar and salt can be added to a dish "to taste" in the final stages of preparation, is a fallacy and will invariably be recognized as such. Attempts to impose such "musical-isms" on a sonification invariably appear as false flavor enhancers, and are more likely to increase blandness and confusion than assist comprehension or palatability. Clarity and integrity will always eclipse pretty or nice. The ear easily palls, and judgement with it, so oscillation between periods of *listening* to an event in its iterative formation in as many modes as feasible (Table 4.2), and silent imagining, will invariably yield more successful results than continual aural saturation. During development, exclusive monitoring with headphones or earbuds will invariably produce unexpected distortions for any sonic environment that will not be eventually delivered through those mediums. Sonification designs are formed by infusing malleable sounds with intractable data, so if clarity and integrity are important, the Grecian/Kantian notion of purposefulness, with a sensibility of active engagement, seems more encompassing and pertinent to the practice that any attempt to accommodate mere matters of taste.

The aesthetics of engagement reflects Dewey's assertion that the important question is not, to paraphrase, "is it music?" but "when is music?" These questions are not ahistorical. In other words, except in as much as the present times inevitably resounds with echoes of the past, and reverberates into them, exploratory composers, and by extension, the restlessness of the Western mind in general, have always actively listened, through their composing, for answers to the second question first. Similarly, what we now consider as the great works of the past cannot meaningfully be interpreted just through contemporary imperatives. In honoring both dimensions of our cultures, the challenge is to try to listen with an understanding of the aesthetic sensibility of the times in which the work was composed. In doing so, the holographic nature of beauty is revealed. This approach is aptly expressed by the perspicuous Buckminster Fuller:

> When I am working on a problem, I never think about beauty. I think only of how to solve the problem. But when I have finished, if the solution is not beautiful, I know it is wrong (Fuller [1980] 1985).

4.7 Summary and Conclusions

Analogical mapping can be described as a mapping or association between elements and descriptions of a source domain and a target domain. Analogy becomes useful in contexts when a listener is familiar with the source domain, and can map familiar elements or relations from the source into unfamiliar (or unknown) elements or relations in the target domain. These mapped elements are analogical inferences and receive varying levels of support from other mapped elements, from knowledge about the target domain, or from more elaborate confirmatory schemes.

As we strive for more flexible, robust and perceptually sustainable and engaging sonifications, what emerges from this chapter is an awareness of the multidimensionally subtle and complex possibilities to be considered when designing. These possibilities are still beyond the best current legacy software tools available from computer music, and realizing them requires a sustained effort in the development of additional capabilities, including the integration of:

1. A database of psychoacoustic algorithms that can used to provide perceptually linear (or other dimensional) transforms for application at the individual sound, sound complex or complete sound environment level, as appropriate.
2. A set of field transforms to produce well-formed convex subsets in all parametric dimension (pitch: tunings, modes; duration: rhythm, tala etc.) and congruently in multi-dimensional parametric spaces (gesture, etc.) (Honingh and Bod 2011).
3. Sets of multidimensional transforms derived from the physical gestures of musical performance, concentrated on understanding and generating the role of extra-notational aspects of music, particularly on emotional expression and affect through micro-gestural affordances (Worrall 2014).
4. Following understandings developed in cognitive science of the pervasive role of embodied conceptual metaphors in almost all of our thinking, the development of tools to integrate metaphorical fames into more tightly integrated analogical bindings.

The next chapter outlines the requirements for such a comprehensive software framework, and proposes an integration of various existing independent components such as those for data acquisition, storage and analysis, together with a means to include new work on cognitive and perceptual mappings, and user interface and control.

References

Ballas JA (1996) Designing sound for an interactive experience. AES, Los Angeles, CA

Barrass S, chaffert N, Barrass T (2010) Probing preferences between six designs of interactive sonification for recreational sports, health and fitness. In: Bresin R, Hermann T, Hunt A (eds) Proceedings of the 3 Rd interactive sonification workshop. KTH, Stockholm, Sweden

Barrass S, Whitelaw M, Bailes F (2006) Listening to the mind listening: an analysis of sonification reviews, designs and correspondences. Leonardo Music J 16(1):13–19

Berleant A (1991) Art and engagement. Temple University Press, Philadelphia, USA

Bjursten LM, Norrsell K, Norrsell U (1976) Behavioural repertory of cats without cerebral cortex from infancy. Exp Brain Res 25(2):115–130

Bornakke T, Due BL (2018) Big–Thick blending: a method for mixing analytical insights from big and thick data sources. Big Data Soc, 1–16

Bowdle GD, Wolff P, Boronat C (2001) Metaphor is like analogy. In: Gentner D, Kokinov BN (eds) The analogical mind: perspectives from cognitive science. MIT Press, Cambridge MA, pp 199–253

Boyd D, Crawford K (2012) Critical questions for big data: provocations for a cultural, technological, and scholarly phenomenon. Inf Commun Soc 15(5):662–679

Bregman A (1994) Auditory scene analysis: the perceptual organization of sound. The MIT Press, Cambridge, MA

Brooks RA (2002) Flesh and machines: how robots will change us. Vintage Books, NY: New York, USA

Burbank J, Steiner P (eds) (1997) On poetic language. In: The word and verbal art: selected essays by Jan Mukarovsky, translated by Burbank J, Steiner P. Yale University Press, New Haven, Connecticut, USA

Byrne JH (1997) Learning and memory. In: Byrne JH, Dafny N (eds) Neuroscience online: an electronic textbook for the neurosciences. Department of Neurobiology and Anatomy, The University of Texas Medical School at Houston (UTHealth), Houston, Texas http://nba.uth.tmc.edu/neuroscience/

Cabrera D, Ferguson S, Schubert E (2007) Psysound3: software for acoustical and Psychoacoustical analysis of sound recordings. In: Proceedings of the thirteenth international conference on auditory display. Montreal, Canada

Campbell D, Fiske D (1959) Convergent and discriminant validation by the multitrait-multimethod matrix. Psychol Bull 56(2):81–105

Cazden N (1980) The definition of consonance and dissonance. Int Rev Aesthet Sociol Music 2:123–168

Chase WG, Simon HA (1973) Perception in chess. Cognit Psychol 4:55–81

Chomsky N (1972) Language and mind, Extended edn. Harcourt, Brace & World, New York

Collins J, Selina H (1998) Heidegger for beginners. Icon Books, Duxford, Cambridge

Cooke D (1959) The language of music. Oxford University Press, New York

Croft W, Cruse D (2004) Cognitive linguistics. Cambridge University Press, Cambridge, New York

Damasio AR (1995) Descartes' error: emotion, reason, and the human brain. Penguin, NY

Damasio AR (1999) The feeling of what happens. Harcourt Brace, NY: New York

De Groot AD (1966) Perception and memory versus thought: some old ideas and recent findings. In: Kleinmuntz B (ed) Problem solving. Wiley, New York

Dewey J (1934) 1952. Art as experience. Perigee Books, The Berkley Publishing Group, New York, New York

Dretske FI (2000) Perception, knowledge and belief. Cambridge University Press, Cambridge

Farnell A (2010) Designing sound. MIT Press, Cambridge MA, USA

Fauconnier G (1985) Mental spaces: aspects of meaning construction in natural language. MIT Press, Cambridge: Cambridge, MA

Fauconnier G, Turner M (1996) Blending as a central process of grammar. In: Goldberg A (ed) Conceptual structure, discourse, and language, 11th ed., 113–29. CSLI, Stanford, CA

Ferguson J, Brewster SA (2017) Evaluation of psychoacoustic sound parameters for sonification. 120–27. Glasgow, Scotland

Ferguson S, Cabrera D, Beilharz K, Song H-J (2006) Using Psychoacoustical models for information sonification. In: Proceedings of the 12 th international conference on auditory display, 113–20. London, UK

Flowers JH (2005) Thirteen years of reflection on auditory graphing: promises, pitfalls, and potential new directions. In: Proceedings of the first symposium on auditory graphs. Limerick, Ireland

Friberg A, Colombo V, Frydén L, Sundberg L (2000) Generating musical performances with Director Musices. Comput Music J 24(3):23–29

Fuller RB (1980) 1985 Beauty. In: Fadiman C (ed) The little, brown book of anecdotes. Little, Brown and Co., NY: New York, USA. Incorporated

Fyk J (1997) Intonational protention in the performance of melodic octaves on the violin. In: Leman M (ed) Music, gestalt and computing: studies in cognitive and systematic musicology. Springer, Berlin

Gaver W (1989) The sonicfinder: an interface that uses auditory icons. Hum Comput Interact 4 (1):67–94
Gaver WW (1988) Everyday listening and auditory icons. Cognitive Science and Psychology Department, University of California, San Diego
Gibson JJ (1966) These senses considered as perceptual systems. Houghton Mifflin Company, Boston, MA
Godøy RI (2006) Gestural-sonorous objects: embodied extensions of schaeffer's conceptual apparatus. Organ Sound 11(2):149–157
Godøy RI, Jensenius AR, Voldsund A, Glette K, Høvin M, Nymoen K, Skogstad S, Tørresen J (2012) Classifying music-related actions. In: Proceeding of the 12th international conference on music perception and cognition and the 8th triennial conference of the european society for the cognitive sciences of music. Thessaloniki, Greece
Godøy RI, Leman M (eds) (2010) Musical gestures: sound, movement, and meaning. Routledge, New York: Bonanza
Gray H (1918) Anatomy of the human body. In: Lewis WH (ed). 20th ed. Lea and Febiger, Philadelphia and New York
Gray JA (1995) A model of the limbic system and basal ganglia: applications to anxiety and schizophrenia. In: Gazzaniga MS (ed) The cognitive neurosciences. MIT Press, Cambridge, MA, pp 1165–1176
Harper D (2001) Aesthetic. In: Online etymology. http://www.etymonline.com/
Heidegger M (1927) 1962 Being and time. Translated by Macquarie J. Blackwell, Oxford: Blackwell. Oxford, UK
Hermann T, Ritter H (2005) Model-based sonification revisited: authors' comments on hermann and ritter, ICAD 2002. ACM Trans Appl Percept 2(4):559–563
Hermann TM (2008) Taxonomy and definitions for sonification and auditory display. In: Proceedings of the 14th international conference on auditory display. Paris, France
Honingh AK, Bod R (2011) Convexity and the Well-formedness of musical objects. J New Music Res 33(4):293–303
Hunt A, Wanderley MM (2002) Mapping performer parameters to synthesis engines. Organ Sound 7(2):97–108
Huron D (2002) Listening styles and listening strategies. In: Proceedings of the society for music theory conference. Columbus, Ohio
Iazzetta F (2000) Meaning in musical gesture. In: Wanderley MM, Battier M (eds) Trends in gestural control of music. IRCAM, Paris
ISO (2014) ISO 12913-1:2014. Acoustics—soundscape—Part 1: definition and conceptual framework. https://www.iso.org/obp/ui/#iso:std:iso:12913:-1:ed-1:v1:en
James W (1907) 1995 Pragmatism: a new name for some old ways of thinking. Dover Publications, Inc
James W (1992) Psyhology: the briefer course. In: James W (ed) Writings 1879–1899. Library of America, New York
Jouvet M (1967) The States of Sleep. Scientific American
Jouvet M, Jouvet D (1963) A study of the neurophysiological mechanisms of dreaming. Electroencephalogr Clin Neurophys 24
Kahn V (1985) Rhetoric, Prudence, and Skepticism in the Renaissance. Cornell University Press, Ithaca and London
Kant I (1790) 2007 Critique of judgement. In: Walker N (ed). Translated by Meredith JC. Oxford University Press. http://philosophy.eserver.org/kant/critique-of-judgment.txt
Keats J (1820) Ode on a Grecian Urn. Annals of the Fine Arts
Kima JH, Demeyb M, Moelantsb D, Leman M (2010) Performance micro-gestures related to musical expressiveness. In: Proceedings of the 11th international conference on music perception and cognition. Seattle, Washington, USA
Kirke A, Miranda ER (2009) A survey of computer systems for Expressivemusic performance. ACM Comput Surv 42(1)

Kohler EC, Keysers MA, Fogassi UL, Gallese V, Rizzolatti G (2002) Hearing sounds, understanding actions: action representation in mirror neurons. Science 297(5582):846–848

Kramer G, Walker B, Bonebright BT, Cook P, Flowers J, Miner PN, Neuhoff J (1999) Sonification report: status of the field and research agenda. Technical report. International Community for Auditory Display

Kranjec A, Lehet M, Woods AJ, Chatterjee A (2019) Time is not more abstract than space in sound. Front Psychol 10(48) https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6367220/

Kubovy M, Van Valkenburg D (2001) Auditory and visual objects. Cognition 80:97–126

Kuhl PK, Conboy BT, Coffey-Corina S, Rivera-Gaxiol M, Nelson T (2008) Phonetic learning as a pathway to language: new data and native language magnet theory expanded (NLM-e). Philos Trans Royal Soc B: Biol Sci 363(1493):979–1000

Lakoff G (2012) Explaining embodied cognition results. Topics Cogn Sci 4:773–785

Lakoff G, Johnson M (1999) Philosophy in the flesh. Basic Books, New York NY

Lakoff G, Johnson M (2003) Metaphors we live by. The university of Chicago Press, London

Lehar S (2000) The function of conscious experience: an analogical paradigm of perception and behavior http://cns-alumni.bu.edu/~slehar/webstuff/consc1/consc1.html

Lehmann AC, Ericsson KA (1997) Research on expert performance and deliberate practice: implications for the education of amateur musicians and music students. Psychomusicology: A J Res Music Cognit 16(1–2):40–58

Lombardo TJ (1987) The reciprocity of perceiver and environment. The Evolution of James J. Gibson's Ecological Psychology. Lawrence Erlbaum, Hillsdale: N. J

Louie K, Wilson MA (2001) Temporally structured replay of awake hippocampal ensemble activity during rapid eye movement sleep. Neuron 29(1):135–156

Mabaquiao NM (2005) Husserl's theory of intentionality. Philosophia: Int J Philos 34(1):24–49

Macionis JJ (2012) Sociology, 14th edn. Pearson, Boston MA

Mayer-Schönberger V, Cukier K (2013) Big Data. Houghton Miffin Harcourt, NY

Merleau-Ponty M (1945) 1962 The phenomenology of perception. Translated by Smith C. Routledge & Kegan Paul, Oxford, UK

Metzger W (2006) Laws of seeing. MIT Press, MA

Miller GA (1956) The magical number seven plus or minus two: some limits on our capacity for processing information. Psychol Rev 63:81–97

Mullane JJ (1973) William James and Phenomenology. Loyola University, Master of Arts, Chicago

Musiek FE (2003) What can the acoustic startle reflex tell us? Hearing J 56(9):55

Narayanan S (1997) Embodiment in language understanding: sensory-motor representations for metaphoric reasoning about event descriptions. Computer Science Division, EECS Department: University of California, Berkley

Neuhoff JG, Heller LM (2005) One small step: sound sources and events as the basis for auditory graphs. In: Proceedings of ICAD 05-eleventh meeting of the international conference on auditory display, 417–19. Limerick, Ireland

Ohman A (1979) The orienting response, attention and learning: an information-processing perspective. In: Kimmel HD, Van Olst EH, Orlebeke JF (eds) The orienting reflex in humans. Erlbaum, Hillsdale, NJ, pp 433–467

Oliveira CJC de (2014) Rise and decline of rhetoric in the European education. Int Res J Qual Educ 1(3):19–23

Palmer SE (1999) Vision science: photons to phenomenology. MIT Press, Cambridge, MA

Petersen SE, Sporns O (2015) Brain networks and cognitive architectures. Neuron 88 (October):207–219

Peterson JB (1999) Maps of meaning: The architecture of belief. Routledge, New York NY

Peterson JB (2017) 2017 Maps of Meaning 05: Story and Metastory (Part 1). https://www.youtube.com/watch?v=ZHmklvx9oJ4

Peterson MA, Skow-Grant E (2003) Memory and learning in figure-ground perception. In: Ross B, Irwin D (eds) Cognitive vision: psychology of learning and motivation. 42:1–34

Piaget J (1936) 1963 The origins of intelligence in the child. Translated by Cook M. International Universities Press, W.W. Norton & Company, New York NY

Plomp R, Levelt WJM (1965) Tonal consonance and critical bandwidth. J Acoust Soc Am 38:548–560

Redman TC (2014) Getting in front of data quality. From Data to Action. Harvard business review insight center

Regier T (1996) The human semantic potential: Spatial language and constrained connectionism. MIT Press, MA

Roads C, Strawn J, Abbott C, Gordon J, Greenspun P (1996) The computer music tutorial. In: Roads C (ed) MIT Press, Cambridge MA

Rock I (1975) An introduction to perception. MacMillan Press, New York

Roddy S (2015) Embodied sonification. PhD Thesis, Trinity College: University of Dublin

Roddy S, Bridges B (2019) Addressing the mapping problem in sonic information design through embodied image schemata, conceptual metaphors, and conceptual blending. J Sonic Stud 17. https://www.researchcatalogue.net/view/558606/558686

Saint-Exupéry A de. (1943) The litte prince. Translated by Woods K. Reynal & Hitchcock, NY. http://www.angelfire.com/hi/littleprince/

Saunders C, Fernyhough C (2016) The medieval mind. Psychologist 29(November):880–883

Savage J, Gibson W, Vaucher G, Sterling L (2012) Punk: an aesthetic. In: Kugelberg J. Rizzoli, New York NY

Schoenberg A (1978) Harmonieliehre (Theory of Harmony). Translated by Carter RE. 3rd (1922) 1st edition 1911. University of California Press, Berkeley, Los Angeles

Sethares WA (2005) Tuning, timbre, spectrum, scale. Science and Business Media. Springer

Sokolov EN, Spinks JA, Näätänen R, Lyytinen H (2002) The orienting response in information processing. Erlbaum, NewJersey: Mahwah. London

Swanson (2000) Cerebral hemisphere regulation of motivated behavior. Brain Res 886:113–16

Sweller J, Ayres P, Kalyuga S (2011) Cognitive load theory. Explorations in the learning sciences, instructional systems and performance technologies. Springer, New York

Todd PM, Loy G (1991) Music and connectionism. MIT Press, Cambridge, MA

Tuuri K, Eerola T (2012) Formulating a revised taxonomy for modes of listening. J New Music Res 41(2):137–152

Tuuri K, Mustonen M, Pirhonrn A (2007) Same sound—different meanings: a novel scheme for modes of listening. In: Proceedings of audio mostly, 13–18. Ilmenau, Germany: Fraunhofer Institute for Digital Media Technology

Ungureanu C, Rotaru I (2014) Philosophy of skillful coping. Motor intentionality vs. representations for action. Procedia Soc Behav Sci 163:220–229

Wertheimer M (1938) Laws of organization in perceptual forms. In: Ellis WD (ed) A source book of gestalt psychology. Hartcourt, Brace and Co, New York

Wertheimer M (1958) Principle of perceptual organization. In: Beardslee DC, Wertheimer M (eds), Readings in perception. D. Van Nostrand Co., Inc., Princeton, NJ

Wilson EO (1999) Consilience: the unity of knowledge. Vintage Books, Random House, Inc.,New York

Worrall DR (2010) Parameter mapping sonic articulation and the perceiving body. In: Proceedings of the 16th international conference on auditory displaproceedings of the 16th international conference on auditory display, 9–15. Washington, D.C, USA

Worrall DR (2014) Can micro-gestural inflections be used to improve the soniculatory effectiveness of parameter mapping sonifications?" Organised sound. Cambridge University Press. 19(1):52–59

Zbikowski LM (2012) Music, language, and what falls in between. Ethnomusicology 56(1):125–131

Zwicker E, Fastl H (2014) Psychoacoustics: facts and models. Springer

Chapter 5
Towards a Data Sonification Design Framework

Abstract The need for better software tools for data sonification was highlighted in the 1997 *Sonification Report*, the first comprehensive status review of the field which included some general proposals for adapting sound synthesis software to the needs of sonification research. It outlined the reasons the demands on software by sonification research are greater than those afforded by music composition and sound synthesis software alone. As its *Sample Research Proposal* acknowledged, the development of a comprehensive *sonification shell* is not easy and the depth and breadth of knowledge, and skills required to effect such a project are easily underestimated. Although many of the tools developed to date have various degrees of flexibility and power for the integration of sound synthesis and data processing, a complete heterogeneous Data Sonification Design Framework (DSDF) for research and auditory display has not yet emerged. This chapter outlines the requirements for such a comprehensive framework, and proposes an integration of various existing independent components such as those for data acquisition, storage and analysis, together with a means to include new work on cognitive and perceptual mappings, and user interface and control, by encapsulating them, or control of them, as *Python* libraries, as well as a wrappers for new initiatives, which together, form the basis of *SoniPy*, a comprehensive toolkit for computational sonification designing.

5.1 Introduction

To date, almost all software for data sonification design has been developed either as standalone applications that have been engineered from first principles, sometimes incorporating third-party low-level audio routines, or as more expansive sonification 'environments' that attempt to encapsulate some general principles and procedures that can be adapted for specific sonification projects as the need arises.

Standalone software applications tend to be designed for individual projects entailing clearly defined tasks such as accurate monitoring (Chafe and Leistikow 2001), or for user-interaction, such as Walker and Cothran's *Sonification Sandbox* (2003). Some applications are designed for a specific type of data or purpose. For

D. Worrall, *Sonification Design*, Human–Computer Interaction Series, https://doi.org/10.1007/978-3-030-01497-1_5

example, *xSonify* is designed as a sonification data analysis tools to meet the need for an effective non-visual approach to displaying science data, particularly space physics data (Candey et al. 2006). Other early examples of stand-alone toolkits are *Listen* (Wilson and Lodha 1996), *MUSE* (Lodha et al. 1997) and SonART (Ben-Tal et al. 2002). In a different vein, *Planetheizer*, aims to provide an interactive synchronized multimodal simulation for planetary systems using data embedded within a planet module that is added to a generic sound-synthesis program (Riber 2018).

In contrast, sonification environments tend to be more generic and expansive projects, often with less deterministic outcomes. They afford greater flexibility than is possible within standalone applications. Some such environments include a cross-platform (*Pure Data*) toolkit for interactive sonification (Pauletto and Hunt 2004) and a multi-institution collaboration *SonEnvir*[1] (de Campo et al. 2004) that encompasses many different sonification research problems, with the aim of developing a generalized sonification environment for scientific data. Such projects seem to have been designed by first choosing a music composition environment and working backwards, perhaps trusting that the data-processing needs at the 'input end' can be adequately handled by the language tools available from within the particular composition system chosen, or managed independently of the sonification process per se. Considering that many software tools for music composition have long gestation periods,[2] this approach is natural and the one assumed in the *Sonification Report* (Kramer et al. 1999).

In addition to scripting environments, *MAX/MSP*[3] and its sibling *Pure Data* (PD)[4] represent another popular approach, which emphasises the primacy of a graphical user interfaces (GUIs) over program code. Whilst the scripting-versus-GUI debate is still active, it is clear from the large user-base and active development of new Patch objects for these platforms that the GUI approach is appealing to some users, and perhaps offers a gentler initial learning curve for exploratory sonification researchers who are visually inclined. In any event, a considerable investment of time is necessary to become proficient in any of these environments and having made the investment, a certain amount of environment 'stickiness' is apparent and understandable.

[1] http://sonenvir.at/.

[2] Supercollider (https://supercollider.github.io/) and *Csound* (http://www.counds.com/) are two such widely used tools. *Csound* has been in active development for more than sixty years (Boulanger 2000). Supercollider began as a community-supported project by its author in 1996, and was re-released as open source under GPL in 2002 (McCartney 2002).

[3] http://www.cycling74.com/.

[4] http://www.pure-data.org. *Pure Data* (PD) PD is Miller Puckette's public release of his Max/MSP. It has been extensively extended by a vibrant community of developers.

5.2 The First Bottleneck: Data

Music composition software that supports the generation and processing of musical metadata, following the algorithmic composition trends and influences outlined in Chap. 1, has tended to concentrate mostly on internally-generated rather than externally-acquired data, or in closed-loop interactive performance contexts where strict control of the data flow is mediated by protocols such as MIDI or OSC,[5] or by real-time direct audio-processing algorithms. For data sonification, while the input data can be thought of as eventually controlling the sound rendering, the transformations this data needs to undergo beforehand can be considerable, particularly in the case of live-streaming applications. Such data processing might reasonably be expected to include multidimensional scaling, filtering and other principal component and statistical analyses, which themselves may become the subject of sonification. Also, each dataset can have potentially unique structural characteristics. Some, such as EEG data, may consist of multiple channels of AC voltage data with a variety of DC-biases and noisiness, as determined by the particular data collection apparatus on a particular patient. Others, such as security 'tick' data flowing from a market trading-engine, will be massively paralleled metadata embedded and multiplexed into a single noisy 'feed'. Difficulties in using such data are compounded when they need to be buffered and streamed in non-real-time, such as when producing multiple overlays of time sequences of different temporal compressions for comparative purposes.

High-level tools for processing such data complexities are seldom found in computer music environments, and are even less likely to be so if the input data is spatial rather than temporal. So, when such tools are necessary, a common response to complex data processing requirements is for a sonifying programmer to 'bite the bullet' and write data-processing routines in the language of the composition environment itself. This was the approach used by *SonEnvir* and *OctaveSC*[6] (Hermann et al. 2006), which are both built on a base of *SuperCollider*, and also Pauletto and Hunt's *PureData*-based *Interactive Sonification Toolkit* (op. cit.). The following exchange on the *SuperCollider* users list epitomises the situation exactly:

> Question: Hi, I'm looking for linear algebra and affine transformation routines for 2D and 3D vector spaces; is there any quark or extension that implement stuff like that?

[5]Open Sound Control. See http://opensoundcontrol.org/. OSC is a communication protocol used for sending information across a computer network. It is promoted as the successor to MIDI (with greatly increased bandwidth and user-customizable representations), and is supported by most modern music software. OSC is the protocol used by SuperCollider's *sclang* to communicate with its sound synthesizer *scsynth*. In most practical uses, OSC is implemented as another layer on UDP connections.

[6]http://www.sonification.de/projects/sc3/index.shtml/.

Answer: The MathLib quark[7] maybe has some useful stuff… or it would be the place to put these.

"…or it would be the place to put these" implies, "if someone else using *SuperCollider's sclang* hasn't erstwhile written them and made them publicly available". The sonification programmer thus has decide whether or not they have the required expertise and time to dedicate to implementing these mathematical tools. While *sclang* is a very elegant and powerful composition language that can support the development of such routines, being unique, it lacks the transportability that more general and widely available programming languages afford. One consequence of this is that, in projects without a dedicated programmer, practical assistance for what are essentially data processing problems may prove more difficult, and perhaps more expensive to obtain than might otherwise be the case. In these circumstances, data is often pre-processed using external tools such as spreadsheets, and then read from generically formatted files by the music composition environment; a process that, whilst it may be appropriate in 'limited data' experiments, is at best susceptible to data corruption, and of no use if the data is coming from a real-time streaming feed or from an interactive dynamic model (Bovermann et al. 2006).

This situation may be characterised as 'data *sonification*' that is, the primary focus is on sound rendering whilst input data is constrained to enable it to be dealt with adequately by the rendering software. The alternative, an emphasis on data-processing tools at the expense of sound rendering flexibility ('*data* sonification') is no more attractive, because the sound palette tends to be small and the range of controls limited; the outcome of which is too-often acoustically mono-dimensional. Some examples of this latter approach include extensions for the *Matlab* numerical environment (Miele 2003), *AVS Visualization Toolkit* (Kaplan 1993), *Excel* spreadsheets (Stockman, Hind and Frauenberger 2005), and the R statistical analysis package (Heymann and Hansen 2002). They provide data handling and processing capabilities but very basic sample-based sound capabilities modelled on a simplistic implementation of the MIDI[8] protocol. The resulting audio experience of such approaches can be overly-mechanistic and unsubtle, resulting in them being difficult to listen to, especially over extended periods of time.

As mentioned, excellent data-processing tools exist in the public domain in multiple languages and are an integral part of much scientific research. Furthermore, they are continually being extended and modernised by global teams of developers. Yet, when the decision is made to use a music composition environment for sonification, these data-processing tools may remain largely inaccessible. So, until sophisticated tools for handling externally-acquired often massively multiplexed datasets and databases, website scrapings etc., are able to be brought to bear on data acquisition, analysis, storage and re-presentation requirements, even

[7]A Quark is a package of SuperCollider classes, helpfiles and C++ source code for unit generators and other code.

[8]http://www.midi.org/.

before any mapping or sound rendering is undertaken, there is limited chance that such music composition software will sufficiently enable the *choosing* [of] *mappings between variables and sounds interactively, navigating through the data set, synchronizing the sonic output with other display media, and performing standard psychophysical experiments* that the *Sonification Report* envisaged.

5.3 A Comprehensive DSDF: Concepts and Requirements

When considering options for a generic well established, mature programming language to integrate DSDF components, one flavor or another of *C++*, *Java* and possibly *Lisp* are obvious candidates. However, when also considering the projected user base and such requirements as ease-of-learning, visual clarity of the code,[9] and rapidity of development, one language stood out above all others: since it was first released in 1991, *Python* has become an extremely popular language in which to think. It is elegant, powerful, very well supported and there is strong evidence that it a highly productive professional programming tool; there does not currently appear to be a serious challenger to it.[10]

5.3.1 Features of the Python Language

The main features[11] of *Python* which make it *the* candidate to perform the vital 'glue' role of bringing all the components of a comprehensive DSDF together are:

Simplicity: It is a relatively simple, minimalistic language. Reading well-written code feels almost like reading formal English. This pseudo-code nature of *Python* encourages programmers to concentrate on solutions rather than syntax. This also makes it relatively easy to learn.
Free and Open Source: It is an example of a FLOSS (Free/Libré and Open Source Software) project. FLOSS is based on the concept of a community which shares

[9]The task-fitness of notation is vital, as Ken Iverson, the inventor of APL, so eloquently expressed (1980).

[10]Without attempting comprehensivity, a few of the places that use Python extensively are Google, the New York Stock Exchange, Industrial Light and Magic, Redhat, Dropbox, Survey Monkey, Uber, and many new media platforms such as Facebook, Instagram, Spotify, Netflix, and Reddit. And then there are organizations using it for scientific research. The list will change, but *Python* is BIG and very mainstream! While notebook extensions (such as Jupyter kernals) and potential rivals (Golang, for instance) appear on the horizon, the nature of the task such a language needs to perform in a DSDF mitigates against being able to make a judgement based on technical specifications alone: a broad and deep user-base is vital.

[11]An operational overview of the language, including a summary of its technical features are outlined in Chap. 6.

knowledge. In simple terms this means that its source code can be freely used, altered and distributed.

A High-level Language: Programmers do not need to concern themselves with low-level details such as memory management. Also, being a dynamically-typed language, every non-null variable name is only bound to an object at execution time by means of assignment statements, and an individual variable name can be bound to objects of different types during program execution.

Portability: Due to it being open-source, it has been ported to multiply many platforms. Because it is interpreted, *Python* code without system-dependent features will work on any platform without any modification.

Interpreted: Unlike compiled languages like C++, there is no separate compilation and execution phases. Programs are just 'run' from the source code. Internally, *Python* converts the source code into an intermediate form called bytecodes (stored in .pyc files), translates this into machine code and then runs it.

Object Oriented: It supports procedure-oriented programming as well as object-oriented programming, whose implementation is simple, compared to languages like *C++* or *Java*, and its ability to use multiple inheritances is powerful.

Extensible: When it is important that code runs very quickly, it can be written in C, and then combined with regular *Python* code. Therefore, one of the most common needs is to call out from *Python* code to a fast, machine-code routine (e.g. compiled using *C/C++* or *Fortran*). The fact that this is relatively easy to do is a big reason why it is such an excellent high-level language for scientific and engineering programming.

Embeddable: It can be embedded within a *C/C++* program to give scripting capabilities to those programs. *Csound* has an embedded *Python*, available through various Opcodes.

Extensive Libraries: The Python Standard Library is extremely large and available wherever *Python* is installed. Besides the standard library, there are various other high-quality libraries such as the Python Imaging Library, a simple image manipulation tool kit.

Scalability: It has a proven scalability track-record as exemplified in Footnote 10.

Interactive Python: Command-line shells for interactive computing. (*IPython*) has the following features:

- Introspection: The ability to examine the type or properties of object at runtime.
- Shell syntax, including command-line completion and history.
- A browser-based notebook interface with support for code, text, mathematical expressions, inline plots and other media. This has subsequently spun-off into Project Jupyter a web-based interactive computational environment for creating, executing, and visualizing Jupyter notebooks.[12]
- Interactive media: presenting rich media content, including interactive data visualization and use of GUI toolkits. This includes, in addition to the standard

[12]http://jupyter.org/.

Python's shell's access to *Tkinter*, the non-blocking interaction with *PyGTK*, *PyQt/PySide* and *wxPython*.
- Flexible, embeddable interpreters.
- Optional tools for many different styles of parallel and distributed computing.

What follows is a summary of the *Sonification Report*'s sonification shell requirements and some occasional comments on the strengths and weaknesses of *Python* as it relates to those requirements as a scripting tool, and as a wrapper for third-party software for the broad range of sonification tasks identified:

Flexibility: Rather than try to be the 'killer application', the Framework should aim to wrap (inherit, or be extended by) the best collection of modules available to it for a particular project. Ideally, although they may be conceptually similar, these modules need to have as few computational interdependencies as possible, thus ensuring that no one of them is indispensable to the DSDF as a whole. There are a collection of several extensively developed, mature and widely used modules that have become almost generic extensions to scientific work with *Python* and these are convenient exceptions to the no-dependencies requirement.[13] Each of these modules has evolved independently over a considerable period of time and are attractive for a DSDF because they have their own ongoing development teams that extend and improve then as well as adapting them to ever-changing hardware and software platforms.
Integrability: As discussed earlier with regard to data, due consideration needs to be taken of the requirements of the various components of the sonification and experimentation process. As is the case with most interdisciplinary ventures, each contributing discipline brings its collection of tools, techniques and standards to the venture and they need to be synergistically integrated. This accommodation needs to be, not only from the perspective of different data domains, but from developments in design science and advances in Human Computer Interaction (HCI). *Python*'s ability to seamlessly 'wrap' independent conformable software modules in ways that permit data to flow between them supports these goals.
Extensibility: In the situation where no *Python* module exists for a particular task, a new one can be designed in the knowledge that it will fit seamlessly within the existing framework. This implies all modules need to be thread-compliant.
Availability: To protect both authors and users, the DSDF needs to be freely available with a minimum of restrictions. There are numerous licensing flavors for public-domain software whose source-code is made generally available, as outlined by the GNU organization.[14] One of *Python's* strengths is that it is based on the concept of a community which shares knowledge (FLOSS), and this is essential if

[13]SciPy is a Python-based ecosystem of open-source software for mathematics, science, and engineering. See http://scipy.org/.

[14]http://www.gnu.org/philosophy/categories.html.

the DSDF is to gain traction with designers. Because it needs to be able to deploy heterogeneous components,[15] the licensing of each component carries through into the DSDF in a way that is standard practice in the software industry. DSDF-specific components need to be issued under the General Public License Version 2,[16] thus encouraging the sharing and free exchange of these tools in the community, at the same time as enabling restrictions to be applied for individual projects as confidentiality agreements demand. A further implication is that the sources and documentation should be freely downloadable from an internet repository.

Accessibility: It is desirable that as many stable, already-known modules be integrated whenever possible. This reduces the learning overhead for all users and enables work that may have already been undertaken with those tools in other contexts to be accommodated within the DSDF.

Maintainability: All software has a maintenance overhead, not the least because of the inevitable changes in the hardware platforms on which it runs. In an open-source module contribution model, it is imperative that compatibility and development standards be both achievable and actively implemented. The Achilles heel of the *Python* independent module model was the maintenance of version compatibility between multiple modules. Fortunately this problem has been greatly reduced by the availability of extensively-tested package management and deployment systems.[17] Community involvement and support in ongoing improvement and development of such projects are not without their own issues, as those who are involved in public-domain projects can attest (Luke et al. 2004).

Portability: The DSDF needs to be able to be instantiated on all major platforms. Furthermore, it is desirable, in certain applications, for modules to be able to be instantiated on different machines, in different locations and networked together: that is, to be heterogeneous. *Python* is available on all major platforms, including portable media devices and single-board microcontrollers such as Arduino and Raspberry Pi.

Durability: The DSDF needs to survive, and this is less likely if it is sustained by the efforts of only a small group of dedicated individuals. While survival can never be guaranteed, in large projects such as this, maximum risk-mitigation is essential. Durability is, in part, related to accessibility, as discussed above, but what is also necessary is a critical mass of developers and users to maintain and develop it. Being built on foundations as broad as *Python* and generic sound-synthesis API such as exists for *Csound*, does build in some resilience. *Csound*, has had a *Python* API since 2006 and continues to improve. It is also likely that *Python* APIs for other new and existing synthesis and control approaches will become available. Over a

---

[15]A heterogeneous computer network connects devices with different operating systems and/or protocols.

[16]http://www.gnu.org/copyleft/gpl.html.

[17]Such as that by Anaconda, a free and open source distribution of the *Python* and *R* programming languages for data science and machine learning related applications. The Anaconda distribution of *Python* has over 6 million users, and includes more than 1000 Modules, including 250 for data science. See http://www.anaconda.com/.

longer timeframe, the community of users can have some confidence that, given that *Python* has an extensive global and still-growing user-base, any move away from it would be heralded in enough time to adapt the DSDF to the new paradigm.

5.3.2 Integration Through Wrapping

Although *Python* comes with an extensive standard library and there is a large resource of external *Python* libraries, as discussed above, we are not limited to using *Python* libraries. A powerful feature of *Python* is its set of well-defined interfaces to other languages. Libraries written in most languages can be integrated through *Python* by 'extending' it with them. The basic principle of a *Python*-based DSDF is to use *Python* to 'wrap' independent software that can be compiled with python-bindings in such a way that data can flow between them. Quite a few tools exist for the (semi-) automatic generation of Python bindings, such as the *Simplified Wrapper and Interface Generator* (SWIG)[18] and *ctypes*, which provides C-compatible data types and can be used to wrap foreign functions, as well as shared and dynamically linked library calls.[19]

Some applications provide *Python* Application Program Interface (API) libraries; other third-party projects may need to have a *Python* API written in order to use them (Lutz 1996, 505). Some software embeds *Python*, either by bundling it as an interpreter or by invoking the *Python* interpreter installed on the user's system as a basic API. *Csound*, for example, has a *Python* interpreter embedded in it and accessible through a series of *Csound-Python* opcodes that are usable in numerous contexts, including in a namespace specific to an individual instance of a *Csound* instrument.[20] This embedding technique, whilst it may be useful in its own right, it does not provide the same flexibility for a DSDF as extending *Python* with external modules or libraries does.

5.4 The *SoniPy* Data Sonification Design Framework (DSDF)

The DSDF framework described here is called *SoniPy*, in-keeping with a naming convention used for frameworks that extend the *Python* programming language. The *SoniPy* design specifies five module sets communicating over three different

[18]http://www.swig.org/.

[19]The recent API *ctcsound.py*, build on the python *ctypes* module, is a considerable improvement on the SWIG-generated *csnd6.py*. It offers much more flexibility and an easier integration with multiple python toolkits.

[20]See https://csound.com/docs/manual/py.html.

networks: the *SonipyDataNetwork* (SDN), the *SonipyAudioNetwork* (SAN) and the *SonipyControlNetwork* (SCN). Modules are grouped according to their role in the data sonification process: Data Processing (DP), Conceptual Modelling (CM), Psychoacoustic Modelling (PM), Sound Rendering (SR) and User Interface (UI). Depending on the dictates of a particular project, modules in a set may be instantiated on different machines. A particular module set may be empty, i.e. contain no modules, or a particular module may belong to more than one module set. Figure 5.1 is a diagrammatic representation of the way the module sets inter-relate, with Table 5.1 providing a key to the various components. It can be observed that *SoniPy*'s modules operate through three networks: Data, Control and Audio. The *SonipyDataNetwork* (SDN) and the *SonipyAudioNetwork* (SAN) are topologically configured as busses, while the *SonipyControlNetwork* (SCN) is a star configuration. This is analogous to the signal and control busses of an automated audio mixing desk. Control routing uses the same network technology as the data, although the destinations may be different. For example, data from a DP module may be sent to a CM module on the SDN bus, under the control of an MF module communicating on the *SonipyControlNetwork*, without the data itself needing to go through an MF module.

SoniPy controls need to be XML compliant and each module set may itself be the hub of a network of processors, with a topology unknown to the SCN router. Rather than sending the sonification itself to display/monitoring devices, the Synth Metadata sub-module sends them sound synthesis parameters, thus enabling individual users to configure sonification configurations and the protection of sensitive data when VPN communication is not viable.

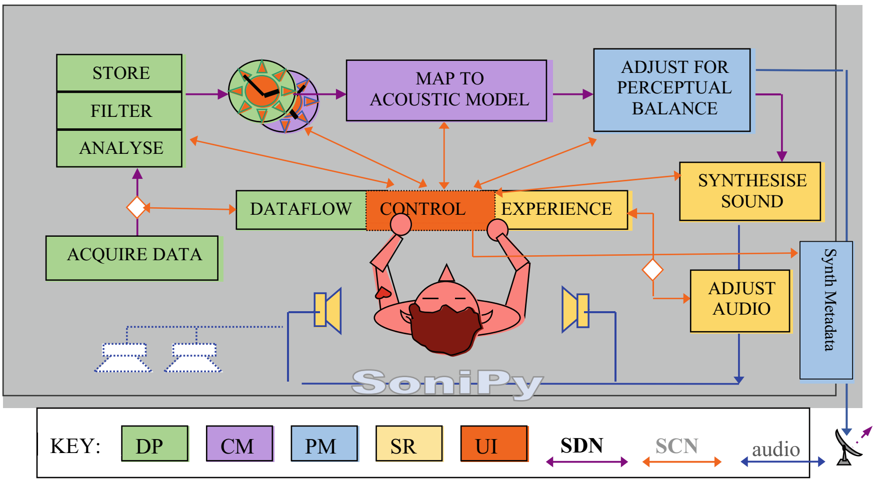


Fig. 5.1 Conceptual flow diagram of *SoniPy*'s five module sets and two networks

5.5 Inter-module Communication: The Three Networks

SoniPy's modular design makes it well suited for the instantiation of all selected modules on a single processor or, in order to take advantage of the computing power that multiple CPUs and machines can afford, the distribution of modules over multiple CPUs, a LAN, the Internet or with Cloud computing-options. Dynamic network configuration for distributed data acquisition and analysis, sound rendering and network audio mixing (using *Netjack*[21] for example) can provide benefits that include trade-offs between performance, including real-time latency, realtime data source sampling and throughput, CPU overhead issues, ease-of-use, maintainability, network reliability, scalability and heterogeneity. *Python*'s strong platform independence aids the implementation of a parallel DSDF, distributed over a heterogeneous network, including, the ability for non-*Python* third-party applications or devices to communicate with *SoniPy* using a range of techniques and protocols including multiple class inheritance, sockets, *OSC* and *MIDI* as well as bespoke protocols (Coulouris et al. 2005).

Other potential uses of a distributed approach include mobile device sound-rendering, and the processing of data remotely under local control, perhaps with the results being sent to another site for mapping and psychoacoustic adjustment before being rendered to sound. Portable devices have become more powerful and web browser technologies more encompassing, so it is now possible to generate sound on the hand-held devices directly under individual user configuration, with central and/or cloud distribution network control. These techniques promise to significantly improve both individual user's ability to select personalized mappings and data security controls, as discussed in the *Netson* project outlined in Chap. 9.

5.6 The Modules

Whenever practicable, local modules are included as Library extensions to the *SoniPy* DSDF. Should a module be instantiated remotely, that instantiation and the control of information to and from it remains under the control of the UI module that initiated the instantiation, where the main application loop, as well as thread and GUI controls, reside. Table 5.1 provides an overview of some key features of the *SoniPy* module sets.

[21]http://jackaudio.org/faq/netjack.html. Python bindings include jackclient-python PyJack.

Table 5.1 Overview of some key features of *SoniPy* module sets

Module sets	Data via *SoniPy* data network (SDN)	Controls via *SoniPy* control network (SCN)
DP	Input: Raw data to be sonified Processes: Analysis, filtration, translation, storage Output: Modified data to be sonified, Metadata	Process selection and OP switching options State (active, waiting, idle)
CM	Input: Data from DP, model selection algorithms Processes: Model selection Output: Selected model(s)	Process selection and output switching options State (active, waiting, idle)
UI	Input: Model control parameters Processes: Psychoacoustic transforms Output: Modified model control parameters	Process interaction and OP switch options Model instantiation map for render(s)) State (active, waiting, idle)
SR	Input: Synthesis models and controls Processes: Render sound Output: Sound	Audio controls: Mute, unmute, gain etc. Audio control changes State (active, waiting, idle)
MF	Input: Login, configuration and process startup Processes: Initiate, route, record activity, monitor usage, system/networks state Output: UI, feedback, log of activity	State (active, waiting, idle, off) Resource usage, UI monitoring

5.6.1 Data Processing Modules

As the principal information processing "activity" in data sonification is taking place in listeners, the role of sonification is to prepare source data in a format that enables them to extract information, and in interactive systems, to take account of their feedback. *SoniPy*'s DP modules consist of a number of object classes which themselves inherit data classes and control classes according to the form and location of the raw data and its intended destination. Class methods include those for

1. Interpolated lookup and mappings, for auditory icons and earcons, for example.
2. Writing data to, and extracting it from, storage (memory, database and/or flat file) for pre-processing, distribution or multi-stream playback.
3. Audification—writing data in formats acceptable as direct input to audio hardware.
4. Simultaneous handling of multiple time-locked streams, such as from biomedical monitors.
5. Deconstruction, analysis and filtering, including of complex meta-tag-embedded multiplexed streams, such as a data feed from a stock-market trading engine.

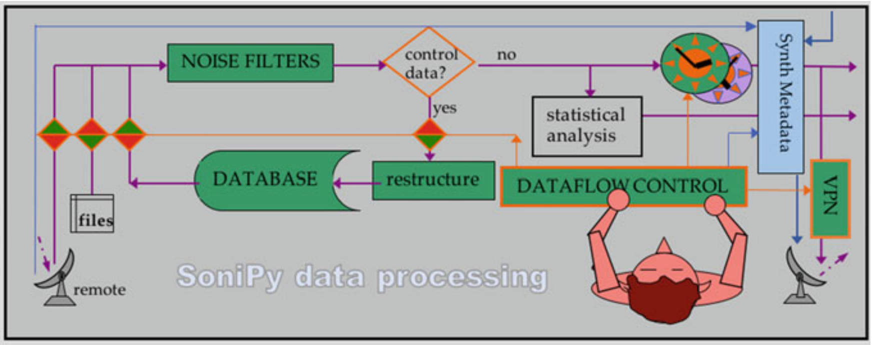


Fig. 5.2 A simplified map of the configuration of *SoniPy*'s Data Processing modules

6. Interaction with model-based sonifications which involves low-level processing of user(s) feedback.
7. Bimodal star network-monitoring when the secure data being sonified is distributed over a virtual private network (VPN) .
8. Simulation of data feeds, including buffering with time compression and expansion.
9. Dynamic control of the scalar field variables and sampling rates variables of wave space sonifications.

Data and control can be manipulated with *SciPy* (Scientific Python), a collection of open-source libraries for mathematics, science, and engineering.[22] The core library is *NumPy*, which provides convenient and fast N-dimensional array manipulation. Figure 5.2 is a diagrammatic representation of one way of configuring *SoniPy*'s data processing modules under *SCN* control.

5.6.2 Scale, Storage, Access, and Persistence

When a dataset used for sonification is finite, not too large, and is acquired in a timely manner, it can all be held in RAM and parsed using the basic data-manipulation tools of any modern programming language. Alternatively, as can be often observed in the auditory display literature, it can often be adequately pre-processed using generic spreadsheet software. As its size, complexity and performance demands increase, which they frequently do with multidimensional real-time datasets, or when what is being sonified is some computed informational

[22]http://www.scipy.org/.

derivative of the raw data, the power and flexibility of an interpreter interface to a software framework reveals its superiority over purely music-oriented software.[23]

In *Python,* direct access to many data management tools is available: Internal lists can be extended to typeless {key : value} dictionaries, then to flat files, perhaps with compression,[24] and *NumPy* array processing, then using framework extension, to commercial relational databases like *Oracle* and the web-prevalent public domain *mySQL*. With very large datasets, such as those regularly employed in astronomy, meteorology and quantitative finance, one has a choice of a number of models, such as *hierarchical*, *network*, *relational*, *entity-relationship*, *object-relational model* and *object*,[25] depending on the structure of the data and the way it is used.

The relational model is characterized by relation, attribute, tuple, relation-value and relation-variable. A relation is, at its heart, a truth predicate about the world, a statement of facts (attributes) that provide meaning to the predicate (Date 2004). Speaking simply, relational systems are structured around relationships between facts, and data retrieval is based on predicate logic and truth statements.[26] Object systems are structured so as to emphasize identity, state, behavior and encapsulation. An *object* has a unique identity that is distinct from its state (the value of its internal fields)-two objects with the same state are still separate and distinct, despite being bit-for-bit mirrors of one another (ibid.).

A well-known programming problem is whether to use an equivalence or identity model to structure data persistence: the object-relational-mapping- or ORM-dilemma.[27,28] The solution depends on both the structure and distribution of the data and the types of enquiry needing to be made of it, and can be the subject of extensive experimentation. Relational databases tend to perform more effectively when the data can be easily structured in a few relatively large tables with a fixed number of fields, and the object model when there are a large number of semi-autonomous instances. These issues are exemplified, with sample code, in the context of data sonification experiments using high-frequency capital market data,

[23]Including *SuperCollider*'s *sclang* probably the closest to *Python* in concept, in that it is text-based, interpretive and contains a dictionary class.

[24]Python has its own simple flat-file data compression module called *Pickle*. Another useful tool, that is being increasingly used for temporary storage and internet transmission of datasets is *JSON*, a lightweight data interchange format: https://docs.python.org/3/library/json.html.

[25]Wikipedia has useful overviews. See http://en.wikipedia.org/wiki/Database_model and http://en.wikipedia.org/wiki/Object-relational_mapping.

[26]In essence, each row within a table is a location for a fact in the world, and a structured query language (SQL) allows for operator-efficient data retrieval of those facts using predicate logic to create inferences from them.

[27]In computer science, the term *persistence* refers to that characteristic of data that outlives the execution of the program that created it. Without persistence, data only exists in RAM, and will usually be 'lost' when the program is terminated.

[28]Wikipedia has a useful introduction.
See http://en.wikipedia.org/wiki/Object-relational_mapping.

in the Chap. 8, where an HDF5-compliant b-tree structure was eventually employed using *pyTables*.[29]

5.6.3 Conceptual Modelling and Data Mapping

In the *SoniPy* DSDF, information mapping is divided into separate cognitive and psychoacoustic stages. The cognitive stage involves the design of 'sound schemas' with semiotics, metaphors and metonyms relating to the task, and aesthetic and compositional aspects relating to genre, culture and palette as discussed in Chap. 4.

Decisions have to be made about functionality, aesthetics, context, learnability, expressiveness, and device characteristics. These decisions are typically drawn from existing knowledge and theories from the relevant sciences, arts and design. The consequences of these compounding decisions are difficult to predict empirically: one of the reasons why sonification is currently more of an heuristic art than a science, and a motivation behind the need to develop a computational design approach.[30] Nevertheless, as different conceptual models are developed, some based in cognition, others culturally determined, they can be integrated into *SoniPy* using the wrapping techniques outlined.

One example might be the construction of an interface to *TaDa*'s methods; a design approach to sonification that provides a systematic user-centered process to address the multitude of decisions required to design a sonification (Barrass 1996b). *TaDa* starts from a description of a use case scenario, and an analysis of the user's task, and the characteristics of the data. This analysis informs the specification of the information requirements of the sonification. *SoniPy*'s support for the *TaDa* method would be through a python-based GUI that captures a user scenario and provides standard *TaDa* fields for semantic analysis. This GUI, connected using the SDN to a *mySQL* database that contains a large number of stories about everyday listening experiences, analyzed using the *TaDa* data-type fields. This database, called *Earbenders*, is a case-based tool for looking up 'sound schemas' at the cognitive design stage (Barrass 1996a). A *Python* interface to *TaDa*'s SonificationDesignPatterns wiki could also be developed as an alternative pattern-language approach for cognitive level design (Barrass 2003).

[29] Hierarchical Data Format (HDF) technologies are relevant when the data challenges being faced push the limits of what can be addressed by traditional database systems, XML documents, or in-house data formats. See http://www.hdfgroup.org/why_hdf/. *PyTables* is a *Python* API for HDF5 data.

[30] This issue is addressed in Sect. 5.9.

5.6.4 Psychoacoustic Modelling

Following metaphorical mapping, the psychoacoustic modelling stage involves the systematic mapping of information relations in the data to perceptual relations in the sound schema (Barrass 1996a). Changes in this mapping cause the automatic remapping of source information through psychoacoustic algorithms (implemented in the fast array processing tools, *NumPy* and *SciPy*) to produce new sounds and/or rendering controls. For example, a change from categorical to ordered information could automatically produce a remapping from a categorical sound (e.g. instrument, object, stream) to an ordered property of a sounding object (e.g. length, excitation, distance).

The concept is that *SoniPy* support this through interactively controlled reconfigurations of the mapping chain from information relations, as translated via the metaphorical mappings, though to auditory relations in the form of psychoacoustic transformations, before audio rendering occurs. As discussed in Chap. 4, such a meta-design schema needs to extend the knowledge summarized in Bregman's *analytical* work (1994) to afford the *synthesis* of virtual sonic environments that exhibit perceptual cohesion, while maintaining aurally differentiable virtual sound sources that solicit listeners to behave in ways that assist them to get enough 'grip' on these sound structures for them to be perceived as cohesive, evolving auditory objects in relation to each other.

5.6.5 Acoustic Modelling

SoniPy provides user access to many more sound-rendering options than if a purely music composition or sound synthesis environment was chosen before beginning the development of other aspects of the DSDF. For low-level audio work, a *Portaudio* module can be used for audification, and as the basis for the development of other such modules, should the need arise (Burk and Bencina 2001). *Portaudio* is used via, the *Python* wrapper *PyAudio* for the audio streaming in the capital market net returns, experiments in Chap. 8. *SndObj* ('sound object') is a middle-level toolkit also immediately available in the same manner (Lazzarini 2000). *SndObj*, like many real-time audio applications today, uses the *Portaudio* interface to the audio hardware. *The STK toolkit* also appears to be wrappable (Scavone and Cook 2005), as does the higher-level *RTCmix C++library*.[31] As already discussed, *Csound* has an embedded *Python* API for writing extended Opcodes and a robust python extension wrapper interface using *ctypes*, called *ctcsound*.

While some high-level applications such as *Max/MSP* and *PD*, are unlikely to become extension toolkits to *SoniPy*, it is still possible to use them by instantiating them independently and communicating with them using computer music protocol

[31] http://www.rtcmix.org/.

specifications *OSC* (Wright et al. 2003) and *MID*I.[32] *SuperCollider3* is a special case because of its inherent modularity: the sound-rendering component, *scsynth*, can be run independent of the scripting language, *sclang*. Communication tests between this 'external' *scsynth* over *OSC* using the *SoniPy* framework as an alternative to *sclang* indicate that this is viable. If very low-latency is a requirement, such as may be the case for interactive sonifications, *pysclang*, originally designed as an alternative to *sclang* for Windows platforms, enables direct communication with an instantiation of the SC language and its internally embedded sound renderer. Although this seems like a circuitous route, it works and is another example of the robustness of the encapsulating framework approach used in *SoniPy*.

Non-operating system dependent text-to-speech synthesis is available through *PySpeak*, an OSX thread-compliant *Python* API to *Espeak*[33,34] or *pyttsx ()* or interfaces to more specialized speech synthesis tools such as the extensive *Festival Speech Synthesis System* (Black et al. 1998) and its *Python* API *Pyfestival* (Robertson 2015).[35]

5.6.6 User Interface, Monitoring, Feedback and Evaluation

User interfaces for monitoring, interaction and feedback of SDN and SCN can happen via the *Python* interpreter. Having access to an interpreter in order to build a complete sonification by iteratively building on small tests is a powerful aspect of *Python*. Heterogeneous connectivity also allows the monitoring of and adjustment to compound, remote design decisions at each stage, thus enabling a better understanding and control of non-linear and emergent effects of embodied interaction, for example, in an overall design.

For cross-platform GUI-building, *wxPython* provides access to *wxWidgets*[36] and *wxGlade*[37] can assist in more-rapid development of GUIs by automatically generating *Python* control code and separating the GUI design and event-handling code. If a relatively consistent interface across all hardware platforms is more desirable, *Tcl* GUI-building tools[38] are available through native *Python* modules such as *tkinter*.

The inclusion of test and survey modules make it possible to design different types of user-based empirical experiments, and conduct and analyze the results within a single framework environment, even using website-based surveys if

[32] http://www.midi.org/.

[33] http://espeak.sourceforge.net/.

[34] The most lauded flexible text-to-speech tool is probably *Festival*.

[35] http://www.cstr.ed.ac.uk/projects/festival/.

[36] http://www.wxwidgets.org/.

[37] http://wxglade.sourceforge.net/.

[38] http://www.tcl.tk/.

desired.[39] This would be a marked improvement on current public-domain experimental psychology software, as there is currently no such tool that permits separately-threaded audio streaming or sound file playback. In addition, a user-contributed library of experiments for evaluating sonification designs could assist in developing standards for measuring the comparative functionality, aesthetics, learnability, effectiveness, accuracy, expressiveness and other aspects of individual design. While the design and implementation of a complete empirical sonification evaluation environment is beyond the means of many researchers in experimental psychology, the contribution to the design and implementation of such community-supported tools may not be. As already evidenced by such projects as *SciPy*, this approach is currently used in some scientific disciplines such as mathematics, physics, chemistry, astronomy, meteorology and genetics, where the design and sharing of experimental tools has long been an integral component of their research endeavors.

Code Example 5.1 is a simple 'high-level' metacode illustration of part of the *SoniPy* framework in action. The first task is to accept a multiplexed meta-tagged data stream from the Australian Securities Exchange (ASX) trading engine. The ASX is medium-sized exchange, on which about 3,500 securities are traded. It generates about 100 MB of Level 1 trading data daily, making it impractical to hold enough data in RAM to do all the calculations necessary.

The data is processed into a *MySQL* database using an Object-Relation Mapping paradigm supported by the *sqlobject* module.[40] This abstracts the handling of the dataset, providing an interface between the tables and indices database paradigm and Python's object-orientation. Other modules (such as *mySQLdb*) are available if direct interaction with the database server in *MySQL* code is more appropriate. A list of securities that meet, or are likely to meet, the criteria necessary for a sonification event to be initiated, is held in RAM and processed as a multidimensional array using the *NumPy* module. When the criteria are met this data is also used as some of the input parameters to the sound renderer. The *pyspeak* module is invoked to synthesize the name of the security being newly rendered. In this example, the sound is rendered by the *SuperCollider3*'s external synthesis engine, *scsynth,* with which the python *scsynth* module bi-directionally communicates using the *OSC* protocol. This permits the use of *SC3*'s *synthdefs* (synthesis definition algorithms) that are capable of responding to the criteria, as established, or as modified in real-time. Other synthesis options, such as the lower-level *pysndobj* or *pyaudio* (the *Python* interface to *portaudio*) are possible, as is the *libsndfile* library for audiofile playback. This metacode example thus illustrates how *SoniPy* combines *Python* code, imported 3rd party modules and user-defined scripts. It is meant here to provide a sense of the immediacy and readability of the approach.

---

[39]See, for example, the free and open-source web framework Django (https://www.djangoproject.com/) and its various survey options: https://djangopackages.org/grids/g/survey-questionnaire/.

[40]http://www.sqlobject.org/.

```
# Metacode Example; Sonipy Main Loop for ASX index sonification
# :::::::::::: import the modules needed :::::::::::::::::::::::

import SonipyUI # master UI class
asxUI= SonipyUI # instantiate master UI class in separate thread
import DataFeeder
import StreamMonitor
# uses matplotlib for graphical presentation
          # of trading engine stats
import DataToDB   # includes import of the sqlobject Module 3
    # which provides OO Class structure of the
    # Relation-to-Object-Mapping necessary for handling DBs.
    # For simple DB calls, use mySQLd
    # to simply pass mySQL commands to the server.
import SecurityMaps
    # mapping OO classes for ASX securities.
    # this Module calls the numpy module for
    # fast multi-dimensional array processing.
import PsychoFilters
    # a set of filters for transforming mapped data
    # Used to adjust Security Maps for psychophysical
    # non-linearities, condition enhancement etc
import SoundRenders
    # A set of synthesis engine interfaces and
    # synthesis definitions (instruments) appropriate
    # for this sonification. Includes OSC and MIDI,
    # and the phoneme synthesiser pyspeak.

# :::::::::::: Sonification threads ::::::::::::::::::::::
ASXFeed=DataFeeder(feedURL='http://localhost')
    # Establish connection to datafeed. The dtafeed
    # URL could be an internet,LAN address or filepath.
    # Uses Python build-in modules for low-level
    # data transfer to RAM buffer.
StreamMonitor.monitorStream (ASXFeed)
    # Monitor data stream.
DataToDB.MultiplexStreamtoDB(ASXFeed, ConnectParams)
    # Start processing ASXFeed into DB. Uses a MySQL DB
    # DB client connection as specified in ConnectParams.
    # Other servers are possible, including remote ones.
ASXalerts=SecurityMaps.Indicator(stocks, UpBollinger(20,9,2))
    # a FIFO of alerts for securities currently trading
    # outside the specified upper Bollinger band. The FIFO
    # is constantly updated whilst DataFeeder is active.
SoundRenderers.BinauralOut='TRUE'
    # Process sound output will for binaural listening.
while (StreamMonitor.StillActive):
    asxUI.Update()                  # update UI settings
    for security in ASXalerts:
        if security in renderList:
            pyspeak(security)       # announce new security
            SoundRenders.scsynth(security, stock.scsynthdef)
```

Code Example 5.1 Metacode example of the *SoniPy* DSDF in action. The task modelled is to accept data streamed from a stock market-trading engine, and use sonification to alert the listener to specific features of the trading activity as given

5.7 Computational Designing: A New Frontier

When a heterogeneous DSDF is sufficiently comprehensive, it will become possible for computationally-literate sonification designers to adapt their representational practices from those of designing objects for auditory engagement to the construction of meta-design systems of formally described relationships that define the ‘state space’ from which numbers of such sonifications can be drawn.[41] This shift from crafting individual sonic objects and streams to defining dynamical space of design possibilities is called *computational designing*.

Approaching the design of auditory displays as computational tasks poses both considerable challenges and opportunities. These challenges are often understood to be technical, requiring scripting or programming skills, however the main challenge lies in computational design thinking which is not best understood as the extension of established designing processes. The intellectual foundations of computational designing rest at the confluence of multiple fields ranging from mathematics, computer science and systems science to biology, psychophysical and cognitive perception, social science, music theory and philosophy. This section outlines the fundamental concepts of computational design thinking based on seminal ideas from these fields and explores how they it might be applied to the construction of models for synthesized auditory environments.

How sonification designers attempt to achieve an effective communication solution with a sonification design task is affected by many things, including the imagined range of possible solutions for the design (the state space), which in turn, is affected by the tools and methodologies employed, and the skills applied in using them. Attempting to evaluate the extent to which the solution is optimal, and how effective it is in achieving a stated goal, remain open questions, because the connection between the individual decisions that are made in the designing process and the way they interact in the solution space are non-linear: at best open to interpretation and at worse a collection of individualized black-box heuristics. Such a description is rarely controversial, even by those calling for robust evaluation and scientific comparison of sonification methods, as it is understood that:

> In the context of data exploration, what can be potentially learnt from a sonification is unknown, or at least not defined properly, and therefore it is very difficult to specify an objective performance measure. (Degara et al. 2013).

Design is a messy business and the relationship between decisions made in the process of ‘tweaking’ the contribution of individual parameters in the final result is rarely the sum of simple linear combinations. So, being able to evaluate the effectiveness of a sonification for a clearly defined purpose is not the same as being able to determine what aspects of the sonification are responsible for, or contribute to, that effectiveness. This is no difference, in form, to constructing a

[41]The state space from which such sonifications are drawn are inaudible, heard only through its instances, or the manifestations of particular *trajectories* through the state space.

neural-network to simulate an observable behavior without being able to understand the complexity of the individual contributions to the network itself.

5.7.1 *Computation Versus Computerization*[42]

The dominant mode of utilizing computers for audio design and production today is computerization or compilation: sonic objects or processes that have been conceptualized in the designer's mind are recorded, manipulated and/or stored in a computer using a Digital Audio Workstation (DAW). Typically, a DAW consists of a computer with sufficient storage and processing speed to be able to mix multiple simultaneous channels of audio to which pre-programmed processes are applied (often in the form of 'plugins') by software (such as *Protools*, *Abelton Live*, *Logic* and *Reaper*), which are used to edit, mix, transform, and store or simultaneously play back those audio channels for direct audition and/or recording.

From a design perspective, this computer-aided or compilation approach does not allow the designer to take advantage of the *computational* power of the computer in the design process itself.

> The manifest form - that which appears - is the result of a computational interaction between internal rules and external (morphogenetic) pressures that, themselves, originate in other adjacent forms (ecology). The (pre-concrete) internal rules comprise, in their activity, an embedded form, what is today clearly understood and described by the term algorithm. (Kwinter 2008).

Expressed simply, computational designing employs computation in the design process itself to *deduce* and place elements, for example, real-time synthesized sounds including microsound responsiveness to situational criteria, user input and the like. Computational design is best understood in contrast to computerized (or computer-aided) design, in which the computer is used to *compile* and arrange fixed design elements such as transforming pre-recorded sound samples to better fit the specific situation in which they are to be used. Computerized design is based on a data model, whereas computational design relies on a procedural model. It involves the processing of information and interactions between elements that constitute a specific environment. The final appearance of these elements, whether they be game objects or sonified information derived from data, is rendered from the processing of intrinsic properties, such as the specific values of data points, important information beacons, or extrinsic properties such as the positional rendering of the object in the acoustic environment in which it is being placed, taking into account the

---

[42]This Section Is Developed Through the Synthesis of Ideas from Multiple Authors in Menges and Ahlquist's Computational Design Thinking (2011).

effect (salience, occlusion[43] etc.) of other objects that have already been or will be placed there. Computation provides a conceptual framework for highlighting the data being rendered according to the importance placed on it at the time by the designer and the interacting user. As a design methodology, whereas computer-aided design begins with *objects* (such as sound samples) and adapt them to specific situations, computational designs start with *elemental properties* of objects (as synthesis parameters) and environmental influences, and use generative rules to formulate (or proceduralize) the specific objects in the specific environment into which they are placed.[44]

Most of the work of the computational designer involves explicitly defining and editing the definition of sets of variables such as psychoacoustic parameters and constraints. In generating specific solutions, logical operations on these sets and their (often dynamically generated) subsets are performed without the designer necessarily being able to conceptualize the full formal implications of their relationships. This can be a positive consequence of such abstraction as it can produce state-space solutions that might not have been intuited, considered or imagined using non-computational approaches. Because the designer is freed from the requirement to produce a single 'masterful' solution, many instantiations can be produced and then evaluated for their effectiveness, in-keeping with the goals of the *SonEx* project (Degara et al. 2013).

5.7.2 Some Advantages of Using a Computational Approach

While there may be some circumstances in which auditory icons and earcons[45] might be sufficient for creating a useful sonic environment, having to rely on them, with some general sonic smudging to attempt to smooth over the cutting-and-pasting, is not a recipe for the development of more sophisticated and responsive sonifications which are clearly needed as the understanding of the psychoacoustic and cognitive correlates of data sonification increases. As the complexity of an environment grows and/or the number of objects in it increases, the computational load of rendering all auditory objects becomes a critical defining feature of how many such objects can be rendered. In a sample-based model, either the library of different samples for an increasing number of sonic objects has to be generated and compiled, or the processing requirements for modifying a smaller subset of each object has to be increased.

---

[43]The salience of a sound is its attention-grabbing or distinctiveness (Delmotte 2012); occlusion refers to the virtual hiding or masking of a sound by others. These characteristics can be altered using signal processing techniques such as filtering and reverberation.

[44]The terms Procedural Design and Computational Design are frequently used interchangeably. Procedural Design is the more general term. When they are computed, Procedural Designs become Computational.

[45]See Chap. 2 for an explanation of these terms and Chap. 4 for an extended discussion.

In a computational design model, the rendered form of both individual auditory objects and general environmental factors-and the ways in which they interact-can be dynamic and highly flexible. This increases variety and reduces the reliance on the modification of decisions that need to be made in advance when using a samples model. For example, whereas sound samples might need to be modified according to situational salience requirements, in a computational model, the salience of objects can be made a feature of the synthesis of the objects themselves. In a sonic environment which is responsive to user-directed interests, for instance, this affords the production of a better balance between local and global reverberation requirements, resulting in a deeper sense of environmental continuity and sound-object integration. Another advantage of using a procedural approach is that, as the level of detail needed and the number of objects increases in the rendered auditory environment, the overall computation time involved relative to that required by manipulating samples, significantly reduces.

Incorporating task-based analysis into design criteria is increasingly understood to be an important step in developing effective data sonifications. At the same time, there is a need to develop larger solution state-spaces with an increased number of dimensions, incorporating both microsound and gestural levels and the intelligence to form multiple mapping solutions from correspondences between them which produce both highly dynamic and situationally responsive individual sound objects and higher-order extra-objective perceptual experiences such as swing, which is produced by systematically modulating the temporal flow.

The ability for designers, and ultimately users, to adapt a data sonification to their aural developing skills is also important. This is supported by Jean Piaget's observation of an aspect of the relationship between representational and perceptual space. During the development of an understanding of a representational space through experimentation (or 'play'), representational activity is 'reflected' or 'projected back' on to perceptual activity (Piaget and Inhelder 1956, 4) as exemplified in the way understanding of musical structures affects the way one perceives musical affect. This observation supports the hypothesis that for a listener, there is dynamic relationship between their critical listening skills, and the sonic complexity and variability of a sound-mapping that can be understood. This emphasizes both the need to have auditory designs which are responsive and/or able to accommodate different and developing listening skills, but also the need to develop the integration of the development of listening and sound-mapping experiences in early childhood education curricula.

If perception is not solely a bottom-up process but incorporates inherited or acquired response biases related to stimulus characteristics and sensory systems that have evolved for multiple, integrated representations of certain features, which are meaningful to the perceiver rather than for just single one-to-one reproductions of the physical world, it makes sense to generate multiple sonifications of any particular data set according to the user's developing sense of meaningfulness. In order to accomplish this, it is necessary for sound designers interested in creating dynamic auditory environments to shift their design representation and thinking efforts from creating bespoke hand-crafted solutions using just their own black-boxes, to a dynamic system model approach in which design activity is supported by extended state spaces that incorporate psychoacoustic and cognitively-informed transforms.

Because of the breadth and depth of detail required, sets of such transforms need to be developed collaboratively using agreed-upon standards. A community-based approach to developing these resources should go some way towards increasing the computational power of the computer as a design tool to *deduce* a wider variety of more effective auditory designs.

5.8 Summary

Sonification research is an interdisciplinary activity and to date, tools for undertaking it have either been other-discipline-specific (e.g. music composition, data-processing, scientific); modified to accommodate the interconnections, *ad hoc* collections, or stand-alone programs developed for a specific sonification task. Because *SoniPy*'s open architecture design can integrate modules conforming to widely accepted inter-process computation standards ('wrappable' libraries), in a non-conflicting way, it has the potential to grow in the directions its user-community needs it to. Instead of sonification researchers trying to piece together a collection of tools they hope can be made to work together, they will be able to choose from a number of possible *SoniPy* framework modules, based on a best-fit-for-the-task evaluation, and rely on the continuity of the framework to provide inter-module integration. Probably the most important feature of the framework approach outlined in this chapter is that, by using tools largely built by and for large communities of uses with a vested interest in their survival independent of data sonification, sonifiers are less likely to be 'held to ransom' by a reliance on the need for the developers of single software applications to continue to respond to the ongoing hardware and operating systems development environments in which they operate, and thus risk the isolation and eventual obsolescence that inevitably follows if they do not. *Python*, *NumPy*, *Matplotlib*, *Portaudio* etc. will continue to evolve and eventually be replaced by improved alternatives, but in ways that are more adaptive than catastrophic.

Because the *SoniPy* DSDF is implemented in a common, 'user-friendly' scripting language, programming assistance, when needed, will be more readily available than if a specialist application were used, and this may, in term, assist individual endeavor and promote better independent evaluation of the empirical results of other research in sonification. There is currently, for example, a dearth of good public-domain software for experimental psychology that uses sound, and the integration of such a module, or set of modules, into *SoniPy* would be a welcome addition to the field. The framework as conceptualized here is far from complete in actuality, nor perhaps, given the openness of the approach, will it ever be. It has been used and has been developing for about ten years, undergoing modifications and re-emphases as needs arise. It has been mainly used for parameter-mapping sonification projects of large multi-dimensional datasets but is extremely flexible and broadly extensible to other sonification techniques. So, it has not been practicable, in the current context, to test all the alternative combinations of modules accessible and useful in undertaking data sonification. The approach taken was to select a number of modules from different domains and test, albeit in a reasonably *ad hoc* way, the viability of their combined use to solve a specific problem that would normally be undertaken, often with some awkwardness, using two or more unrelated tools. Once library encapsulation had been effected, especially since the adoption of the Anaconda *Python* platform, module conflict has virtually become a thing of the past.

While the project is in its infancy, a major concern of the overall design has been the means by which it can be implemented and maintained by a small team of coordinating developers. Whether or not this occurs is dependent on to the vagaries of public forces but there is some evidence that, since the original papers were published (Worrall et al. 2007; Worrall 2008), others have begun to build on this initial work, for example, for the *Linux* operating system (Fabbri and Chiozo 2008). It may, however, find its place in the public sphere as a part of a larger project, such as *SciPy*, or the more tightly defined *python(x,y.)*.[46]

There are some clear, and some not-so-clear distinctions between software *design* and software *engineering* as designer and computer musician Bill Buxton emphasizes (2003). Interpretive languages tend to afford individual users the opportunity to do both, but this relies on strong design principles being well implemented, both in the underlying language (*Python* in this instance) as well as extensive framework extensions such as *SoniPy*. By establishing an open-source project based on such design principles, it is hoped that the initial work presented here will be taken up by others who, in turn, will contribute an evolving framework that is useful to the wider sonification community; that data sonification software design 'escapes' from the engineers to become a more widely accepted part of what it means to 'do' sonification than is currently the case.

If perception is not solely a bottom-up process, but incorporates inherited or acquired response biases related to stimulus characteristics and sensory systems that have evolved for multiple, integrated representations of certain features which are meaningful to the perceiver rather than for just single one-to-one reproductions of the physical world, it makes sense to generate multiple sonifications of any particular data set according to the user's developing sense of meaningfulness. In order to accomplish this, it is necessary for sound designers interested in creating dynamic auditory environments to shift their design representation and thinking efforts from creating bespoke hand-crafted solutions using just their own black-boxes, to a dynamic system approach in which design activity is supported by extended state spaces that incorporate psychoacoustic and cognitively-informed transforms. Because of the breadth and depth of detail required, sets of such transforms need to be developed collaboratively using agreed-upon standards.

[46] http://www.pythonxy.com/.

A community-based approach to developing these resources should go some way towards increasing the computational power of the computer as a design tool to *deduce* a wider variety of more effective auditory designs.

A critical appraisal and consequential deeper understanding of the requirements arising from the physical, perceptual and cognitive issues outlined in Chap. 4, together with the data, computational and interface issues outlined in this Chapter, has led to a model for a heterogeneous software framework for data sonification, based on the integration of two existing and mature software platforms: *Csound* and *Python*. Named *Sonipy*, in recognition of its inheritance, the challenge remains to develop and test the capability of the framework as extensively and rigorously as possible. It is hoped that the somewhat modest beginning to this development in Part II of the book encourages others to take up this challenge, or develop strongly reasoned alternatives.

References

Barrass S (1996a) EarBenders: using stories about listening to design auditory interfaces. In: Proceedings of the first Asia-Pacific conference on human computer interaction, APCHI 1996. Information Technology Institute, Singapore

Barrass S (1996b) TaDa! Demonstrations of auditory information design. In: Proceedings of the third international conference on auditory display, ICAD 1996. Xerox PARC, Palo Alto, California

Barrass S (2003) Sonification from a design perspective. In: Proceedings of the ninth international conference on auditory display, Boston, USA, pp 6–9

Ben-Tal O, Berger J, Cook B, Scavone G, Daniels M, Cook P (2002) SONART: the sonification application research toolbox. In: Proceedings of the eighth international conference on auditory display, Kyoto, Japan, pp 2–5

Black A, Taylor P, Caley R (1998) The festival speech synthesis system. http://www.cstr.ed.ac.uk/projects/festival.html

Boulanger R (ed) (2000) The csound book: perspectives in software synthesis, sound design, signal processing, and programming. MIT Press, Cambridge, MA, USA

Bovermann T, Hermann T, Ritter H (2006) Tangible data scanning sonification model. In: Proceedings of the twelfth international conference on auditory display, London, UK, pp 20–23

Burk P, Bencina R (2001) PortAudio–an open source cross platform audio API. In: Proceedings of the ICMC, La Habana, Cuba

Buxton W (2003) Performance by design: the role of design in software product development. In: Proceedings of the second international conference on usage-centered design, 26–29 Oct 2003, Portsmouth, NH, pp 1–15

de Campo A, Frauenberger C, Höldrich R (2004) Designing a generalized sonification environment. In: Proceedings of the tenth meeting of the international conference on auditory display, Sydney, Australia, pp 6–9

Candey RM, Schertenleib AM, Diaz Merced WL (2006) Xsonify sonification tool for space physics. In: Proceedings of the 12th international conference on auditory display, London, UK

Chafe C, Leistikow R (2001) Levels of temporal resolution in sonification of network performance. In: Hiipakka J, Zakharov N, Takala T (eds) Proceedings of eighth international conference on auditory display, Espoo, Finland, pp 50–55

Coulouris G, Dollimore G, Kindberg T (2005) Distributed systems: concepts and design. Addison-Wesley, Boston, MA

Date CJ (2004) An introduction to database systems, 8th edn. Addison-Wesley

Degara N, Nagel F, Hermann T (2013) Sonex: an evaluation exchange framework for reproducible sonification. In: Proceedings of the international conference on auditory display, ICAD 2013. Łódź University of Technology, Łódź, Poland, pp 167–174

Delmotte V (2012) Computational auditory saliency. PhD thesis, Atlanta, Georgia: Georgia Institute of Technology

Fabbri R, Chiozo MF (2008) 'Figusdevpack': a python-based framework for audio and multimedia newbies. In: Proceeding of the 2008 linux audio conference, Cologne

Goudazi V (2018) Sonification and HCI. In: New directions in third wave human-computer interaction, vol 1-Technologies

Goudazi V, Vogt K, Höldrich R (2015) Observations on an interdisciplinary design process using a sonification framework. In: Proceedings of the international conference on auditory display, ICAD 2015, Graz

Hermann T, Baier G, Stephani U, Ritter H (2006) Vocal sonification of pathologic EEG features. In: Proceedings of the twelfth international conference on auditory display, London, UK

Heymann M, Hansen M (2002) A new set of sound commands for R; sonification of the HMC algorithm. In: Proceedings ASA (ed) Statistical computing section. American Statistical Association, Alexandria, VA, pp 1439–1443

Iverson KE (1980) Notation as a tool for thought. The 1979 ACM turing award lecture. Commun ACM 23. http://portal.acm.org/citation.cfm

Kaplan B (1993) Sonification in AVS. In: AVS 1993, Walt Disney World, Lake Buena, Vista, FL, pp 24–26

Kramer G, Walker B, Bonebright BT, Cook P, Flowers J, Miner PN, Neuhoff J (1999) Sonification report: status of the field and research agenda. Technical report. International Community for Auditory Display

Kwinter S (2008) Far from equilibrium: essays on technology and design culture. Edited by Davidson C. Rice University, Actar

Lazzarini V (2000) The sound object library. Organ Sound 5:35–49

Lodha SK, Beahan J, Heppe T, Joseph A, Zane-Ulman B (1997) MUSE: a musical data sonification toolkit. In: The fourth international conference on auditory display, Palo Alto, CA, pp 1119–1134

Luke R, Clement A, Terada R, Bortolussi D, Booth C, Brooks D, Christ D (2004) The promise and perils of a participatory approach to developing an open source community learning network. In: Proceedings of the conference on Participatory Design, Toronto, Canada

Lutz M (1996) Programming python. O'Reilly & Associates, Sebastopol, CA

McCartney J (2002) Rethinking the computer music language: supercollider. Comput Music J 24 (4):61–68

Menges A, Ahlquist S (eds) (2011) Computational design thinking. Wiley

Miele JA (2003) Smith-Kettlewell display tools: a sonification toolkit for Matlab. In: Proceedings of the Ninth international conference on auditory display, Boston, USA, pp 6–9

Pauletto S, Hunt A (2004) A Toolkit for interactive sonification. In: Proceedings of the tenth meeting of the international conference on auditory display, Sydney, Australia, pp 6–9

Piaget J, Inhelder B (1956) The child's conception of space. Translated by Langdon FJ, Lunzer JL. Routledge & Kegan Paul, London, New York

Riber AG (2018) Planethesizer: approaching exoplanet sonification. In: Proceedings of the 24th international conference on auditory display, ICAD 2018. Michigan Technological University

Robertson P (2015) Pyfestival. https://pypi.org/project/pyfestival/

Scavone GP, Cook P (2005) RtAudio and a synthesis toolkit (STK) update. In: Proceedings of the 2005 international computer music conference, Barcelona, Spain

Stockman T, Hind G, Frauenberger C (2005) Interactive sonification spreadsheets. In: Proceedings of the eleventh meeting of the international conference on auditory display, pp 6–9. Limerick, Ireland

Walker BN, Cothran JT (2003) Sonification sandbox: a graphical toolkit for auditory graphs. In: Proceedings of the ninth international conference on auditory display, Boston, USA, pp 6–9

Wilson CM, Lodha SK (1996) Listen: a data sonification toolkit. In: Proceedings of the international conference on auditory display, Palo Alto, California

Worrall DR (2008) Overcoming software inertia in data sonification research using the *SoniPy* framework. In: Proceedings of the inaugural international conference on music communication science. University of NSW, Sydney, Australia, pp 5–7

Worrall DR, Bylstra M, Barrass S, Dean RT (2007) *SoniPy*: the design of an extendable software framework for sonification research and auditory display. In: Proceedings of the 13th international conference on auditory display, Montréal, Canada, pp 26–29

Wright M, Freed A, Momeni A (2003) Open sound control: state of the art 2003. In: Proceedings of the 2003 conference on new interfaces for musical expression. National University of Singapore, Singapore

Part II
Praxis

Chapter 6
The Sonipy Framework: Getting Started

Abstract Having established the design criteria for a comprehensive heterogeneous data sonification software framework (DSDF) in the previous chapter, this chapter introduces two pillars of such a framework, the *Python* and *Csound* programming languages, as integrated through a *Python*-*Csound* Application Programming Interface (API). The result is as a mature, stable, flexible, and comprehensive combination of tools suitable for real and non-realtime sonification, some of the features of which are illustrated in the examples of subsequent chapters.

6.1 Two Pillars of the *SoniPy* DSDF: *Python* and *Csound*

Many of the issues raised in Chap. 5 are generic to software in general. Anecdotally, 'late delivery' and 'over-budget' are as common as they are notorious characteristics of the commercial software industry because, for all but the largest applications, it is difficult for them to maintain hardware and operating-system currency: two of the essential activities necessary to meet their obligations to paying customers if they hope to survive. Some do, of course,[1] but increasingly, developers and users who are more interested in producing a sustained effort than engaging in the perfectly valid experimental practice of exploring new software ideas, are turning to open-source, community-supported tools to sustain ongoing viability in the medium-to-longer-term.

The need for a serious solution to the issues outlined was first undertaken by the author in 2003, as one of a group of sonifiers that encountered multiple difficulties in trying to sonify single large multidimensional EEG datasets (Barrass 2004; Barrass et al. 2005; Dean et al. 2004). While all researchers involved were computationally and technically literate, it became apparent that if the convoluted impediments experienced were any indication, it would be very difficult for almost everyone to achieve consistent, repeatable results under the same conditions with anything but the simplest datasets. This led the author to develop a specification for

[1] *SuperCollider*, for example. McCartney (2002).

D. Worrall, *Sonification Design*, Human–Computer Interaction Series, https://doi.org/10.1007/978-3-030-01497-1_6

an experimental software sonification framework, beginning with the requirements identified in the *Sonification Report*.

The framework is called *SoniPy*, in-keeping with a naming convention used for frameworks that extend the *Python* programming language and is available for all widely-available hardware platforms, as is the sound synthesis language decided upon, *Csound*. These choices were not arbitrary. *Python* possesses all the features of a modern modular programming language that are considered essential for an experimental development environment. It is

> a general-purpose programming language … which may also serve as a glue language connecting many separate software components in a simple and flexible manner, or as a steering language where high-level Python control modules guide low-level operations implemented by subroutine libraries effected in other languages (Watter et al. 1996).

Other descriptors include: simple, but not at the expense of expressive power, extensible, embeddable, interpreted, object-oriented, dynamically typed, upwardly compatible, portable and widely and freely available.[2]

Some reasons for selecting *Csound* as the sound-synthesis base for *SoniPy* are that it is an extremely flexible and mature text-based programming language that has a strong community and is well maintained on very many platforms. *Csound* is

> A powerful and versatile software synthesis program….. Csound transforms a personal computer into a high-end digital audio workstation-an environment in which the worlds of sound-design, acoustic research, digital audio production and computer music composition all join together in the ultimate expressive instrument. (Boulanger 2000, 5).

The nearest alternative to *Csound* for power and flexibility for data sonification sound synthesis is probably Supercollider's *scsynth*. It *is* possible for *Python* to communicate with *scsynth* via the OSC networking protocol, however, compared to the maturity of the *Python-Csound* interface, that interface remains relatively unexplored.

The design details of the *SoniPy* framework was explored in more detail in the previous chapter. This chapter provides a basic practical introduction to the *Python* and *Csound* programming languages, and their *Python* API integration, resulting in a more grounded understanding the foundational basics of the framework and the coded examples with this part of the book.

6.2 A Brief Introduction to *Python*

Python is a high-level, scripting language that is relatively easy to learn. This overview is meant as a brief introduction to the basic concepts for those who do not know the language, and so that the description of the *Python Csound* API discussed

---

[2]http://www.python.org/ A more detailed description of these characteristics are outlined in the previous chapter (Sect. 5.9).

later will be understandable. There is a plethora of study material available, depending on your previous experience with other languages. Unwise as it is to recommend any one text for all readers, *Think Python: How to Think Like a Computer Scientist* (Downey 2012) is a good, reasonably comprehensive beginner's resource, and the online interactive version[3] aids rapid understanding of the basic concepts.

6.2.1 The Python *Interpreter*

Many operating systems come packaged with a basic *Python* interpreter. The simplest way to check if an interpreter is installed, is to open an operating system shell and type `python` <return> at the prompt, which we symbolize here with a '`$`' sign:

```
$ python
```

If there is a *Python* installed on the computer, an interpreter prompt (>>>) will appear, in which case the ideas in this section can be explored. While the default system version may be adequate for learning the basics, it is inadequate for a range of operations necessary for data sonification. So, in order to accommodate these needs, we currently recommend Anaconda's *Python*, version 2 or version 3 (preferred).[4] For a fuller discussion, see the Modules Libraries and Namespaces Sect. (6.2.13), later in this chapter.

In its simplest form, *Python* can be understood as a program that interprets commands passed to it. The program can be run in two modes: interactive and non-interactive. In non-interactive mode it can be run from the operating system's (command line) shell, with a text file of *Python* code commands as input to it. So, if the text file `myPythonCode.py` contains the single line

```
print("Hello world, G'day!")
```

Running *Python* from the command-line shell causes *Python* to execute all the commands in the file:

```
$ python myPythonCode.py
```

[3]http://interactivepython.org/runestone/static/thinkcspy/.

[4]https://www.anaconda.com/download/.

Resulting in the display of the text

```
Hello world, G'day!
```

after which the *Python* interpreter quits. In interactive mode, the *Python* interpreter is run from the shell without a `.py` file of *Python* commands as input:

```
$ python
(python will print some system information here)
>>>
```

The appearance of the *Python* interpreter prompt (>>>) indicates that it is it is ready to accept commands. Let's print the result of adding two integers together:

```
>>> print(3 + 5)
8
>>>
```

The *Python* interpreter prints the result of the addition followed by another prompt (>>>), indicating it is ready to accept more input. You can quit the interpreter and return control back to the operating system's shell, like this:

```
>>> quit()
$
```

The following sections provide a succinct overview of many details of the *Python* language, with minimal commentary. Readers unfamiliar with the language are encouraged to use them as a basis for their own interactive exploration.

6.2.2 Values, Variables, Types and Expressions

```
>>> myVariable ="hello" # variable assignment
>>> type(myVariable)
<class 'str'>
>>> myVariable = 2+3j

>>> type(myVariable)
<class 'complex'>
```

6.2.3 Built-In Types

Built-in types: `string, int, float, tuple, list, range, dict` `of which three are numeric`

```
Integer     <class 'int'>
Floating Point    <class 'float'>
Complex     <class 'complex'>
```

6.2.4 Arithmetic Operators

In the following table, the operators are listed in precedence group order. Operators of the same precedence are evaluated left-to-right.

Symbol	Operation
`**`	Exponentiation
`//`	Integer truncation
`*`	Multiplication
`/`	Division
`%`	Modulus reduction
`+`	Addition
`−`	Subtraction

6.2.5 Boolean Comparison Operators

Expression	Operation
`True`	Truth (==1)
`False`	Falsehood (==0)
`==`	Equivalent to
`>`	Greater than
`<`	Strictly less than
`>=`	Greater than or equal to
`<=`	Less than or equal to
`!=`	Not equal to
`is`	Object identity
`is not`	Negated object identity

6.2.6 Variable Names

1. Can contain both letters and numbers, but have to begin with a letter (including underscore).
2. Can be arbitrarily long.
3. Are case sensitive.
4. Cannot be a *Python* keyword.

The following are *Python* keywords:

```
False, None, True, and, assert, break, class,
continue, def, del, elif, else, except, exec, fi-
nally, for, global, if, import, in, is, lambda,
nonlocal, not, or, pass, print, raise, return,
try, while, with, yield.
```

6.2.7 Variable Assignments

Symbol	Operation
`=`	Assign to
`+=`	Reassign increment
`-=`	Reassign decrement
`*=`	Reassign multiplicative increment
`/=`	Reassign divisive increment
`%=`	Reassign modulus reduction

6.2.8 Ordered Sets: Sequences

Known generically as *sequences* (because order matters), there are four types: lists, tuples, ranges and text strings.

6.2.8.1 Strings

A string is a quote-enclosed, sequence of a characters. Strings can be sliced but are immutable, i.e. they cannot be changed. Single-quoted and double-quoted strings are treated equivalently.

```
>>> txt="The quick brown fox jumps over the lazy dog."
>>> print(txt[26:-1])
over the lazy dog
```

Some arithmetic operators can be applied to strings:

```
>>> aStringy = "abcdefg123"
>>> bStringy = " Goldfish"
>>> print(aStringy *2)
abcdefg123abcdefg123

>>> print(aStringy[:4].upper() + bStringy*4)
ABCD Goldfish Goldfish Goldfish Goldfish
```

Strings can be formatted using the `%` operator,[5] in a way similar to the `printf` function in the C standard library.

```
>>> stringy = "1 min = %d secs.\n1 sec = %.3f msec." \
% (60, 1.0/1000)
>>> print(stringy)
1 min = 60 secs.
1 sec = 0.001 msec.
```

Triple-quoted strings can be spread across multiple contiguous lines. For example:

```
"""Triple-quoted strings allow multi-line comments.
They are also used immediately after the first line in
a function declarations to produce comments delivered
by the Python help() function. See the example in the
section 6.4.10 on functions, below. NB: All strings
(including triple-quoted strings can be defined with
matching ' or " characters. """
```

6.2.8.2 Tuples

A tuple is a left-right parenthesis-enclosed, comma-separated sequence of a data of the same or different types. Tuples can be sliced but are immutable:

```
>>> aTup=(1,2,3,[11,22,"thirty-three"])
>>> print(aTup[3][2])
thirty-three
```

[5]Not to be confused with the modulus operator, as above (Sect. 6.2.4).

6.2.8.3 Lists

A list is left-right square-bracket-enclosed, comma-separated sequence of a datum of the same or different types. Lists are assignable, sliceable and mutable (i.e. they can be expanded, and reduced by addition, deletion, insertion and replacement.

```
>>> myList=[1, 2.1, 1/3, 2/3.0, 'golly', (2,3,4), \
     ["listy", 'This is a \n single-quoted string.']]
```

6.2.8.4 Ranges

The range function takes three arguments: start integer (optional, defaults to 0), the end integer, and the increment (optional, defaults to 1):

```
>>> myRange = range(3,13,2)
>>> type(myRange)
<class 'range'>
>>> print(myRange)
range(3, 13, 2)
```

6.2.9 *Dictionaries*

A dictionary is a left-right brace-enclosed, comma-separated sequence of a `key:value` elements pairs in a comma-separated `list`. A dictionary can use any immutable type as a `key`. For example:

```
transDict = {
1: ['one','un','eins','et','yksi','kubili'],
2: ['two','deux','zwei','to'],
}
>>> type(transDict)
<class 'dict'>
>>> print (transDict [1])
['one', 'un', 'eins', 'et', 'yksi', 'kubili']
>>> print (transDict [1][5])
'kubili   # zulu!
```

New entries in the dictionary can be made by direct assignment:

```
>>> transDict[3] =
['three','trois','drei','tre','kuthathu']
```

A dictionary's keys are available as a list:

```
>>> transDict.keys()
dict_keys([1, 2, 3])
```

A dictionary's values are available as a list:

```
>>> transDict.values()
dict.values([['one', 'un', 'eins', 'et', 'yksi',
'kubili'], ['two', 'deux', 'zwei', 'to'], ['three',
'trois', 'drei', 'tre', 'kuthathu']])
```

6.2.10 Functions

Functions take the following form:

```
def function-name(argument-list):
"Descriptive help string for the function."
—function-body—
return (result)
```

The comma-separated argument-list is optional, as is the `return (result)` statement, which is only necessary if the function needs to return an assignable result on execution. `result` is a Tuple and can be expressed with or without left-right parentheses, i.e. as

```
return a, b
```

Names in the `argument-list` are objects passed by reference. The body of the function has to be indented to a higher level (i.e. to the right). For example:

```
def sum(a, b):
"""Returns the sum of numerics a and b."""
   return ("Sum of %g and %g is %g" % (a, b, a+b))

>>> res = sum(2,3.4)
>>> print(res)
The sum of 2 and 3.4 is 5.4
```

6.2.11 The Scope of Variables

Unless declared otherwise, the scope of names in the `argument-list` of a function is local to the function itself. Also, when an object is assigned a new name within a function, the scope of that binding in local. It must be passed by reference (using the `return` statement) if the value of it is to be retained.

Objects that are in the global namespace (see later) can be accessed directly from within a function if that name is not (re)defined in the function definition itself. A name can be made global using the `global` keyword.

```
>>> AGLOB=[1,2,3]
>>> global AGLOB
```

In order to easily distinguish them from local variables, many programmers follow the convention of using capital letters for global variables. To modify the value of a global variable within a function, it is necessary to declare it as global in the function definition:

```
def myFunc(anInt):
   "Shows a glob var being modified in function."
   global AGLO
   AGLOB += AGLOB + [anInt]
```

If a global variable is used but not modified within the function, it is not necessary to declare it. For example:

```
def myFunc2(aFloat):
   "Shows use of a glob var not declared in function."
   x = aFloat * AGLOB
   return (x)
```

6.2.12 *Flow of Execution: Looping, Iteration, and Flow Control*

Statements are executed one at a time, in order from top to bottom. Defining a function is not the same as executing it. Execution always begins at the first statement of a file and, when it is called, a function. So function calls are like detours in the flow of execution. In the process of executing a function call, that function may itself call other functions and so on.

Conditional Execution is controlled by five types of flow control statements: `while`, `for`, `if-else`, `with` and `try-except`.

6.2.12.1 `while`

```
myInt = 0
while(myInt <10):
   print (myInt)
   myInt +=1
print("FIN")
```

6.2.12.2 `for x in y`

```
listy = [1,2,3,4,5,6,7,8,9]
for item in listy:
   print(item)
```

When the item's ordinal number is required, use the range construct:

```
for counter in range(len(listy)):
   print(counter, listy[counter])
```

6.2.12.3 `if-elif-else`

The simplest form of the `if` statement is a simple conditional test:

```
if x > 0:
print("x is positive")
```

This can be extended into a chain of else-if (`elif`) conditional tests, concluding with a final `else` (which can be omitted if it is unnecessary): Note the indentation. The first un-indented statement marks the end of the conditional block:

```
if x < y:
   print("%g is less than %g" % ( x, y))
elif x > y:
print("%g is less than %g" % ( x, y))
else:
print("%g and %g are equal" % ( x, y))
```

Conditionals can be nested to any depth:

```
if (conditional):
   …
   if (conditional): …
   else …
else: …
```

6.2.12.4 **with**

The `with` keyword is a context manager. It is a condensed form of the older-style `try-except-finally` block. For example, the following code opens and reads the contents of a file `file.txt` and concatenates the lines of the file in a list of strings:

```
with open("file.txt") as filePointer:
    lines = filePointer.read().split("\n")
```

6.2.12.5 **try-except**

The `try-except` conditional structure is used to execute an operation that may cause an exception, but if it does, the program should handle it and continue. In the following code snippet, the user is prompted for a filename. If the file cannot be found, an exception is raised and handled non-abortively.

```
fileFound = False
while not fileFound:
   fname=input("Enter file name (or "q" to quit: ")
   try:
      fp= open(fname, "r")
      fileFound = True
   except:
      if fname=="q": fileFound = True
      else: print("There is no file named",fname)
```

The `try` statement executes the statements in the first block. If no exceptions occur, it ignores the `except` statement. If an exception occurs, it executes the statements in the `except` branch and then continues.

6.2.13 Namespaces, Modules and Libraries

A namespace is an abstract container created to hold a logical grouping of unique identifiers (i.e. names). An identifier defined in a namespace is associated only with that namespace. The core *Python* function `dir()` is used to list the contents of namespaces. Upon loading the Python interpreter, a list of the content of the core namespace can be obtained using it[6]:

```
>>> dir()
['__annotations__','__builtins__','__doc__','__loader__
','__name__','__package__','__spec__'] 6
```

Python core is extended by a modules, including a variety of specialized, often quite extensive or libraries. Some of them, such as *NumPy* (a numerical array processor) and *matplotlib* (a mathematical plotting library) are very widely used. These library modules are made available from central repositories such as python.org, or through independent package managers.

Any user-defined file can be thought of as simple module and imported into *Python* using the import command. So, a file `myModule.py` contains a variable `myVar` and a function `myFun()` can be imported into *Python* with the `import` command. As a result of the import, a new namespace is created for the contents of that imported module, and the objects that belong to it are available through the module's namespace.

```
>>> import myModule # ".py" is omitted when importing
>>> dir()
['__annotations__','__builtins__','__doc__','__loader__
','__name__','__package__','__spec__','myModule']
```

The content of the `myModule` namespace is distinct, and accessed as such:

```
>>> dir(myModule)
['__builtins__','__cached__','__doc__','__file__',
'__loader__','__name__','__package__','__spec__',
'myFun','myVar']

>>> print(myModule.myVar)
2.718281828
```

The `import` function can be used in several forms:

```
>>> import myModule as mym
```

[6]The exact contents of this list depends somewhat on the version of *Python* being used.

Or the contents of the file can be read into the current namespace and executed using the format

```
>>> from myModule import *    # * means "everything"
```

To `import` a single function into the current workspace, use

```
>>> from myModule import myFun
```

6.2.14 Object Orientation: Classes Variables and Methods

This section provides a brief introduction to the way Object-Orientated Programming (OOP) is supported in *Python*. It is not meant as a comprehensive guide to OOP, rather, as a conceptual overview which should make the coded examples in the chapters following more understandable.

All objects in *Python* are of one class or another. For example, a text-string is an instance of the class `str`:

```
>>> type("A bass is a fish and a musical instrument.")
<class 'str'>
```

Python supports the creation of a user-defined object types using the `class` keyword. Code Example 6.1 illustrates some features of a simple class.

```
>>> type(Audio)
<class 'type'>
```

Once a new class has been defined, an instance of it can be created (a process known as instantiation) and assigned to a variable. For example:

```
>>> myAudio = Audio()
```

Class definitions often include new methods, which are defined like functions, to operate on future instances of the class. The first argument to a class method is always the class object, by convention called `self.` A special method named `__init__()` is executed immediately on instantiation to perform custom initialization, such as, in the example below, of associating a instance variable with a particular audio filename.

Here's how to instantiate the `Audio()` class and perform actions with it using the `normalize()` and `bandpassFilter()` methods:

```
>>> myAudio = Audio("notDrowning.wav")
>>> myAudio.normalize(-3.1)
>>> myAudio.bandpassFilter([80, 9000])
```

Classes can contain both *class variables*, that are shared by all instances of the class, and *instance variables*, which are used to support instance-specific details. In Code Example 6.1, `__audiofileFormats` is a class variable and `audiofile` is an instance variable. The reason for differentiating between class and instance variables in the example, is in order to distinguish between all the generic audio waveform file formats which our `Audio` class can process and which are independent of any particular instance of it, and a particular audiofile associated with an instance of the class which can be unique to that instance. Instance variables are assigned or reassigned directly through the instance:

```
>>> myAudio.headroom = -1.5   # allow dB headroom
```

6.2.14.1 Class Inheritance

In OOP, the term *inheritance* is used when a class uses code that it derives from another class. By default, new classes are derived from a basic `type` called `object`, but can also be derived (or inherit from) an existing class, or multiple classes. A child (or sub) class inherits methods and variables from parent (or base) classes. To illustrate this, we'll use the `Audio` class defined earlier as one parent class, and create `Data`, another parent class, and then a new child class of these two parents called `Sonify()`.

In addition to defining its own variables and methods, a child class inherits those of its parent(s), and this allows instances to access the methods and variables of the parent classes as if they were their own. A child class can also override such variables and methods by redefining them (Code Example 6.2).

Here's how to instantiate the `Sonify()` child class and use it perform actions using the `Audio()` and `Data()` parent class structures as well as those unique to the `Sonipy` child class.[7]

```
>>> myson = Sonify("swimrates.csv", "drowning.ogg")
>>> myson.addUIdevice("laptop")
>>> print (myson.UIdevices)
['laptop']
>>> myson.dataStats()
Performing statistical analysis of the data …
>>> print(__audiofileFormats)
["wav", "aiff", "mp3]
```

[7]Copies of the code in all Code Examples is available for download from the book's repository.

```
class Audio():   # The "()" are optional
    """For applying DSP effects on audio files."""
    __audiofileFormats=["wav","aiff","mp3] # a class variable

    def __init__(self, audiofile):
        """ Initialize instances of the Audio class."""
        self.audiofile = audiofile    # an instance variable
        self.headroom = 0.0
        self.bandpassLimits = [0,0]
        # do other stuff on instantiation

    def normalize(self, headroomInDB):
        """A method to amplitude scale an audio file."""
        self.headroom = headroomInDB
        print ("Normalizing %s to %1.3f dB limit of max."
        % (self.audiofile, self.headroom))

    def bandpassFilter (self, lowerFreq, upperFreq):
        """A method to bandpass filter an audio file"""
        self.bandpassLimits = (lowerFreq, upperFreq)
        # stuff to do
```

Code Example 6.1 A simple *Python* class definition: Audio()

Other more esoteric operations on class inheritance not discussed here, include the `super()` function, which enables access to parent methods overwritten in a child class definition, and polymorphism, which provides the ability to leverage the same interface for different underlying forms.

6.3 A Brief Introduction to *Csound*

Csound is a sound and music computing system which was originally developed by Barry Vercoe in 1985 at MIT Media Lab. Since the 1990s, it has been enhanced by an international group of core developers. A wider community of volunteers contribute examples, documentation, articles, and takes part in ongoing *Csound* development with bug reports, feature requests and discussions with the core development team.

The *Csound* language follows earlier prescient work on the Music "N" lineage of acoustic compilers developed in the 1960s by Max Mathews, Joan Miller, F. Richard Moore, John Pierce and Jean-Claude Risset (Mathews 1969). These compilers connect together low-level operational building blocks, called *opcodes*. The *Csound* language is compiled into an efficient set of signal-processing

```
class Data():
    """For handling data files and streams."""
    __datafileFormats=["txt", "csv", "xls", "xlsx","zip", "dhf5"]

    def __init__(self, dataFile):
        """ Initialize instances of the Data class."""
        self.dataFile = dataFile # instance variable
        # do other stuff on instantiation

    def dataStats(self):
        """Produces basic statistical analysis of the data"""
        print("Performing statistical analysis of the data …")
        # stuff to do

from datetime import datetime
class Sonify(Audio, Data):
    """Child class to handle combining parents of a
sonification."""
    def __init__(self, datafile, audiofile):
        """ Initialize instances of the Sonify class."""
        Audio.__init__(self, audiofile) # init Audio() parent
class
        Data.__init__(self, datafile)   # init Data() parent class
        self.sonTime = datetime.now()   # log the sonification
time
        self.bandpassLimits = (80,9000) # uses the Audio method
        self.UIdevices = []             # init Sonify() variable

    self.audiofile = "backflip.wav" # init.an Audio()instance var

def addUIdevice(self, deviceString):
    """ Define sonification output device."""
    if deviceString.strip() not in self.UIdevices:
        self.UIdevices += [deviceString]
    else:
        print(%s already in UI the device list" % deviceString)
```

Code Example 6.2 Multiple class inheritance in *Python*: Data() and Sonify()

operations that are then performed by its audio engine. With over a 1300 such opcodes available to the programmer, *Csound* is a very mature sound-synthesis language. But don't let its historical importance fool you: it is still being actively developed and extended, making it a compilation and real-time performance tool that is arguably unrivalled in the breadth of its sound generation and processing capabilities! The *Csound* compiler can be run

1. From the operating system's shell command line, in which case the configurations of the signal processing operations and their controls are read from text files–often .csd, .orc and/or.sco files.

2. From its various frontends.[8]
3. Through a high-level *Csound* Application Programming Interface (API). The *Python-Csound* API is discussed later in this chapter.

Although *Csound* has a strong tradition as a tool for composing electroacoustic music, it is used for any kind sound that can be made with the help of the computer. *Csound* was traditionally used in a non-interactive score-driven context, but nowadays it is mostly used in real-time. It can run on a multitude of different platforms including all major operating systems, Android and iOS, as well as is a web browser. *Csound* can also be called through other programming languages such as *Python*, *Lua*, *C/C*++ and *Java*.

6.3.1 A Basic Overview of the Csound Language

6.3.1.1 Instruments

The basic structuring unit in *Csound* is the instrument, which is used to configure a series of operators called unit generators that are defined in the *Csound* language by combinations of various operation codes or *opcodes*. An instrument is a model of a sound-processing/generating object that *Csound* will instantiate any number of times necessary, according to the demands of a score. Code Example 6.3 is an example of a very simple instrument definition, as coded in a `.csd` file.

6.3.1.2 Variable Types

Csound has a number of variable types, defined by their initial letter: g, i, k, and a. These types define not only the content type of the data, but the action time. When an instrument is run, there are two distinct action times:

1. Initialization time, a single pass through the code which is used to initialize variables and unit generators, and to execute code that is designed to run once per instance. By default, scores are also time-sorted.
2. Performance time, when instruments process signals, looping over the unit generator code. Performance time has three different time “frames”:
    1. Instance time (init, i-time). Uses scalar data during pre-performance or initialisation time.
    2. Control time (k-time). Uses scalar data during performance.

[8]Available from the *Csound* website: http://csounds.com/.

```
<CsoundSynthesizer>
<CsOptions>
-odac                   ; write output to built-in DAC
;-Wo "testy.wav"        ; write output to file. ";" is a comment
</CsOptions>
<CsInstruments>
sr=48000                ; sampling rate
ksmps=32                ; nr of samples in an audio buffer
nchnls=2                ; 2 output channels
0dbfs=1 ; set the max db value to 1
instr 1 ; the beginning the instrument definition
iamp = 0.5              ; line 2
icps = 440              ; line 3
asig oscili iamp, icps  ; line 4
aout vco2 0.5, 220      ; line 5
outs asig, aout         ; write aout to a stereo buffer
endin                   ; the end instrument definition 1
</CsInstruments>
<CsScore>
i1 0 1 1 0.2 205
i1 0.75 0.5 0.3 500
i1 1.25 0.5 0.4 700
e
</CsScore]
</CsoundSynthesizer>
```

Code Example 6.3 Contents of a simple *Csound*`.csd` file

3. Audio time (sample, or a-time). Uses vector data during performance to produce the audio signal that is sent to the output

In Code Example 6.3, lines labelled `line 2` and `line 3` will execute (once) at i-time, followed by the initialization of the oscillator (at `line 3`). At performance time, line 4 and 5 run in a loop for the required duration, filling and consuming audio signal vectors `asig` and `aout`. The size of these vectors is determined by the system constant `ksmps`, and defaults to 10. The sampling rate is also set by the constant `sr` and its default value is `44,100`. A third quantity, the control rate (`kr`) is set to `sr/ksamps`, determining how often each control (scalar) variable is updated. This also determines the processing loop period (how often the processing code is repeated).

The scope of most variables is local to a single instrument. Global variables can be created by attaching a ‘g’ to the start of the variable name (e.g. `gi1`, `gkNew`, or `gasig`). A common usage of global audio variables is to provide memory locations which can be used as audio buffers for bus-like operations such as signal routing for global reverberation.

6.3.1.3 Parameter Numbers

Instruments can contain arguments that are initialized inside the instrument definition itself. These are identified by parameter numbers `p1`…`pN` and appear on instrument score statements as "p-fields". The first three are fixed:

`p1` The instrument number of the event.
`p2` The start time of the event.
`p3` The duration of the event in seconds.[9]

Subsequent p-fields (`p4, p5` …) can be freely used by instruments.
For instance:

```
instr 1
iamp = p4
icps = p5
asig = oscili, iamp, icps
out (asig)
endin
```

6.3.1.4 Control Variables

Control variables are used to make parameters change over time. For example, for in the instrument code below, a trapezoidal envelope generator `linen` is used to shape the amplitude by gradually ramping the amplitude to and from zero, thus eliminating the clicks at the beginning and end of sounds that would result if the instrument's amplitude was turned on or off instantaneously. For memory efficiency and execution speed, such controls are best specified with k-time rather than a-time variables, as illustrated by `ksig`:

```
instr 1
idur = p3
endin
iamp = p4
icps = p5
ksig = linen iamp,0.1,idur,0.2
asig = oscili ksig, icps
out(asig)
```

[9]If p3 is set to −1, the note will hold as long as Csound is open, unless a "F 0 z" statement is added to the score.

6.4 The *Python-Csound* API (`ctcsound.py`)

The *Csound* API provides complete, stable access to the *Csound* application—from configuration and code compilation to audio sound synthesis and control. This set of functions and classes has been "wrapped" for various languages. In the case of *Python*, the most current wrapper is `ctcsound.py`[10] which is automatically installed on the host computer in the *Csound* installation process. It can be used with both *Python* versions 2.7.x and 3.x.y. It can be imported into a running *Python* instance with the command:

```
>>> import ctcsound
```

Python will produce an error if it cannot find the `ctcsound.py` file. The most common cause of such an error is that the location of `ctcsound.py` is not in the `PYTHONPATH`.

Most of the work will be done with the instantiation of two classes: `Csound()`, which is used to creating a `Csound()` object and `CsoundPerformaceThread()` for performing it, either to file or in real-time as requested by the output options.

A typical use of the `ctcsound` API involves the following steps:

1. Creating a `Csound()` object
2. Setting up options to control the compiler and engine
3. Compiling code
4. Starting and running the audio engine
5. Interacting with the compiled instruments

For ease of reference, and to provide a broad perspective on the system as a whole, an Appendix of variables and methods of the `Csound()` and `CsoundPerformaceThread()` classes is provided at the end of this chapter, together with instructions on how to obtain functional information about them through the *Python* `help()` function. In all cases of conflict, the code in the latest version of `ctcsound.py` is the "One Source of Truth."

6.4.1 The Csound *Class Csound()*

With *Csound* running in a separate thread, new instruments can be run and controls set using the methods of the `Csound()` class, such as `setOption()`. For example, the following code creates a `Csound()` object, assigns it to output audio to the soundcard, compiles some code, and executes it:

[10] The "`ct`" prefix refers to the fact that it is made with the *Python* `ctypes` module.

```
>>> myOrc = """
instr 1
idur = p3
iamp = p4
icps = p5
ksig = linen(iamp,0.1,idur,0.2)
asig = oscili(ksig, icps)
out(asig)
schedule(1,0,1,0dbfs/2.5,440)
schedule(1,1,1,0dbfs/2.5,550)
schedule(1,2,1,0dbfs/2.5,660)
endin"""
>>> cs = ctcsound.Csound ()
>>> cs.setOption("-odac")
>>> cs.compileOrc(myOrc)
>>> cs.start()
>>> cs.perform()
```

Executing the *Python* command `ctcsound.Csound()` creates an instance of (i.e. “instantiates”) the `Csound()` class. It gets an opaque pointer that must be passed to most `Csound()` API functions. This pointer can be accessed from the `Csound()` instance that is passed to callback routines and is also used in instantiating the `CsoundPerformanceThread()` class (see below). Example:

```
>>> cs = ctcsound.Csound ()
--Csound version 6.11 (double samples) May 11 2018
[commit: 25b2e8e53bc924526eaad34e0768a5e866638e94]
libsndfile-1.0.28

>>> type (cs)
<class 'ctcsound.Csound'>

>>> cs.csound()
cs.csound()
140280412919296
>>>
```

6.4.2 *The Csound Class* `CsoundPerformanceThread()`

When running the `Csound` engine, it is often simpler to start a new thread for performance, so that the user can interact with it. For this purpose, the `CsoundPerformanceThread()` class takes a *Csound* object and allows it to be played:

```
>>> cs = ctcsound.Csound()
>>> cs.setOption ("-odac")
>>> cs.compileOrc(code)
>>> cs.start()
>>> perf=ctcsound.CsoundPerformaceThread(cs.csound())
>>> perf.play()
```

The `ctcsound.CsoundPerformanceThread()` class performs a score in a separate thread until the end of score is reached. The playback (which is paused by default) is stopped by calling `stop()`, or if an error occurs. The constructor takes a `Csound()` instance pointer as argument; it assumes that the *Csound* compiler was called successfully (using `csound.compile()`) before creating the performance thread. Once the playback is stopped for one of the above mentioned reasons, the performance thread calls `csound.cleanup()` and returns. Extending the previous code example:

```
>>> import ctcsound
>>> cs = ctcsound.Csound()
--Csound version 6.11 (double samples) May 11 2018
[commit: 25b2e8e53bc924526eaad34e0768a5e866638e94]
libsndfile-1.0.28


>>> cs.csound()
cs.csound()
140280412919296
>>> perf=ctcsound.CsoundPerformanceThread (cs.csound())
>>> type (perf)
<ctcsound.CsoundPerformanceThread object at
0x107464ac8>
>>>
```

6.4.3 Real-Time Event Generation

Code Example 6.4, our final example, illustrates a simple complete integration of *Csound* and *Python* vi the API. Read from a.py file, it puts it all together, generates new score events in real-time. The sound-generating instrument is an enhancement of the `pluck` unit-generator: a simple stereo Karptlus-Strong plucked-string model. Two classes are instantiated: (a) `Csound()` which is responsible for managing the sound-synthesis of events generated by *Python* which passes them in realtime via the `makeCSevent()` function to the `csound.ReadScore()` method, and (b) `CsoundPerformanceThread()` which performs/synthesizes the events in

```
def makeCSevent (*args):
    """Make a csound event and play it in an already active
      Performance thread."""
    eventStr = "".join(map(str, args))
    print (eventStr)                   # delete if necessary
#   csPerf.InputMessage(eventStr+"\n") # alternative push method
#   cs.ScoreEvent(eventStr)            # alternative push  method
    cs.ReadScore (eventStr)            # push the score
# --------- BEGIN csound instrument definition string ------------
myOrcOptions="""
sr     = 48000
kr     = 4800
ksmps  = 10
nchnls = 2
0dbfs  = 1.0"""
pluckIt="""
instr 2
  instrNr      = p1
  istartTime   = p2
  kloc  = p9 ; location
  kbeta = 0  ; spread … for ambi out
  ifn   = 0  ; random
  idur  = p3
  ifreq = p4
  imeth = p6
  imethP1      = p7
  imethP2      = p8
  asig pluck ifreq, p5, p5, ifn, imeth, imethP1, imethP2
  kres line 1, idur, 0  ; so tail does not leave DC offset hanging.
  asig = asig * kres
  if (nchnls == 2) then    ; stereo out
    aL, aR pan2 asig, (kloc % 180) * 180 ; pan stereo
(0=hardL,1=hardR)
    outs aL, aR
   else out asig ; mono out
   endif
endin """
# import ctcsoud API ; instantiate & initialize the Csound() class
import ctcsound
cs = ctcsound.Csound ()       # instantiate the Csound() class
cs.setOption("-odac")         # DAC output
cs.SetOption("-b 4096")       # output buffer size
cs.setOption("-d")            # suppress displays ; test!
# --------- instantiate the RT perf thread ------------------
csPerf = ctcsound.CsoundPerformanceThread(cs.csound())
cs.readScore("f 0 z \n") # keep Csound running for a long time...
# -------- sound "fonts" for "rendering" ---------------------
cs.readScore("f 20 0 0 1 \"../../sounds/agogo1-H.wav\" 0 0 0 \n")
cs.compileOrc(myOrcOptions)  # SR, NrChans
cs.start()                    # start CS synth
csPerf.play()                 # start separate performance thread
cs.CompileOrc(pluckIt)        # compile the pluckIt instrument
def passCallback():
    pass                      # suppress all terminal output from
CS
if not TRACE:
    cs.setMessageCallback(passCallback())
```

Code Example 6.4 Illustration of a simple complete integration of *Csound* and *Python* via the API a separate thread until the end of score is reached. This dual-threaded process permits the realtime generation of events and their synthesis to proceed without them interfering each other.

6.5 Summary

An understanding of the necessity for including features for a comprehensive software sonification shell that lie outside, or on the periphery of, computer music composition tools as they are currently envisaged, resulted in choosing *Python* and its encapsulation of *Csound* as a firm foundation for sound synthesis and data processing for sonification. But this is only the foundation of a flexible and powerful platform, that not only integrates sound synthesis and data acquisition, storage and analysis, but acts as a complete heterogeneous software framework for data sonification research and auditory display design.

Appendix: The Main *Python-Csound* API Methods

This appendix consists of a collection of tables of the Methods (i.e. functions) of the two substantive classes in the *Python-Csound* API: Csound() and CsoundPerformanceThread(). There is extensive help for their Methods available through the *Python* `help()` command.

Python help for `ctcsound` is cumbersome, as is likely the help for the `Csound()` and `CsoundPerformanceThread()` classes. However, descriptions of the Methods listed here are available as follows:

```
>>> import ctcsound
>>> cs = ctcsound.Csound() # instantiation of Csound() class.

>>> help (cs.version)
>>> myPerf = ctcsound.CsoundPerformanceThread(cs.csound())
>>> help (myPerf.status)
```

By convention, documentation of *Python* objects such as functions and classes, is emplaced as a triple-quoted string on the line following its definition. The *Python* help() function accesses this string as object __doc__ (sometimes called a "double-underscore-doc" string). Sometimes, such as when assisting vision-impaired users, it is convenient to access it directly so it can be passed to a text-to-speech converter. This can be achieved, in extension of the above example, as follows:

```
>>> docum = ctcsound.Csound.version.__doc__
>>> docum = cs.version.__doc__    # equivalent to the previous line
>>> print (docum)
"Returns the version number times 1000 (5.00.0 = 5000)."
>>>
```

Python-Csound *API Error Codes*

These global error codes are defined in the API and used by the other classes described here (Table 6.1).

There is no specific help available for them in *Python*. They can be retrieved explicitly using `ctcsound.Code`, For example:

```
>>> import ctcsound
>>> ctcsound.CSOUND_MEMORY
-4
>>>
```

Python-Csound *API Csound() Methods*

Tables 6.2, 6.3, 6.4, 6.5, 6.6, 6.7, 6.8 and 6.9.

Table 6.1 The Csound API error codes

Error code	Val	Explanation
CSOUND_SUCCESS	0	Completed successfully
CSOUND_ERROR	-1	Unspecified failure
CSOUND_INITIALIZATION	-2	Failed during initialization
CSOUND_PERFORMANCE	-3	Failed during performance
CSOUND_MEMORY	-4	Failed to allocate requested memory
CSOUND_SIGNAL	-5	Termination requested by SIGINT or SIGTERM

Table 6.2 The Csound() class instantiation methods

Instantiation methods
__init__(hostData = None, pointer_ = None)
__del__()
csound()
version()
APIVersion()

Table 6.3 The Csound() class performance and input/output methods

Performance methods	General input/output methods
parseOrc(orc)	outputName()
compileTree(tree)	setOutput(name, type_, format)
deleteTree(tree)	outputFormat()
compileOrc(orc)	setInput(name)
evalCode(code)	setMIDIInput(name)
compileArgs(*args)	setMIDIFileInput(name)
start()	setMIDIOutput(name)
compile_(*args)	setMIDIFileOutput(name)
compileCsd (csd_filename)	setFileOpenCallback(function)
compileCsdText (csd_text)	setRTAudioModule(module)
perform()	module(number)
performKsmps()	inputBufferSize()
performBuffer()	outputBufferSize()
stop()	inputBuffer()
cleanup()	outputBuffer()
reset()	spin()
sr()	addSpinSample(frame, channel, sample)
kr()	spout()
ksmps()	spoutSample(frame, channel)
nchnls()	rtRecordUserData()
nchnlsInput()	rtPlaydUserData()
get0dBFS()	setHostImplementedAudioIO(state, bufSize)
currentTimeSamples()	audioDevList(isOutput)
sizeOfMYFLT()	setPlayOpenCallback(function)
hostData()	setRtPlayCallback(function)
setHostData(data)	setRecordOpenCallback(function)
setOption(option)	setRtRecordCallback(function)
setParams(params)	setRtCloseCallback(function)
params(params)	setAudioDevListCallback(function)
debug()	
setDebug(debug)	

Table 6.4 The Csound() class realtime MIDI I/O methods

MIDI I/O methods	
setMIDIModule (module)	setExternalMidiOutOpenCallback (function)
setHostImplementedMIDIIO (state)	setExternalMidiWriteCallback (function)
midiDevList (isOutput)	setExternalMidiOutCloseCallback (function)
setExternalMidiInOpenCallback (function)	setExternalMidiErrorStringCallback (function)
setExternalMidiReadCallback (function)	setMidiDevListCallback (function)
setExternalMidiInCloseCallback (function)	

Table 6.5 The Csound() class score handling, messages and text methods

Score handling methods	Messages and text methods
readScore(sco)	message(fmt, *args)
scoreTime()	messageS(attr, fmt, *args)
isScorePending()	messageLevel()
setScorePending(pending)	setMessageLevel(messageLevel)
scoreOffsetSeconds()	createMessageBuffer(toStdOut)
setScoreOffsetSeconds(time)	firstMessage()
rewindScore()	firstMessageAttr()
setCscoreCallback(function)	popFirstMessage()
	messageCnt()
	destroyMessageBuffer()

Table 6.6 The Csound() class channels, control and events methods

Channels, control and events methods	
channelPtr(name, type_)	setInputChannelCallback(function)
listChannels()	setOutputChannelCallback(function)
deleteChannelList(lst)	setPvsChannel(fin, name)
setControlChannelHints (name, hints)	pvsChannel(fout, name)
controlChannelHints(name)	scoreEvent(type_, pFields)
channelLock(name)	scoreEventAbsolute(type_, pFields, timeOffset)
controlChannel(name)	inputMessage(message)

(continued)

Table 6.6 (continued)

Channels, control and events methods	
setControlChannel(name, val)	killInstance(instr, instrName, mode, allowRelease)
audioChannel(name, samples)	registerSenseEventCallback(function, userData)
setAudioChannel(name, samples)	keyPress(c)
stringChannel(name, string)	registerKeyboardCallback(function, userData, type_)
setStringChannel(name, string)	removeKeyboardCallback(function)

Table 6.7 The Csound() class tables and function table display methods

Table methods	Function table display methods
tableLength(table)	setIsGraphable(isGraphable)
tableGet(table, index)	setMakeGraphCallback(function)
tableSet(table, index, value)	setDrawGraphCallback(function)
tableCopyOut(table, dest)	setKillGraphCallback(function)
tableCopyIn(table, src)	setExitGraphCallback(function)
table(tableNum)	
tableArgs(tableNum)	
isNamedGEN(num)	
namedGEN(num, nameLen)	

Table 6.8 The Csound() class opcodes methods

Opcode methods	destroyThreadLock(lock)
namedGens()	createMutex(isRecursive)
newOpcodeList()	lockMutex(mutex)
disposeOpcodeList(lst)	lockMutexNoWait(mutex)
setYieldCallback(function)	unlockMutex(mutex)
createThread(function, userdata)	destroyMutex(mutex)
currentThreadId()	createBarrier(max_)
joinThread(thread)	destroyBarrier(barrier)
createThreadLock()	waitBarrier(barrier)
waitThreadLock(lock, milliseconds)	sleep(milliseconds)
waitThreadLockNoTimeout(lock)	spinLock(spinlock)
notifyThreadLock(lock)	spinUnlock(spinlock)
appendOpcode(opname, dsblksiz, flags, thread, outypes, intypes, iopfunc, kopfunc, aopfunc)	

Table 6.9 The Csound() class miscellaneous methods

Miscellaneous methods	queryGlobalVariable(name)
runCommand(args, noWait)	createGlobalVariable(name,nbytes)
realTime(timerStruct)	queryGlobalVariableNoCheck(name)
CPUTime(timerStruct)	destroyGlobalVariable(name)
rand31(seed)	createCircularBuffer(numelem, elemsize)
randomSeedFromTime()	initTimerStruct(timerStruct)
seedRandMT(initKey)	env(name,withCsoundInstance = True)
randMT(state)	readCircularBuffer(circularBuffer,out, items)
utilityDescription (name)	peekCircularBuffer(circularBuffer,out, items)
runUtility(name,args)	writeCircularBuffer(circularBuffer,in, items)
setLanguage(lang_code)	flushCircularBuffer(circularBuffer)
listUtilities()	destroyCircularBuffer(circularBuffer)
closeLibrary(library)	openLibrary(libraryPath)
setGlobalEnv (name,value)	getLibrarySymbol(library,symbolName)

Python-Csound *API CsoundPerformanceThread() Methods*

Table 6.10.

Table 6.10 The CsoundPerformanceThread() class methods

CsoundPerformanceThread() methods	
__init__(csp)	togglePause()
__del__()	stop()
isRunning()	record(filename, samplebits,numbufs)
processCB()	stopRecord()
setProcessCB(function, data)	scoreEvent(absp2mode,opcod, pFields)
csound()	inputMessage(s)
status()	setScoreOffsetSeconds(timeVal)
play()	join()
pause()	flushMessageQueue()

References

Barrass S (2004) Listening to the mind listening call for submissions. www.icad.org/websiteV2.0/Conferences/ICAD2004/concert_call.htm

Barrass S, Whitlaw M, Potard G (2005) Listening to the mind listening. In: Media international Australia incorporating culture and policy, February 2005

Boulanger R (ed) (2000) The Csound book: perspectives in software synthesis, sound design, signal processing, and programming. MIT Press, Cambridge

Dean RT, White G, Worrall DR (2004) Mind your body. A collaborative sonification. In: Presented at the international conference on auditory display, the studio, Sydney Opera House

Downey AB (2012) Think Python: how to think like computer scientist. O'Reilly Media, Sebastopol. https://greenteapress.com/wp/think-python/

Mathews MV (1969) The technology of computer music. The MIT Press, Cambridge

McCartney J (2002) Rethinking the computer music language: supercollider. Comput Music J 24 (4):61–68

Watter A, van Rossen G, Ahlstrom J (1996) Internet programming with Python. M&T: Books, New York

Chapter 7
Audification Experiments: Market Data Correlation

> The flexibility of money, as with so many of its qualities, is most clearly and emphatically expressed in the stock exchange, in which the money economy is crystallized as an independent structure just as political organization is crystallized in the state. The fluctuations on exchange prices frequently indicate subjective-psychological motivations, which, in their crudeness and independent movements, are totally out of proportion in relation to objective factors (Simmel 1979, 326).

Abstract Despite intensive study, a comprehensive understanding of the structure of capital market trading data remains elusive. The one known application of audification to market price data reported in 1990 that it was difficult to interpret the results, probably because the market does not resonate according to acoustic laws. This chapter illustrates some techniques for transforming data so it *does* resonate; so audification may be used as a means of identifying auto-correlation in trading- and similar-datasets. Some experiments to test the veracity of this process are described in detail, along with the computer code used to produce them. Also reported are some experiments in which the data is sonified using a homomorphic modulation technique. The results obtained indicate that the technique may have a wider application to other similarly structured time-series datasets.

7.1 Introduction

There are many reasons, both sociological and technical, why capital markets, are an interesting application—domain for sonification.[1] Sociologically, they have become a powerful, some might say almost religious, contemporary force, even as

[1]The nomenclatures used for various kinds of markets is somewhat convoluted and are often misused. A stock market is for trading equities, namely ownership interests in companies. A financial market is an institutional structure or mechanism for creating or exchanging financial assets, namely those that are real such as land, buildings, equipment, patents. The term capital market is used generally to refer to markets for stocks, bonds, derivatives and other investments and it is this term we adopt here, or the less formal The Market. Both are generic terms signifying all regulated exchanges and their activities.

D. Worrall, *Sonification Design*, Human–Computer Interaction Series, https://doi.org/10.1007/978-3-030-01497-1_7

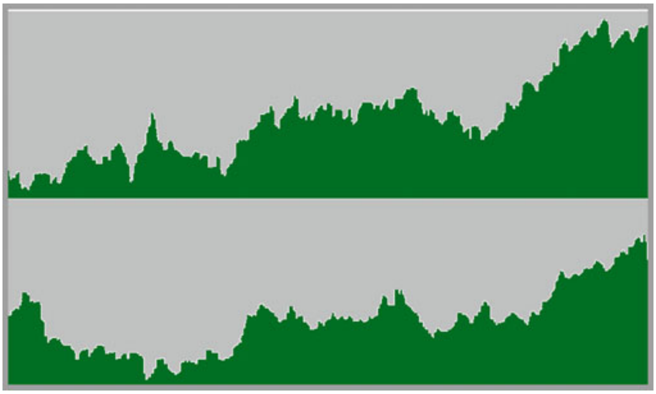


Fig. 7.1 Only one of these two graphs is of a real market. Which one is it?

their overtly emotional expressive open-outcry marketplaces, have become virtualised gatherings of disembodied screen traders, and/or algorithmic trading systems. Sonification of the activities of these markets can be thought of as a form of re-embodiment.

While such sociological considerations form the background and motivations for the experiments discussed in this chapter, the focus here is technical. The statistical analysis of trading data, and stochastic modelling using that analysis, is an area of ongoing quantitative research in finance. Two principal concerns are (1) to find techniques to accurately describe the way prices are distributed statistically, and (2) to what extent auto-correlation exists and can be detected, even pre-empted, as market prices evolve. Understanding the first is important for the risk analysis of various trading instruments in the longer term, and understanding the second is important in attempts to predict future, especially catastrophic, events.[2]

The power of visual representation to enhance and deepen an understanding of phenomena through their data abstractions is undisputed. However, as with many time-domain processes, a visual representation does not always reveal the structure of the data. Mandelbrot, arguably the most well-known Quant,[3] asserts that is not possible to *visually* distinguish graphs of real market data, and those of Brownian-motion-generated data (Mandelbrot and Hudson 2004, 16–18) (Fig. 7.1).[4]

[2]The bottom chart. The top chart is derived from a random-walk model.

[3]The term 'Quant' is used in the field to identify those who quantitatively analyze capital market data, or use such analysis to construct investment portfolios with a specific risk profile. See Sect. 2.2.

[4]Brownian motion is an independent (that is uncorrelated) random walk in which the size and direction of the next (price) move is independent of the previous move(s). A *statistical* analysis of time series data is concerned with the distribution of values without taking into account their sequence in time.

This leads to a data sonification question, which the experiments described in this chapter seek to address: Can trading data be presented in a way that permits its auto-correlation to be aurally distinguished from statistically similar, but uncorrelated, data?

These experiments were conducted in the context developing *SoniPy*, the heterogeneous software framework for the sonification of large multidimensional datasets which is discussed in Chap. 5, so the primary emphasis was tool development and the perceptual testing was informal rather than empirically verified.[5]

7.2 The Data

The importance of understanding the data itself before attempting to sonify it has been emphasized by many authors. Sarah Bly enunciated the imperative:

> Whatever the technique, a first concern is to understand as much as possible about the data itself…With multivariate data, it is important to know not only the relationships between various variables but also which variables are predominant in carrying the information about that data (Bly 1994).

And more than a decade later, John Flowers again emphasized the need to proceed judiciously rather than hoping to 'strike it lucky':

> There may be some auditory analogies to visual plot clustering (auditory scatterplots offer a primitive example) but it will take a great deal of display engineering, and judicious variable selection to make this principle apply to more complex multivariate data. Insightful data groupings are not likely to pop out by simply submitting a multivariate set to a sonification engine and "plotting everything." (Flowers 2005)

This idea is intensified by the understanding, as discussed in Chaps. 3 and 4, that is the *intended* effect that ultimately identifies information in data and perceptual truth relies on frameworks of meaning. As will be discussed later, it is important in this particular circumstance, to distinguish between real trading data, simulated trading data, as they usually, arise out of different processes and can have quite different structural characteristics.

7.2.1 Context

Capital markets are (increasingly virtual) places where companies and traders converge around an exchange to raise new investment capital, and investors and speculators trade exchange-registered securities such as stocks (shares), bonds, currencies, futures and their derivatives. These exchanges have strict government-regulated

[5]See Chap. 3, Table 3.1 for a description of different ways of acquiring knowledge.

mechanisms for such activity and the community of freely-participating individuals around them communicate more-or-less informally with each other, and formally through exchange-registered brokers, who themselves provide information to their clients about specific trading activity, as well as about other more general environmental (financial, political, meteorological etc.) conditions that may affect an individual's trading decisions. Such decisions, enacted by the brokers, cause excitations of the trading system, known colloquially as a *trading engine*, which in turn produces data records of its activities. Some of that data, and various summaries of it, are fed back for the information of market participants, and some of which results in the exchange (buying and selling) of securities. In turn, these marketplaces, operate as systems within national and global economies. Each exchange's trading system is thus designed to be acephalously appropriate for the types of securities that are traded on it.

Trading engines need to be fast, efficient and accurate.[6] They generate large quantities of data, reflecting the moment-to-moment shifting situation of the order book of each of their trading securities as potential buyers and sellers adjust their declared positions, and then sometimes undertake trades. Security Trading datasets are sets of time-ordered trading events having a number of empirical dimensions, such as price and volume,[7] depending on the particular type of security being traded (share, futures, options etc.) and from which other data may be derived.

7.2.2 *Quantitative Analysis*

As a discipline in finance, quantitative analysis begins with Bachelier's speculation that the market is efficient, and thus price movement must be a random walk (Bachelier 1967). At the time of his speculation in 1900, the mathematics of random walks was well known and an analysis of markets, in terms of it, enabled the construction of portfolios of stocks with defined risk profiles. Benoît Mandelbrot's study of price action in cotton led him to question this received wisdom, and to develop another mathematics, which he called *fractal*, to accommodate his analytic findings. Fractal mathematics has become a widely applicable tool in many fields, both analytic and generative, and continues to be a basis for contemporary quantitative analysis.

[6]It is somewhat ironic that, in an enterprise that relies on 'clean' data, financial data often requires considerable 'washing' before its sonification can be undertaken. This situation is exacerbated by the trait that, with datasets over a certain size, the use of metadata tagging is uncommon, principally because it significantly increases the overall size of the dataset, even though the omission increases the likelihood of error. In any event, any cleaning has to be undertaken algorithmically and so it is expedient to have the tools for doing so integrated with the sonification software being used.

[7]The term 'volume' is used throughout to mean 'trading volume' not a psychoacoustic parameter.

Quantitative analysts generally use statistical analysis and stochastics to model the risk profiles of market indices, segments and individual securities as accurately as possible so as to assist in the construction of investment portfolios with certain characteristics, such as risk exposure. To underestimate the risk of a portfolio is to court calamity, while overestimating it invites lower returns than might have otherwise been possible. Quantitative analysis, especially of high-frequency data, remains an area of active research (Farmer et al. 2005) and readable introductions are available in the popular science press (Mandelbrot and Hudson 2004; Peters 1991; Skiena 2004; Sornette 2003).

7.3 Review of Previous Work

7.3.1 Survey of the Sonification of Financial Data

The first users of technology-enabled financial market data sonification were probably the *bucket-shop* traders in the early years of the twentieth century, who were reputed to be able to differentiate the sounds of stock codes, and prices that followed, from the sounds made by their stock-ticker machines as they punched recent trading information, telegraphed from an exchange, into a strip of rolling paper tape (Lefèvre 1923). Janata and Childs suggest that Richard Voss may have been the first to experiment with the sonification of historical financial data: stock prices of the IBM corporation (2004). This is possible, as Voss and Mandelbrot, were research collaborators in fractal mathematics at IBM's Thomas J. Watson Research Center and Voss played an early seminal role in the visualization of fractal structures, and in the analysis of the fractal dimensions of music and speech (Voss and Clarke 1975, 1978).

Kramer and Ellison used financial data in the early 1990s to demonstrate multivariate sonification mapping techniques (Kramer and Ellison 1991). This work was later summarized and published with sound examples (Kramer and Ellison 1991). The trading data used included four-and-a-half years of the weekly closing prices of a US stock index, a commodity futures index, a government T-bond index, the US federal funds interest rates, and value of the US dollar. Mappings were to pitch, amplitude and frequency modulation (pulsing and detuning), filter coefficients (brightness) and onset time (attack). Mapping concepts included redundant mapping and datum highlighting (beaconing).[8]

Ben-Tal et al. sonified up to a year's end–of–day data from two stocks simultaneously by mapping them to perceptually distinct vowel-like sounds of about one second duration (Ben-Tal et al. 2002). A single trading day was represented as a single sound burst. The closing price for the day was mapped to the center frequency, and the volume of trade to the bandwidth. These values were scaled such

[8]See Chap. 2 for a discussion of common sonification techniques.

that the parameters for the last day of trade in each period corresponded to a reference vowel. Closing price was mapped to the number of sound bursts and volume (the number of trades) to duration. They informally observed that they could categorize high volume, high price trading days as loud, dense sounds, while low volume, low price days were heard as pulsed rhythmic sounds.

Brewster and Murray tested the idea that traders could use sounds instead of line-graphs to keep track of stock trends when they are away from the trading floor (Brewster and Murray 2000). Using Personal Digital Assistants with limited screen space over a wireless network, one month of (presumably intraday) price data for a single share was mapped to pitch via MIDI note numbers. Participants, all students whose previous trading experience was unreported, were required to try to make a profit by buying and selling shares while monitoring price movement using either line or sound graphs. As trade transaction costs appear not to have been factored into the calculations, profits or losses were presumably gross. The experimental results showed no difference in performance between the two modes, but participants reported a significant decrease in workload when they used the sonification, as it enabled them to monitor the price aurally while simultaneously using the visual display to execute trades.

Nesbitt and Barrass also undertook a multimodal sonification and visualization study of market depth, to test whether subjects could predict the price direction of the next trade (2002).[9] They used real data from a single security's order book. The visualization used a landscape metaphor in which bid and ask orders (to buy and sell), were 'banked' on either side of a river, the width of which thus represented the size of price gap between the highest bid and the lowest ask, known as the 'bid—ask spread'. A wider river implied slower flow (fewer trades) and so on. The sonification employed the metaphor of an open-outcry market. A sampled male 'buy' and a female 'sell' voice displaying a discretely partitioned dataset (price, volume, price-divergence) was mapped into a discretely partitioned three-dimensional 'importance space' (pitch, loudness, stereo-location). This experimental design illustrates how sonification can be used to assist the apprehension of data segmentation such as where the trajectory of a parameter under focus changes.

Janata and Childs developed *Marketbuzz*, as an add-on to conventional trader's terminals, such as those by Bloomberg, for the sonification of real-time financial data (2004). They used it to evaluate tasks involving the monitoring of changes in the direction of real-time price movements, with and without auditory or visual displays. A significant increase in accuracy using auditory displays was reported, especially when traders were visually distracted by a simultaneous diversionary "number-matching" task. Further, Childs details the use of sonification to highlight

[9]Market depth is a term used to denote the structure of potential buy and sell orders clustered around the most recently traded price. Traders are able to withdraw or restructure their orders at any time before they are "met" by an order from a trader taking the other side of the trade.

significant price movements relative to opening price, as well as continually changing features of Stock Options (2005).

Mezrich, Frysinger and Slivjanovski developed a dynamic representation, employing both auditory and visual components, for redundantly displaying multiple multivariate time-series (1984). Each data variable was represented by a particular timbre. The values of the variable were mapped to pitch. The analyst could focus on a subset of the data by interactively brightening or muting individual variables and could play the data both forwards and backwards. Subsets of the data could be saved and juxtaposed next to each other in order to compare areas where the data might be similar. In almost all cases, the sonified data performed as well as, or better, than the static displays.

In the 1990s Frysinger experimented with playing back market price, data directly as a sound waveform. He reported that he found that “the results proved difficult to interpret, probably because the stock market does not follow physical-acoustic resonance laws” resulting in natural or ‘ecological’ sounds that can be understood from everyday listening, experience (Frauenberger et al. 2007; Frysinger 1990, 2005). There appears to be no further reports of security data audification prior to the work reported here. Hayward also suggested that another reason audification fails for arbitrary data such as stock market figures, or daily temperatures, is the amount of data required: even at low sampling rates, it is difficult to make a sound with a duration long enough to reveal valuable information to the listener (Hayward 1994). We return to the specific concerns expressed by Frysinger and Hayward later in this chapter and the next. Other applications of audification to time series data are discussed in Chap. 2.

Two other types of sonifications of securities data demonstrate different motivations but are mentioned here for completeness. The first is Ciardi’s *sMax*, a toolkit for the auditory display of parallel internet-distributed stock-market data (2004). *sMax* uses a set of *Java* and *Max* modules to enable the mapping and monitoring of real time stock market information into recognizable musical timbres and patterns. The second is Mauney and Walker’s rendering of dynamic data specifically for peripheral auditory monitoring. The system reads and parses simulated real-time stock market data that it processes through various gates and limiters to produce a changing soundscape of complementary ecological sounds (2004).

There are a number of studies, principally those whose purpose was the study of parameter-mapping and auditory graphs, which have been omitted from this survey because it is not clear that there is anything in the findings specific to the structure of financial data, principally because, unless it is generated using very advanced modelling techniques, fictional data is unlikely to exhibit the same structural characteristics as real financial time series data.

7.3.2 Sonification of Stochastic Functions

Aside from their use in algorithmic music composition, stochastic functions have received scant attention in sonification research. Perhaps the first was a study of the use of parameter-mapping and physical model sonification in a series of experiments monitoring the performance of Markov chain Monte-Carlo simulations for generating statistical data from higher dimensional probability density functions (Hermann et al. 2001) and parameter ‘tuning’ in the Hybrid Monte-Carlo algorithm, following the addition of some basic sound-generating tools in the statistical package *R* (Heymann and Hansen 2002). The informal auditing of a technique to sonify, using amplitude modulation, cross-correlations in irregularly spiking sequences that resemble a Poisson process, led to the postulation that the use of sonification for time series analysis is superior to visualization in cases where the intrinsic non-stationarity of an experiment cannot be ruled out (Baier et al. 2005). Time series data was generated by shaping a uniform distribution (white noise) with a cumulative probability density function, in a differentiation study of the perceptualization of some statistical properties of time series data generated using a Lévy skew alpha-stable distribution.[10] This distribution is of interest to modelers of financial time series (Skiena 2004). The study found no evidence that skewness in their data was perceivable, but participants were able to distinguish differences in kurtosis, which correlated with roughness or sharpness of the sound. This research provided empirical support for a part of the initial informal findings of the experiments outlined below. Since those studies, there has apparently been no published work on the sonification of stochastic functions except one study on the sonification of experimentally-observed Brownian motion organized into optical structures (McKell 2016) and an experiment in translating the complex, multi-dimensional data from simulations of biomolecules into intuitive knowledge by using Markov models to map features of a biomolecular energy landscape to sonic parameters, structural animations and free energy diagrams (Arbon et al. 2018). Neither of these explorations report any independent empirical validation of the techniques.

7.4 Experiments: Audification of Security Index Returns[11]

The three experiments described here sought to (a) discover an effective way to directly audify a capital market trading dataset, that preserved its autocorrelation characteristics and (b) ascertain informally whether such a dataset can be aurally discriminated from an audification of a statistically equivalent uncorrelated dataset. The null hypothesis in each case was that no distinction could reliably be made

[10]Similar to that used by Xenakis for his ST series of compositions (Xenakis 1971, 136–43).

[11]The audio examples, data and *Python* code used to calculate, plot and sonify examples in this chapter are available from the online repository for this book.

between the audifications. Whilst a *statistical* analysis of time series data is concerned with the distribution of values without taking-into-account their sequence in time, changing the sequence of values in a time series, known as decorrelation, completely destroys the frequency information while keeping the statistical properties invariant.

7.4.1 The Dataset

The dataset chosen consists of twenty-two years of the daily closing price of All Ordinaries Index (ticker XAO) of the Australian Securities Exchange (ASX),[12] as illustrated by the plot in Fig. 7.2. The highlight and enlargement insert is of the market movement, over 500 days leading up to and following the dramatic market action on "Black Tuesday" (20 October 1987); the largest one-day percentage decline in the history of western markets.[13]

The first task was to find a way to overcome the non-resonance problem referred to earlier, as discussed by Hayward (op. cit.); one that transformed the dataset to be suitably oscillatory, while preserving its correlational integrity. An equivalent problem is to be found in quantitative analysis:

> The *price* of an asset as a function of time is perhaps the most natural financial time series, but it is not the best way to manipulate the data mathematically. The price of any reasonable asset will increase *exponentially* with time, but most of our mathematical tools (e.g. correlation, regression) work most naturally with linear functions. The *mean* value of an exponentially-increasing time series has no obvious meaning. The *derivative* of an exponential function is exponential, so day-to-day changes in price have the same unfortunate properties (Skiena 2004).

The Net Return, or simply, the Return, is a complete and scale-free summary of contiguous investment performance that oscillates from positive values (increase) to negative values (decrease) around zero (no change). For an asset whose price changed from p_t at time t to $p_{t+\partial t}$ at time $t + \partial t$, the simple linear return R_{lin} is defined as $R_{lin} = p_{t+\partial t} - p_t$. Because prices tend to move exponentially over longer timeframes, that is, in percentage terms, a better measure than R_{lin} is the ratio of successive price differences to the initial prices.

$$R_{net} = \frac{p_{t+\partial t} - p_t}{p_t}$$

[12]The XAO is the broad Australian market indicator, a composite of the 500 largest companies listed on the exchange, weighted by capitalization. Contextual details are available at https://www.asx.com.au/products/indices.htm.

[13]See http://en.wikipedia.org/wiki/Black_Monday_(1987).

These are known as Net Linear Returns (Sornette 2003, 36). The XAO dataset was converted to Returns. The statistical properties of these returns, summarized in Table 7.1, clearly shows that they are not a normally-distributed.

Figure 7.3 is a plot of these Net Returns. The insert is of the first 500 samples, similar to the insert in Fig. 7.2.

The difference between the largest minimum Return in October 1987 ("Black Tuesday"), as highlighted in Figs. 7.2 and 7.3, and the second-largest minimum Return is 62% of the total returns space. This is shown in the histogram of Fig. 7.4, which illustrates the frequency of Net Returns. The single minimum and second-largest minimum are circled, but barely visible at this scale.

So, despite its anecdotal interest, an 'audacious' clipping of the largest minimum sample to that of the second-largest minimum was performed so as to provide a less clustered sample space for the dataset as a whole. The resulting returns are plotted in Fig. 7.5.

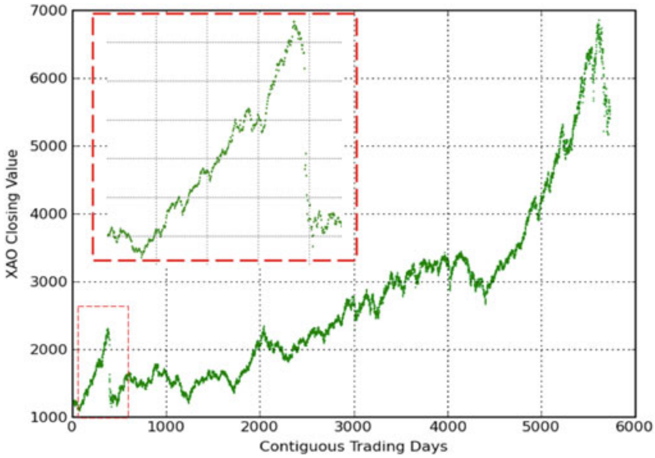


Fig. 7.2 A plot of 22 years of daily closing values of the ASX's XAO, highlighting the market action on "Black Tuesday" (20 October 1987)

Table 7.1 Statistical properties of the XAO Net Returns under study

Statistic	Value
Number of samples	5725
Minimum sample	0.333204213
Maximum sample	0.05886207483
Arithmetic mean	2.1748845e04
Variance	9.5685881e-05
Skewness	7.6491182
Kurtosis	241.72988

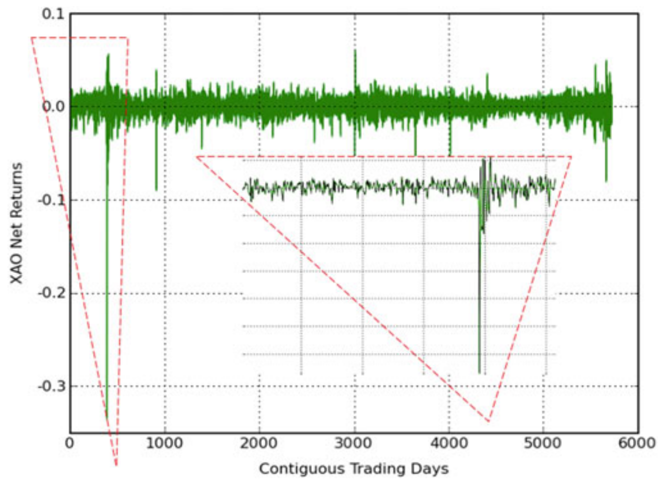


Fig. 7.3 A plot of Net Returns of the XAO dataset

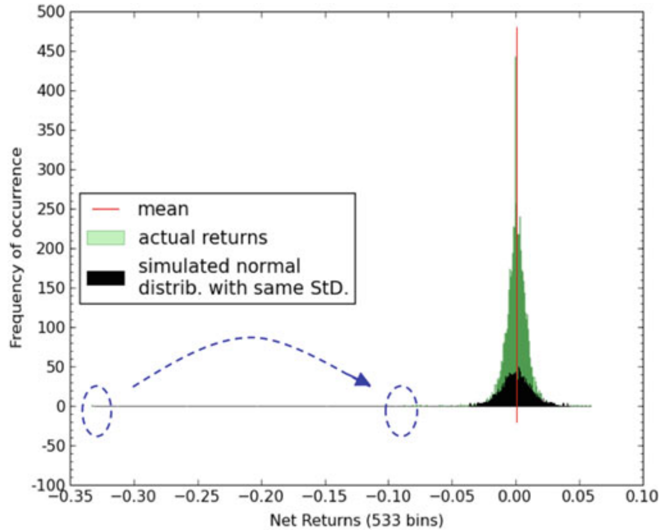


Fig. 7.4 A histogram of Net Returns of the XAO dataset

Its histogram, show in Fig. 7.6, illustrates both the skewness and kurtosis of the dataset, when compared to a normal distribution. The size of the sample-bins of this histogram is kept constant to those of the Fig. 7.4 histogram by decreasing the number of bins. Of interest is the asymmetry of the outliers (the data at the extremities): there are more of the negative variety than positive, and negative ones exist further from the mean; even more so when considering that this is the clipped dataset.

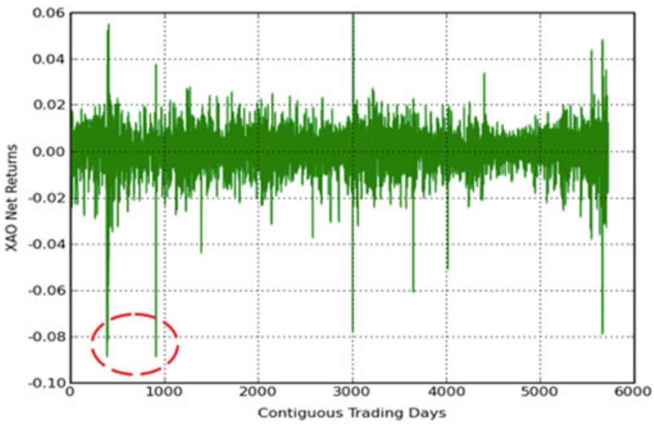


Fig. 7.5 A plot of Net Returns, highlighting the clipping of the largest negative return

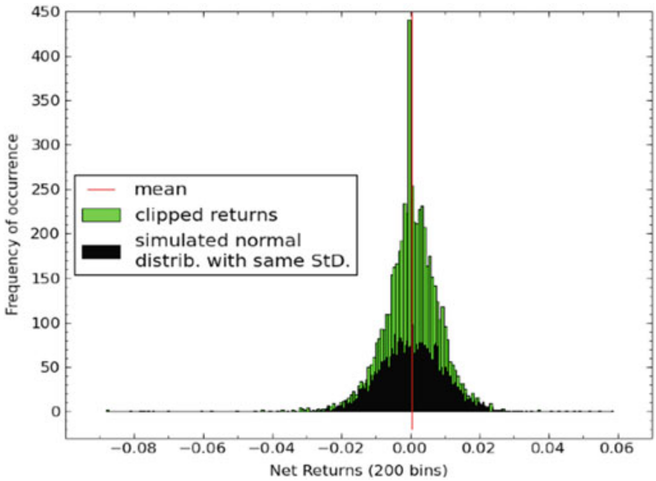


Fig. 7.6 A histogram of the Net Returns, illustrating both the skewness and kurtosis of the dataset. The most extreme positive and negative returns define the gamut of the abscissa

7.4.2 Experiment 1: Can Market Correlation Be Heard?

The purpose of this experiment was to determine whether or not the audification of Returns, could be aurally distinguished from various statistically-related datasets, such as uniform—and Gaussian-distributed samples. A subsequent question was whether or not the Net Returns could be distinguished from its decorrelated form. For visual comparison, Fig. 7.7 shows the plots of both the correlated and decorrelated datasets. A number of features are visually apparent. Both have long tails but they

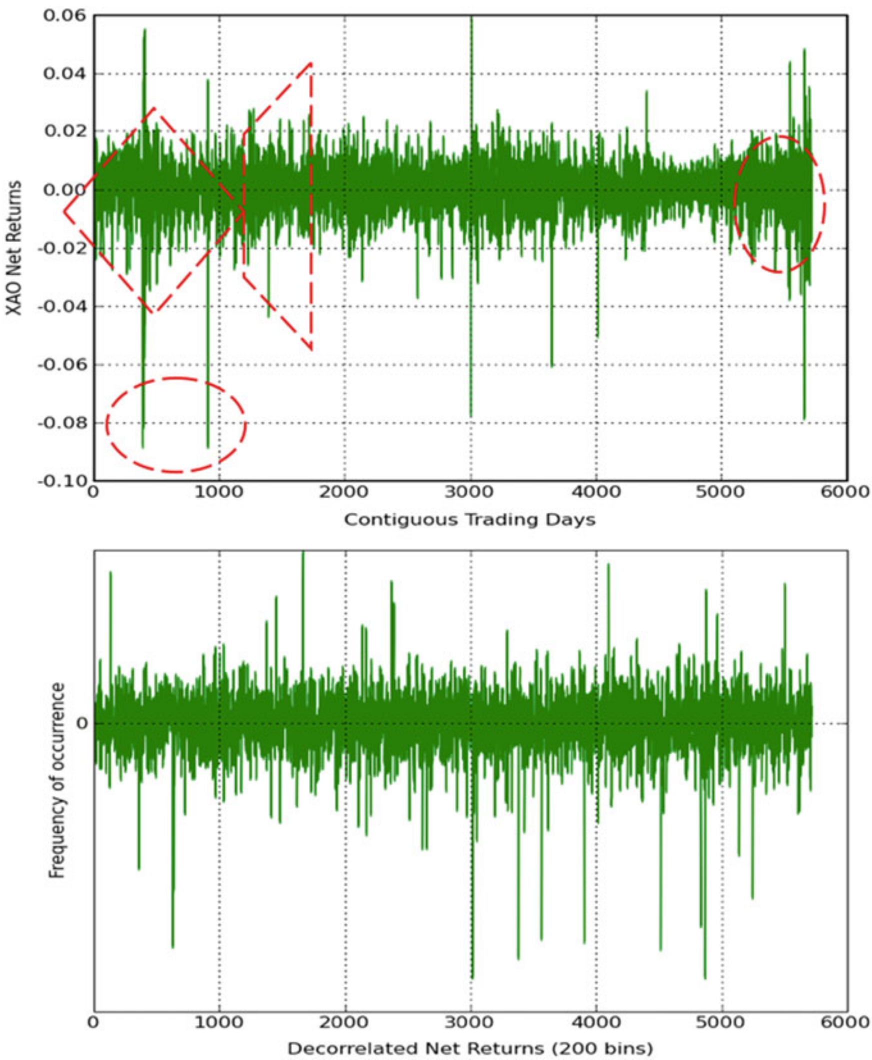


Fig. 7.7 A comparison of the correlated and decorrelated Returns

appear more evenly distributed throughout the decorrelated dataset, contributing to its more chaotic visual appearance, whilst the correlated dataset appears to have periods of increasing (trapezoid), low (circle), and ramped (diamond) volatility.

In order to test whether the audifications of these datasets could be distinguished, a number of chunks of audio were prepared with the same number of samples as the Net Returns. In Audio Example 7.1, these four chunks can be heard four times at a sample rate of 8 kHz in the order A-B-C-D. Each audio chunk is of approximately 0.7 s duration. There is a one-second gap between each chunk and a further

Filename	Description
audio7-1.wav	A. Uniform-Distributed Stochastic B. Gaussian-Distributed Stochastic C. Decorrelated Returns D. Raw Returns

Audio Example 7.1 An auditory comparison of four noise distributions

one-second gap between repeats. The following informal observations can be made[14]:

- The uniformly distributed noise (A) is clearly distinguishable from the Gaussian (B). This distinction is unsurprising: in electronic music parlance, it is that between white and band-limited noise. As would be expected, the uniformly random noise sounds "brighter" because of the comparatively greater prevalence of higher frequencies.
- The raw and decorrelated returns (D and C) are clearly distinguishable from A and B: Qualitatively, they sound rougher or grainier, and have less evenly distributed spectral energy than A and B. As reported in empirical studies, this may be interpreted as a result of the increase in kurtosis (Baier et al. 2005; Frauenberger et al. 2007).
- The 8 kHz sampling rate was settled on after some initial heuristic experimentation with higher and lower values. There appears to be an optimal compromise between durations long enough for possible temporal patterns to be perceptible and sampling rates high enough to make shorter-term correlations perceptible. No formal method for determining the optimization seems to be currently known, yet the choice clearly influences the perceptibility of pattern, as was also observed by Dombois in his seismic audification studies (Dombois 2002a, b).

7.4.3 Experiment 2: Correlation and Decorrelated Compared

Having ascertained that the Net Returns were clearly distinguishable from uniform and Gaussian noise, a second experiment was conducted to ascertain whether or not the raw Returns and the decorrelated Returns could be aurally distinguished from each other. An additional decorrelated Return (E) was generated in the same manner described for C, in Experiment 1, and three files were prepared with the

[14]All Audio Examples referred to can be directly downloaded from the book's online repository.

Audio Example 7.2 Three sequences each of three audio chunks. C & E are decorrelated versions of the Net Returns (D)

Filename	Description
audio7-2_DCE.wav	Sequence: D-C-E
audio7-2_CDE.wav	Sequence: C-D-E
audio7-2_CED.wav	Sequence: C-E D

following sequences in which the original Net Returns (D) was placed in first second and third place respectively (Audio Example 7.2).

The listening task, on multiple presentations of these audio files in random order, was to try to determine, in each case, which one of the three chunks sounded different from the other two. The informal findings of several listeners, all of who had musical training, can be summarized as follows:

- The task was a more cognitively demanding than those in Experiment 1.
- Distinguishability was dependent on a narrower band of sampling rates.
- Above 8 kHz the characteristics described earlier seem to disappear. Below 3–4 kHz the roughness created by the individuation of large-valued samples meant that the principal means of identifying the raw returns was probably more by its invariance across all chunk presentations than by direct chunk comparison.
- Between 4 and 8 kHz sampling rate, a distinct, though subtle, amplitude modulation was observable in the Net Return chunks that seems not to be present in the decorrelated ones. This amplitude modulation effect required attentive listening, probably, in part, due the relatively short duration of the audio chunks (less that 700 ms).

This last observation indicated the need for more data to enable longer durations, or the application of a technique other than audification that enables a slower sample presentation rate. As no intraday data was available for the dataset in question, the latter approach was chosen in Experiment 3.

7.5 Experiment 3: Homomorphic Modulation Sonification

This experiment was designed to test a simple proposition: That the four datasets A, B, C and D of Experiment 1 could be distinctly identified under homomorphic mapping into a pitch-time auditory space. A homomorphic mapping is one in which the changes in a dimension of the auditory space track changes in a variable in the dataset, with only as few mediating translations as are necessary for comprehension (Kramer 1994). A narrow interpretation, here called Homomorphic Modulation Sonification is used, in which time in the dataset was mapped to time in the auditory display, and sample value was mapped to pitch deviation (both positive and negative) from a center frequency, as illustrated in Fig. 7.7, a reproduction of Fig. 2.1.[15]

[15]For a more detailed description of the technique, see Chap. 2, Sect. 2.2.2.3.

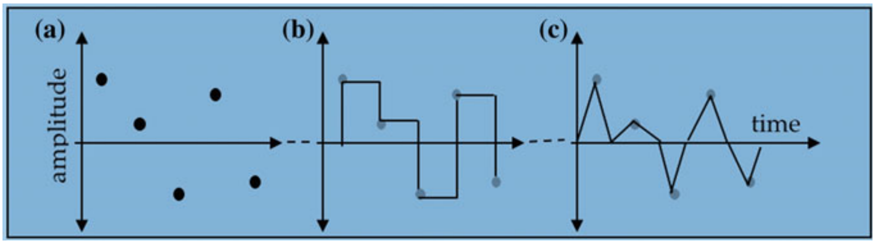


Fig. 7.8 The different amplitude profiles of **a** the sample the samples **b** being realized with amplitude modulation, and **c** individually enveloped events

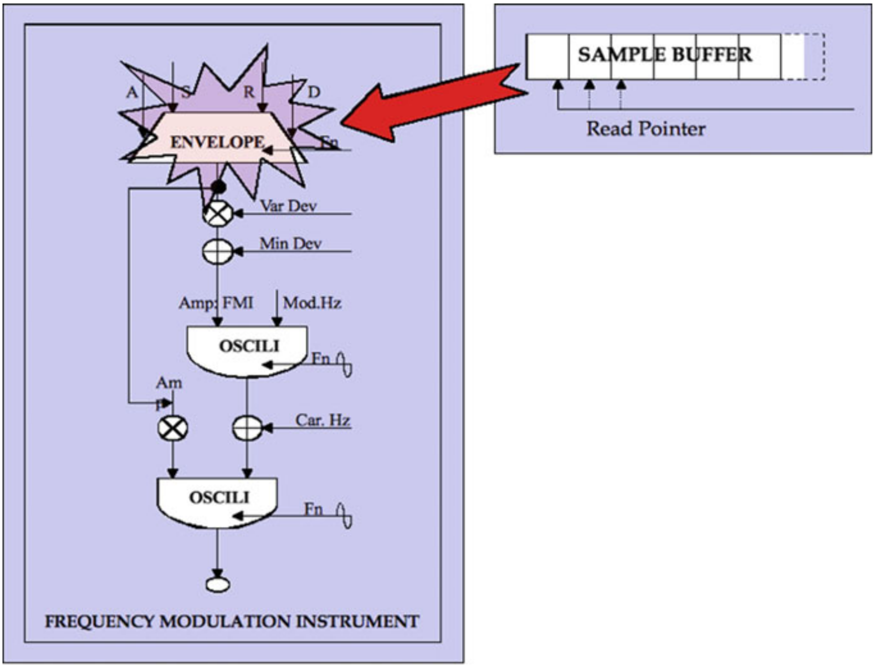


Fig. 7.9 Block diagram of the *Csound* instrument used for the homomorphic mappings

There is a subtle but important distinction between Homomorphic Modulation Sonification and the type of parametric mapping in which each datum is played as, or contributes to, a separate tone with its own amplitude envelope. In the separate-tones case, the audio-amplitude profile of the resulting audible stream fluctuates from—and—to zero, resulting in a sequence of auditory objects individuated by more—or—less rapid onset transients. With modulation however, a single continuous pulsed waveform results, affording the opportunity for the

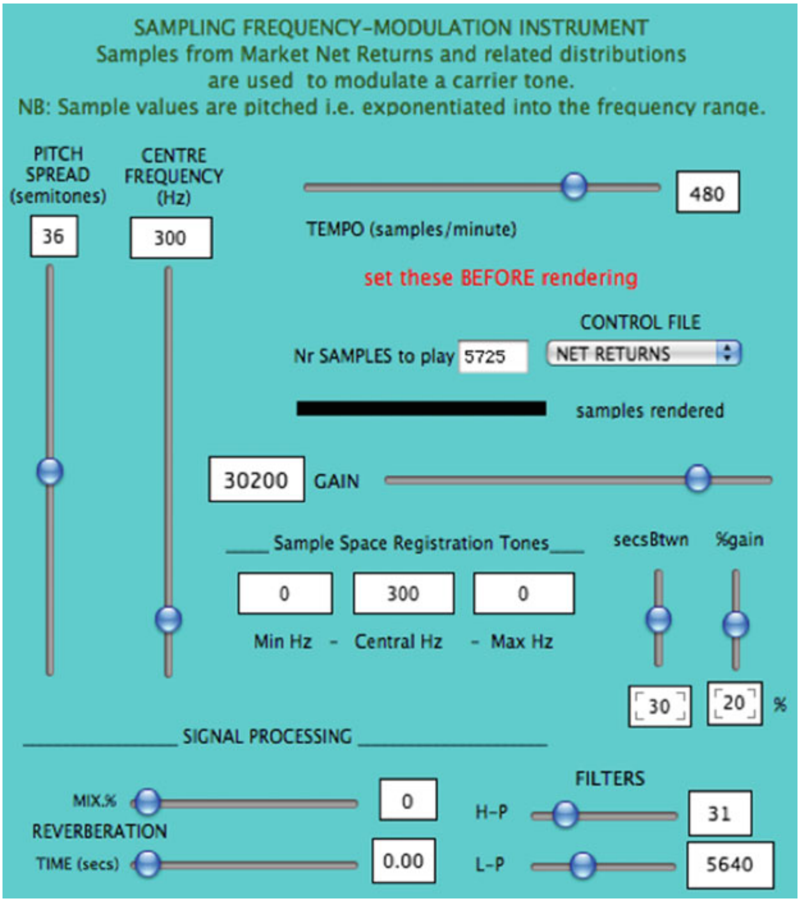


Fig. 7.10 User interface to the sampling frequency modulation instrument

amplitude formant to be held relatively constant, resulting in a lower perceptual loading (Patterson 1982; Best et al. 2008) (Fig. 7.8).

Figure 7.9 is a block diagram for a *Csound* instrument used to produce this homomorphic mapping. Control was implemented to facilitate heuristic investigation of the perceptual space in *SoniPy* (via MIDI and Csound's *Python* API) and in *MacCsound*. Figure 7.10 shows a controller interface developed using Matt Ingalls' *MacCsound* (Ingalls 2011).[16] The setting shown is for frequency-modulating a 300 Hz tone according to the values of successive samples within a range of three octaves at the rate of 480 modulations per minute (8 Hz) from the Net Returns distribution.

[16] *MacCsound* is no longer available. However, it would be possible to write a similar controller in *Sonipy*—using the minimal *Python* `tkinter` module, for example.

Structurally, this is a basic frequency modulator in which an ADSR for controlling modulation index is replaced by a sample buffer of the audio samples. These samples, which can be read directly from an audio file, are then used in

```
;----------- SAMPLING FREQUENCY-MODULATION INSTRUMENT -------------
; Reads an audio file into table so samples can be used as FreqMods
sr = 44100               ; sample rate  / second
kr = 4410                ; control rate / second
ksmps  = 10
ga1 init 0               ; init global audio receiver/mixer
instr 1 ; ----------------------- INSTRUMENT 1  -------------------
; ------------- controller inputs/outputs ------------------------
  kgain      invalue      "gain"              ; at render time-if possible
  kspread    invalue "spread"    ; get pitchSpread from slider
  kCentreHz  invalue      "centreHz"    ; centre freq
  ktempo     invalue      "tempo"             ; get tempo from slider
  itempo     = 60/i(ktempo)/2    ; get tempo from slider
; -------------------- FM component ------------------------------
  kNoiseOffset      = 3                ; listIndex of noisetypes
  kNoiseType  invalue "noiseType" ; choose type to render
  kNoiseTypet= kNoiseType + kNoiseOffset ; weird indexing
  iNoiseType = 3 + i (kNoiseType)
  kNrSamples  invalue "nrSamples" ; get Nr samples to render
  iNrSamples = 2 * i(kNrSamples) ; Nr samples in the file
  kmixpercent invalue "reverb"    ; NB also fed 2 reverb instr.
                         ; 2 control OP % of dry sig
; -------------------- FM component ------------------------------
  kCnt       = 0
  ; Create index into GEN01sampleArray to be ++ed, accord to tempo
  kcps init 1/(iNrSamples * itempo)     ; Nrsamples in the file; 5725
  kndx phasor kcps         ; kndx=(0-1) norm(sampTable (index))
  kCnt = kCnt +1                  ; pick new samp every 'beat'
  printks "\nknsx = %f, ", 10, kndx
  ktabOP tab kndx, iNoiseType, 1
  kFreq = kCentreHz * powoftwo (kspread/2 * ktabOP/12)
  printks "Cnt=%6.1f tabVal=%5.3f Fq=%5.3f\n",kCnt,1,ktabOP, kFreq
  a1 oscili kgain, kFreq, 1      ; Sine: tableValss to FMcentreHz.
  out a1 * 1-(kmixpercent/100)    ; and output the result
  ga1 = ga1 + a1                  ; OP to global buff
endin
instr 2 ; --------------- INSTRUMENT 2  --------------------------
; ------------------ controller inputs/outputs ---------------------
  kgain      invalue "gain"                   ; gain of FM samples
  kgainPcnt  invalue "regisGain" ; %FM samp.gain 4tick regs
  kregisGain = kgain * kgainPcnt/100
  kregisRate invalue      "regisRate"  ; repeat rate of ticks
  kspread    invalue "spread"    ; get pitch spread in ST
  kCentreHz  invalue      "centreHz"    ; centre freq
  outvalue   "lowerHz", kCentreHz - powoftwo (kspread/24)
  outvalue   "upperHz", kCentreHz + powoftwo (kspread/24)
  a1 oscil kregisGain, kregisRate,1     ; FM kCentrHz by set amount
  a1 oscili a1, kCentreHz, 1            ; a sine, using centreHz.
endin
instr 99 ; ----------- INSTRUMENT 99 -- REVERB  ------------------
  kreverbtime  invalue "reverbTime"
  kmixpercent  invalue "reverb"
  a3 reverb ga1, kreverbtime         ; reverberate what is in ga1
  out a3 * (kmixpercent/100)         ; and output the result
  ga1 = 0                            ; empty receiver for next pass
endin ;----------------------- FIN -------------------------------
```

Code Example 7.1 The *Csound* instruments used to implement sampling frequency modulation

Filename	Description
audio-7-3_snippets.wav	4 x 20 seconds of homomorphic modulation sonifications of the datasets A (uniform) B (Gaussian), C (Decorrelated Returns), and D (Net Returns).

Audio Example 7.3 Homomorphic modulation sonifications of four datasets

Filename	Description
audio-7-4_uniform.wav	HMS Uniform-Distributed
audio-7-4_gaussian.wav	HMS Gaussian-Distributed
audio-7-4_decorrelated.wav	HMS Decorrelated Returns
audio-7-4_ netReturns.wav	HMS Net Returns

Audio Example 7.4 Full dataset versions of snippets in Audio Example 7.3

sequence to control the frequency deviation from a user-defined center 'reference' carrier frequency. The corresponding code is detailed in Code 7.1.

With this controller, it was possible to dynamically adjust the pitch spread and center frequency while auditing. In addition to the sample modulator, the sonification also uses has a simple audio 'tick' registration tone generator that acts as an auditory reminder of the upper, lower and central limits of the current render. Both its *secsBtwn* (frequency-of-occurrence) and relative loudness (*%gain*) is adjustable.

The files listed in Audio Examples 7.3 and 7.4 provide various sonic realizations of the homomorphic modulations of the Net Returns samples generated for Experiments 1 and 2, but using the homomorphic mapping sonification technique. For consistency, all are rendered with the settings illustrated in Fig. 7.10.

Audio Example 7.3 is a series of twenty-second 'snapshots' of each of the four sample sets A, B, C and D.

The files listed in Audio Example 7.4 are extended, full-length (6' 40") examples of each of the for snippets of Audio Example 7.3, thus better revealing the different characteristics of each of the datasets.

7.5.1 Observations on Experiment 3

The informal findings of Experiment 3 can be summarized as follows:

1. The difference between homomorphic mapping sonifications of A, and B is easily noticeable, at least to a musically trained listener, as it was in the audifications of Experiment 2.

2. The homomorphic mapping of A can be observed to be evenly spread across the pitch gamut, while the Normal distribution of B can be observed to be more closely clustered around the center of the gamut, the mean, with fewer outliers.
3. Again, C and D are noticeably different to A and B. Whilst both C and D appear to have short trending auto-correlative sequences, those of D (the Net Returns) appear more consistently, and when they do, they appear to last for longer periods of time. This is particularly noticeable in the sequences of consecutive zero or small Net Returns, a characteristic consistent with the observation by chartists and technical analysts that securities prices frequently shift to a new price 'zone' quite quickly interspersed with longer times consolidating those zones before moving again (Murphy 1999).
4. The characteristics discussed appear invariant under both translation (transposition) and linear scaling of the gamut if the standard psychophysical nonlinearities, such as Fletcher-Munson equal loudness contours (Cook 1999) are taken into account.
5. An attempt to 'enhance' the homomorphic mappings by smoothing the sample transitions using bandpass filtering and reverberation, added noisy artifacts that were distracting, not enhancing. This had been anticipated on the basis that such signal processors are designed for inputs with complex spectra, not simple sine-tone modulated signals, but it was useful to confirm that the 'starkness' of the carrier sine-tone with acoustic transients created by the transitions between the modulating data samples, created a superior sounding sonification with less perceptual dissonance, than the results of using of a complex timbre, as a conventional musical aesthetic might suggest.

7.6 Conclusions

This investigation was conducted during the development of software solutions in the *SoniPy* environment for the sonification of large multidimensional datasets (see Chaps. 5 and 6). As such, the primary emphasis was tool development and perceptual testing was informal.[17] It is apparent from these experiments that the simple technique of using net returns is applicable to the sonification of capital market trading data and so it may be possible to apply the same techniques to other similarly structured time-series datasets from such applications as electroencephalography, trans-synaptic chemical transmitters, and the hierarchical networks arising from social affiliations. A further interesting extension study would be the application of the techniques to the functional simulations of financial-market-like time-series, an active field of econometric investigation in which large datasets can be generated when needed.

[17]A well-commented version of the script developed for these experiments, along with the audio examples, is available for download from the book's online repository.

Although other techniques can be applied, the directness of audification makes it appealing, and controlled empirical experiments to determine which features of the dataset are perceivable under those conditions would be worthwhile. The observation of amplitude modulation in the Net Returns in Experiment 2 suggests that an empirical study to isolate the aural characteristics of cross-correlation, such as the spectral modulation suggested by the study, may be useful. This would require the preparation of additional, unrelated, raw returns datasets of the same sample size.

The choice of sampling rate clearly influences pattern perceptibility, as was also observed in seismic audification studies (Speeth 1961) but, apart from limiting the resulting frequency band imposed, no reliable formal optimization method is known and this deserves empirical attention, perhaps by using the using the fractal dimension of the dataset as a potential correlation index.

The effect of the size of the dataset on the sampling rate also needs to be isolated, as, whether or not higher sampling rates on larger datasets reveal other distinguishing features is currently not known.

References

Arbon RE, Jones AJ, Bratholm LA, Mitchell T, Glowacki DR (2018) Sonifying stochastic walks on biomolecular energy landscapes. In: Proceedings of the 24th International Conference on Auditory Display (ICAD 2018). Michigan Technological University

Bachelier L (1967) The theory of speculation. In: Cootner PH (ed) The random character of stock market prices, translated by A.J. Boness. MIT Press, Boston, MA

Baier G, Hermann T, Müller M, Lara OM (2005) Using sonification to detect weak cross-correlations in coupled excitable systems. In: Proceedings of the Eleventh Meeting of the International Conference on Auditory Display. Limerick, Ireland, pp 6–9

Ben-Tal O, Berger J, Cook B, Scavone G, Daniels M, Cook P (2002) SONART: the sonification application research toolbox. In: Proceedings of the Eighth International Conference on Auditory Display. Kyoto, Japan, pp 2–5

Best V, Ozmeral EJ, Kopco N, Shinn-Cunningham BG (2008) Object continuity enhances selective auditory attention. Proc Natl Acad Sci USA 105(35):13174–13178

Bly S (1994) Multivariate data mappings. In: Kramer G (ed) Auditory Display: sonification, audification, and auditory interfaces. Proceedings, Santa Fe Institute studies in the sciences of complexity, vol XVIII. Addison Wesley, Reading MA, pp 405–416

Brewster S, Murray R (2000) Presenting dynamic information on mobile computers. Pers Technol 4(209–212):2

Childs E (2005) Auditory graphs of real-time data. In Proceedings of The Eleventh Meeting of the International Conference on Auditory Display. Limerick, Ireland, pp 6–9

Ciardi FC (2004) 'SMAX'. A multimodal toolkit for stock market data sonification. In: Proceedings of the Tenth Meeting of the International Conference on Auditory Display. Sydney, Australia, pp 6–9

Cook PR (ed) (1999) Music, cognition, and computerized sound: an introduction to psychoacoustics. The MIT Press, Cambridge, MA

Dombois F (2002a) Underground sounds. An approach to earthquake prediction by auditory seismology. Geophys Res Abstr 4:02–02741

Dombois F (2002b) Auditory Seismology on Free Oscillations, Focal Mechanisms, Explosions and Synthetic Seismograms. In: Proceedings of the Eighth International Conference on Auditory Display. Kyoto, Japan, pp 2–5

Farmer JD, Patelli P, Zovko II (2005) The predictive power of zero intelligence in financial markets. In: Proceedings of the National Academy of Sciences of the United States of America, 102. Nr 6:2254–59. http://www.pnas.org/content/102/6/2254.full.pdf

Flowers JH (2005) Thirteen years of reflection on auditory graphing: promises, pitfalls, and potential new directions. In: Proceedings of the First Symposium on Auditory Graphs. Limerick, Ireland

Frauenberger C, de Campo, Eckel AG (2007) Analysing time series data. In: Proceedings of the Thirteenth International Conference on Auditory Display. Montréal, Canada, p 26

Frysinger SP (1990) Applied research in auditory data representation. Extracting meaning from complex data: processing, display, interaction. SPIE-The International Society for Optical Engineering, Santa Clara, CA

Frysinger SP (2005) A brief history of auditory data representation to the 1980s. In: Proceedings of the first symposium on auditory graphs. Limerick, Ireland

Hayward C (1994) Listening to the earth sing. In: Kramer G (ed) Auditory Display: sonification, audification, and auditory interfaces. Proceedings, Santa Fe Institute studies in the sciences of complexity, vol XVIII. Addison Wesley, Reading MA, pp 369–404

Hermann TM, Hansen H, Ritter H (2001) Sonification of markov chain monte carlo simulations. In: Proceedings of the Seventh International Conference on Auditory Display. Helsinki University of Technology, pp 208–216

Heymann M, Hansen M (2002) A new set of sound commands for R: Sonification of the HMC algorithm. In: Proceedings ASA (ed) Statistical computing section. American Statistical Association, Alexandria, VA, pp 1439–1443

Ingalls M (2011) MacCsound controller interface. Macinosh. http://www.csounds.com/matt/MacCsound/

Janata P, Childs E (2004) 'Marketbuzz': Sonification of real-time financial data. In: Proceedings of the Tenth Meeting of the International Conference on Auditory Display. Sydney, Australia, pp 6–9

Kramer G (1994) An introduction to auditory display. In: Kramer G (ed) Auditory Display: sonification, audification, and auditory interfaces. Proceedings, Santa Fe Institute studies in the sciences of complexity, vol XVIII. Addison Wesley, Reading MA, pp 1–77

Kramer G, Ellison S (1991) Audification: the use of sound to display multivariate data. In: Proceedings of the International Computer Music Conference. San Francisco, C.A: ICMA, pp 214–221

Lefèvre E (1923) Reminiscences of a stock operator. Wiley, New Jersey

Mandelbrot BB, Hudson RL (2004) The (mis)behaviour of markets. Basic Books, New York

Mauney BS, Walker BN (2004) Creating functional and livable soundscapes for peripheral monitoring of dynamic data. In: Proceedings of the Tenth Meeting of the International Conference on Auditory Display. Sydney, Australia, pp 6–9

McKell C (2016) Sonification of optically-ordered brownian motion. In: Proceedings of the 42nd International Computer Music Conference 2016. HKU University of the Arts, Utrecht, pp 525–529

Mezrich JJ, Frysinger SP, Slivjanovski R (1984) Dynamic representation of multivariate time-series data. J Am Stat Assoc 79:34–40

Murphy J (1999) Technical analysis of the financial markets. New York: Institute of Finance, New York

Nesbitt K, Barrass S (2002) Evaluation of a multimodal sonification and visualisation of depth of market stock data. In: Proceedings of the Eighth International Conference on Auditory Display. Kyoto, Japan, pp 2–5

Patterson RD (1982) Guidelines for auditory warning systems on civil aircraft. Civil Aviation Authority, London

Peters EE (1991) Chaos and order in the capital markets. Wiley, New York

Simmel G (1979) The philosophy of money. In: Frisby D (ed) Translated by Bottomore T, Frisby E. Routledge, New York

Skiena SS (2004) Financial time series data. http://www3.cs.stonybrook.edu/~skiena/691/lectures/lecture13.pdf

Sornette D (2003) Why stock markets crash. Critical events in complex financial systems. Princeton University Press, Princeton, N.J

Speeth SD (1961) Seismometer Sounds. J Acoust Soc Am 33:909–916

Voss RF, Clarke J (1975) 1/f noise music and speech. Nature 258(November):317–318

Voss RF, Clarke J (1978) '1/f Noise' in Music: music from 1/f noise. J Acoust Soc Am 63(1):258–263

Xenakis I (1971) Formalized Music: thought and Mathematics in Music. Indiana University Press, Indiana

Chapter 8
Parameter-Mapping Sonification of Tick-Data

Abstract The previous chapter explored the use of audification of a numerical series, each member representing the daily closing value of an entire stock market, to observe the cross-correlation (trending) within the market itself. This chapter employs parameter-mapping sonification to study the perceptual flow of all trades in individual stock groupings over a trading day by applying various filters and selection methodologies to a detailed 'tick' dataset. It outlines the use of a size/mass metaphorical model of market activity and the simultaneous use of two opposing conceptual paradigms without apparent conceptual contradiction or cognitive dissonance to demonstrate, given conducive conditions, the power of intention over perception and sensation, in auditory information seeking.

8.1 Introduction[1]

Can the sonification of intra-day data from the capital markets[2] provide data domain experts and others with insights into activity in these markets not afforded by other display techniques? The experiments outlined in this chapter were conducted in the context of an exploratory and broadly speculative project in partnership with the Capital Markets Cooperative Research Centre.[3] A second parallel motivation was also the advancement of tools and techniques for sonification design through the incremental development of *Sonipy*, a heterogeneous software framework for the

[1]For an overview of previous work on the sonification of stock market data, see Chap. 7, Sect. 7.3.1.

[2]The nomenclatures used for various kinds of markets is somewhat convoluted and are often misused. A *stock market* is for trading equities, namely ownership interests in companies. A *financial market* is an institutional structure or mechanism for creating or exchanging financial assets, namely those that are real such as land, buildings, equipment, patents. The term *capital market* is used generally to refer to markets for stocks, bonds, derivatives and other investments and it is this term we adopt here, or the less formal The Market. Both are generic terms signifying all regulated exchanges and their activities.

[3]http://www.cmcrc.com/.

D. Worrall, *Sonification Design*, Human–Computer Interaction Series, https://doi.org/10.1007/978-3-030-01497-1_8

sonification of large multidimensional datasets, as discussed in Chaps. 5 and 6. The primary emphasis is thus more tool development than perceptual testing. That is, it is more propositional and procedural than empirical.[4]

8.2 The Dataset[5]

The dataset for the experiments reported in this chapter contains one financial quarter of a time-sequenced trading engine log of a generic mid-sized market: approximately 5 Gigabytes of ASCII characters, one flat-file per day. Structurally, this dataset is a multiplexion of all the on-market TRADEs with the incremental (Bid/Ask) adjustments to the orderbooks of approximately 3000 trading instruments. It had to be extensively cleaned before being statistically analyzable in preparation for sonification; a very time-consuming task, but essential to achieving a workable database. Such a dataset is far too large to correct 'by hand' so such a process also provides something of an opportunity to become procedurally acquainted with the data itself before attempting to sonify it; the importance of which was also discussed in Chap. 7. This dataset suffered from the not uncommon comma-separated values (CSV) field corruption malaise: even though commas were used as field separators, they were also embedded in some data fields, resulting in the destruction of field integrity. This can happen inadvertently in data copying and preparation, or within some European datasets, where the comma and period are also used in an alternate way—1.234,56 instead of 1,234.56, for example. In such circumstances, a more detailed pattern-matching approach is required to reconstruct each TRADE record with the requisite number of fields. Fortunately the *Python* regular expression operator `re` contains powerful pattern-matching capabilities to support such undertakings. The data-cleaning approach for this project was to iteratively develop and modify a series of active filters built around these regular expression-matching capabilities, until one day's data parsed correctly. Most of the data from other days then required only minor adjustments to the filters; however, one particular day's data was especially corrupt, and required extensive 'personal' attention. The example code does not reveal the process used to arrive at it, nor the superiority of an interpreted rather than compiled language for undertaking such incremental fastidiousness, but is included as an example of a data-cleaning method.

Some of the limitations imposed to make the dataset generic ruled out certain types of internal comparisons, such as day-of-week, GICS[6] market sector or index membership, meaning that it was not possible to group these securities other than

[4]See Chap. 3, Table 3.1 for a descriptive summary of different ways of acquiring knowledge.

[5]The audio examples, data and python code used to calculate, plot and sonify examples in this chapter are available from the book's online repository.

[6]The Global Industry Classification Standard) consists of 10 Sectors aggregated from 24 Industry Groups, 67 Industries, and 147 Sub-Industries. See https://en.wikipedia.org/wiki/Global_Industry_Classification_Standard.

by abstracted statistical characteristics such as value and frequency of trades. The resultant dataset's metadata specification is outlined in Table 8.1 and field descriptors in Table 8.2 in the Appendix at the end of this chapter.

8.3 Experiment 4: $value at Time

The aim of Experiment 4 was to provide a means by which a musically-naïve listener[7] could aurally observe something of the nature and extent of the way value, measured in monetary terms, changed ownership during a trading day. There is no simple relationship between the traded *price* of an individual unit of a security and that of an individual unit of another security.[8] However, one way to meaningfully compare securities being traded is

> **The sonification task:** To sequentially sonify the total $value of all `TRADES` at occurred each second in the trading day.

by comparing the monetary value of trades ($value hereafter). To effect this sonification, intraday TRADE[9] data was adapted to a model of $value; that is, one that reflects the relative importance of trades involving various amounts of a finite resource (money). Without wishing to overstress the point, the task is thus of the sonification of (human-valued) information.

8.3.1 Size Matters

Perceptually, the size of an object is habitually related to its mass, as evidenced by the surprise experienced when picking up a large piece of pumice. Portentous objects and events in the natural world are more usually associated with lower-pitched, less frequent sounds. For example, there appears to be many more small birds than large ones, and so on.[10] A basic statistical analysis of the TRADE data reveals proportionally fewer high-$value trades than low-$value trades, indicating that $value is in line with the principles outlined, as expected. This is symbolically illustrated in Fig. 8.1.

---

[7]That is, to use sound relationships which do not require musical training for them to be able to be identified.

[8]Because the actual $price is related to the number of units of a security a company has on issue, not just to the value of each of such units.

[9]Capitalized words are Order Types and words enclosed with <and> are Order Type Fields as defined in Tables 8.1 and 8.2 in the Appendix.

[10]For an expanded description of these phenomena as they relate to film, see Branigan (1989).

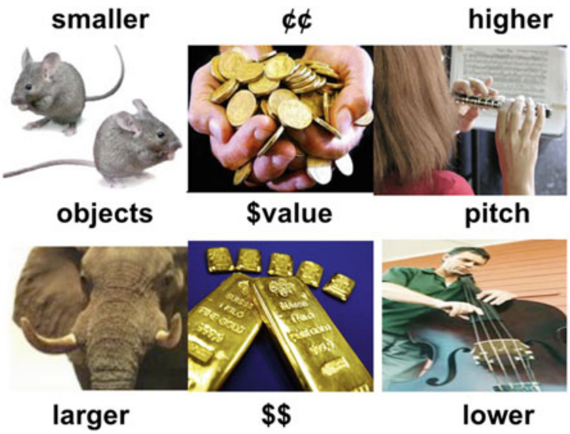


Fig. 8.1 Symbolic representation of the relationship between size, value and pitch

Similarly, there appears to be an inverse relationship between the *duration* of an event type and it's approximate *frequency* (of occurrence). When compared to smaller events, larger events in the same domain occur less often and last for a longer period of time than smaller ones: the snap of a twig is more frequent and shorter than the crash of a tree, the scurry of mice more frequent than a stampede of elephants etc. So, when sonifying larger $valued trades at lower frequencies, it is necessary to increase their relative duration, comparatively, in order to provide perceptual balance.

These generalised relationships are also in evidence in the evolution of the logarithmic 'mappings' between psychophysical, physical and psychological phenomena. For example, it takes longer for the pitch of sound in the range of 30–50 Hz to be perceptually determined than it does for a pitch in the 300–500 Hz range. A similar psychological relationship can be observed in changes in perceived financial value; accounting for why $value is measured in the percentage gain or loss rather than in absolute amounts. This exponential relationship is also in evidence in the distribution of $value *between* securities in the market as a whole, as illustrated in Fig. 8.2.

So, for the reasons just outlined, in this experiment, a proportional *inverse* mapping was applied between sound frequency and the total $value of all trades at each moment in the trading day. The psychoacoustic mapping between pitch and duration as are illustrated in Figs. 8.3, and in 8.4, between a tone's onset—time and pitch.

A further psychoacoustic adjustment made was a very basic 'inverse Fletcher–Munson curve of equal loudness'; mapped to counter—balance the known psychoacoustic phenomena that the centre of the pitch gamut of human hearing is more amplitude—sensitive than the extremes (Mathews 1999; Roederer 1973, 71–82). A more detailed illustration of this relationship is shown in Fig. 8.5.

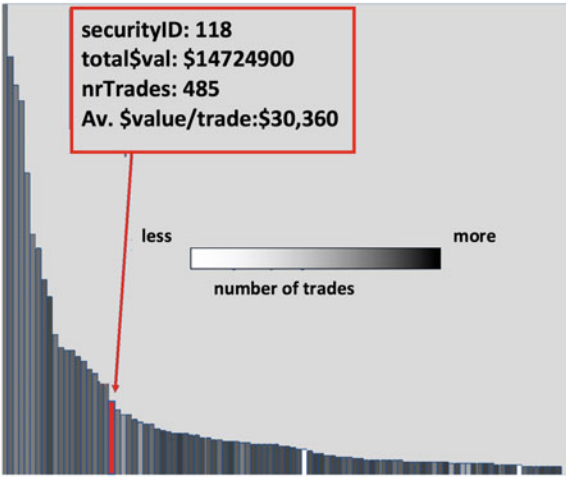


Fig. 8.2 The distribution of $value traded for individual securities in a single trading day. Shading indicates the number of TRADEs for each security

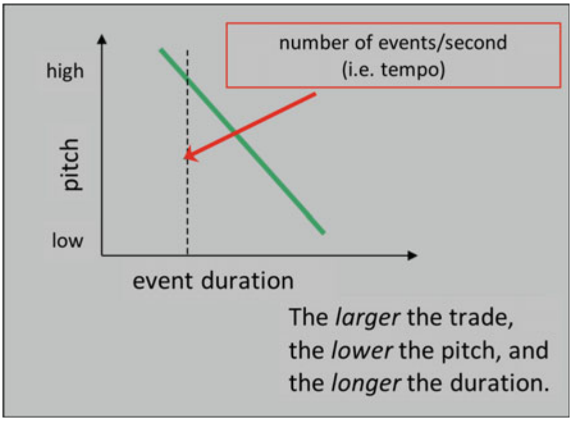


Fig. 8.3 The principle information mappings $value is inversely–proportional to pitch and the lower the tone the longer the duration. Notice that the pitch (green line) is linear, implying an exponential frequency scale

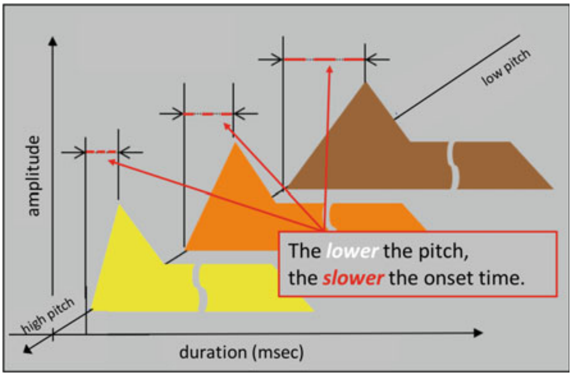


Fig. 8.4 A second psychoacoustic adjustment: larger $value trades (lower–pitched) have slower onset–times, in keeping with physical characteristics of material resonators

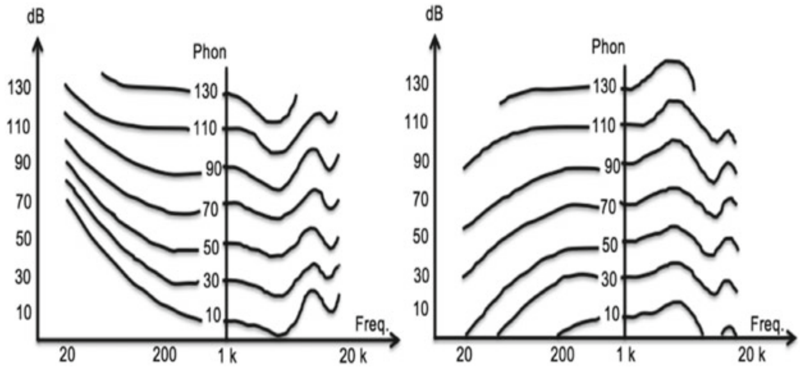


Fig. 8.5 The Fletcher–Munson curves of equal loudness (left) and its inverse. Used for frequency-dependent adjustment of amplitude to counterbalance this hearing non-linearity

8.3.2 Data Subset and Sound Rendering

The dataset for Experiment 4 is subset of a single metadata type (TRADE) of the complete dataset. This subset is used sequentially, day-at-a-time, second-at-a-time so the data can be extracted from the flat-file format in a single pass, thus not requiring the need for a separate, more sophisticated, persistence (database) model. This criterion would not change if the technique were to be applied to a buffered real-time data feed.

The *Csound* instrument used to render this new dataset to sound was designed for straightforward fine adjustment of the acoustic–psychoacoustic relationships. For the purpose of these experiments no graphical user—interface was implemented, though the one used in Experiment 2 in Chap. 7 could be easily adapted. Code Example 8.1 shows the shows the instrument itself and the head of the control (score) file, including the first 30 s of data to be rendered. Loading the instrument with the psychoacoustic adjustment tasks, rather than using an inflexible renderer and a pre-calculated score, affords rapid iterative use of an essentially very basic instrument: the emphasis being on clarity of articulation rather than complexity of timbre.

The examples Audio Example 8.1 are of the same TRADE data, but with different temporal compressions. The first three are each of 30 s duration with the same data, from Market Open. The fourth, audio-8.4d.wav, is a rendering of the entire day's TRADEs at a ratio of 1:60. That is, one second represents 1 min of TRADEs. 1:60 was chosen because (a) it is under the temporal 'fission boundary' for all frequencies of human hearing (Bregman 1994, 58–73) and (b) it is easy to identify the temporal occurrence of characteristics by watching a clock: one minute of audio is equivalent to one hour of trading. A trading day is six hours (10 am to 4 pm), so one day's trading data can be listened to in six minutes.

```
;-score for playing VAT data.Compilied by CsoundAPI in sumVAT.py--
sr  = 44100            ; sampling rate
kr  = 4410             ; conrol rate
ksmps  = 10
nchnls = 1             ; mono output is sufficient

instr 1 ; ------- simple sine instrument with ADSR ---------
  ifreq      = 2048
  kampScale  = 25      ; post 1stPB-adjust for max Playback amp.
                ; val used = EOFoverallAmps=28148.7
  p5 = p5/4            ; post 1stPB, push freq spect+Maj3rd
  kmaxamp = kampScale * (1200-(1100*p5/ifreq)); p5 = freq
  p3 = (p3*0.9) + (0.1*ifreq/p5) ; p3 = dur
  idur = 1/p3          ; freq=1/dur
  kenv linseg 0, 0.01,1, p3-.01, 0
                ; for softer attack use f2=1st 1/2 sine
                ; kenv oscil 1, idur, 2
  asig oscili kenv, p5, 1 ; sine
  out asig * kmaxamp endin
endin
; -----------------------------------------------------------
; scorefile for sumVAT001. sums$value@time.Generated by sumVAT.py
; choose one tempo (in BPM)
; t 0   60     ;  1 x time compression: 5 sec = 1 second
; t 0   720    ; 12 x time compression: 5 sec = 1 minute
t 0 1800       ; 30 x time compression: 2 sec = 1 minute
; t 0   3600   ; 60 x time compression: 1 sec = 1 minute

f 1 0   16384 10 1   ; sine
f 2 0   9 0.5 1 0     ; 1/2 sine 4soft envelope
i 1  0 1.0 1.0   633.43160780
i 1  7 1.0 1.0  7874.78575632
i 1  8 1.0 1.0  8162.64188795
i 1 10 1.0 1.0  8162.64188614
i 1 13 1.0 1.0  8075.19729843
i 1 16 1.0 1.0  8104.24089335
i 1 17 1.0 1.0  8104.24088033
i 1 18 1.0 1.0  8162.64187037
i 1 19 1.0 1.0  7988.68936790
i 1 20 1.0 1.0  8133.38893143
I 1 23 1.0 1.0  8046.25763285
i 1 24 1.0 1.0  8162.64184986
i 1 25 1.0 1.0  8162.64184804
i 1 26 1.0 1.0  8133.38890612
i 1 27 1.0 1.0  8162.64184414
i 1 29 1.0 1.0  8075.19713194
e ; end of score ; ------------- FIN -----------------------
```

Code Example 8.1 *Csound* instrument for rendering $value TRADEs for Experiment 4

Filename	Duration	Compression ratio
audio-8-1a.wav	30 secs	1:1 (realtime)
audio-8-2b.wav	30 secs	1:12
audio-8-3c.wav	30 secs	1:30
audio-8-4d.wav	6 hrs (1d)	1:60 (1 sec=1min)

Audio Example 8.1 Audio examples of the $value sum of the total market, moment-to-moment

8.3.3 Observations on Experiment 4

Described informationally, audio-8.4d.wav presents a fluctuating, almost continuous stream of small $value TRADEs (high frequencies) interspersed occasionally with those of larger $value. There are five high $value TRADEs, at the beginning of the day spaced five minutes (five sonification seconds) apart. At the end, there are two, after a few seconds silence, the first of which is very $value large.[11]

Each of the five large $value indications occur on Market Open. That is, the market is sequentially opened for trading in five groups, five minutes apart.[12] The initial TRADEs occurring simultaneously on a group's opening are a combination of those TRADEs identified in the orderbooks as Market-On-Open TRADEs and those that are initiated by brokers' trading software 'triggering stops' in the pre-market—open price—discovery phase.

Following Market Close, a trading engine 'price-determination' process is undertaken in which the closing prices of all securities are calculated. Of the two large $value indicators that occur at the end of the day, the first represents the value of those TRADEs marked as <Market-On-Close>. In addition to the trades that occur 'on-market', brokers are permitted to trade between themselves without going through the exchange's trading engine, with the requirement that such trades must be registered through the engine within a short period of the market closing. The last $value indication is of the registration of those off-market trades.

While there was an expectation of an exponentially greater proportion of lower $value TRADEs than higher, the proportion of low—valued trades seemed out of proportion even to those expectations. The reasons for this phenomenon prompted a further, investigation, Experiment 5.

8.4 Experiment 5: Sonification of Market Volatility

An initial attempt was made to simultaneously sonify the TRADE data of individual securities by using the $value of each security's TRADE data exactly when the trade occurred. This resulted in each TRADE contributing a component frequency to the overall spectrum comprised of all such TRADEs for that moment in time.[13] Audio-8-2a.wav presents the results. Audio-8-2b.wav applies the same techniques,

---

[11]So large that the (low) frequency may not be perceptible using headphones.

[12]This 'orderly open' approach is to mitigate against the risk of a trading engine crash when opening in a volatile environment. On the ASX, while an individual security always belongs to the same group, the order of opening is randomized in an attempt to suppress the exploitation of this 'feature'.

[13]These preliminary experiments, as exemplified by Audio Example 8.2, were actually conducted prior to the commencement of Experiment 4, but they are reported here as they contributed to the results of Experiment 5.

Filename	Dur.	Compression ratio
audio-8-2a.wav	30 secs	1:60 (1 sec=1min)
audio-8-2b.wav	30 secs	1:60($value <$2000)

Audio Example 8.2 Using simultaneous $value TRADEs to contribute to the composition of an evolving spectrum

except that all trades lower in value than $2000 value were filtered out, which accounts for the relative decrease of higher pitches (Audio Example 8.2).

As was the case for Experiment 4, the results were difficult to interpret, and appeared unsatisfactory for the same reasons. What was needed was a more detailed understanding of the structure of an order becoming a TRADE. To undertake the analysis to effect that, the data had to be made accessible in a manner appropriate for the analysis intended. It was not clear whether or not such analysis would have to include data other than TRADEs and so it was decided to store all the records in a single database. The first approach was to use a relational model with a local *mySQL* server accessed via *pySQL*, the *Python* wrapper API. However, the process was difficult to manage as search times were unacceptable, given the time compressions being employed. The standard technique for this kind of problem in relational-database management is to create additional tables of indices (key, pointers) to reduce the query-response time (Coronel 2000, 89). These additional tables increase the size of the database as a whole. The technique was applied progressively in order to optimise the query-time/database-size configuration, but the model could not deliver query-times low enough to permit real-time rendering of the 1:60 time compression required. Perhaps, with continual refinement, a workable solution may have been found, but the relational approach was eventually abandoned for an object-oriented b-tree structure, using HDF5[14] through the *pyTables* API.[15]

A TRADE order type contains a number of fields.[16] Using the now more—easily—accessed data, analysis revealed long sequences of the following behaviour (in multiple securities simultaneously): Many small—$value TRADEs were initiated as Market Orders, by a single trader through the same broker, either for accumulating or distributing holdings of security at the BID or ASK price. A plausible explanation for this pattern is that by so 'hitting the bid' or 'ask', these TRADEs remain relatively

[14]Hierarchical Data Format library version 5, (HDF5) is a versatile, mature scientific software library designed at NCSA supercomputing facility for the fast, flexible storage of enormous amounts of data. It provides a robust way to store data, organized by name in a tree-like fashion. With HDF5, extremely large datasets (hundreds of gigabytes in size) are organized in a filesystem-like hierarchy using containers called "groups", and accessed using the traditional POSIX/path/to/resource syntax.

[15]Code for translating the data to HDF5 is available in the Repository.

[16]See Table 8.2 in the Appendix for a description of the metadata structure of the trading—engine data.

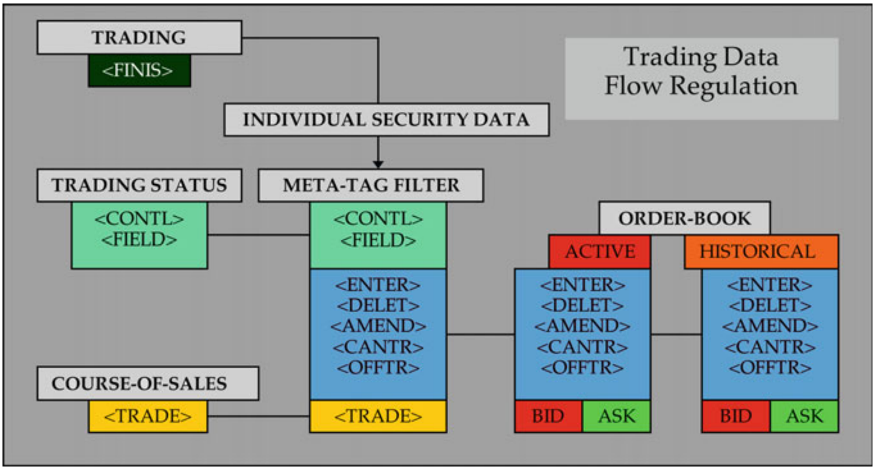


Fig. 8.6 A graphic illustration of part of the HDF5 file structure used to trace the movement of TRADE orders

undetected or at were least are considered insignificant.[17] When another trader entered the market 'significantly', on the same side, or when the price, or activity in the orderbook changed significantly, the sequence was often broken, sometimes to resume, sometimes not. Figure 8.6 graphically illustrates the structure through which the TRADEs were traced in order to arrive at this explanation.

The activity described has characteristics of computer-based algorithmic trading; in small amounts, presumably by broking firms that are not paying for each individual TRADE, in an attempt to avoid detection by the 'suppliers' of the other side of the TRADE; such knowledge in some situations, being sufficient to trigger a repositioning of their TRADEs position in the market's order Book, to their better advantage.

8.5 Experiment 6: Sonification of Price Accumulations

In order to provide a better representation of the way prices fluctuate, the small TRADE 'dispersion' characteristic discussed in the previous section needed to be contained, or managed—whether or not the explanation given is accurate. So, an accumulation filter was applied, as follows: All contiguous trades in a security at the same price were accumulated in a buffer and the accumulated $value was sonified only when the price changed. Thus, this sonification became, in effect a monitor of market volatility,: For multiple-day monitoring, it was also necessary to implement

[17]Because, being 'At Market' TRADEs, they can pass into and out of the Order Book almost instantaneously, without significantly shifting the BID or ASK price. An auditory alert would be an ideal way of drawing attention to such events.

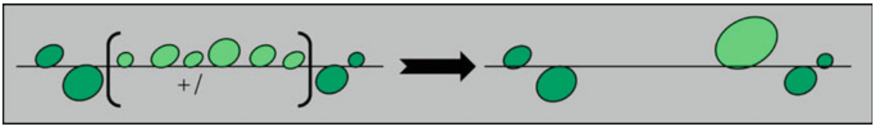


Fig. 8.7 A graphic representation of the filter applied to TRADE data for the Experiment 5 sonifications. $value TRADEs in a security are accumulated until the price changes, at which point the accumulated value is sonified. On the LHS, the dark-green circles represent trades sonified without accumulation because price changed. The smaller light-green circles represent TRADEs that are accumulating (±). The RHS illustrates the overall result

Filename	Description	Compression ratio
audio-8-3a.wav	All trades (No $value filter)(full day)	1:60 (1 sec=1min)
audio-8-3b.wav	All trades of $value>$10K (30min)	1:60
audio-8-3c.wav	All trades of $value>$50K (30min)	1:60

Audio Example 8.3 Three examples of cumulative TRADE data, with $value filters

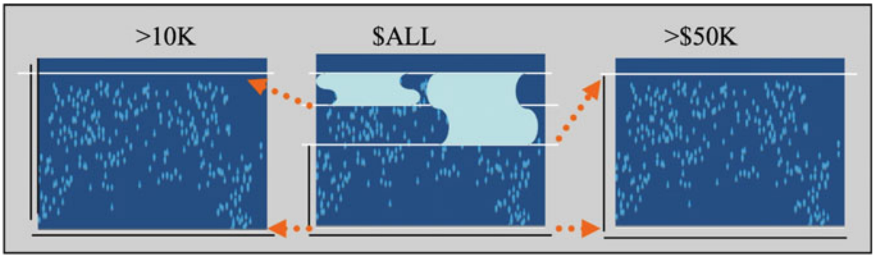


Fig. 8.8 A graphic representation of the filtering of cumulative TRADEs below $10 K and $50 K and rescaling the results to the same pitch gamut before rendering to audio

inter-day buffers to accommodate this type of situation. Figure 8.7 graphically illustrates the process.

Audio Example 8.3 demonstrates the rendered results of a single day's trading, with various settings of the filter. Figure 8.8 is a graphic representation of the relationship between these renderings, including the pitch gamut adjustments necessary to enable a direct comparison of the distribution of $value in each of the ranges.

8.5.1 Observations on Experiment 6

The market opening, is clearly audible, although, because of the accumulation technique employed, the five Market-On-Open, events, discussed in Sect. 8.3.3, are less pronounced. There is clearly less activity in the middle of the day—lunchtime, for the big (market moving) traders perhaps. The increased activity and volatility towards the end of the trading day is quite noticeable, especially in the last fifteen or twenty minutes of trading. On some days there are two Market-On-Close, events, as discussed, however on the day presented in this example, there is only one. There is still a preponderance of smaller—$value Trades, but they appear more realistically weighted than in the previous experiments. The volatility of the market can be easily inferred from this mapping, and overall, the technique seems to work well.

8.6 Experiment 7: Using Conflicting Conceptual Mappings

One further experiment was undertaken with the cumulative mapping technique employed in Experiment 6. A sonification event only occurs when the price changes so, to each event a second pitch was appended to indicate whether the TRADE that triggered the sonification event was a downward or upward movement in <price>. The pitch of that tone was made *higher* when the <price> increased and *lower* when it decreased; the size of the interval between the two tones indicating the extent of the price difference. Renderings in Audio Example 8.4 demonstrate the technique.

8.6.1 *Observations on Experiment 7*

The metaphor of a rising pitch to indicate a rise in the trading price is somewhat in conflict with the overall metaphor of larger $value trades being represented by lower pitches, as outlined in Sect. 8.3.1. It is interesting to note that there appears to be no difficulty in cognitively separating the two opposing mapping paradigms

Filename	Description (all $value>$50K)	Compression ratio
audio-8-4a.wav	First hour of trading	1:60 (1 sec=1min)
audio-8-4b.wav	Middle-of-Day (12h30-13h30)	1:60
audio-8-4c.wav	Last hour of trading	1:60
audio-8-4d.wav	Full day	1:60

Audio Example 8.4 Four examples of cumulative TRADE data with a shift in pitch to indicate whether the <price> of the following trade increased or decreased and to what extent. A higher pitch indicates that the price rose. Audio rendering occurs when cumulative $value >= $50 K. Time compression 1:60 (1 sec = 1 min) when they are superimposed. Having had the concepts explained to them, several listeners were asked informally if, or not, they could describe whether a specific event was higher or lower in $value than the previous and whether the TRADE following was higher or lower in price. None appeared to have any difficulty in separating the two paradigms. In fact, when it was pointed out that the two superimposed conceptual mappings were opposed, psychoacoustically, most of those tested expressed surprise, indicating that they hadn't noticed.

This informal experiment illustrates that it is possible that superimposed concepts may be easily separated, cognitively, even when mapping of the two paradigms into the same psychoacoustic space is opposed: intention, that is information, can take precedence over perception and sensation when given the conditions to do so. In linguistics, this context-sensitivity of meaning is known as *deixis* or after C.S. Peirce, *indexical*.

8.7 Summary

The opportunity to work with a relatively large high-frequency data set was useful in testing the application of the *SoniPy* framework paradigm in the domain that traditional music composition software is the weakest; namely, the handling of external multivariate datasets. While the tools and techniques for handling large multiplexed datasets is not uniquely a sonification problem, the special skills that it requires need to be acquired for such sonifications to be possible. These experiments served to emphasise the flexibility of using an interpreted language when searching for solution rather than just coding a well-defined algorithm. The extensive collection of data handling tools *Python* provided access to worked extremely well, both individually and in tandem. In fact, the exploratory techniques needed to clean, search and manipulate such a dataset would have been prohibitively arduous without them.

In the $value experiments, the conceptual mapping techniques employed using high levels of psychoacoustic redundancy seemed to work well, especially those sonifying market volatility. The lack of cognitive dissonance when two opposing conceptual paradigms were superimposed in Experiment 7 is a clear indication of the power of intention, that is, given conducive conditions, goal-directed information seeking, eclipses perception and sensation.

Appendix Market Trading Order Types

See Tables 8.1 and 8.2.

Table 8.1 Metadata description of market-trading order types

Order type	Fields	Description
<date code>.CSV files contain one message per line (separated by a newline character (<\n>). Lines in each file are sorted in the order that the messages were sent from the trading engine. This should be in the same order as the <timestamp> attached to each message, but this has to been verified		
ENTER	<timeStamp>, ENTER, <securityID>, Bid|Ask, <orderID>, <price>, <disclosedVolume>, <dclosedValue>, <flags>, U, <brokerTraderRefs>	A new order, or the untraded volume & undisclosed volume of an order carried forward from the last trading day (will have 00:00:00 timeStamp).
DELET	<timeStamp>, DELET, <securityID>, <orderID>, B|A	Deletion of the untraded volume and untraded undisclosed volume of an order. i.e. an order cancellation. B for a Bid order, A for an Ask order.
AMEND	<timeStamp>, AMEND, <securityID>, Bid|Ask, <oldOrderID>, <newOrderID>, X, <newPrice>, <newDisclosedVolume>, <newDisclosedValue>, <flags>, <traderID>	An amendment to an order. After an order is amended, it is referred to by its new ID, not its old ID. The traderID is that of the trader who amended the order.
TRADE	<timeStamp>, TRADE, <securityID>, <tradeID>, <price>, <disclosedVolume>, <disclosedValue>, <flags>, <bidorderID>, <askorderID>	An on-market trade.
OFFTR & CANR	<timeStamp>, OFFTR|CANTR, <securityID>, <offtrID>, <executeTimeStamp>, <price>, <disclosedVolume>, <disclosedValue>, <flags>, <bidBrokerTraderRefs>, <askBrokerTraderRefs>	An off-market trade, or a cancelled trade (the orders involved are not re-inserted into the order-book).
FIELD	<timeStamp>, FIELD, <securityID>, <Security> <[1-5]>	An update to the group number of a security. The last field contains the group number.
CONTL	<timeStamp>, CONTL, <securityID> <statusText>	Pre-open/opening/open/suspend/close/adjust period/trading halt for a particular security or a particular group or all securities.
FINIS	<timeStamp>, FINIS	Signifies that there are no more messages for this day (can probably be ignored—only useful in real-time data). <timeStamp> is usually 24:00:00.

Table 8.2 Field descriptors for the market order types

<table>
<tr><th>FIELD</th><th>DESCRIPTION</th></tr>
<tr><td><timeStamp></td><td>Message time stamp reported by the trading engine. HH:MM:SS</td></tr>
<tr><td><securityID></td><td>Same as <securityID> in securityinfo.txt file. If this is “-” then there is a mapping problem in the data. In this case, we suggest you ignore these messages.</td></tr>
<tr><td><orderID></td><td>Order ID. A natural number. Unique for the current day.</td></tr>
<tr><td><price></td><td>Price to exactly 3 decimal places, prefixed by ‘$’.</td></tr>
<tr><td><disclosedVolume></td><td>Disclosed volume. A natural number.</td></tr>
<tr><td><disclosedValue></td><td>Disclosed value, rounded to exactly 2 decimal places.</td></tr>
<tr><td><flags></td><td>Containing a list of flag codes. If there are no flags a double-quoted blank string is emplaced. Refer to market.cfg for descriptions of flags.</td></tr>
<tr><td><brokerTraderRefs></td><td>Broker reference and trader reference of the order (if they exist).
()|(@<brokerID><“”>&<traderID>)|(@<brokerID>)|(&<traderID>)</td></tr>
<tr><td><newOrderID></td><td>New order ID of an amended order. After an order is amended, it is referred to by its new ID(not its old ID).</td></tr>
<tr><td><newPrice></td><td>New price of an amended order, to exactly 3 decimal places, prefixed by ‘$’.</td></tr>
<tr><td><newDisclosedVolume></td><td>New disclosed volume of an amended order. A natural number.</td></tr>
<tr><td><newDisclosedValue></td><td>New disclosed value of an amended order, rounded to exactly 2 decimal places.</td></tr>
<tr><td><brokerID></td><td>Broker ID. A natural number prefixed by ‘#’.</td></tr>
<tr><td><traderID></td><td>Trader ID. A natural number prefixed by ‘#’.</td></tr>
<tr><td><tradeID></td><td>Trade ID for an on-market trade. A natural number.</td></tr>
<tr><td><bidorderID></td><td><orderID> of the Bid order.</td></tr>
<tr><td><askorderID></td><td><orderID> of the Ask order.</td></tr>
<tr><td><offTrID></td><td>Trade ID for an off-market trade. A natural number</td></tr>
<tr><td><executeTimeStamp></td><td>TimeStamp for the time the trading engine reports an off-market trade as happening (will probably be different to <timeStamp>). HH:MM:SS|H:MM:SS</td></tr>
<tr><td><bidBrokerTraderRefs></td><td>Broker reference and trader reference of the bid order (if they exist).
B()|B(@<brokerID><“ ”>&<traderID>)
|B(@<brokerID>)|B(&<traderID>)</td></tr>
<tr><td><askBrokerTraderRef></td><td>Broker reference and trader reference of the ask order (if they exist).
A()|A(@<brokerID><“ ”>&<traderID>)
|A(@<brokerID>) |A(&<traderID>)</td></tr>
<tr><td><statusText></td><td>Control message state for this security, and whether it was initiated for a particular security (in this case no control type is stated), a security group or all securities. <status>[<“ ”><controlType>]</td></tr>
<tr><td><status></td><td>The status of the security, can be either: Open, Closed or Suspended
Opening: Not open for trading nor order entry, because opening-initiated trading is occurring
PreOpen: Open for order entry but not trading. Adjust-Period
Trading-Halt</td></tr>
<tr><td><controlType></td><td>Whether the control was initiated for a particular security group (“group”), or for all securities (“system”).</td></tr>
</table>

References

Branigan E (1989) Sound and Epistemology in Film. The Journal of Aesthetics and Art Criticism 47:311–324.

Bregman A (1994) Auditory scene analysis: the perceptual organization of sound. The MIT Press, Cambridge, MA

Coronel PRC (2000) Database systems: design, implementation and management, 4th edn. Course Technology (Thomson Learning), Cambridge, MA

Mathews MV (1999) What is loudness? In edited by Cook PR and PRMusic. The MIT Press, Cambridge, MA, pp 71–78

Roederer JG (1973) Introduction to the physics and psychophysics of music. Springer Verlag, New York

Chapter 9
Polymedia Design for Network Metadata Monitoring

Abstract The design of a realtime monitor for an organization's digital network can produce several significant design challenges, both from the technical and human operational perspectives. One challenge is how to capture network data with minimal impact on the network itself. Also, from an operational perspective, sounds need to perform *en suite* over long periods of time while producing only minimal listener fatigue. This chapter describes two related network data sonification projects which resulted in a set of audiovisual "concert" compositions (*Corpo Real*), an immersive installation, and a perceptual monitoring tool (*Netson*). This tool uses both sonification and visualization to present monitoring humans with features of data flow that allow them to experience selectable operational network characteristics. In doing so, it can be used to assist in the peripheral monitoring of a network for improved operational performance.

9.1 Introduction

Data sonification can be simply described as the systematic and reproducible transformation of data into sound for non-linguistic interpretation by listeners.[1] It occupies a space in the aural gamut similar to that occupied by the landscape in visual art: Just as some landscapes are more realistically representational than others, in the sonic arts, there is a gamut between rhetorical, forms such as word-painting, and abstract musical styles[2] and also between physically-embodied gestures on acoustic and virtual instruments, and procedurally-generated soundscapes. Today, data sonification techniques are used in both culturally descriptive and informationally pragmatic ways. So, some sonifications amplify structures in data in perceptually explicit ways, while others distort or disguise these structures in the service of cultural discourse through extra-representational ideals such as musical forms.

[1]Chapter 2 provides an overview of the field, including a more comprehensive definition.
[2]This issue is discussed in more detail in Chap. 1, Sects. 1.5 and 1.6.

D. Worrall, *Sonification Design*, Human–Computer Interaction Series, https://doi.org/10.1007/978-3-030-01497-1_9

The explorations described in this chapter present some approaches to large-scale sonification design that involve Big Data,[3] distributed platforms and audio-visual polymedia design techniques.[4] *Corpo Real* is a digital network perceptualization project that was developed to reveal aspects of the temporal structure of computer network data flows in a large organization. Its development began in 2014 as an Art and Technology project for Fraunhofer-Institut für Integrierte Schaltungen (IIS hereafter) in Erlangen, Germany[5] and included a large mural, some 120 wall-mounted images throughout the Institute (Fig. 9.1) and three audio-visual sonification. The title is metaphoric: a corporation is considered as a living organism disclosed through perceptualized aspects of its communal activity in interior (intranet) and exterior (internet) representations: a play on the idea of revealing the *corpor*ation's corporeal (bodily) existence through the connective neural "tissue" of its digital networks.

NetSon is a perceptual monitoring tool that employs sound and moving image to reveal features of an organization's digital network. It functions on the pragmatic side of the fuzzy boundary between transcendental expression and an analysis and problem-solving tool. Its pragmatic use is to assist in the peripheral monitoring of the activity of individual parts of a network, and of the network as a whole system. It is designed to assist people in understanding the operational characteristics of the network in real time while they are also engaged in other tasks. It can be configured in a number of different ways: as an IT operational support tool, as a multichannel installation in an institution's foyer, as an alternative for a telephone system's "on-hold" music, and as a streaming audiovisual webcast, and in private or public versions for intranet and internet websites. The techniques used are equally applicable to other Big Data monitoring and exploration challenges.

9.1.1 Big Data

There is no accurate account of how many gigabytes of data flow thorough the IIS networks during any 24 h period; even just to count and categorize that data would place such an unacceptable load on the network's operation as to render the process unviable. The monitoring and exploration of such large amounts of data require a new generation of tools and methods to convey abstract information in algorithmically reproducible ways. Furthermore, the numbers of sub-networks that operate

[3]At the time of writing, there seems to be no agreement on whether the term Big Data is capitalized or not. This text errs on the side of capitalization to accommodate other meanings.

[4]*Polymedia* is a descriptive term introduced by the author into artistic practice in 1984. Following the notion of polyphony, a polymedia composition is one in which the simultaneous integration of different media are employed to transmit a polymodal "message". VR (Virtual reality) works are inherently polymedia as they *rely* on the functional integration of two or more of the aural, visual, tactile, kinesthetic, proprioceptive, etc. senses (Worrall 1989, 2003).

[5]https://www.iis.fraunhofer.de/.

Fig. 9.1 Some of the wall images hung during the 18 months of the polymedia exhibition *Corpo Real* (*Photo* Udo Rink, 2015)

in such an environment: some private, some virtual as well as various kinds of connections (cable, WIFI, Bluetooth, satellite etc.) also suggests that another means of monitoring network traffic is required. There are many other challenges of Big Data, as recognized by the European Commission's Research and Innovation activities, including

> …collaborative projects to develop novel data structures, algorithms, methodology, software architectures, optimization methodologies and language understanding technologies for carrying out data analytics, data quality assessment and improvement, prediction and visualization tasks at extremely large scale and with diverse structured and unstructured data.[6]

Techniques such as information perceptualization (principally visualization and sonification) employ human perceptions to take advantage of the broad pathways between the brain and sensory organs to allow users to hear, see, explore, and intuitively understand large amounts of information in as short a period of time, and with the minimum cognitive load possible. Like most mammals, humans have evolved as polymodal perceiving beings. Sight and hearing, for example, are sensitive in different ranges of the frequency spectrum and can be used to attend to spatial and temporal information with different levels of discrimination. So, while our aural perception is clearly superior in differentiating temporal sequencing and frequency discrimination, and better adapted to detecting movement in our surrounding environment, our visual perception has well evolved to creating stable

[6]http://ec.europa.eu/research/participants/portal/desktop/en/opportunities/h2020/topics/9084-ict-16-2015.html.

"scenes" that assist us in understanding the spatial relationships between perceptual objects in our field of vision. The impetus for a polymodal approach to network perceptualization comes from an understanding that while the separation of the senses can be useful for deconstructive analysis, and we still have a lot to learn about how to use them for detecting information in data, they do not work independently, and certainly not in isolation (Merleau-Ponty [1945] 1962).

9.1.2 Data Transfer Protocols: TCP and UDP

In order to design a realtime network-monitoring tool, it is necessary to understand the fluctuations in the rate of data flow between programs and devices. Data flow in a modern digital network is established and maintained via the Transmission Control Protocol (TCP), which is a foundational protocol of the Internet Protocol Suite which specifies how data should be packetized, addressed, transmitted, routed and received.[7] In order to understand the sonification of data flow explored in the projects of this chapter, we first provide a simplified description of TCP. Readers desiring a more detailed description are encouraged to begin with the Wikipedia article.[8]

TCP is a bidirectional connection-oriented protocol, which means a connection between two computers is established and maintained until source and destination (server and client) programs at each end have finished exchanging a message. The protocol determines how to break the message into packets that networks can deliver, sends packets to and accept packets from the network, manage flow control, and—because the protocol is meant to provide error-free data transmission—handle retransmission of missing or garbled packets and the receiver's acknowledgement of all packets that have arrived.

Although each packet in the transmission of a single message will have the same source and destination addresses, packets may be sent via multiple routes. The destination device (client) waits until the packets have arrived, acknowledges those it receives (from the server) and requests the retransmission of any it does not receive within a certain time period. It then sorts and assembles all the packets into the original message. The retransmission and reordering of packets after they arrive introduces latency in a TCP message stream. So, especially in an environment where machines with diverse network speeds communicate with each other, a mechanism for flow control is essential to improve transmission efficiency. The TCP has such control mechanisms, including delaying the transmission of some packets during periods of congestion. To achieve this, it uses a sliding-window end-to-end flow control protocol so as to avoid having the sender send too much data for the receiver to receive and process it reliably in the time

[7]The TCP is defined by the Internet Engineering Task Force, https://www.ietf.org/.

[8]https://en.wikipedia.org/wiki/Transmission_Control_Protocol/.

available. In acknowledging receipt of a "chunk" or segment of a message, a receiver specifies the amount of additional data that it is (next) willing to accept. The sender can send only up to that amount of data before it obtains another acknowledgement from the receiver.

The TCP client-server handshaking process does ensure error-free message transmission, but it is time-costly and necessarily includes more data and potentially slower delivery times. In time-sensitive situations where latency and variations in latency (jitter) caused by the TCP handshaking are likely to be unacceptable, such as streaming video, multichannel audio and voice-over-internet (VoIP), a simpler User Datagram Protocol (UPD) can be employed instead. UPD is mono-directional (sender to receiver only), and each packet is about sixty-percent smaller than a TCP packet. It has no congestion control and does not concern itself with reordering packets or requests for the retransmission of lost packets. In the *Netson* project, the image rendering/display software is controlled via UDP packets, which it receives from the sonification software in realtime, thus ensuring a tightly-integrated audio-visual display with ostensibly no delay between the appearance of the sound and image components of a perceptualized network event.

9.1.3 The Sonipy *Network Flow Meter*

One monitoring tool that IIS network administrators use to study such a large data set is *sflow*, which employs a sampling technique: a data packet or small group of data packets are "plucked" from the stream at a known sampling rate as they pass through a switch.[9] This random collection of packets is "wrapped" with a meta-packet that identifies such things as the time of creation, and the source and destination of the packets (Bejtlich 2013). Because the sampling rate is fixed (but configurable), the time difference between *sflow*-generated metadata packets is the amount of time (in microseconds) between successive sampled packets, thus providing information of the network load: The inverse of the flow-rate of the system at the time of sampling.

The *Sonipy* network flow meter was designed to aurally monitor the rate of flow of message packets as they pass through a network switch.

Flow is represented as a mapping of inter-event time difference to pitch: Fig. 9.2 is a simplified illustration of this process.

The pitch (dark green cell) rises and falls in proportion to duration between message events arriving at the network switch from the sender. A higher pitch indicates an increased flow rate which is determined by a decrease in the arrival times of successive packets. The system might be imagined to be like the time-trace of a single channel of a mixing console's fluctuating digital VU meter in which higher voltage signals are registered as higher vertical peaks. Another metaphor of

---

[9] https://sflow.org/.

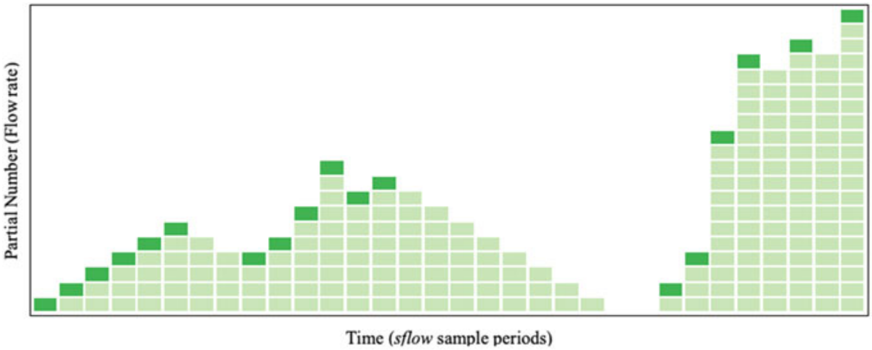


Fig. 9.2 A graphical representation of the *Sonipy* network flow-rate meter in which pitch is governed by time-differences between *sflow*-sampled packets arriving at the network switch

the process is of a charging and discharging capacitor that receives a charge each time a sample (*sflow*) packet is received and plays a tone whose pitch is proportional to the charge on the capacitor. Following the playing of the tone, the charge begins to dissipate such that if the duration of time that elapses until the next packet is received is *greater* that a specified Δt, the next pitch will be lower in proportion to the value of Δt, or higher otherwise. This charge dissipation or "leaking" process is represented in the figure by a light green-colored column without dark green cell at the top. Thus, if the network is unencumbered, the data will flow quickly and the pitch will remain steady or rise. Else, the pitch will rise and fall in proportion to the rate of packet flow. Pauses, delays, and other obstructions to unencumbered flow will cause delays in packet delivery and thus lower pitches. Prior to this project, the nature of this load fluctuation was perceptually unclear to the administrators and it was thought, following earlier work (Gaver et al. 1991), that such a tool might enable those monitoring the network to learn to detect network malfunctions earlier than they otherwise might. The resulting melodic gestures, which can be "tuned" to the flow-rate, appear not unlike a solo instrumental improvisation,[10] so a plucked-string algorithm (Karplus and Strong 1983) was employed.

9.2 The *Corpo Real* Art and Technology Project

Corpo Real is a somewhat elaborate digital network audio-visual perceptualization that was developed to portray aspects of the temporal structure of computer network data flows in a large organization. The aesthetic impulse for the *Corpo Real* animations was decidedly musical, or meta-musical, in that they are abstract, even

[10]When "tuning" these improvisatory gestures, the goal was to seek settings that maximally revealed the digital *Duende* of the network's performance (Lorca 1960).

transcendentally or other-worldly transportative.[11] This approach is contrasted with that of *NetSon* (discussed in Sect. 9.3) which is more pragmatic.[12] However, in what may be considered a feature of the expressive practice, the distinction between musical abstraction and practical representation is not considered essential or categorical because, to paraphrase Dewey, the important question is not "is it music?" but "when is music?" Further discussion of this aesthetic is explored in Chaps. 1 and 4.

The initial *Corpo Real* experiments specifically took a "concert model"[13]; a decidedly artistic approach, that is, unconstrained by the early need to produce useful practical outcomes, at the same time as remaining designerly (Cross 2006). This undertaking provided a greater opportunity than a purely pragmatic approach would have, to discover unusual features of the network data flow; to explore some novel perceptualization approaches and extreme constraints. But also, to learn, as a multilingual interdisciplinary group of artists and engineers who had never worked together before, how to function as an effective team to create better understandings and, hopefully, sounder insights into the information in the data.[14] The importance of an initial period of free-form creative exploration of potential sonification scenarios cannot be over-emphasized as it provides opportunities to discover unusual, potentially useful, features of the dataset before pragmatic design constraints are applied.

The accompanying image software[15] is controlled by sound-rendering parameters which it receives from the sonification software in real-time in a simple UDP format. When a sonification event is initiated, the sonification engine sends a UDP signal to the visualization engine which analyzes the audio stream produced by the audio renderer, generating a basic Fourier transform (frequency separation) of the audio signal which it then uses to control the image generation algorithms.

9.2.1 Short Description: net–path–flow

This study perceptualizes the rate of flow of data across the entire network. As described earlier, pitch and time are used to represent the flow rate: The pitch rises and falls as the duration between network events decreases and increases: its

[11] After Baumgarten ([1783] 2013) and Kant ([1790] 2007).

[12] After Dewey ([1934] 1952), James ([1907] 1995), and Merleau-Ponty (1964).

[13] Nomenclature is often difficult in interdisciplinary fields and sonification research is no exception. The term "concert model" is used here to describe a type of presentation in which the listener is a non-interacting recipient of a presentation (Walker and Kramer 1996).

[14] The image-making and visualization programming were undertaken by Udo Rink and Tebjan Halm.

[15] vvvv—a multipurpose toolkit http://vvvv.org/. Animations for *Corpo real,* created by Tebjan Halm and Udo Rink.

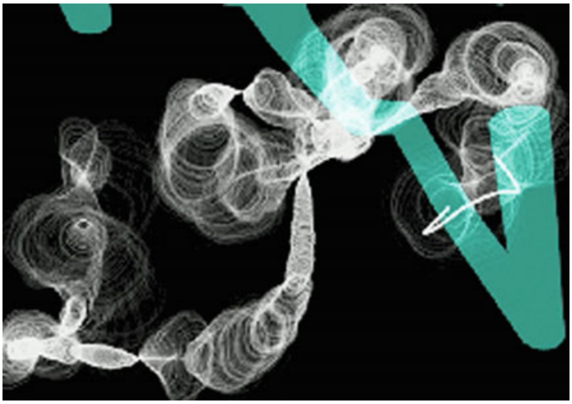


Fig. 9.3 A screen-captured image from the animation *net-flow-path*, the first study of the polymedia composition *Corpo Real* (Rink and Worrall 2015)

melodic contour helps us to hear the structure of the temporal flow of network data which is being temporally-dilated to 100 times slower than real-time, making thus the time jittering aurally perceptible.

The animation contains two ring forms. The network address structures are represented in green and the sound structures in white. The larger polygonal (green) form is chosen as one path through the network topography and the smaller, wandering rings expand and contract in relation to the intensity of the flow. The brush diameter is controlled by the (RMS) sound level, in two simultaneous scales: The brush size become thinner with low intensity, and thicker with high intensity tones (Fig. 9.3).

A spiraling movement restricts the image paths to a canvas slightly larger than that rendered. The spiral was chosen as an allusion to the shape of the cochlea in the inner ear (Rink 2017). The animations can be viewed online and other material associated with the project is available on the author's website.[16]

9.2.2 *Short Description:* in-cooperation

This study is the most abstract of the three. The tight integration of the sound and shimmering image forms highlight the subtle temporal interactions of these primary senses. It perceptualizes the whole of the network at IIS sorted into various departments, sonified and visualized according to tone color (*klangfarbe*) and spatial projection (Fig. 9.4).

[16]http://www.avatar.com.au/netson/.

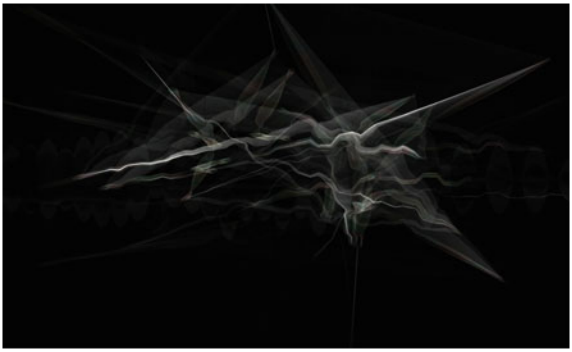


Fig. 9.4 A screen-captured image from the animation *in cooperation*, the second study of the polymedia composition *Corpo Real* (Rink and Worrall 2015)

9.2.3 Short Description: 3am chill

This study is a perceptualization of the network traffic at 3am local time, when the majority of the network traffic consists of the servers transferring existing data (here represented by a plucked string algorithm) together with other less frequent traffic such as email and after-hours international activity. Again, departments and operational activities are represented timbrally (Fig. 9.5).

The aim is to represent the overall status of the network: "all-is-well", or "virus attack" etc. in realtime. On the evening of this recording, the servers are "humming along" in gentle equilibrium. The green boxes and the various tone colors spatialize the flowing data packet groups, projected in space according to their origins, receding into the past through a single vanishing point.

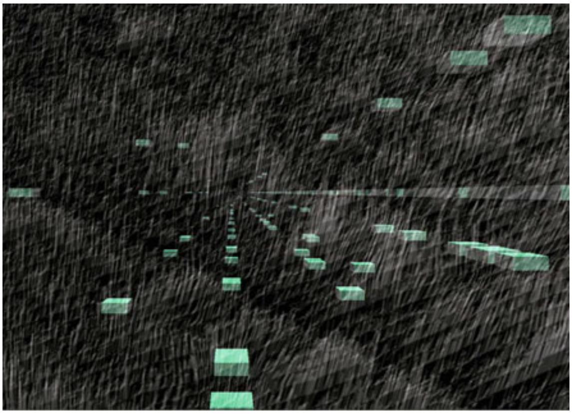


Fig. 9.5 A screen-captured image from the animation *3am chill*, the third study of the polymedia composition *Corpo Real* (Rink and Worrall 2015)

9.3 The *Netson* Network Monitor

The *NetSon* network monitor was designed to represent two kinds of network information: individual and group events (for example, activity in Department X, or on Server P) and overall network processes, such as network load, as previously discussed. The project emerged from the success of the *Corpo Real* project. The open design brief was to try to discover if something useful might emerge from applying the results of investigations made in that project for more pragmatic purposes. This is not the same as asserting that the client did not know what they wanted,[17] rather that they were aware of some of the inadequacies of the network monitoring techniques then in use, and had an open mind to exploring alternative approaches. During the design and construction phases, its potential uses were articulated and refined through iterative consultations. Several unique applications emerged and continued to produce useful understandings about the operation of the particular network under study, as well as potential applications in a wider sphere of operations from the monitoring of inner-city traffic congestion to the operation of sparse-array telescopes.

9.3.1 Review of Related Work

Surveillance, network monitoring and Big Data are currently very topical, and the inadequacy of purely computational approaches are well recognized (Axelsson 1999). Before discussing the details of *Netson*, an overview of prior and subsequent related work is in order.

Studies in the use of auralization and visualization of software behavior for the design, analysis, tuning and documentation of algorithms (Brown 1992) influenced Francioni and Jackson's auralization of parallel program behavior (1993). Gilfix and Couch's *Peep*, represented the operational state of a computer system or network with a sonic environment (2000). Chafe and Leistikow sonically adapted *ping*, the tool commonly used for timing data transmission over the Internet, to sonify the delay times between the send and returned acknowledgement of Internet Protocol (IP) packets to determine network load characteristics (2001). This work also existed as a sound and animation installation (Chafe and Niemeyer 2001). The *Stetho* Network Sonification System generated MIDI events corresponding to network traffic (Kimoto and Ohno 2002). Also aimed at system administrators, *NeMoS* (Network Monitoring with Sound) was designed as a network sonification system in which user-defined network events are associated with MIDI tracks. The system was designed to allow simultaneous monitoring of different parts of a

[17]"If I'd asked customers what they wanted, they would have told me, "A faster horse!" People don't know what they want until you show it to them…". has (apparently incorrectly) been attributed to Henry Ford. See Sherrinton (2003, 33).

potentially large network system, with a single musical flow representing the whole state of the part of the system in which the user was interested (Malandrino et al. 2003). The *SysSon* platform was developed as realtime network sonification engine capable of addressing network traffic data based on network traffic metadata (Rutz et al. 2015). Garcıa-Ruiz, with others, investigated the application of sonification as a educational tool for network intrusion-detection. Sonification prototypes for the mapping of log-registered attacks into sound include animal noises (auditory icons) and piano tones (earcons). Informal testing confirmed earlier findings that the earcons were more easily identifiable, while the auditory icons were superior for specific recall or specific conditions (Garcia-Ruiz et al. 2007; El Seoud et al. 2008; Garcia-Ruiz et al. 2010).[18]

Ballora, with others, explored human aural and visual recognition abilities to detect intrusion patterns in realtime multichannel sound and 3D visualization of network traffic using IP addresses and return codes (2010), and information such as the date and time of exchanges and the sender's and receiver's IP addresses and port numbers (2011). The interactive soundscape *InteNtion* (Interactive Network Sonification) mapped network activity into MIDI messages that are sent to dedicated synthesisers to generate mixed sounds, resulting in an interactive musical soundscape (Giot and Courbe 2012). Wolf and Fiebrink developed a coded interface for sonifying network data. Their *SonNet* extracts information at various levels from individual packets to overall network state information (2013).

The dematerialization of computer networks into WiFi hotspots, the Cloud and the increasingly more powerful "always-on" personal computing devices, has contributed to both their ubiquity and their vulnerability to abuse, such as intrusion and denial-of-service attacks as well as other illegal activities, such as insider-trading in securities, commercial-in-confidence breaches and espionage (Di Pietro and Mancini 2008), not to mention the growing plethora of domestic "smart" devices. While the importance of monitoring network data traffic is well understood, there is no clear consensus about the pattern of such activities and consequently how any but the simplest security breaches might be detected (Jajodia et al. 2010). Not only does the sheer volume of data and metadata passing through large networks create a challenge to effective monitoring, but the addition of data created by the monitoring tools themselves exacerbates the difficulty. A number of shortcomings of existing visualization-only process-monitoring systems include the limited number of visual properties onto which data can be mapped, and the requirement that operators dedicate their full attention to displays in order to ensure that no vital signals are missed. When the processes in question are temporal and/or realtime in nature, this requirement is particularly stressful as it restricts the ability of those monitoring to perform other tasks simultaneously (Hildebrandt and Rinderle-Ma 2015), rendering them prone to produce false positives and false negatives (Axelsson 1999). Flexible tools that afford the presentation of this data in realtime, or near realtime are essential and it has been demonstrated that sonification

[18]This issue is discussed in more detail in Chap. 2, specifically Sect. 2.2.1.5.

can provide a potential solution to the challenges of such circumstances (Hildebrandt et al. 2014; Axon et al. 2016, 2017). Tools with such potential have multiple wider applications, including securities trading, industrial process monitoring (Gaver et al. 1991), and surveillance more generally.

Suggestions that combining sonification with visualization techniques in order to tackle the disadvantages of current monitoring and analysis tools of security data is supported by the work described earlier in this Chapter. However, the assertion that sonification is more suitable for data that changes over time, while visualization is more suitable for spatial information such as network topology maps, and non-time-based data, is an over-simplification. As also demonstrated by the *Netson* system described here, the integrated polymedial use of these two primary senses modalities can provide temporal and spatial support in manner neither of them achieve independently.

Vickers has engaged with multiple others in various network sonification projects, including the detection of unwanted behavior by studying the inherent self-organized criticality[19] in network traffic and situational awareness scenerios such as combatting cyber-attacks (Fairfax et al. 2014). Continuing the situational awareness paradigm, Debashi addresses the question of how sonification can be used in real-time to provide the protocol flow granularity required to understand the network environment as behaviors develop. In the process, with Vickers, he developed *SoNSTAR, a* monitoring *tool* to be used by network administrators for the purpose. Their solution involved identifying certain Internet Protocol (IP) flow types between recognized source and destination IP addresses within defined time periods of time (Debashi and Vickers 2018). A flow event is identified as a change in the behavior or operation of a flow. A single such event represents a combination of features of a flow while a set of events represents flow behavior which, in turn, represents the state of the network traffic. *SoNSTAR* uses such flow events to trigger sound synthesis and control events.

9.4 Technical Overview of the *Netson* System

The operational state of the network being monitored is represented in *Netson* as a polymedia environment. An environment is not simply an abstract space, but, as discussed in Chap. 4, a symbiotic coalition of places where objects are situated and events occur. This section discusses the major operational components of *NetSon.* The subsection numbering follows the order of the red-circled numerical labels in the Fig. 9.6 schematic.

[19]Self-organized criticality (SOC) is a property of dynamical systems that have a critical point as an attractor. In a phase transition, the spatial and/or temporal scale-invariance characteristic of the critical point ensures that there is no need to precisely tune control parameters, because the system effectively tunes itself as it evolves towards criticality.

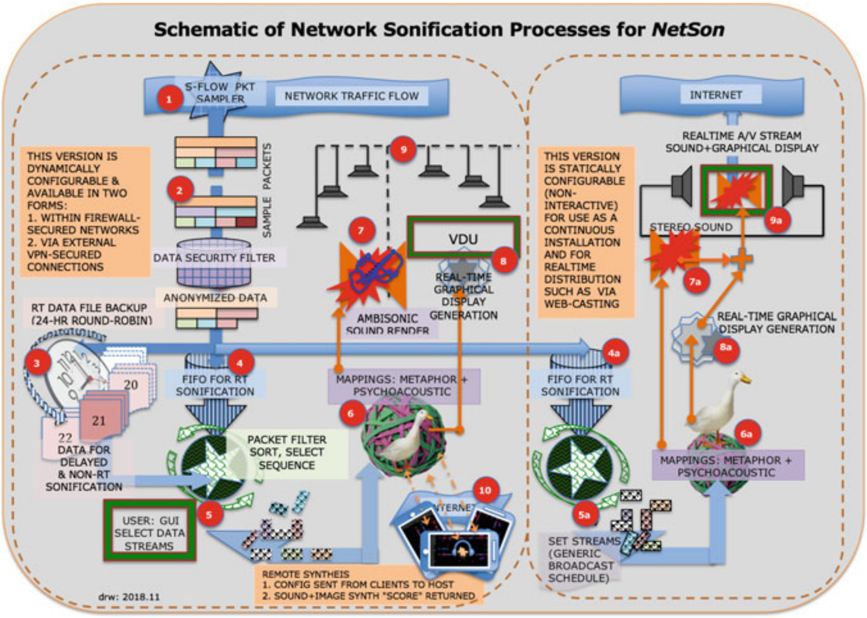


Fig. 9.6 *Netson* operational schematic

Examples of code used to implement this schematic are available from the repository. The system has been extensively tested over several years both in-house and in several installations and demonstrations in Europe, S.E. Asia and USA during which it performed stably and to specification. Some partially-implemented additions to improve both data security and multiple individualized monitoring are discussed at the end of the chapter.

9.4.1 Data Handling

Some major issues in capturing something of the character of modern institutional networks include the sheer volume of data involved, and how to capture this data with minimal impact on the network itself, have already been discussed. The conceptual structure of *sflow* proved to be superior to more traditional alternatives (Fig. 9.6 ①).

9.4.2 Data Security

Exposing any aspect of an organization’s data network to scrutiny, particularly when using virtual (non-fixed IP) access, presents potential security risks, and needs to be undertaken with a great deal of caution, not the least when the organization's viability is dependent on its intellectual property. The general security procedures applied included using fixed IP addresses through secure networks and portals whenever possible, use of encrypted and poll-monitored VPN, and ensuring that the broadest packet discretization necessary for the monitoring task be applied (including, where possible only metadata (e.g. sflow packets). All source and destination IP addresses are stripped of their least significant byte (LSB) before being transferred to the perceptualization system. While anonymizing data impacts on the ability to identify and thus sonify for specific machines and/or locations, it has the benefit of reassuring individuals that their personal activity is not under surveillance. This was an individualized solution to the particular (IIS) implementation, but not necessary to the functional operation of the system. A fast, efficient, proprietary encryption protocol that will enable transmission across public networks with impunity is under development (Fig. 9.6 ②).

9.4.3 Data Storage Repository

Filtered data is backed-up into a 24 h round-robin repository for offline use as required for such activities as further analysis, adjustments to the mapping model, sound rendering experiments and the non-realtime exploration of newly identified features. This repository employs a set of 24 h files, reused daily. An additional external back-up process of these daily file-sets enables *Netson*-independent longer time-frame studies (Fig. 9.6 ③).

9.4.4 Data Filtering

The data is then made available to the sonification systems via (several, if necessary) FIFOs that are read and cleared as required by the individual instantiation (in-house, web, VPN distribution etc.). A configuration file provided, and regularly updated by, the network engineers ensues the address-mapping remains current. The security of this (read-only) file is critical: for maintaining network integrity; to assist in intrusion detection, and in order to re-classify IP addresses for specific tasks. While this config-file structure leads to another layer of abstraction, the advantage is that *Netson* can automatically and instantaneously adjust to any network changes without requiring re-initialization (Fig. 9.6 ④ and 4a).

9.4.5 Data Stream Selection

This individually-instantiated user-interface operation enables flow control and reassignment prior to inverse-psychphysical mapping. For example, if server <=> backup operations require human-monitor cognitive attention, they can be solo- or group-selected here before any perceptualization rendering occurs. Once all groups have been selected, metaphorical mapping routines can be optionally adjusted to maximize their perceptual intelligibility. This type of n-dimensional "zooming" can include parameter-gamut expansion such as pitch-range and greater timbral disambiguation. Such operations would lend themselves to automation in a computational design model (Fig. 9.6 5 and 5a).

9.4.6 Metaphorical Information Mapping

Given that some versions of *Netson* are run in public places, the soundfield needs to be able to support a diversity of distinguishable event types while not being "overtly annoying or distracting" (Vickers 2011). That is, able to be heard, listened to for long periods of time, when necessary, with a minimum of fatigue. For this reason, in contradistinction to much parameter-mapping sonification, "melodic" pitch structures are used very sparingly in favor of a relatively simple, psychophysically speaking, diverse timbral palette (Fig. 9.6 6).

Pschophysical simplicity is not the same as parametric simplicity. A sine tone, for example, often referred to as "a simple sine tone", excites a very small portion of the ear's tectorial membrane and related cochlea nerve fibers, resulting is greater auditory and mental stress than, all other things being equal, the distribution of a "load" across a broad range of frequencies. The types of sounds used underwent considerable longitudinal testing. They were selected so as to combine to form an auditory environment in which the individual sounds can be clearly differentiated by attentional listening. Attentional listening can occur when, supported by a conducive environment, the nature of the sounds affords a listener being able to be only peripherally aware of them. That is, the sounds minimally impose themselves, but are "available" if the listener attends to them. This is discussed in Chap. 4, together with other modes of listening. Attentional listening occurs at a pond where frogs and insects produce a chorus of acoustic communication that is essential for their survival in both defense of territory and in attracting mates. Audio Example 9.1 provides a particularly noisy nighttime example.

In *Netson's* resulting sonic ecology, each sound or sound-class, represents some part of the network operation. So the listener is able to discern singular important events from a large collection of (sometimes subtly different) sounds, at the same time as being able to contextually interpret *changes* in the state of the system as a whole. As discussed in Chap. 4, there is a trade-off between using highly differentiable mono-frequency events which produce high cognitive and psychophysical

Filename
audio 9-1.wav

Audio Example 9.1 An example of the attentional listening. A nighttime pond chorus (Tomakin 2013)

load, and acoustic noises, which, while lower on psychophysical load, need time for them to be disambiguated.

Only a few identifiable "real-world" sounds are used. One is an assemblage of small bells, which are used to identify activity on the printer network. The reason for the use of such bells was simple: Mechanical typewriters rang a right margin-bell to allow the typist to prepare for a new line, and the combination of key-clatter and bell-ringing seemed pleasant enough, given that they were also of a different class of sound than others used. In any event, of the hundreds of people surveyed, even those who had not experienced the historical margin-bell, had no difficulty, once the mapping was revealed to them, in identifying printer network events. Some astute listeners, working in the environs of the sonification, maintained they could identify distinct organizational events such as certain group meetings, by the frequency and timbral clustering of the resulting "carillon". In one instance, I was reliably informed that a colleague would be late for a lunch appointment because they needed to print some documents in preparation for a post-prandial rendezvous. While hearing the bell-signal that a print run had begun, and the subsequent appearance of the individual, is hardly a verifiably causal event, the delight it produced for those "in the know" was one of the simple features that endeared the *Netson* project to its audience.

Whilst a lot of network traffic is between known entities (fileservers and departments, for example) some traffic, such as email, website or repository data arrived from, or was sent to, unidentified addresses. A simple glissando (gliding tone) was incorporated into a load monitor, as described in Sect. 9.1.3 to signify such traffic. Network packets leaving IIS *to* an unknown address, caused the pitch of the tone to rise, and fall, if arriving *from* an unknown address. This metaphorical mapping mirrors the structure of extensor-flexor reflexes of our bodies: larger/higher/further is out/up and vice versa. Extending this metaphor, the distance travelled is indicated by the tone length: the further the distance, the longer the duration of the tone. Audio Example 9.2 provides examples of the mappings discussed.

Filename
audio9-2.wav

Audio Example 9.2 Examples of some of the mapping strategies employed in the *Netson* network monitor

A simple color-coding that shows the source and destination of the data enables *Netson* to display some features of the data visually, others aurally, and still others, both. For example, because a redshift occurs when a light source moves away from an observer and a blueshift occurs when a light source moves towards them, red and blue are used indicate whether data is being received *at* (blue) or sent *from* (red) IIS. By varying the transparency of the simple graphical shapes, it was possible to overlay several of them to create visual clusters of various densities that enhance the aural perception of flow dynamics within the event groups.

The stability of the virtual spatial location of sound sources in sonification is integral to the metaphor within which the technique is used. Informal user-testing indicates that these features are easily learned and remembered over several days, and within a few minutes, users are able to make observations about the activity of the network. It may be that the dilation/compression (red/blue) Doppler-shift metaphor would not work with a less scientifically literate audience. In which case, a blue-sky/setting sun could be preferred. This "sounding objects at fixed locations" approach assists the listener to differentiate the various identities as they appear over time.

9.4.7 Sound Rendering and Playback

The acoustics of audio playback environs can vary considerably–from full bandwidth 3D VR systems to earbuds–and so are not controlled-for within *Netson*. It was anticipated that the ambient aural environment of the rendering was of sufficient quality to support the identification and differentiation of its objects and processes. That being said, it is observable that the more extended a dynamic range an acoustic signal is, the more it commands attention, and so in quiet environments, or in situations demanding extended periods of auditing, a reduced (compressed) dynamic range is conducive to reducing startle reflex reactions and listener fatigue (Fig. 9.6 ⑦).

Sound rendering and psychoacoustic transforms are achieved entirely using the *Sonipy* methods outlined in Chap. 6; distributed between *Csound* instrument configurations (for low-level, high load computation) and the Python *Csound* score generated in realtime by *Python*. Also included in the *Csound* orchestra are various selectable output options. Experiments with stereo and vector-based-amplitude panning (VBAP), though forceful, never produced as convincing an environmental ambience as ambisonic rendering into two, four or eight channels for different venues (Elen 2018). Perhaps this is hardly surprising, given the use of the environment metaphor.

9.4.8 Image Rendering

The image-renderer, *VisSonification,* was written in *Unity3D*[20] by Wolfram Nitsch, the IIS network engineer who also maintained the necessary data security and access controls during its various international excursions, as discussed earlier. An image-rendering of the data being sonified assisted users to easily identify the names and institutional locations of the objects and events being sonified. On receipt of a simple UDP packet from the *Csound* sound-renderer, two transparency-adjustable rectangular "cells" are drawn on the left-hand-side of the image at the appropriate vertical locations: One cell is drawn for the source (shaded blue) and one for the destination (shaded red). As the image is scrolled (left-to-right, speed adjustable), cells may overlap. High-density dataflows thus appear as brighter regions on the display, caused by the overlapping, semi-transparent, cells. Other shadings are possible, as illustrated in Fig. 9.7, where orange is assigned to the printer server network, and cerise to all unidentified internet locations (Fig. 9.6 ⑧ and ⑨).

The sound-to-image UDP slave approach to image rendering has two advantages. Firstly, it places the burden of selecting items to be rendered on the sonification renderer, thus enabling the more computationally intense image rendering process to be undertaken without recourse to graphic enhancement. Secondly, it ensues a perceptually stable connection between the appearance of the sound and image events. Should there be any disruptions to data, there is no danger of audio-visual discrepancies (jitter).

Like an instrument listing in an orchestral score, the object/event keys on the left hand side are maintained in strict order by the rendering software. If there has been no further events for an object when a rendered event scrolls off the right-hand (i.e. the scoreline is now blank), the rendering software drops that scoreline until another related event occurs, leaving visual space for a new/different scoreline to appear. One hears a sequence of events, identifies them through a combination of remembering similar previous aural events, visually locating their identity in the dynamic legend and then observing how such a sequence relates to previous sequences. Because perception involves both short term remembering and anticipation, observation during this monitoring process relies on subtle ways in which sometimes the ears lead the eyes and in others, the reverse.

9.4.9 Internet Streaming and Extensions

A live video stream to the intranet was provided to enable a private version of *NetSon* within IIS. Following the success of the early explorations was made available on the internet website, redacted of information considered a potential

[20]https://unity3d.com/.

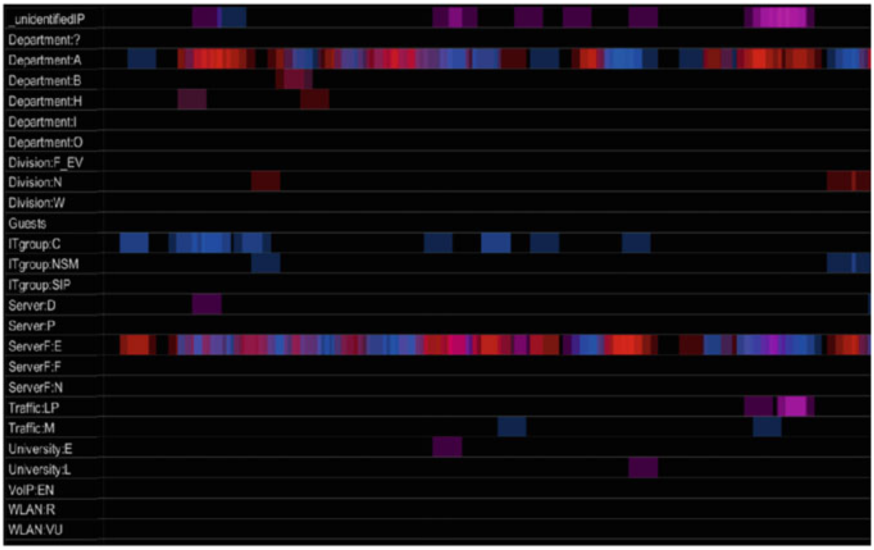


Fig. 9.7 Illustration of the primary mode of the graphical display

security risk. Several video are available which demonstrate the operation and function of *Netson* are available through the book's repository (Fig. 9.6 ⑩).

The power and perceptual simplicity of *Netson* has encouraged users to suggest a range of applications than originally envisaged. The increasing computational power of hand-held devices, and the generic support for extended and virtual reality technologies, including the implementation of browser-based 3D audio, and increasingly more sophisticated user-interfaces, has encouraged the design of an improved multi-user individualized monitoring approach, with both robust data security and improved performance. This will render it more widely useful in both organization and domestic situations (Fig. 9.6 10a).

References

Axelsson S (1999) The base-rate fallacy and its implications for the difficulty of intrusion detection. In: Proceedings of the 6th ACM conference on computer and communications security (ACM 1999), Singapore, pp 1–7

Axon L, Creese S, Goldsmith M, Nurse JRC (2016) Reflecting on the use of sonification for network monitoring, pp 254–261

Axon L, Nurse JRC, Goldsmith M, Creese S (2017) A formalised approach to designing sonification systems for network-security monitoring. Int J Adv Secur 10(1 and 2)

Ballora M, Hall DL (2010) Do you see what I hear: experiments in multi-channel sound and 3D visualization for network monitoring? 7709:77090J–7709-7. https://doi.org/10.1117/12.850319

Ballora M, Giacobe NA, Hall DL (2011) Songs of cyberspace: an update on sonifications of network traffic to support situational awareness. In: Braun JJ (ed) Multisensor, multisource information fusion: architectures, algorithms, and applications, vol 8064. International Society for Optics and Photonics

Baumgarten AG ([1783] 2013) Metaphysics: a critical translation with kant's elucidations, selected notes, and related materials. Translated by C.D. Fugate and J. Hymers. Bloomsbury Publishing, London and New York

Bejtlich R (2013) Practice of network security monitoring: understanding incident detection and response. No Starch Press, San Francisco

Brown M (1992) An introduction to zeus: audiovisualization of some elementary sorting agorithms. In: Proceedings of the ACM CHI 92 human factors in computing systems conference. Addison-Wesley, Monterey, CA, pp 663–664. http://www.acm.org/pubs/articles/proceedings/chi/142750/p663-brown/p663-brown.pdf

Chafe C, Leistikow R (2001) Levels of temporal resolution in sonification of network performance. In: Hiipakka J, Zakharov N, Takala T (eds) Proceedings of eighth international conference on auditory display. Espoo, Finland, pp 50–55

Chafe C, Niemeyer G (2001) Ping sound installation. Documentation. Ping Sound Installation. http://crossfade.walkerart.org/ping/

Cross N (2006) Designerly ways of knowing. Springer, London

Debashi M, Vickers P (2018) Sonification of network traffic flow for monitoring and situational awareness. PlonS ONE 13(4)

Dewey J ([1934] 1952) Art as experience. Perigee Books, The Berkley Publishing Group, New York

Di Pietro R, Mancini LV (eds) (2008) Intrusion detection systems. Springer

Elen R (2018) Ambisonic.Net. Where surround-sound comes to life. https://www.ambisonic.net/

El Seoud S, Garcia-Ruiz M, Edwards A, Aquino-Santos R, Martin M (2008) Auditory display as a tool for teaching network intrusion detection. Edited by D. May. Int J Emerg Technol in Learn (IJET) 3(2):59–62

Francioni J, Jackson JA (1993) Breaking the silence: auralization of parallel program behavior. J Parallel Distrib Comput, June

Garcia-Ruiz MA, Edwards A, Aquino-Santos R, Martin MV, El Seoud SA (2007) Using auditory display to teach network intrusion detection. Kassel University Press: Villach, Austria

Garcia-Ruiz M, Vargas M, Martin MV, Kapralos B, Tashiro J, Acosta-Diaz R (2010) Best practices for applying sonification to support teaching and learning of network intrusion detection. In: Herrington J, Hunter B (eds) Proceedings of the world conference on educational multimedia, hypermedia and telecommunications. Association for the Advancement of Computing in Education: Toronto, Canada, pp 752–757

Gaver WW, Smith RB, O'Shea T (1991) Effective sounds in complex systems: the Arkola simulation. In: Robertson G, Olson G, Olson J (eds) CHI 1991 conference on human factors in computing systems. ACM Press/Addison-Wesley, New Orleans, pp 85–90

Gilfix M, Couch A (2000) Peep (the network auralizer): monitoring your network with sound. In: 2000 LISA XIV, New Orleans, pp 109–117

Giot R, Courbe Y (2012) InteNtion–interactive network sonification. In: Nees MA, Walker BN, Freeman J (eds) Proceedings of the 18th international conference on auditory display (ICAD 2012). Georgia, USA, pp 235–236

Hildebrandt T, Rinderle-Ma S (2015) Server sounds and network noises. IEEE, pp 45–50

Hildebrandt T, Hermann T, Rinderle-Ma S (2014) A sonification system for process monitoring as secondary task. In: 2014 5th IEEE conference on cognitive infocommunications (CogInfoCom), pp 191–196

Jajodia S, LIu P, Swarup V, Wang C (eds) (2010) Cyber situational awareness. Springer

James W ([1907] 1995) Pragmatism: a new name for some old ways of thinking. Dover Publications, Inc.

Kant I ([1790] 2007) Critique of judgement. Edited by N. Walker, Translated by J.C. Meredith. Oxford University Press. http://philosophy.eserver.org/kant/critique-of-judgment.txt

Karplus K, Strong A (1983) Digital synthesis of plucked string and drum timbres. Comput Music J 7(2):43–55

Kimoto M, Ohno H (2002) Design and implementation of stetho network sonification system. In: Proceedings of the 2002 international computer music conference. Goteborg, Sweden, pp 273–279

Lorca FG (1960) The theory and function of the duende. In: Selected poems. Translated by J.L. Gili. Penguin

Malandrino D, Mea D, Negro A, Palmieri G, Scarano V (2003) Nemos: network monitoring with sound. In: Proceedings of the 2003 international conference on auditory display, Boston, USA, pp 251–254

Merleau-Ponty M (1964) The primacy of perception and other essays on phenomenological psychology, the philosophy of art, history and politics. Edited by J. M. Edie, Translated by W. Cobb. Studies in phenomenology and existential philosophy. Northwestern University Press, Evanston

Merleau-Ponty M ([1945] 1962) The phenomenology of perception. Translated by C. Smith. Routledge & Kegan Paul, Oxford, UK

Rink U (2017) Der Algorithmus ist der Pinsel - Abstrakte Kunst durch Programmierung. Brandenburg University of Technology, Brandenburg, Germany

Rink U, Worrall D (2015) “Netson.” Youtube channel. Netson. 2015.https://www.youtube.com/channel/UCDm

Rutz HH, Vogt K, Höldrich R (2015) The sysson platform: a computer music perspective of sonification. In: Vogt K, Andreopoulou A, Gallese V (eds) ICAD 15: proceedings of the 21st international conference on auditory display. Institute of Electronic Music and Acoustics, University of Music and Performing Arts, Graz, Austria, pp 188–196

Sherrinton M (2003) Added value: the alchemy of brand-led growth. Palgrave McMillan

Vickers P (2011) Sonification for process monitoring. In: Hermann T, Hunt A, Neuhoff JG (eds) The sonification handbook. Logos Publishing House, Berlin, Germany, pp 455–491

Walker, Bruce N, Kramer G (1996) Human factors and the acoustic ecology: considerations for multimedia audio design. In: Audio engineering society convention 101. http://www.aes.org/e-lib/browse.cfm?elib=7436

Wolf KE, Fiebrink R (2013) SonNet: a code interface for sonifying computer network data. In: 13th international conference on new interfaces for musical expression (NIME-13). Daejeon + Seoul, Korea Republic, pp 503–506

Worrall DR (1989) A multichannel performance space for presenting experimental polymedia compositions. Chroma J Aust Comput Music Association 1(3):3–6

Worrall DR (2003) Virtual performance: performance effected in virtual space or by non-human performers. In: Scott-Maxwell A, Whiteoak J (eds) Compendium of music and dance in Australia. Currency Press, Sydney, Australia

Index

D. Worrall, *Sonification Design*, Human–Computer Interaction Series, https://doi.org/10.1007/978-3-030-01497-1